JMIR Public Health and Surveillance

Impact Factor (2023): 3.5
Volume 12 (2026) ISSN 2369-2960 Editor in Chief: Travis Sanchez, PhD, MPH

Contents

Viewpoints

Reframing Pandemic Preparedness: Why Positive-Pressure Exposure Control Merits Formal Public Health Evaluation (e90211)
Yusaku Fujii. 14

Reimagining Data for Equity: Reaching the Unreached Map Explorer for Localized Public Health Intelligence in the Western Pacific Region (e93324)
Sumbul Hashmi, Anne-Laure Budts, Vincent Meurrens, Jessica Stow, Fukushi Morishita, Rajendra-Prasad Yadav. 21

Beyond Training: Systems Framework for Sustainable Health Informatics Investment in Africa (e89482)
Eric Ikoona, Lucy Namulemo, Ronald Kaluya, Rebecca Ikoona, Foday Sahr. 30

Integrating AI Into Governmental Public Health Decision Making: Challenges, Considerations, and a Path Forward (e88470)
Elizabeth Campbell, Oluremilekun Oyefolu, Sarah Gillani, Hannah Goodtree, Alison Kelly, Caitlin Rivers, Crystal Watson. 40

Bias-Mitigated AI as a Foundation for Resilient and Effective Health Systems (e88457)
Jarbas Barbosa da Silva Jr, Maureen Birminghamm, Ana Rivière Cinnamond, Eldonna Boisson, Mary Valdez, Sebastian Garcia Saiso, Joao Souza, Myrna Marti, Leah-Marie Richards, Javier Guzman, Jennifer Nelson, Karina Pesce, Ana Haddad, James Fitzgerald, Ernesto Bascolo, Marcelo D'Agostino. 51

The 100-Miles for QVOICES (Queer Voices in Oncology Igniting Change Through Empowered Stories of Survivorship) Campaign: Overview of the Benefits and Costs of Crowdfunding as a Funding Mechanism for Research (e80079)
Viktor Clark, Milena Insalaco, Charles Kamen, Lauren Ghazal. 57

Reviews

Everyday Digital Technology Use and Youth Health: Scoping Review of Longitudinal Studies (e85094)
Preetika Banerjee, Louise Holly. 66

Digital Tools' Effectiveness on Physical Activity Outcomes in Children and Adolescents: Umbrella Review (e75769)
Garden Tabacchi, Antonino Scardina, Antonella Amato, Marta Giardina, Giulia Accardi, Valentina Di Liberto, Giuseppe Giglia, Sonya Vasto, Monica Frinchi, Paolo Boffetta, Walter Mazzucco, Marianna Bellafiore. 84

XSL•FO
RenderX

The Effectiveness of Contact Tracing to Reduce Transmission of Infectious Diseases During Epidemic or Pandemic Response: Rapid Systematic Review (e84805)
Michael Nunns, Samantha Febrey, Kieran Becker, Morgan Weiland, Jill Buckland, Rebecca Abbott, Rebecca Whear, Alison Bethel, Liz Shaw, Kate Boddy, Serena Carville, Tamsyn Harris, Jo Thompson Coon, G Melendez-Torres. 112

Original Papers

Examining the Factor Structure of Objective Health Literacy and Numeracy Scales: Large-Scale Cross-Sectional Study (e71701)
Chihiro Moriishi, Keisuke Takano, Takeyuki Oba, Naoki Konishi, Kentaro Katahira, Kenta Kimura. 133

Data Completeness and Acceptability of Annual Feedback Reports and Tuberculosis Surveillance in Spain, 2018-2024: Mixed Methods Analysis (e93719)
Allegra Chatterjee, Alvaro Roy, Rocio Amillategui, María Sastre, Christine Schwarz, Rosa Cano, Laura Herrera-León, Zaida Herrador, Spanish Working Group on Tuberculosis. 148

Evaluating Tuberculosis Surveillance Using Global Standards and Benchmarks in the Philippines: Mixed Methods Study (e77058)
Minal Ahson, Donna Gaviola, Stephanie O'Connor, Andro Gutierrez, Marvin Apas, Maria Eguia, Devon Pacial, Arianne Zamora, Romel Lacson, Alethea De Guzman, Patrick Moonan. 160

Global Heated Tobacco Product User Estimates, 2014-2024: Descriptive Surveillance Study Using Manufacturer Disclosures (e88761)
Giorgi Mzhavanadze, Gerry Stimson, Tomasz Jerzy ski. 167

Systematic Assessment of Flavor Cues and Additives in Cigarettes and Heated Tobacco Products in Korea: Cross-Sectional Surveillance Study (e87537)
Geon Heo, Doyoon Kim, Yehyun Kim, Ara Kim, Nagyeong Cho, Naeun Kang, Jungmi Park, Eunsil Cheon, Seulgi Kim, Susan Park, Sungmin Park, Sung-il Cho, Heewon Kang. 182

Epidemiological Shifts and Trends From 2009 to 2023 in Hospital Admissions for Ruptured Aortic Aneurysms and Aortic Dissections in Austria: Retrospective, Population-Wide Study (e81702)
Fadi Taher, Maria Leinweber, Afshin Assadian, Amun Hofmann. 201

Twenty-Five Years of Firearm Homicides and Suicides Among US Children and Young Adults, 1999 to 2024: Intersectional Surveillance Analysis (e96931)
Chuka Emezue, Tipparat Udmuangpia. 212

Temporal Trends in Type 2 Diabetes Burden and Care Outcomes in Taiwan From 2012 to 2021: Nationwide Population-Based Cohort Study (e86356)
Ming-Yen Lin, Yi-Ling Wu, Hsiaoting Wu, Chih-Cheng Hsu, iH3 Research Group. 231

Broadband Access and Ophthalmologist Density in the United States: Cross-Sectional Questionnaire Study (e88473)
Patrick Akarapimand, John Lin, Ingrid Scott, Sally Baxter. 269

Digital Health Center App for Community and National Malaria Surveillance in Cambodia: Implementation Case Study (e85881)
Pengby Ngor, Huy Rekol, Eng Thyda, Ry Rith, Sok Kimleng, Yoem Rattana, Ou Vunsokserey, Hem Vanna, Chang Pumeoung, Pascal Ringwald, Richard Maude, Lisa White, Siv Sovannaroth. 279

XSL•FO
RenderX

Health Care Providers' Perspectives on Early Warning Systems for Acute Respiratory Infections in Canada: Qualitative Study (e85244)
Jinfan Qiang, Kevin Maruthananth, Maaha Farrukh, Andrew Pinto, Banafshe Hosseini. 300

Requirements and Implementation Challenges of Automated Outbreak Detection Systems in the European Union: Mixed Methods Study (e96576)
Paulina vom Felde genannt Imbusch, Ann Vietor, Inessa Markus, Michaela Diercke, Alexander Ullrich. 313

The Barcelona Injury Surveillance System (BISS) for Safer Cities: Observational, Descriptive Study Using Routine Health and Police Information Databases (e82079)
Katherine Pérez, Elena Santamariña-Rubio, Mònica Cortés-Albaladejo, Lucía Artazcoz, Adnan Hyder, Carme Borrell. 324

Establishing One Health Surveillance Platform for Electronic Integrated Disease Surveillance and Response in Malawi: Action Design Research Study (e72029)
Tsung-Shu Wu, Matthew Mvula, Edward Chado, Blessings Kamanga, Daniel Mapemba, Rajab Billy, Louis Nyirongo, Annie Mwale, Evelyn Banda, Matthew Kagoli, Tiwonge Manda, Gunnar Bjune, Jens Kaasbøll. 383

Interactive Digital Visualization Counseling for Lifestyle Change in Patients at Risk of Cardiovascular Diseases: Randomized Controlled Trial (e83488)
Adrijana Svenšek, Mateja Lorber, Zalika Klemenc-Ketiš, Lucija Gosak, Dominika Muršec, Gregor Štiglic. 401

Using Natural Language Processing to Examine State Child Maltreatment Policies and Associations With Outcomes: Multistate Cross-Sectional Study (e99643)
Zhidi Luo, Richard Epstein, Nethra Sambamoorthi, Lutfiyya Muhammad, Neil Jordan. 418

Effectiveness of Online Training on Cardiopulmonary Resuscitation and Automated External Defibrillator Use in Simulations With Rural Populations: Quasi-Experimental Community Intervention Study (e80359)
Elena Taverna-Llauradó, Sara Martinez-Torres, Alba Roca-Biosca, Meritxell Pallejà-Millán, Francisco Martín-Luján, Cristina Rey-Reñones. . . 4 3 5

Occupational Fatigue and Multidimensional Traffic Risk Outcomes Among Motorcycle-Based Food Delivery Workers: Cross-Sectional Study (e92667)
Sookyung Kim, Chang Park, Mi-So Shim. 445

Population-Level Impact of the Enterovirus A71 Vaccination Program on Hand, Foot, and Mouth Disease: Ecological Time-Series Study (e85604)
Ye Tong, Huan Fan, Jing Wang, Xinru Zeng, Xiaoqing Cheng, Changjun Bao, Liguo Zhu, Hong Ji, Xiang Huo. 459

The Impact of US Food and Drug Administration's Advisory and Enforcement Actions on US Sales of Esco Bars Products in 2023: Synthetic Control Approach (e81033)
Hyungsik Shin, Jonathan Hannings, Zachary Cahn, Xin Xu, Robin Toblin, Matthew Farrelly, Brian King. 468

Uncovering the Reasons Behind Maternal Care Dropout in Bangladesh: Cross-Sectional Study (e85875)
Syeda Alam, Plabon Sarkar, M Rifat, Sumaiya Jahan, Rokibul Islam, Israt Jahan, Sanjib Saha. 480

Exploring Motives for Reducing Alcohol Consumption Among Users of an Alcohol Reduction App: Content Analysis (e88992)
Lucille Seppi, Olivia Maynard, Dimitra Kale, John Cunningham, Claire Garnett. 495

Mediation Analysis of the Relationship Between Health Literacy and the French General Population's Opinions on Hepatitis B Vaccination: Representative Cross-Sectional Survey of the SLAVACO Project (e82496)
Bakary Cissé, Sylvie Boyer, Jeremy Ward, Julien Mancini. 509

XSL•FO
RenderX

Association Between Gestational Blood Pressure Trajectories and Postpartum Normotension Recovery in Hypertensive Disorders: Retrospective Cohort Study (e89295)
Zhijiang Liang, Huili Wei, Xiaojun Xu, Rong Xu, Xinyue Yang, Jinlian Dong, Xiaoyan Fan, Yukun Chen, Meiling Feng, Xin Zhou, Yijun Cai, Lijuan Lv. 525

Assessment of Knowledge, Attitudes, and Behaviors Toward Others' Mental Disorders to Improve Mental Health First Aid: Cross-Sectional Study in a French Workplace (e93207)
Matthieu Lebrat, Florence Carrouel, Benjamin du Sartz de Vigneulles, Michel Lamure, Corélie Salque, Roger Salamon, Catherine Massoubre, Nicolas Franck, Claude Dussart. 541

Psychosocial Correlates of Sedentary Behavior Reduction Readiness Relevant to Zero-Time Exercise Among Chinese Adults in Sedentary Occupations: Cross-Sectional Study (e95707)
Ziyan Xiong, Xiangyuan Li, Lu Chai, Xiongqiang Duan, Xiang Hu, Yuan Xie, Li Cong. 560

Clinical Epidemiology of Cancer in People Living With HIV in Germany: Retrospective, Observational, Multicenter Federated Claims Data Analysis (e81092)
Bastian Reiter, Stefanie Andreas, Linnea Schumann, Bernd Bender, Isabel Schnorr, Andreas Heidenreich, Christoph Stephan, Andrea Prunotto, Andrea Laukhuf, Julius Wehrle, Matthias Müller, Fabio Aubele, Marlien Hagedorn, Fady Albashiti, Ulrich Seybold, Julia Roider, Melanie Stecher, Daniel Maier, Jörg Vehreschild. 577

Acceptability and Feasibility of Longitudinal Sampling for Sexually Transmitted Enteric Infections in Gay, Bisexual, and Other Men Who Have Sex With Men (GBMSM): Prospective Cohort Pilot Study Conducted in 2022 in South East England (e73762)
Holly Fountain, Katie Thorley, David Reid, Dana Ogaz, Daniel Richardson, Hannah Charles, Kate Baker, Claire Jenkins, Caisey Pulford, Gwenda Hughes, Nigel Field, Catherine Mercer, Xavier Didelot, Noel McCarthy, Hamish Mohammed, Holly Mitchell. 589

Moderating Role of Condom-Use Inertia on the Association Between Status Quo Bias and Pre-Exposure Prophylaxis Resistance Intention Among Chinese Men Who Have Sex With Men: Cross-Sectional Study (e88806)
Min Zhao, Shanshan Li, Xiang Zhao, Jincheng Hu, Zhuoheng Yin, Fanpu Ji, Weiming Tang, Lei Zhang. 606

Stigma, Food Insecurity, and Limited Social Support as Psychosocial Correlates of Antiretroviral Therapy Adherence Among Pregnant Women Living With HIV: Cross-Sectional Study (e84914)
Setor Sorkpor, Ibrahim Yigit, Rachel Thompson, Brittany Lane, Geoffrey Anguyo, Jerry Ouner. 622

Preferences for mHealth Features to Support Engagement in the HIV Preexposure Prophylaxis Cascade Among Men Who Have Sex With Men in Peru: Cross-Sectional Online Survey (e84982)
Jorge Gallardo-Cartagena, Dora German-Quiñones, Kelika Konda, Susan Buchbinder, Jorge Sanchez, Frederick Altice. 634

Profiles of Willingness to Use Pre-Exposure Prophylaxis Modalities and an HIV Vaccine Among Sexual and Gender Minority Individuals in Brazil, Mexico, and Peru: Cross-Sectional Online Survey (e75753)
Jazmin Qquellon, Kelika Konda, Oliver Elorreaga, Hamid Vega-Ramirez, Centli Guillén-Díaz-Barriga, Dulce Díaz-Sosa, Brenda Hoagland, Juan Guanira, Marcos Benedetti, Cristina Pimenta, Beatriz Grinsztejn, Carlos Caceres, Valdilea Veloso, Thiago Torres. 654

Impact of a Culturally Tailored Educational Intervention on Human Papillomavirus Knowledge and Vaccination Intentions Among Women in Saudi Arabia: Quasi-Experimental Study (e104529)
Ebtesam Elsayed, Fathia Mersal, Heba Mohamed, Ala Aljehani, Najla Aljehani. 669

Prevalence, Trends, and Associated Factors of HIV, Syphilis, and Anti–Hepatitis C Virus Seropositivity Among Attendees at Voluntary Counseling and Testing Clinics in Zhejiang Province, China, From 2022 to 2024: Retrospective Quantitative Study (e96584)
Jinlei Zheng, Lin Chen, Lin He, Tingting Jiang, Hui Wang, Wanjun Chen, Yapin Yao. 688

XSL•FO
RenderX

A Critical Health Literacy Podcast to Counter Health Misinformation at Scale: Randomized Controlled Trial (e78003)
Vanesa Mora Ringle, Amanda Jensen-Doss. 702

Legal Infoveillance of Unlicensed Medical Practices in South Korea Through Criminal Court Decisions Using Machine Learning: Retrospective Observational Study (e92591)
Min Kim, Min Choi. 717

Encouraging Adults at Risk for Type 2 Diabetes to Enroll in Diabetes Prevention Programs Through a Media Campaign in Hawai'i: Cross-Sectional Study (e90880)
Kara Saiki, Eunjung Lim, Blythe Nett, Gail Ogawa, Lance Ching, Riana Kawasaki, Nianest Barreto, Katherine Inoue, Lola Irvin, Kiana Kaupiko, L Keliikoa, Meghan McGurk. 728

The Impact on Audience Engagement of Coordinating a Public Health Campaign on Antimicrobial Resistance Through a Network of Health Content Creators: Longitudinal Observational Study (e86587)
Fangyue Chen, Jack Cooper, Amish Acharya, Simon Dryden, Ara Darzi, Kate Grailey. 747

A Social Media Campaign to Promote COVID-19 Vaccination: Cost-Effectiveness Analysis (e84540)
Michael Long, Jeffrey Bingenheimer, Khadidiatou Ndiaye, Dante Donati, Nandan Rao, Selinam Akaba, Sohail Agha, William Evans. 757

Physical Activity Indicators Among Children and Adolescents in Lebanon, Qatar, and the United Arab Emirates: Comparative Synthesis of Active Healthy Kids Report Card Data From 1998 to 2022 (e85998)
Ameneh Baghestani, Lina Majed, Merilyn Lock, Ali Alrahma, Patrick Abi Nader, Suzan Sayegh, Abdulla Al-Mohannadi, Javaid Nauman, Salomé Aubert, Mark Tremblay, Tom Loney. 767

Depressive and Anxiety Symptoms and Their Comorbidity Across Industries: Multi-Industry Cross-Sectional Study in China (e91782)
Lei Cao, Ruihong Ran, Huadong Zhang, Heling Bao. 782

Associations of Health Care System Deficiency–Related Patient Deterioration With Burnout, Depressive Symptoms, and Sleep Problems Among Emergency Medical Service Providers: Cross-Sectional Study (e90130)
Min Choi, Soo Kim, Yonghwan Moon, Hyekyung Woo. 797

Impact of Different COVID-19 Pandemic Phases on Emergency Department Length of Stay and Intensive Care Unit Admission Risk: Single-Center Retrospective Cohort Study (e87597)
Zehua Wang, Xiangmei Zhao. 811

Genetic Subtypes, Pretreatment Drug Resistance, and Associated Factors Among Individuals Newly Diagnosed With HIV-1 in Zhejiang Province From 2022 to 2024: Cross-Sectional Study (e96139)
Wei Cheng, Qin Fan, Jiafeng Zhang, Jun Jiang, Jiezhe Yang, Yun Xu, Lin Chen, Chengliang Chai, Xiaohong Pan. 823

Crimean-Congo Hemorrhagic Fever Disease Awareness Scale: Scale Development, Validity, and Reliability Study (e93724)
Gülay Yilmazel, Ay e Çalmaz, Emre Kele 837

Antimicrobial Resistance Across Four Iraqi Governorates: A Multicenter Cross-Sectional Surveillance Analysis With Implications for Antimicrobial Stewardship Policy and WHO AWaRe–Aligned Public Health Strategies (e94436)
Hemn Omer, Omeed Darweesh, Gulbahar Karim, Baref Rashid, Rozhgar Ahmed, Andrew Seaton, Brian Godman, Amanj Kurdi. 858

Adolescent Self-Reported Recovery for Substance Use in Illinois: Statewide Representative Epidemiological Study (e82792)
Douglas Smith, Crystal Reinhart, Shahana Begum, Tarjani Bhatt. 875

XSL•FO
RenderX

Determinants of Willingness to Receive Health Information From Neighborhood Food and Beauty Establishments: Cross-Sectional Study (e86435)
Seema Aithal, Huiling Guo, Zoe Hildon, May Lwin, Angela Chow. 883

Willingness to Undergo Human Papillomavirus Testing Among Men Who Have Sex With Men in China Based on the Information-Motivation-Behavioral Skills Model: Online Cross-Sectional Study (e85543)
Xin Liu, Xuelin Yang, Mohan Lyu, Zhenwei Dai, Shu Jing, You Xin, Fei Yu, Shenglan Tang, Xiaoyou Su. 897

The Association Between Family Health and Proactive Health Risk Management With the Mediating Role of Health Literacy: Nationwide Cross-Sectional Study (e73659)
Shangfeng Tang, Yilin Wei, Heng Dong, Kangkang Zhang, Chunying Wang, Jia Song, Hua Qing, Yibo Wu. 913

Prevalence and Associated Factors of Renal Disease in Saudi Residents Attending Primary Health Care Centers in Riyadh, Saudi Arabia: Cross-Sectional Study (e81849)
Saad Alshahrani, Ashraf El-Metwally, Awad Alshahrani, Badr Al-Khateeb, Aljohrah Aldubikhi, Khadijah Angawi, Amani Alharthy, Lubna Alnaim, Amal Zaidan, Raed Aldahash. 928

Trends and Core Competence Shifts in Nurses' Infectious Disease Emergency Response Competence Across COVID-19 Pandemic Phases: Repeated Cross-Sectional Survey and Network Analysis (e83281)
Jiagen Xiang, Qiuxia Liang, Yufei Lu, Meizhen Lin, Juan Liu, Hongman Li, Jingjing Cai, Yan Wu, Bingbing Li, Wenting Liu, Ming Yang, Xiaoning Sun, Yan Luo, Yibing Tan, Hairun Liu, Jiaying Li, Liming Lu, Zengjie Ye. 942

Gestational Diabetes Mellitus and Associated Factors in Rural Areas of Northern Vietnam: Cross-Sectional Survey (e79688)
Duc Le, Thi Nguyen, Thi Nguyen, Thi Tran, Xuan Luong, Thi Do, Xuan Tran, Van Ngo, Thi Bui, Tien Nguyen, Van Hoang, Thanh Vu. 954

Factors Affecting Acceptance of Life Education in Mainland China: National Cross-Sectional Study (e78844)
Xinyue Zhang, Nana Huang, Yuyang Yi, Yinlin Wang, Wai-Kit Ming, Yibo Wu, Xiaohong Ning, Chen Wang. 965

Prevalence and Factors Associated With Acute Stress Disorder Among Adults Ever Infected With COVID-19 During the Ending Phase of the Pandemic in 7 Chinese Cities: Cross-Sectional Study (e73002)
Ziying Yang, Yanqiu Yu, Hui Lu, Xu Wang, Yong Xu, Junqiang Ying, Xianying Wen, Lei Luo, Meng Wang, Muwen Liu, Xingyi Geng, Xuchong Zhao, Biyu He, Tao Liu, Remina Maimaitijiang, Jing Gu, Joseph Lau. 991

Regional Disparities in Factors Associated With Subjective Health Among Older Adults in Aging and Super-Aged Areas of Korea: Nationwide Cross-Sectional Study (e80189)
Gi-Moon Nam, Ye-Rin Lee, Ji-Hye Mun, Su-Hyeun Cho, Hee-Sook Lim. 1008

HIV Risk Profiles and Pre-Exposure Prophylaxis and Postexposure Prophylaxis Uptake Among Men Who Have Sex With Men in China: Cross-Sectional Latent Class Analysis (e81405)
Ziwei Wu, Xinrui Zhang, Xingliang Zhang, Junfang Xu. 1023

Association Between Health Literacy and Prehypertension in South Korean Adults: Cross-Sectional Study Using the 2023 Korea National Health and Nutrition Examination Survey (e82684)
Jiyeon Chun, Dan Kim, Suk-Yong Jang, Eun-Cheol Park. 1041

The Gendered Mediation Effects of Social Support on Fertility Intentions Among Childless Adults of Reproductive Age in China: National Cross-Sectional Study (e77484)
Xinyu Xu, Peiyu Liu, Peihua Ren, Wenyan Shi, Sha Lai. 1059

Survey Modalities and COVID-19 Vaccine Uptake in Vietnamese Americans: Cross-Sectional Study (e77520)
Celine Nguyen, Alexander Le, Paul Yeh, Ben King, Angelica Nguyen, Jannette Diep, Bich-May Nguyen. 1075

XSL•FO
RenderX

Knowledge, Attitudes, Practices, and Vaccination Willingness Toward Mpox (Monkeypox) Among Chinese Medical Students: Cross-Sectional Study (e86981)
Yang Liu, Yuehui Jia, Honglong Li, Jie Ge, Yunfeng Han, Zhiping Xie, Jiaxin Chen. 1088

Digital Exclusion Among People Experiencing Homelessness and Residents of Urban Communities in Brazil: Cross-Sectional Study (e77124)
Ariela Tártaro, Dulce Gomes, Thaís Berra, Reginaldo Tavares, Yan Alves, Letícia Ferezin, Antônio Ramos, Nathalia Zini, Maria Pelodan, Marcela Popolin, Ricardo Arcêncio. 1100

Drug-Facilitated Sexual Assault Pornography and Sexual Violence While Partying: Cross-Sectional Study (e80110)
Pablo Prego-Meleiro, Guadalupe Pastor-Moreno, Irantzu Recalde-Esnoz, Luis Sordo. 1113

Problematic Alcohol Use Among Adolescents in Germany: Representative Cross-Sectional Study (e87800)
Rebekka Schröder, Tim Hamer, Ralf Suhr, Lars König. 1125

Association of Vitamin D and Weight Status With Neurodevelopmental Outcomes in a Large Pediatric Population: Cross-Sectional Study (e89756)
Ying Shen, Meiying Gao, Lejing Guan, Lidan Sun, Yanfen Xu, Bethel Daniel, Kyla Zhang, Guannan Bai. 1139

Social Networks and Their Influence on the Choice of Unassisted Smoking Cessation: Cross-Sectional Study in Six Cities in China (e74147)
Sihui Peng, Tingzhong Yang, Lijing Li, Weifang Zhang, Randall Cottrell. 1151

Urinary Sodium Excretion and the Risk of Prevalent Anemia: Nationwide Population-Based Cross-Sectional Study (e88408)
Sang Suh, Dohyeon Lee, Seong Ma, Sunyong Yoo, Soo Kim. 1167

Long-Term Exposure to Particulate Matter 2.5 and Ozone and the Risk of Acute Respiratory Infections: Community-Based Prospective Cohort Study (e88045)
Xiaolei Duan, Xiao Yu, Hongyu Liang, Zexuan Wen, Wenyong Zhou, Ying Ma, Weibing Wang, Yaxu Zheng, Huanyu Wu. 1179

Longitudinal Population-Level Shifts in Infectious Disease–Specific Health Literacy During and After the COVID-19 Pandemic in Zhejiang, China: Six-Round Cross-Sectional Study (e89958)
Zhao Yusui, Dingming Yao, Yue Xu, Xiujing Hu, Qingqing Wu, Lei Wang, Xuehai Zhang, Zhen Jiang. 1192

Socioeconomic and Regional Disparities in Long-Term Mortality Among Intensive Care Unit Survivors in South Korea: Nationwide Cohort Study (e89127)
Nayeon Choi, Jaehoon Oh, Kyung Yoo, Juncheol Lee, Sang Lee. 1209

Triglyceride-Glucose Index and Risk of End-Stage Kidney Disease in Young Adults: National Population-Based Cohort Study (e73750)
Jin Kim, Yu Lee, Sungjin Chung, Hyeon Hwang, Kyungdo Han, Hye Yoon. 1226

Alcohol Consumption and Age-Specific Risk of Esophageal Cancer: Prospective Cohort Study (e92949)
Yongbo Yang, Xing Xing, Zhen Qin, Chunyang Han, He Zhu. 1245

Determining the Association Between Hearing Disability and Injury Risk in Older Adults Using Propensity Score Matching: Quasi-Experimental Study (e78826)
Woo-Ri Lee, Sungyoun Chun, Youyoung An, Hyun Choi. 1259

Association Between Low Chinese Visceral Adiposity Index Values and Chronic Lung Disease in Older Adults: Cohort Study (e78627)
Xiaorong Zhu, Hui Tian, Zhi Liu, Zhoubin Huang, Miaomiao Yuan, Zhangshun Tu, Kaiwang Cui, Jianping Liu, Xiangwen Gong. 1274

XSL•FO
RenderX

Postpandemic Change in Demographic and Clinical Features of Patients With Omicron Who Were Hospitalized: Territory-Wide Retrospective Repeated Cross-Sectional Study in Hong Kong (e75635)
Christie Ching, Sunny Chan, Teddy Lee, Hugo Pui, Bosco Leung, Man Wong, Tafu Yamamoto, Chak Tong, Cantian Wang, Timothy Rainer, Abraham Wai. 1285

Developmental Trajectories of Positive Expectancies of Cannabis Use Effects Among Early Adolescents: Longitudinal Observational Study Using Latent Class Growth Analysis (e85652)
Weisiyu Qin, Dong-Chul Seo, Wura Jacobs, Sijia Huang, Kit Elam. 1306

Longitudinal Analysis of Variations in Daily Step Counts and Long-Term Implications of COVID-19 Waves and Restriction Phases in Qatar's Step Into Health Program: Mixed Methods Study (e76860)
Lina Majed, Suzan Sayegh, Feriel Dalansi, Abdulla Al-Mohannadi, Marco Cardinale, Abdulaziz Farooq. 1327

Long-Term Survival and Life-Sustaining Device Use in Survivors of First and Second Out-of-Hospital Cardiac Arrest: Retrospective Cohort Study (e90416)
Chih-Hsien Lin, Cheng-Yi Fan, Edward Huang, Chih-Wei Sung. 1343

Real-World Mapping of Multiple Primary Carcinoma Combinations and Survival Outcomes in Shanghai, China: Retrospective Registry-Based Study (e82355)
Jianwei Shi, Yichen Chen, Chen Chen, Xiaopan Li, Hua Jin. 1354

Epidemiological Distribution Characteristics of Tuberculosis Among Older Adults in Chongqing (2020-2024): Spatial-Temporal Analysis (e89671)
Bojie Gao, Yu Xin, Wenping Liao, Lin Shi, Chengguo Wu, Jun Fan, Qingya Wang, Shanrong Huang, Xinyuan Yi, Yong Li, Wen Zhang, Chuan Pu. 1367

The Evolving Burden of Stroke in China's 832 Poverty-Alleviated Counties (2019-2024): Nationwide Spatiotemporal Analysis (e91487)
Mengdi Chen, Dong Xia, Wenping Li, Jie Wang, Zhiyu Lv, Jiapeng Chen, Lulu Zhang. 1381

Detecting Spatiotemporal Patterns of Newly Diagnosed HIV Infection in the China-Myanmar Border Region, 2010 to 2022: Longitudinal Observational Study (e81767)
Jin Yang, Qiyu Zhu, Pei Liu, Jibao Wang, Xing Duan, Yikui Wang, Shijiang Yang, Yuecheng Yang, Runhua Ye, Peng Xu, Song Duan, Cong Jin. 1396

Disparities in Antiemetic Prophylaxis Care Processes Predicted by Patient Neighborhood: Retrospective Cohort and Geospatial Analysis (e69133)
Jiuying Han, Neng Wan, Cameron Jacobson, Nathan Pace, Cade Kartchner, Alexander Hohl, Robert Schonberger, Douglas Colquhoun, Richard Dutton, Michael Andreae, John Pearson. 1409

Risk Factors Associated With Tuberculosis Diagnostic Delay in the Jiangsu Province, China (2011-2021): Spatiotemporal Database Analysis Study (e80052)
Yifan Tang, Cheng Chen, Mingming Chen, Kai Wang, Sifan Wang, Yi Lin, Qiao Liu, Chengxiu Ling, Tenglong Li, Limei Zhu. 1424

The Potential Impact of Federal Funding Cuts on Access to Pre-Exposure Prophylaxis in Atlanta, Georgia: Geographic Modeling Study (e89473)
Noah Mancuso, Patrick Sullivan. 1436

Gestational Hypertension as a Mediator of Prenatal Ozone Exposure and Term Low Birth Weight: Birth Cohort Study (e81412)
Yuxin Zhang, Quanjun Liu, Mengwei Song, Qiulin Huang, Qingbo Zhao, Qingqing Tao, Chenchen Zhang, Qin Li, Qing Wang. 1443

Projected Health and Economic Benefits of Air Quality Targets in China: Modeling Study (e84809)
Siyuan Wang, Zhiwei Xu, Gian Di Tanna, Yawen Jiang, Mingsheng Chen, Laura Downey, Stephen Jan, Lei Si. 1458

XSL•FO
RenderX

Assessing the Effectiveness of Crowdsourced Data to Detect Established Tick Populations in Quebec, Canada: Retrospective Ecological Study (e90734)
Ariane Dumas, Jade Savage, Kirsten Crandall, Jules Koffi, Patrick Leighton, Nicholas Ogden, Erin Rees, Catherine Bouchard. 1471

Prevalence of Early Rheumatic Heart Disease Among Asymptomatic Students in Underserved Communities in Ethiopia: Cross-Sectional Observational Study (e87039)
Amsalu Chuma, Desalew Kassie, Jens-Uwe Voigt, Melkamu Asmare, Ahmed Youssef, Carolina Varon, Rik Willems, Michelle Yates, Adane Sikamo, Yidnekachew Birhan, Chunzhuo Wang, Bart Vanrumste. 1494

Population Estimates and Hypertension and Diabetes Prevalence: Cross-Sectional Quantitative Study Comparing Electronic Health Record–Derived Counts, Census, and Centers for Disease Control and Prevention Population Level Analysis and Community Estimates (e86337)
Tyler Winkelman, Kelly Bergmann, Peter Bodurtha, Alanna Chamberlain, R Dudley, David Haynes, Steven Johnson, Thomas Kottke, Karen Margolis, Gabriela Benitez, Devon Nerstad, Patrick Olson, James Peacock, Nayanjot Rai, Stephen Waring, Bjorn Westgard, Paul Drawz. . . . 1 5 0 7

Early Detection Intervals for Evaluating Event-Based Surveillance System: Reference Dataset Development Study (e87030)
Yannan Shen, Philip AbdelMalik, Russell Steele, David Buckeridge. 1559

Endemic Channel Parametrization in Dengue Surveillance: Methodological Assessment of Retrospective Windows, Outbreak Trends, and Zero-Case Periods in Colombia (e79914)
Juan Umaña, Juan Montenegro-Torres, Julian Otero, Maria Tavera-Cifuentes, Natalia Niño-Machado, Catalina González-Uribe, Juan Cordovez, Mauricio Santos-Vega. 1573

Active Surveillance for COVID-19 Vaccine Safety Using Sequential Analysis in Korea: Population-Based Retrospective Observational Study (e75094)
Na-Young Jeong, Haerin Cho, Heehyun Won, Suvin Park, Joongyub Lee, Hyesook Park, Nam-Kyong Choi. 1585

Multimodal Data Approaches for Examining the 2024-2025 Highly Pathogenic Avian Influenza Outbreak in the United States: Descriptive Study (e86209)
Juliana Sopko, Aimee Han, Jacqueline Powers, Jacquelin Sauer, Mansi Avunoori, Stanislaw Zakrzewski, Allison Krugman, Abhishek Dasgupta, Kara Sewalk, Autumn Gertz, Benjamin Rader, James Sheldon, Brennan Klein, Jessica Rivera, Moritz Kraemer, Samuel Scarpino, John Brownstein. 1601

Predicting Tuberculosis Outcomes Using Routine Surveillance Data in Chiang Mai, Thailand: Retrospective Cohort Study (e86495)
Porramat Saksaen, Ekkarat Boonchieng, Aksara Thongprachum, Surasak Maotheuak, Sineenart Chautrakarn, Waraporn Boonchieng. 1620

Comparison of Measured 24-Hour Urinary Salt Excretion With Spot Urine and 24-Hour Dietary Recall Estimates Among Adolescents and Parents: Cross-Sectional Study (e85549)
Sandeep Kaur, Akashdeep Chauhan, Rajesh Kumar, Manmeet Kaur. 1631

Comparison of In Vitro Metrics With Real-World Risk of Drug-Induced Parkinsonism Due to Antipsychotic Drugs: Retrospective Cohort Study (e81876)
Woo-Taek Lim, Hyun Lee, Seungyeon Kim, Kwangsoo Kim, Yong Ahn, Minseok Hong, Yun Yu, Ha Jang. 1642

Texting to Recruit Unassisted Tobacco Users at a Large Safety Net Health System Into Quitline Service for a Medicaid Value-Based Care Program: Pragmatic Randomized Controlled Trial (e83269)
Cindy Valencia, Melanie Dove, Karen Kim, Paul Giboney, Hal Yee Jr, Carrie Kirby, Christopher Anderson, Shu-Hong Zhu, Elisa Tong. 1656

Multiyear Wastewater Monitoring of High-Risk Substances Surrounding an Annual Electronic Dance Music Festival, Las Vegas, Nevada, USA, 2023-2025: Surveillance Study (e102679)
Daniel Gerrity, Casey Barber, Rebecca Trenholm, Brett Vanderford, Katherine Crank, Ching-Lan Chang, Michael Moshi, Anil Mangla, Cassius Lockett, Edwin Oh. 1665

XSL•FO
RenderX

Performance of an Electronic Sentinel Surveillance System for Bacterial Gastroenteritis: Mixed Methods Evaluation Study (e102373)
Inmyung Song, Yerin Heo, Minsol Jo, Jung-yeon Park, Sukhyun Ryu, Dong-Sook Kim. 1681

Age-Related Patterns in Health Literacy and Health Information–Seeking Behavior in a Korean Population: Cross-Sectional Study (e86992)
Da Kwon, Young Kwon. 1694

Characteristics Influencing Adherence to SARS-CoV-2 Serial Antigen Testing Following a Federal Safety Communication: Cohort Study (e92002)
Micajah McGarity, Catherine Pollack, Elizabeth Currier, Andrew Godoshian, Zachary Smith, Marcus Siddall, Gregg Tavolacci, Reynolds Salerno, Thomas Tsai, Michael Iademarco, Joseph Miller. 1713

Hand, Foot, and Mouth Disease Surveillance in Anhui, After Introduction of the Enterovirus 71 Vaccine and During the COVID-19 Pandemic: Ecological Interrupted Time Series Study (e99663)
Kangguo Li, Jiadong Wu, Wanwan Ma, Tao Chen, Yunzhi Zenghuang, Jia Rui, Zeyu Zhao, Jiabing Wu, Tianmu Chen. 1742

Social Determinants of Childhood Vaccination Coverage in the United States Using National Immunization Survey Data From 2010 to 2023: Cross-Sectional Study (e81746)
Ramya Majji, Yan Zhuang, Felix Pabon-Rodriguez. 1755

Prevalence and Determinants of Blood Exposure Accidents Among Health Care Workers in Kalemie, Democratic Republic of the Congo: Cross-Sectional Study (e97009)
Joe Kabamba Dibwe, Fiston Ilunga Mbayo, Papy Mukalay Muvumbu, Gloire Kasongo Lupitshi, Jean Banze Kabwe Ilunga, Didier Sashila Mukebo, Dieudonné Mpunga. 1770

Linking Women's Empowerment to Mental Health Symptoms: Population-Based Study (e85839)
Maggie Craig, Hannah Walter, Masum Ali, Shahriar Faruque, Sanjib Saha. 1783

Analgesic Prescribing Trends of Opioids, Nonsteroidal Anti-Inflammatory Drugs, and Paracetamol in Taiwan (2001 2020): Age-Period-Cohort Cross-Sectional Study (e92534)
Chun-Ji Lin, Chia-Shuan Chang, Chih-Yi Wu, Jiun-Yi Wang, I-Ching Lin, Hung-Yi Chiou. 1796

Experts' Opinions on the Sustainable Use of Digital Health Tools for Effective Future Pandemic Preparedness and Response: Questionnaire Study (e84164)
Kabelo Mauco, John Holmes, Anthony Luberti, Badisa Mosesane. 1810

Transforming a Pandemic Tool Into a Public Health Asset: Modified Delphi Study (e77763)
Mohsen Malekinejad, Cesar Mendez-Lizarraga, Nicole Skehan, Alyssa Bercasio, Rakhee Palekar, William Lober. 1820

Evaluation of the Finnish Exposure Notification App "Koronavilkku" According to Key Performance Indicators: Cross-Sectional Study (e89472)
Wioleta Kitowska, Mika Pihlajamäki, Sara Wickström, Kaija Tolmunen, Otto Helve, Juha Mykkänen, Timothée Dub, Lotta Siira. 1834

Automated Daily Phone Surveys in Older Adults: Feasibility Study in the Multi-Ethnic Study of Atherosclerosis (e86333)
Sally Kim, Cyanna McGowan, Shaina Alexandria, Grace Achepohl, Lauren Hoffer, Norrina Allen, Mandy Wong, Kiarri Kershaw. 1855

Multidimensional Discrimination Toward Single-Parent Families and Its Association With Depressive Symptoms of Parents: Cross-Sectional Study in South Korea (e83771)
Seong-Uk Baek, Jin-Ha Yoon. 1866

XSL•FO
RenderX

Barriers and Opportunities to Include Underrepresented Population Groups in Vaccine Trials: Cross-Sectional, Observational, Online Survey Study From the VACCELERATE Research Network (e89025)
Mine Tanriover, Dimitrios Poulimeneas, Grammatiki-Christina Tsopela, Ioannis Kopsidas, Christos Argyropoulos, Romina Di Marzo, Olena Valdenmaiier, Stine Jakobsen, Paula Valle Simon, Antonio Carcas-Sansuan, Ole Olesen, Oliver Cornely, Zoi Pana, Theoklis Zaoutis, Murat Akova. 1877

Physical Therapy Provision After COVID-19 in the Netherlands: Retrospective Cohort Study (e88633)
Renee Veldkamp, Peter Spreeuwenberg, Arie Verburg, Thomas Hoogeboom, Rob de Bie, Willemijn Meijer. 1891

Trends in Antimicrobial Consumption in Pakistan (2016 2028): Retrospective Observational Study With Forecasting (e81288)
Zikria Saleem, Brian Godman, Zunaira Akbar, Abdul Haseeb, Muhammad Qamar, Anees Rehman, Mahmoud Elrggal. 1909

Implementation Effect of a Pilot Urban Primary Health Care Model on the Usage of Health Care From Medically Trained Providers for Chronic Diseases Among Slum Populations in Bangladesh: Findings From an Implementation Research Study (e76290)
Md. Rabbani, Zakir Hossain, Khadija Tisha, Kamrun Nahar, Zillur Sakin, Towhida Nasrin, Mohammad Ahmed, Md. Hasan, Chandrasegarar Soloman, Margub Jahangir, Maya Vandenent, Shehrin Mahmood. 1918

COVID-19 Information Sources and Vaccination Status Among Californian Adults by Generation Using the 2022 California Health Interview Survey: Cross-Sectional Study (e85904)
Julia Zabala, Melissa Sablik, Gina Finical, Victoria Keeton, Janice Bell. 1940

Effectiveness of Inactivated COVID-19 Vaccination Against COVID-19–Related Hospitalization and Severe Outcomes in Adults ≥80 Years During Omicron Circulation in Beijing, China: Retrospective Cohort Study (e82915)
Dan Zhao, Ying Ma, Juan Li, Xiaomei Li, Zhiqiang Cao, Wei Yao, Jiang Wu, Luodan Suo. 1953

Investigating the Impact of Online Abortion Myths on Health Care Providers and Advocates: Interview Study (e77050)
Rachel Moran, Julia Simoes, Taylor Agajanian, Izzi Grasso, Anna Beers, Amanda Swarr, Emma Spiro, Anna Swan. 1978

Perceived Difficulty Accessing Primary Health Care and Its Associations With Sociodemographic Factors and Exposure to Violence in a Violent Urban Community in Rio de Janeiro, Brazil: Cross-Sectional Study (e94922)
Ana Cunha, Jurema da Mota, Francisco Bastos. 1991

Factors Influencing Response Rates in National Internet-Based Questionnaires: Analysis of a Large National Survey (e89761)
Søren Birkeland, Anders Haakonsson, Sören Möller. 2006

Examining Artificial Intelligence Chatbots' Responses in Providing Human Papillomavirus Vaccine Information for Young Adults: Qualitative Content Analysis (e79720)
Alfu Laily, Laura Schwab-Reese, Megan Davish, Emily Cahue, Kathryn LaRoche, Natalia Rodriguez, Robert Duncan, Randolph Hubach, Monica Kasting. 2017

Protocols

Adapting the Tobacco Pack Surveillance System Protocol to Assess Electronic Cigarette Packaging: Protocol for a Content Analysis (e89325)
Reiley Hartmuller, Jennifer Brown, Hannah Barker, Qinghua Nian, Kevin Welding, Joanna Cohen, Katherine Smith. 241

XSL•FO
RenderX

Integration of VectorCam, an AI-Enabled Mosquito Surveillance Tool, Into Uganda's National Malaria Elimination Program: Protocol for a Mixed Methods Hybrid Type III Implementation Study (e99717)
Mei-Li Hey, Marina Rincon Torroella, Diwakar Mohan, Sunny Patel, Winnifred Kansiime, Catherine Maiteki-Sebuguzi, Aryaman Shodhan, Youseph Yazdi, Peter Waiswa, Soumyadipta Acharya. 254

Corrigenda and Addendas

Correction: Awareness of Sexual Partner's HIV Status Among Men Who Have Sex With Men in China: Cross-Sectional Survey Study (e101200)
Qian Ma, Tingting Jiang, Wanjun Chen, Shaoqiang Jiang, Jinlei Zheng, Hui Wang, Lin He. 1517

Correction: Cross-Sectional Study on Oral Nicotine Product Sales Trends in Scandinavia From 2018 to 2025 (e94157)
Marina Murphy, Diane Henenberg, Lindsay Reese. 1518

Correction: Diagnostic Testing Preferences in Rural and Vulnerable Populations During a Pandemic: Discrete Choice Experiment (e87714)
Eline van den Broek-Altenburg, Jamie Benson, Yvonne Jonk, Abimbola Leslie, Jan Carney, Gary Stein. 1520

Tutorials

Building Public Health Data Dashboards: Tutorial Playbook (e83157)
Elwin Wu, Raymond Balise, Benjamin Katz, Daniel Harris, Matthew Bullard, Naleef Fareed, Marc Larochelle, Jennifer Villani. 1522

Bayesian Models to Generate Small Area Estimates of Population Health: Tutorial for Using Rate Stabilizing Tools and Their Output (e83498)
David DeLara, Ryan Zomorrodi, Harrison Quick, Joshua Tootoo, Ruiyang Li, Justan Baker, Jihyeon Kwon, Michele Casper, Adam Vaughan. . 1 5 3 9

Research Letters

Cross-Sectional Study on Oral Nicotine Product Sales Trends in Scandinavia From 2018 to 2025 (e85490)
Marina Murphy, Diane Henenberg, Lindsay Reese. 1963

Indications for Medical Cannabis and Their Alignment With US-Wide Disease Prevalences in the US' Largest Medical Cannabis Program: Cross-Sectional Study (e89716)
Sebastian Jugl, Nicole Smolinski, Tansu Aydo an, Almut Winterstein. 1967

Intersecting Sex Work and Substance Use Risk Among Sexual and Gender Minoritized Individuals Recruited Online in San Francisco, California: Survey Results (e83555)
Sean Arayasirikul, Jarett Maycott, Angela Olivares. 1972

XSL•FO
RenderX

Retractions

Retraction: Secular Trends in Gastric and Esophageal Cancer Attributable to Dietary Carcinogens From 1990 to 2019 and Projections Until 2044 in China: Population-Based Study (e101187)
JMIR Publications Editorial Office. 2033

Retraction: Arterial Stiffness and Obesity as Predictors of Diabetes: Longitudinal Cohort Study (e108369)
JMIR Publications Editorial Office. 2034

XSL•FO
RenderX

Reframing Pandemic Preparedness: Why Positive-Pressure Exposure Control Merits Formal Public Health Evaluation

Yusaku Fujii, BE, ME, PhD

Graduate School of Science and Technology, Gunma University, 1-5-1 Tenjin-cho, Kiryu, Gunma, Japan

Corresponding Author:
Yusaku Fujii, BE, ME, PhD
Graduate School of Science and Technology, Gunma University, 1-5-1 Tenjin-cho, Kiryu, Gunma, Japan

Abstract

Powered air-purifying respirators (PAPRs) are respiratory protective devices that actively supply filtered air and maintain positive pressure in the breathing zone, thereby limiting the inward leakage of unfiltered ambient air through gaps. Their principles and technology are already established. Positive-pressure PAPRs have primarily been used by health care workers in high-risk environments. However, their potential as an individual-level infection-control measure applicable to the general public has not been formally positioned for evaluation within the context of pandemic preparedness. This viewpoint aims to demonstrate the rationale for and need to position exposure control using positive-pressure PAPRs designed for the general public as a subject of systematic public health evaluation. Unlike existing measures that depend primarily on behavioral change, spatial separation, or acquisition of immunity, this intervention directly reduces aerosol inhalation exposure in an individual's breathing zone—the physical process that precedes airborne infection. Existing engineering studies increasingly make it possible to consider a public-oriented positive-pressure PAPR not merely as an idea but as a concrete intervention model with the potential for manufacture and implementation. However, the extent to which engineering measures of aerosol protection translate into individual-level prevention of infection and population-level infection control under real-world conditions must be established empirically. Individual-level preventive effectiveness, population-level infection-control effectiveness, engineering performance, feasibility of implementation, sustained use, social acceptability, cost, and equity must therefore be evaluated comprehensively by integrating medical and engineering assessments within a public health framework. This article does not advocate the immediate introduction of positive-pressure PAPRs for the general public. Rather, it identifies an institutional gap: established principles and technologies have not been incorporated into formal public health evaluation as a population-level infection-control measure. Evaluating positive-pressure exposure control for the general public is necessary to determine whether it could broaden policy options in future respiratory infectious disease pandemics and reduce excessive reliance on measures with major societal consequences.

(JMIR Public Health Surveill 2026;12:e90211) doi:10.2196/90211

KEYWORDS

pandemic preparedness; powered air-purifying respirator; PAPR; positive-pressure exposure control; population-level intervention; public health evaluation; privacy-preserving governance

Introduction

Two broad strategies played major roles in the global response to the COVID-19 pandemic [1]. The first combined behavioral interventions, such as mask mandates and physical distancing [2], with mass vaccination intended to establish population immunity [3]. The second was lockdown—the broad restriction of social and economic activity when health systems faced the threat of collapse [4].

Although lockdowns helped curb the spread of infection [4], the COVID-19 pandemic and the accompanying broad societal restrictions placed severe burdens on economies, education, and social life [5,6]. Lessons from the COVID-19 pandemic have reaffirmed the importance of infection surveillance, early-response systems, international cooperation frameworks, and national pandemic preparedness plans [7,8]. At the same time, reconciling epidemic control with the continuation of social activity requires not only the continued improvement of existing measures but also the exploration and evaluation, during nonemergency periods, of new interventions that could complement them.

Existing measures include means of physically reducing inhalation exposure, such as masks. However, measures intended for the general public that actively supply filtered air, maintain positive pressure in the breathing zone, limit the inward leakage of unfiltered ambient air, and control inhalation exposure with a high protection factor have not been established as formal public health options. During respiratory infectious disease pandemics, inhalation of infectious aerosols can be a major route of transmission [9,10]. Because infection is influenced by the infectious dose reaching the respiratory tract, substantially reducing that dose below the threshold could, in theory, prevent infection regardless of the ambient pathogen concentration [11].

XSL•FO
RenderX

A potential third axis is an engineering intervention that directly and physically reduces inhalation exposure. One concrete means is the powered air-purifying respirator (PAPR). A positive-pressure PAPR actively supplies filtered air and maintains positive pressure in the breathing zone, thereby limiting the inward leakage of unfiltered ambient air through gaps. Its principles and technology are already established, and PAPRs have primarily been used in high-risk clinical settings, such as intensive care units and infectious disease isolation environments. However, their potential as an individual-level infection-control measure applicable to the general public has not been formally evaluated as a public health option for pandemic preparedness [12].

The author has named the concept of exposure control using positive-pressure PAPRs for the general public "PAPR for Everyone" (PFE). The concept was proposed at the beginning of the pandemic in 2020 [13], after which the engineering and epidemiological evidence underlying it was accumulated in stages [14-18]. Since 2024, findings concerning manufacturability and ease of implementation, required use rates, respiratory-state estimation, real-time leak detection, and time-managed use have been published in journals including *JMIR Biomedical Engineering* and *Scientific Reports*. These engineering and theoretical studies have brought positive-pressure PAPRs for the general public to a stage at which they can be considered a concrete, testable intervention rather than merely a conceptual proposal.

This article aims to demonstrate the rationale and need for positioning exposure control using positive-pressure PAPRs for the general public as a subject of formal public health evaluation in pandemic preparedness. It summarizes the engineering and theoretical evidence accumulated for PFE and examines why this option has been overlooked. Rather than calling for its immediate introduction, we propose beginning with field studies during seasonal influenza epidemics and systematically evaluating individual-level preventive effectiveness, population-level infection-control effectiveness, engineering performance, feasibility of implementation, sustained use, social acceptability, cost, and equity by integrating medical and engineering assessments within a public health framework.

Current Evidence

Dose-Based Infection Control and the Significance of PAPRs

Infection by airborne pathogens requires inhalation of a sufficient quantity of viable pathogens—an infectious dose [11,19]. Interventions that reduce this dose below the relevant threshold reduce the probability of infection [11,19]. Accordingly, a means of directly and reliably reducing inhalation exposure could, in principle, be a highly powerful infection-control intervention [11,19].

High-efficiency particulate air (HEPA) filters have a collection efficiency of at least 99.9% for 0.3 µm test particles, while the most penetrating particle size (MPPS) varies around 0.3 µm depending on the filter medium, flow rate, and other conditions [20]. A PAPR uses a blower to actively supply air cleaned through a HEPA filter to the breathing zone and prevents the ingress of unfiltered ambient air by maintaining positive pressure in that zone [14,21]. The assigned protection factor (APF) for a helmet- or hood-type PAPR is generally 25; however, an APF of 1000 may be applied to models for which the manufacturer demonstrates performance of at least 1000 based on a workplace protection factor (WPF), simulated workplace protection factor (SWPF), or equivalent test [22]. This article focuses on positive-pressure PAPRs with an APF of 1000 that meet these testing requirements. Under specified conditions of fit and operation, they are expected to reduce the concentration of inhaled aerosols to approximately one-thousandth of the ambient concentration [22]. A reduction in inhaled infectious dose of this magnitude would be expected to substantially reduce the probability of infection in ordinary living environments [11,19].

This protective mechanism differs fundamentally from that of passive masks and negative-pressure respirators. In such devices, inhalation creates negative pressure inside the device, and inadequate facial sealing may allow unfiltered ambient air to enter through gaps. By contrast, a PAPR actively maintains positive pressure in the breathing zone and therefore limits inward leakage through gaps in principle. Comparing an APF of 10 for a negative-pressure respirator with an APF of 1000 for a positive-pressure PAPR, inhalation exposure under specified conditions of fit and operation differs by approximately 100-fold, or about 2 orders of magnitude [22]. This substantial reduction in inhaled infectious dose could be expected to have a decisive effect, bringing the probability of infection close to zero in ordinary living environments.

Cost is one barrier to widespread use. Medical-grade PAPRs are expensive, although their basic structure consists only of a nonwoven filter, pump, and battery [21,23]. The author demonstrated that a prototype providing equivalent performance could be assembled from commercially available components costing approximately US $40 in total [14]. Mass production is highly likely to reduce costs to a level suitable for public health procurement.

Findings From the Author's Research Group

The author's research group has progressively developed the engineering and epidemiological foundations of PFE. Four studies make important contributions.

First, the rate of PAPR use required for it to function as an alternative to lockdown was quantified using an epidemiological mathematical model based on the effective reproduction number [15]. Simulations showed that continuous use by approximately 55% or more of the population could reduce the effective reproduction number from 2 to <0.9. They also showed that, if everyone used a PAPR, an approximately 55% reduction in each individual's probability of infection would produce an equivalent population-level infection-control effect (R_t).

Second, a detailed fluid-flow model of airflow inside a PAPR was developed [16]. A method was proposed for estimating respiratory state from a differential-pressure sensor and a pulse-width modulation (PWM) control signal, enabling

XSL•FO
RenderX

respiratory-phase-adaptive control intended to provide pressure-supported breathing through reverse-pressure control.

Third, validation with a physical prototype demonstrated that the wearer's respiratory flow could be estimated accurately in real time and that leakage associated with improper fit could be detected using only 1 differential-pressure sensor and the pump drive signal [17]. These findings established a technical basis for respiratory-state-responsive airflow control and the implementation of pressure-supported breathing.

Finally, the PAPR Wear-Rate Management System Network (PWS-NET) was proposed and prototyped [18]. This system combines real-time use and leak detection using a differential-pressure sensor, GPS location reporting through a smartphone, and a rule-based web server. It introduces the concept of saved allowable time (SAT), defined as the amount of time during which an individual is permitted to remove the PAPR. Prototype validation confirmed that SAT updating, violation detection, and real-time feedback to users functioned as intended.

PWS-NET is a framework that simultaneously enables population-wide monitoring of use rates and individual selection of SAT [18]. In this context, an AI-output governance framework comprising the Guide to the Expression of Legitimacy of Output (GLO) and the Verifiable Record of AI Output (VRAIO) becomes important [24]. GLO is a common language that treats an output candidate generated by an AI system as a "claim" and uses structured metadata to express whether its purpose and content comply with publicly disclosed rules. In this article, the same concept is envisaged as applying to the use of information and system decisions in PWS-NET. VRAIO provides the institutional and technical infrastructure needed to ensure the authenticity and verifiability of GLO declarations through compliance verification by an independent recorder, tamper-resistant records, and subsequent independent audits [24].

The more effective PWS-NET becomes, the greater its capacity to collect and use personal information. The GLO-VRAIO framework could impose verifiable institutional constraints on that capacity and offer a promising governance option for implementing PFE not as an instrument of intensified surveillance, but as public health infrastructure premised on privacy protection. In other words, a 3-layer structure—physical protection by PAPRs, use-rate management by PWS-NET, and governance of information use and system outputs through GLO-VRAIO—would seek to reconcile effective infection control with a high level of protection for citizens' rights.

Logical Implications of the Current Evidence

Four implications follow from the aforementioned analysis. First, during pandemics dominated by airborne transmission, PAPRs could partially substitute for lockdown and may make it possible to manage infection risk without halting societal functions. Second, positive-pressure exposure control using PAPRs directly controls the physical process of inhalation exposure to infectious aerosols, irrespective of pathogen type. It may therefore be available immediately after the emergence of a new respiratory pathogen, without waiting for pathogen-specific vaccines or therapeutics. If inhalation exposure can be reduced sufficiently, it may be possible to control transmission without relying solely on acquisition of immunity and to secure time for pathogen identification, characterization of transmission, and development of vaccines and therapeutics. Third, PAPRs give individuals the option to reduce their risk of infection, and voluntary uptake could contribute to society-wide epidemic control. Finally, combining PWS-NET with the GLO-VRAIO framework may enable the information needed for infection control to be used while making it possible to verify the purposes for which information on use status, movement histories, and biometric information is collected and analyzed, as well as compliance with privacy protection rules [18,24]. This could provide an institutional basis for reconciling infection control grounded in the principle that "what cannot be measured cannot be controlled" with the protection of citizens' privacy and rights.

These implications indicate that PFE could become an additional axis of pandemic response, linking individual-level inhalation-exposure control to population-level infection control [14,15,18].

Scope for Improvement: PAPR Technology Has Yet to Be Fully Optimized

Current PAPRs are used primarily in specialized hazardous environments in medical and industrial settings. They are produced in small quantities for limited markets and are subject to rigorous performance testing, quality assurance, and certification. Consequently, they tend to be expensive, and there has been little product-development competition concerning the lightness, comfort, functionality, design, and affordability required for routine use by the general public [23,25]. At the same time, their basic functional components are a nonwoven filter, an electric fan, and a battery, while their hoods or helmets are composed mainly of plastic, fabric, or nonwoven materials [14,21,23]. Current prices and forms therefore partly reflect conventional applications and market structures rather than inherent technological limitations.

Because an integrated helmet-type PAPR contains a battery, electric fan, control circuitry, and supporting structure, it also offers a suitable platform for integrating sensors that measure respiratory state and leakage, communication functions, augmented-reality displays, and smartphone functions. However, the required protective performance and comfort must be reconciled while evaluating the effects of added functions on weight, power consumption, heat generation, safety, and privacy. In addition to helmet-type devices, possible implementations of positive-pressure exposure control include personal booth systems for offices and restaurants [14,15] and ventilation-unit replacement systems that cover entire interiors of homes, offices, hotels, and other buildings.

This article has focused primarily on positive-pressure PAPRs with an APF of 1000 as a benchmark for high individual protection. However, the performance required of PAPRs for the general public need not uniformly be an APF of 1000. A

XSL•FO
RenderX

lower protection factor may suffice if it can achieve the individual-level objective of keeping the inhaled infectious dose below the level required to establish infection or the public health objective of maintaining the effective reproduction number below 1 across the population. The required performance level will vary with the pathogen's infectious dose, environmental concentration, duration of exposure, rate of PAPR use, and continuity of use. Future empirical studies should therefore verify the performance of APF 1000-class devices while also identifying the protection level that is necessary and sufficient to achieve individual- and population-level infection-control objectives.

PFE can be deployed without PWS-NET; PWS-NET is a future option intended to reconcile use-rate management with individual choice of SAT [18]. However, the use of wearing histories, location information, proximity information, and biometric data could permit inferences about an individual's patterns of daily life. By using GLO to make structured declarations of the purposes of data acquisition and analysis and VRAIO to verify, record, and audit compliance with privacy protection rules, information use by PWS-NET could be handled as a verifiable institutional procedure [24].

If positive-pressure PAPRs for the general public become a formal subject of public health evaluation and future demand becomes clear, mass production, component standardization, and competition among products may advance, promoting lower prices and designs suitable for daily use. The present generation of prototypes is a starting point, not a finished product. By continuously incorporating public health needs and evaluation results into engineering development, PAPRs must evolve into lighter, more comfortable, and safer devices while maintaining the required protective performance.

In the longer term, if PAPRs become sufficiently comfortable and convenient, people may voluntarily choose them to breathe purified air regardless of whether governments request their use. Just as people routinely use hygienically managed water, this concept envisages a future in which individuals can also choose purified air to breathe. When PAPRs shift from specialized protective equipment for emergencies to routine devices for controlling inhalation exposure, control of airborne infectious diseases may reach a major turning point—from a framework dependent on behavioral restrictions and requests to wear protective devices to 1 grounded in voluntary control of individual exposure.

A Structural Oversight: Why Has an Apparently Obvious Option Never Been Evaluated?

The filtration performance of HEPA filters, the prevention of ambient-air ingress through positive pressure, and the dose-based understanding of infection described here are not new principles; each is grounded in existing research and standards [11,20,22].

No controlled trial has measured the effect of population-level PAPR use on infection rates, and prospective studies evaluating real-world effectiveness, adherence, and adverse events are scarce [12]. World Health Organization (WHO) preparedness and resilience for emerging threats (PRET) module 1 provides an integrated planning framework for respiratory pathogen pandemics [26]. However, positive-pressure exposure control for the general public is not explicitly identified as an independent candidate intervention. The absence of an institutional conceptual category may help explain why the option has not been evaluated despite the existence of its underlying principles and technology.

PAPRs have primarily been studied and used as respiratory protection for health care workers exposed to highly infectious diseases. The systematic review by Licina et al [12] found that existing PAPR research mainly addressed the protection of health care workers from highly infectious viral diseases and identified the higher cost of PAPRs relative to N95 respirators as an implementation disadvantage. A narrative review by the same research group showed that PAPRs have been considered protective equipment for health care workers performing high-risk procedures—particularly aerosol-generating procedures (AGPs)—on patients with suspected or confirmed COVID-19 pandemic [27]. Elkington et al [28] further reported that existing certified PAPRs typically cost £600 to £1100 (GBP £1=US $1.35, as of August 14, 2026) and explicitly identified this high price as one of the main reasons why PAPRs were not widely adopted worldwide during the COVID-19 pandemic, even among health care workers.

Thus, PAPRs were developed within the conceptual category of high-performance, expensive, specialized protective equipment for health care workers facing a high risk of infection. Face masks, by contrast, were treated as an adjunctive behavioral intervention [29]. The idea of redesigning and evaluating PAPRs—which were not widely adopted even among health care workers—as a population-level infection-control measure for routine use by the general public therefore appears to have remained effectively outside conventional research and institutional assumptions. The concept of a full-fledged engineering exposure-control device worn routinely by the general public fit none of the conventional categories and occupied an academic and institutional no-man's-land. Without a category, a concept cannot even be placed on the agenda for evaluation.

The cost of this structural omission became real during the COVID-19 pandemic. When an unknown pathogen emerged, rapidly deployable options were effectively limited to behavioral interventions plus vaccination, or lockdown. A third option—one with the potential to reduce inhalation exposure by orders of magnitude regardless of pathogen type—existed in principle but was absent in practice. Closing this gap is precisely what should be done now to prepare for the next pandemic.

PFE should therefore be examined by combining medical and engineering assessments within a public health framework. Engineering assessment should use an APF of 1000 as a high-performance benchmark and establish aerosol-protection performance, maintenance of positive pressure, removal of exhaled carbon dioxide, stability of device performance, and applicable conditions of use in ordinary daily life. Medical

XSL•FO
RenderX

assessment should determine how an engineering-measured reduction in exposure translates into inhaled infectious dose and individual-level prevention of infection, while also evaluating the safety of prolonged use. Public health assessment should determine how individual-level preventive effectiveness, use rates, and fit or use status translate into population-level infection-control effectiveness and identify the protection levels and use rates necessary and sufficient to achieve infection-control objectives. Feasibility of implementation, social acceptability, cost, and equity should also be evaluated comprehensively. If networking through PWS-NET or a similar system is used, the effect of the PAPR alone should be distinguished from the additional effect of networking, and the effects on privacy and citizens' rights should be treated as independent evaluation outcomes.

To translate the rapid deployment feasibility of PAPRs into an actual pandemic response, beginning development only after a crisis occurs would be insufficient. During nonemergency periods, performance requirements, safety, sustained use, and conditions of use for public-oriented models should be evaluated, while the number of units required, manufacturing capacity, standardization of components and consumables, and systems for procurement, stockpiling, and distribution should be considered in advance. Current PAPRs are bulky and may appear excessive in daily life; nevertheless, models certified with an APF of 1000 can provide high aerosol-protection performance under specified conditions of fit and operation. A realistic approach is to prepare practicable models first and then progressively reduce weight and noise and ensure adequate airflow during physical activity on the basis of evaluation and operational experience.

Conclusions

Another respiratory infectious disease pandemic is likely to occur in the future. The question is whether we will search for new countermeasures only after the next crisis begins or evaluate concrete, testable candidate interventions in advance during the current nonemergency period.

Exposure control using positive-pressure PAPRs for the general public directly reduces aerosol inhalation exposure in the individual breathing zone—the process that precedes airborne infection—regardless of pathogen type. Its principles and technology are already established, and existing engineering and theoretical research has brought it to a stage at which it can be considered a candidate intervention with the potential for manufacture and implementation. Nevertheless, it has not been positioned as a subject of formal public health evaluation as a population-level infection-control measure.

This viewpoint does not advocate the immediate introduction of PFE or a universal mandate for its use. Existing engineering and theoretical evidence does not establish the effectiveness of PFE; rather, it justifies beginning its formal evaluation. Although the roles of engineering, medicine, and public health overlap in practice, engineering assessment should verify aerosol-protection performance, maintenance of positive pressure, removal of exhaled carbon dioxide, and stability of device performance in ordinary daily life conditions. Medical assessment must determine how engineering-measured reductions in exposure translate into individual-level prevention of infection and safety during prolonged use. Public health assessment must examine how individual-level preventive effectiveness, population use rates, and continuous and proper use at the necessary times and locations translate into population-level infection-control effectiveness. Feasibility of implementation, social acceptability, cost, and equity must also be evaluated comprehensively.

An APF of 1000 provides a benchmark for evaluating high individual protection, but the performance required of PAPRs for the general public need not uniformly be an APF of 1000. The protection level that is necessary and sufficient to achieve individual- and population-level infection-control objectives should be determined through public health evaluation incorporating medical and engineering evidence. Networking through PWS-NET or similar systems should also be distinguished from evaluation of the PAPR alone; its additional infection-control effect and its implications for privacy and citizens' rights must be evaluated independently.

To advance this evaluation institutionally, we propose that WHO and national pandemic preparedness plans explicitly position positive-pressure exposure control for the general public as a "candidate intervention for which evidence of effectiveness needs to be accumulated." Field studies should begin during nonemergency periods to evaluate engineering performance, individual-level preventive effectiveness, population-level infection-control effectiveness, and implementation conditions in stages, while examining the feasibility of manufacturing, procurement, stockpiling, and distribution. Positive-pressure exposure control for the general public should be added to the scope of formal public health evaluation before the next pandemic arrives.

Acknowledgments

The author thanks the colleagues and collaborators who provided valuable discussion during the development of this work. The author used generative AI assistants (ChatGPT [GPT-5.6 Sol; OpenAI] and Claude [Claude Opus 5; Anthropic PBC]) to support literature searches, review the logical consistency of the arguments, refine the manuscript's language, and assist with English translation. These tools were not used to generate research data, perform analyses, or draw scientific conclusions. The author critically reviewed and verified all AI-assisted content and assumes full responsibility for the integrity and accuracy of the final manuscript.

XSL•FO
RenderX

Funding

This work was supported by the Japan Society for the Promotion of Science (JSPS) *Kagaku Kenkyūhi Josei Jigyō* (KAKENHI; grant 25K00735).

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Conflicts of Interest

None declared.

References

1. Sachs JD, Karim SS, Aknin L, et al. The Lancet Commission on lessons for the future from the COVID-19 pandemic. Lancet 2022 Oct 8;400(10359):1224-1280. [doi: 10.1016/S0140-6736(22)01585-9] [Medline: 36115368]
2. Chu DK, Akl EA, Duda S, et al. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. Lancet 2020 Jun 27;395(10242):1973-1987. [doi: 10.1016/S0140-6736(20)31142-9] [Medline: 32497510]
3. Anderson RM, Vegvari C, Truscott J, Collyer BS. Challenges in creating herd immunity to SARS-CoV-2 infection by mass vaccination. Lancet 2020 Nov 21;396(10263):1614-1616. [doi: 10.1016/S0140-6736(20)32318-7] [Medline: 33159850]
4. Flaxman S, Mishra S, Gandy A, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature 2020 Aug;584(7820):257-261. [doi: 10.1038/s41586-020-2405-7] [Medline: 32512579]
5. Betthäuser BA, Bach-Mortensen AM, Engzell P. A systematic review and meta-analysis of the evidence on learning during the COVID-19 pandemic. Nat Hum Behav 2023 Mar;7(3):375-385. [doi: 10.1038/s41562-022-01506-4] [Medline: 36717609]
6. Vardavas C, Zisis K, Nikitara K, et al. Cost of the COVID-19 pandemic versus the cost-effectiveness of mitigation strategies in EU/UK/OECD: a systematic review. BMJ Open 2023 Oct 31;13(10):e077602. [doi: 10.1136/bmjopen-2023-077602] [Medline: 37907290]
7. Fauci AS, Folkers GK. Pandemic preparedness and response: lessons from COVID-19. J Infect Dis 2023 Aug 16;228(4):422-425. [doi: 10.1093/infdis/jiad095] [Medline: 37035891]
8. Harrington WN, Kackos CM, Webby RJ. The evolution and future of influenza pandemic preparedness. Exp Mol Med 2021 May;53(5):737-749. [doi: 10.1038/s12276-021-00603-0] [Medline: 33953324]
9. Greenhalgh T, Jimenez JL, Prather KA, Tufekci Z, Fisman D, Schooley R. Ten scientific reasons in support of airborne transmission of SARS-CoV-2. Lancet 2021 May 1;397(10285):1603-1605. [doi: 10.1016/S0140-6736(21)00869-2] [Medline: 33865497]
10. Meyerowitz EA, Richterman A, Gandhi RT, Sax PE. Transmission of SARS-CoV-2: a review of viral, host, and environmental factors. Ann Intern Med 2021 Jan;174(1):69-79. [doi: 10.7326/M20-5008] [Medline: 32941052]
11. Zhang X, Wang J. Dose-response relation deduced for coronaviruses from coronavirus disease 2019, severe acute respiratory syndrome, and Middle East respiratory syndrome: meta-analysis results and its application for infection risk assessment of aerosol transmission. Clin Infect Dis 2021 Jul 1;73(1):e241-e245. [doi: 10.1093/cid/ciaa1675] [Medline: 33119733]
12. Licina A, Silvers A, Stuart RL. Use of powered air-purifying respirator (PAPR) by healthcare workers for preventing highly infectious viral diseases-a systematic review of evidence. Syst Rev 2020 Aug 8;9(1):173. [doi: 10.1186/s13643-020-01431-5] [Medline: 32771035]
13. Fujii Y, Takita A, Hashimoto S. A helmet type mask “distancing-free mask”: an engineering solution that eliminates the lockdown. J Mech Elect Intel Syst 2020;3(3):1-7 [FREE Full text]
14. Fujii Y. An engineering alternative to lockdown during COVID-19 and other airborne infectious disease pandemics: feasibility study. JMIR Biomed Eng 2024 May 14;9:e54666. [doi: 10.2196/54666] [Medline: 38875692]
15. Fujii Y. Examination of the requirements for powered air-purifying respirator (PAPR) utilization as an alternative to lockdown. Sci Rep 2025 Jan 7;15(1):1217. [doi: 10.1038/s41598-024-82348-0] [Medline: 39774613]
16. Fujii Y, Takita A, Hashimoto S, Amagai K. Estimation of respiratory states based on a measurement model of airflow characteristics in powered air-purifying respirators using differential pressure and pulse width modulation control signals-in the development of a public-oriented powered air-purifying respirator as an alternative to lockdown measures. Sensors (Basel) 2025 May 7;25(9):2939. [doi: 10.3390/s25092939] [Medline: 40363375]
17. Fujii Y. The real-time estimation of respiratory flow and mask leakage in a PAPR using a single differential-pressure sensor and microcontroller-based smartphone interface in the development of a public-oriented powered air-purifying respirator as an alternative to lockdown measures. Sensors (Basel) 2025 Aug 28;25(17):5340. [doi: 10.3390/s25175340] [Medline: 40942768]
18. Fujii Y. Time-managed PAPR use enables a balanced approach to infection control and personal freedom. Sci Rep 2026;16:2878. [doi: 10.1038/s41598-025-32625-3]
19. Prentiss M, Chu A, Berggren KK. Finding the infectious dose for COVID-19 by applying an airborne-transmission model to superspreader events. PLoS One 2022;17(6):e0265816. [doi: 10.1371/journal.pone.0265816] [Medline: 35679278]

XSL•FO
RenderX

20. Nazarenko Y. Air filtration and SARS-CoV-2. Epidemiol Health 2020;42:e2020049. [doi: 10.4178/epih.e2020049] [Medline: 32660218]
21. Kothakonda A, Atta L, Plana D, et al. De novo powered air-purifying respirator design and fabrication for pandemic response. Front Bioeng Biotechnol 2021;9:690905. [doi: 10.3389/fbioe.2021.690905] [Medline: 34552915]
22. 1910.134 - respiratory protection. Occupational Safety and Health Administration. URL: https://www.osha.gov/laws-regs/regulations/standardnumber/1910/1910.134 [accessed 2026-08-14]
23. Khoo D, Yen CC, Chow WT, et al. Ultra-portable low-cost improvised powered air-purifying respirator: feasibility study. Br J Anaesth 2020 Aug;125(2):e264-e266. [doi: 10.1016/j.bja.2020.04.082] [Medline: 32446500]
24. Fujii Y. No trust without trust infrastructure: the Extended Kelvin Principle and its application to AI output governance. AI 2026;7(6):218. [doi: 10.3390/ai7060218]
25. Ng I, Lee K, Kave B, et al. HALO CleanSpace PAPR evaluation: communication, respiratory protection, and usability. Infect Control Hosp Epidemiol 2023 Feb;44(2):295-301. [doi: 10.1017/ice.2022.71] [Medline: 35361300]
26. Preparedness and resilience for emerging threats module 1: planning for respiratory pathogen pandemics. World Health Organization. 2024. URL: https://www.who.int/publications/i/item/9789240084674 [accessed 2026-08-14]
27. Licina A, Silvers A. Use of powered air-purifying respirator (PAPR) as part of protective equipment against SARS-CoV-2-a narrative review and critical appraisal of evidence. Am J Infect Control 2021 Apr;49(4):492-499. [doi: 10.1016/j.ajic.2020.11.009] [Medline: 33186678]
28. Elkington PT, Dickinson AS, Mavrogordato MN, et al. A personal respirator to improve protection for healthcare workers treating COVID-19 (PeRSo). Front Med Technol 2021;3:664259. [doi: 10.3389/fmedt.2021.664259] [Medline: 35047921]
29. Feng S, Shen C, Xia N, Song W, Fan M, Cowling BJ. Rational use of face masks in the COVID-19 pandemic. Lancet Respir Med 2020 May;8(5):434-436. [doi: 10.1016/S2213-2600(20)30134-X] [Medline: 32203710]

Abbreviations

AGP: aerosol-generating procedure
APF: assigned protection factor
GLO: Guide to the Expression of Legitimacy of Output
HEPA: high-efficiency particulate air
MPPS: most penetrating particle size
PAPR: powered air-purifying respirator
PFE: Powered air-purifying respirator for Everyone
PRET: preparedness and resilience for emerging threats
PWM: pulse-width modulation
PWS-NET: Powered air-purifying respirator Wear-Rate Management System Network
SAT: saved allowable time
SWPF: simulated workplace protection factor
VRAIO: Verifiable Record of AI Output
WHO: World Health Organization
WPF: workplace protection factor

Edited by A Mavragani; submitted 23.Dec.2025; peer-reviewed by W Lindsley, X Niu; revised version received 20.Jul.2026; accepted 05.Aug.2026; published 27.Aug.2026.

Please cite as:
Fujii Y
Reframing Pandemic Preparedness: Why Positive-Pressure Exposure Control Merits Formal Public Health Evaluation
JMIR Public Health Surveill 2026;12:e90211
URL: https://publichealth.jmir.org/2026/1/e90211
doi:10.2196/90211

© Yusaku Fujii. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 27.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Reimagining Data for Equity: Reaching the Unreached Map Explorer for Localized Public Health Intelligence in the Western Pacific Region

Sumbul Hashmi[1], MDS; Anne-Laure Budts[1], MSc; Vincent Meurrens[1], BSc; Jessica Sarah Stow[2], BSc; Fukushi Morishita[3], MSc; Rajendra-Prasad Yadav[3], MD

[1]EPCON, Antwerp, Belgium

[2]EPCON, Workshop 17 146 Campground Rd, 3rd floor, Snakepit Building, Newlands, Cape Town, Western Cape, South Africa

[3]World Health Organization Regional Office for the Western Pacific, Manila, Philippines

Corresponding Author:
Jessica Sarah Stow, BSc
EPCON, Workshop 17 146 Campground Rd, 3rd floor, Snakepit Building, Newlands, Cape Town, Western Cape, South Africa

Abstract

Geographic information systems (GISs) and digital dashboards have become central to epidemic monitoring and crisis response, yet their broader potential for routine, equity-focused public health decision-making remains underused. As health systems increasingly prioritize reaching underserved and vulnerable populations, mapping platforms must evolve from static visualization tools to dynamic decision-support systems. This paper reflects on the development and strategic role of the Reaching the Unreached (RTU) map explorer, a regional geospatial decision-intelligence platform developed in collaboration with the World Health Organization Regional Office for the Western Pacific. The RTU map explorer integrates high-resolution subnational data across 36 countries and areas in the Western Pacific Region. It consolidates demographic, socioeconomic, health access, and vulnerability indicators into a unified geoportal that enables visualization at resolutions ranging from the national level to the village level. Drawing on open-access spatial repositories and modeled datasets, the platform applies population-based spatial clustering methods to settings where administrative units are too large to capture meaningful local variation. Built using an open-source architecture, the platform offers interactive functionalities, including layer activation, filtering, prioritization tools, and location-specific mapping, to support localized planning and resource allocation. By combining contextual and health-related data within a single interface, the RTU map explorer facilitates a shift from reactive surveillance to proactive, equity-oriented decision-making. Policymakers can compare vulnerabilities across regions, program managers can prioritize high-need districts or communities, and local implementers can identify specific locations for targeted interventions. Although challenges remain, including data recency, modeled data uncertainty, and the need for field validation, the platform illustrates how integrated geospatial intelligence can move public health beyond documenting disparities to actively addressing them through localized, data-driven action.

(JMIR Public Health Surveill 2026;12:e93324) doi:10.2196/93324

KEYWORDS

geospatial intelligence; digital health; vulnerability mapping; public health surveillance; health equity

Introduction

Geographic information systems (GISs) and mapping platforms have long been used in public health, with the most frequent application being epidemic monitoring and disease surveillance [1,2]. A near-global use of such tools was observed during the COVID-19 pandemic, during which mapping platforms and live dashboards provided intuitive outputs and were extensively leveraged [3]. However, as the epidemic curve plateaus, the use of these highly valuable applications also tends to diminish [4]. Such solutions possess considerable potential to address numerous public health questions and support routine decision-making when properly implemented [3]. These tools are beneficial for a range of stakeholders, including policymakers, program managers, and local implementers such as community-based organizations.

To ensure a map's utility for all stakeholders and its relevance to routine decision-making, the required spatial resolution must transcend the national or district level [5]. A well-designed visualization can address critical questions such as the following: "Which villages lack basic Water, Sanitation, and Hygiene (WASH) services?" "Where are pregnant women least likely to deliver in a maternity facility?" "Where are children dropping out of school?" "Where do children require additional nutritional assistance?" "Which areas have poor connectivity to primary healthcare?" or "Which communities are lagging in childhood

XSL•FO
RenderX

immunization coverage?" [6]. Answering these questions requires a comprehensive understanding of the local context; therefore, integrating this information into a single platform is a logical approach [7].

The Map Explorer for Reaching the Unreached: A Regional Decision-Intelligence Platform

One such platform, the Reaching the Unreached (RTU) map explorer for the Western Pacific Region, encompasses 36 countries and areas within the World Health Organization (WHO) Western Pacific Region. Indonesia, which transitioned from the WHO South-East Asia Region to the Western Pacific Region in May 2025, was not included because all analyses were completed prior to this reassignment. The Pitcairn Islands were also excluded due to their small population size. This paper reflects on our experience developing the platform, which is a regional initiative by the WHO to strengthen data-driven decision-making. We discuss its evolution as a regional intelligence tool and how it can support the transition toward localized, equitable, and continuous decision-support systems in public health. The platform was designed for use by regional, national, and subnational public health decision-makers, including WHO teams, ministries of health, program managers, funding partners, and nongovernmental or community-based implementers seeking to identify underserved populations and guide more localized, equitable action.

The platform was developed by EPCON in September 2024 in collaboration with the WHO Regional Office for the Western Pacific to support the implementation of the WHO Regional Framework for RTU [8]. The overarching goal of this project was to develop a platform capable of supporting data-driven decision-making for common public health concerns and of assisting in the identification of vulnerable populations at scales extending down to the village or community level.

The RTU map explorer hosts high-resolution, subnational data spanning a range of indicators, including demographic characteristics, socioeconomic development, overall human development, health care access, and population vulnerability. The data are available across all administrative levels, from the national level down to a resolution of a few kilometers (village or neighborhood level). This granular data make the platform suitable for diverse public health programs, such as those focused on disease control (eg, tuberculosis [TB], HIV, malaria, and other epidemics) as well as programs related to population welfare, education, immunization, WASH, and health promotion, with a focus on underserved or unreached populations [9,10]. Given the constraints of limited resources and a high vulnerability burden, the efficient use of resources is paramount [11]. Investing in advanced tools that facilitate context-driven decision-making can ensure that interventions reach the populations most in need and optimize intervention implementation [12].

The RTU map explorer is built on a geoportal, the web-based mapping component of the Epi-control platform (EPCON). The Epi-control platform is an AI-driven "digital twin" that is deployed in several countries to support specific disease control programs, primarily TB and HIV [13,14]. Within the Epi-control platform, data from health programs can be augmented with high-resolution subnational contextual data to train predictive models that identify high-transmission areas or disease "hotspots." The "digital twin" is a dynamic model that updates in near real-time, or as required, to forecast new hot spots based on programmatic progress and planning. Community-based interventions are guided by a comprehensive, data-driven process, moving beyond sole reliance on case notification rates. This approach has yielded improved intervention outcomes and helped direct resources to priority areas, thereby maximizing program impact [14,15]. However, the implementation of this approach and its outcome evaluation in the Western Pacific Region have yet to be evaluated. The RTU map explorer functions as a data integration and visualization platform, providing the static contextual intelligence layer upon which, when combined with programmatic or case notification data, programmatic decision-making can be grounded.

Transforming the Role of Maps

The first step was a thorough search for subnational high-resolution data. Relevant data sources were identified from open online repositories that host survey or modeled indicators, including the Demographic and Health Surveys (DHS) spatial repository, the Institute for Health Metrics and Evaluation (IHME), WorldPop, the Humanitarian Data Exchange (HDX), and country government websites. The data sources were selected based on the criteria of resolution, recency, and extent; a complete inventory is provided in Multimedia Appendix 1. Such high-resolution raster data were extracted and aggregated within each administrative boundary using zonal statistics. Pixel values were either summed (for count-based data, eg, population counts) or averaged (for continuous or rate-based data, eg, prevalence data). To aggregate values up the administrative hierarchy, population-weighted averages were used for indicators that vary with population size (eg, vaccination coverage), and mean values were used for those that do not vary with population size (eg, distance to the nearest health care facility).

On the basis of the availability of high-resolution data and high population density, the smallest administrative units in some countries were further subdivided into smaller units. This was achieved using a recursive weighted k-means clustering algorithm [16]. For each administrative unit, the algorithm iteratively assigned population-weighted centroids using k-means clustering, from which "Thiessen" (Voronoi) polygons were constructed and their populations computed [17]. If the population of a generated polygon exceeded a target threshold of approximately 10,000 individuals, the unit was iteratively subdivided. The 10,000-person population threshold was determined in collaboration with project stakeholders to make administrative units comparable across the 36 countries and areas of the Western Pacific Region.

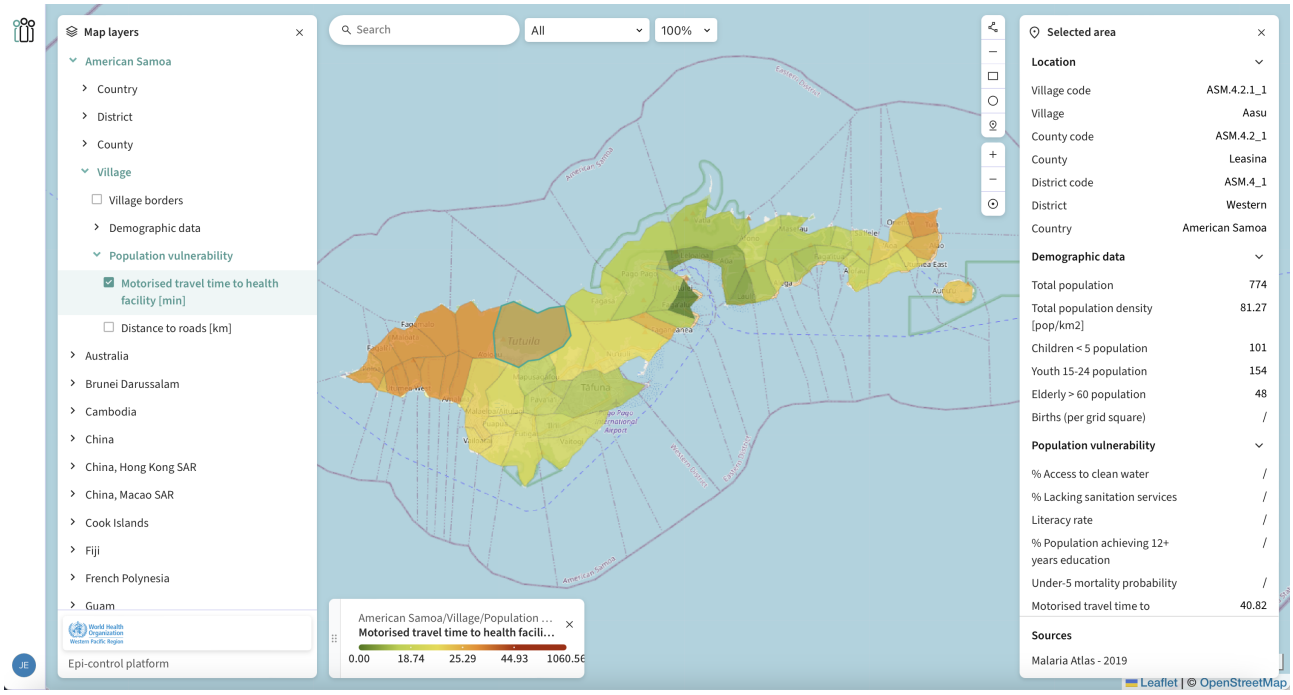
The weighted k-means algorithm partitioned the population into clusters, after which new Voronoi polygons were generated around the resulting cluster centroids, and oversized polygons

XSL•FO
RenderX

were replaced with these smaller units. This process was repeated recursively until all polygons fell within the target population range. Each polygon was then spatially clipped to fit within its original administrative boundary. The final output comprised a table of Thiessen polygons, each associated with a population count and clipped geometry. The map explorer is implemented using proprietary Python code (version 3.12.3; Python Software Foundation) and leverages open-source software components.

The data layers are organized separately for each of the 36 countries and areas, followed by a comprehensive "Western Pacific Region" tab that visualizes all countries simultaneously for intercountry comparisons. For each country, the data layers are subdivided based on the respective administrative levels, commencing at the national level, followed by district, county, and village levels. Figure 1 demonstrates the selection of a layer to visualize data at the highest available resolution for American Samoa (village).

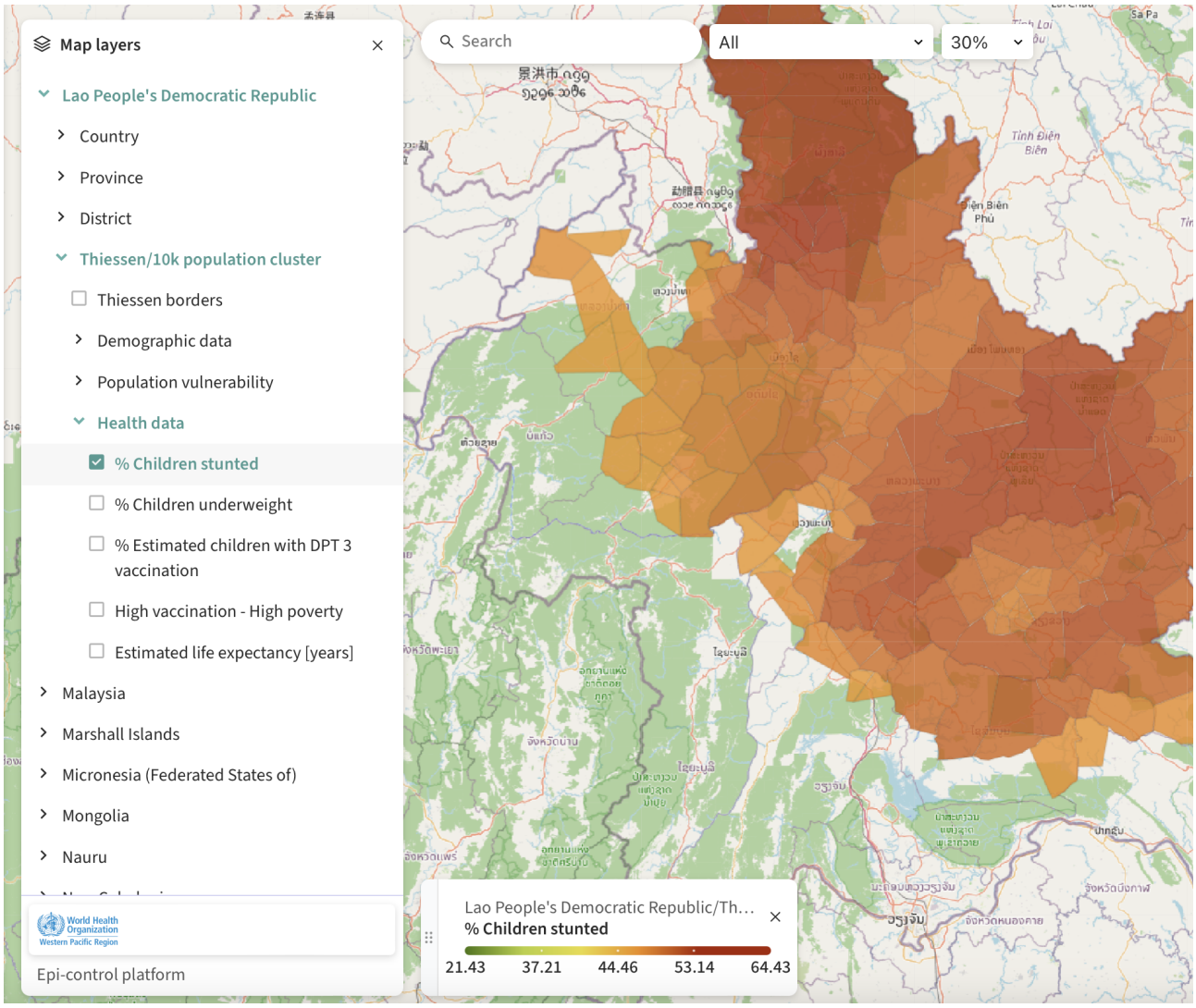
Figure 1. Arrangement of data layers for American Samoa, showing motorized travel time to health facilities at the village level. The color scale ranges from green (low travel time to health facilities) to orange red (high travel time to health facilities).



The map explorer provides several functionalities for seamless exploration of data layers. To commence exploration of a particular country, a data layer must be selected to activate the map. The "focus" tool facilitates direct zooming onto the country whose layer has been activated. The "location pointer" tool allows users to place a "pin" directly on the map and provides a Google Maps link to that specific location. This link can be shared or used to explore the neighborhood within the village or community, proving particularly useful for local implementers or nongovernmental organizations (NGOs) working closely with communities. It aids in identifying a specific landmark or village community for outreach purposes. The "filter" option enables the shortlisting of priority locations for targeted interventions. Figure 2 demonstrates the filtering capability of the platform to identify at-risk communities with high prevalence of child stunting. This is valuable for both program management and community-based organizations, allowing users to filter priority districts, wards, villages, or communes for specific resource allocation or intervention planning.

XSL•FO
RenderX

Figure 2. Enlarged view of the data layers available under different categories at the Thiessen-level (subdistrict-level) visualization of children stunting prevalence in the Lao People's Democratic Republic. The values are filtered to display the top 30% Thiessens with the highest prevalence of child stunting. The color scale ranges from green (low percentage of children stunted) to orange red (high percentage of children stunted). Thiessen borders are shown in lighter lines, and active layers are displayed in green on the left-side layer panel.



Bridging the Gap Between Data and Decision

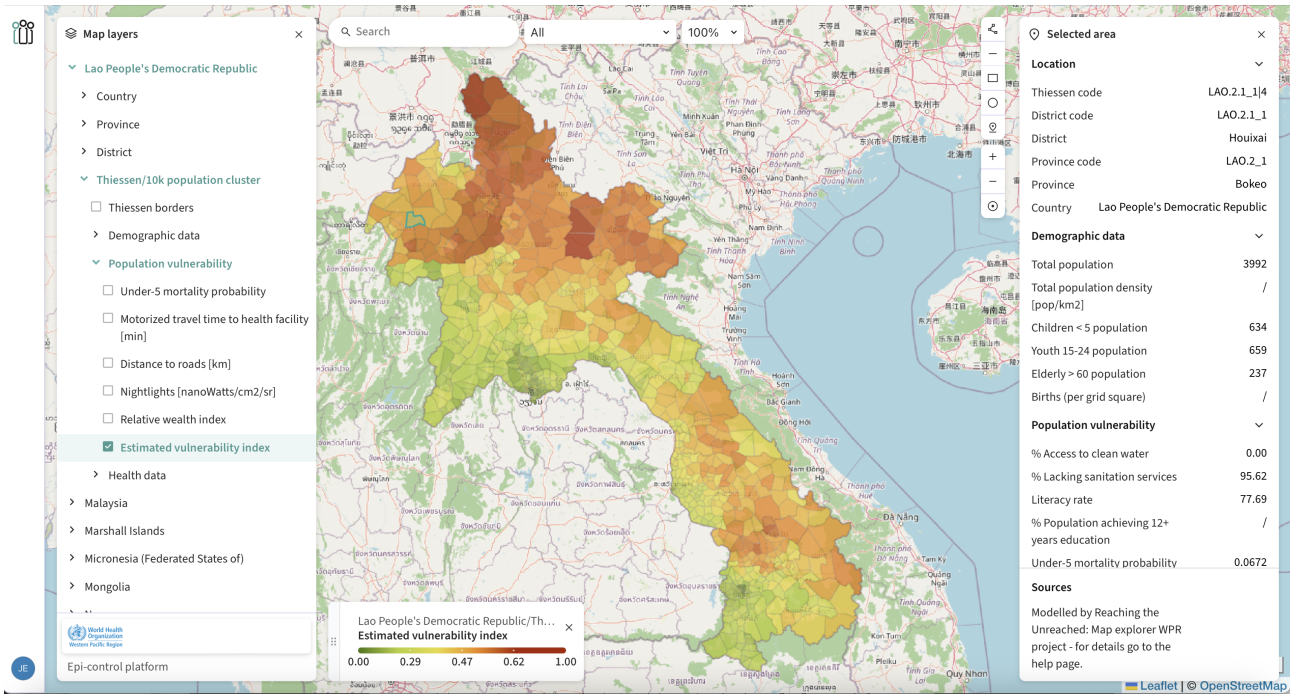
The platform is designed to support public health decision-making, the planning of local programs, and disease surveillance. The extensive data library can facilitate horizontal program integration. Sociodemographic conditions and local context are determinants not only of health and disease conditions but also of human behavior, local practices, and the design of interventions [18]. The platform offers an opportunity to gain a comprehensive understanding of the local context from the national to the village level, making it suitable for designing integrated programs such as TB and nutritional support interventions, maternal and child health (MCH) and immunization programs, and education and financial support initiatives.

To further support decision-making across multiple administrative units, an estimated vulnerability index was constructed using 11 indicators spanning 3 domains: health access (travel time to the nearest health center, distance to roads, and vaccination coverage), demographic and health outcomes (population density, elderly population aged ≥60 years, child mortality, and the prevalence of underweight children), and socioeconomic conditions (male and female literacy, nightlight intensity, access to water, and the prevalence of open defecation). For each indicator, spatial units were ranked relative to one another, with the direction of ranking assigned so that a higher rank consistently represented greater vulnerability. Each unit's rank was then normalized by dividing it by the maximum rank observed for that indicator, producing a score between 0 and 1 for every indicator across all spatial units. The vulnerability index for each spatial unit was computed as the mean of its normalized ranks across all available indicators. This formulation treats each indicator with equal weight and accommodates spatial units with missing data by computing the mean over observed indicators only, ensuring that partial data does not systematically bias the index. Figures 3 and 4 demonstrate the visualization of the vulnerability index at high resolution (Thiessen). For visualization at less granular administrative levels, unit-level index scores were spatially aggregated using the median to ensure robustness to outlying

XSL•FO
RenderX

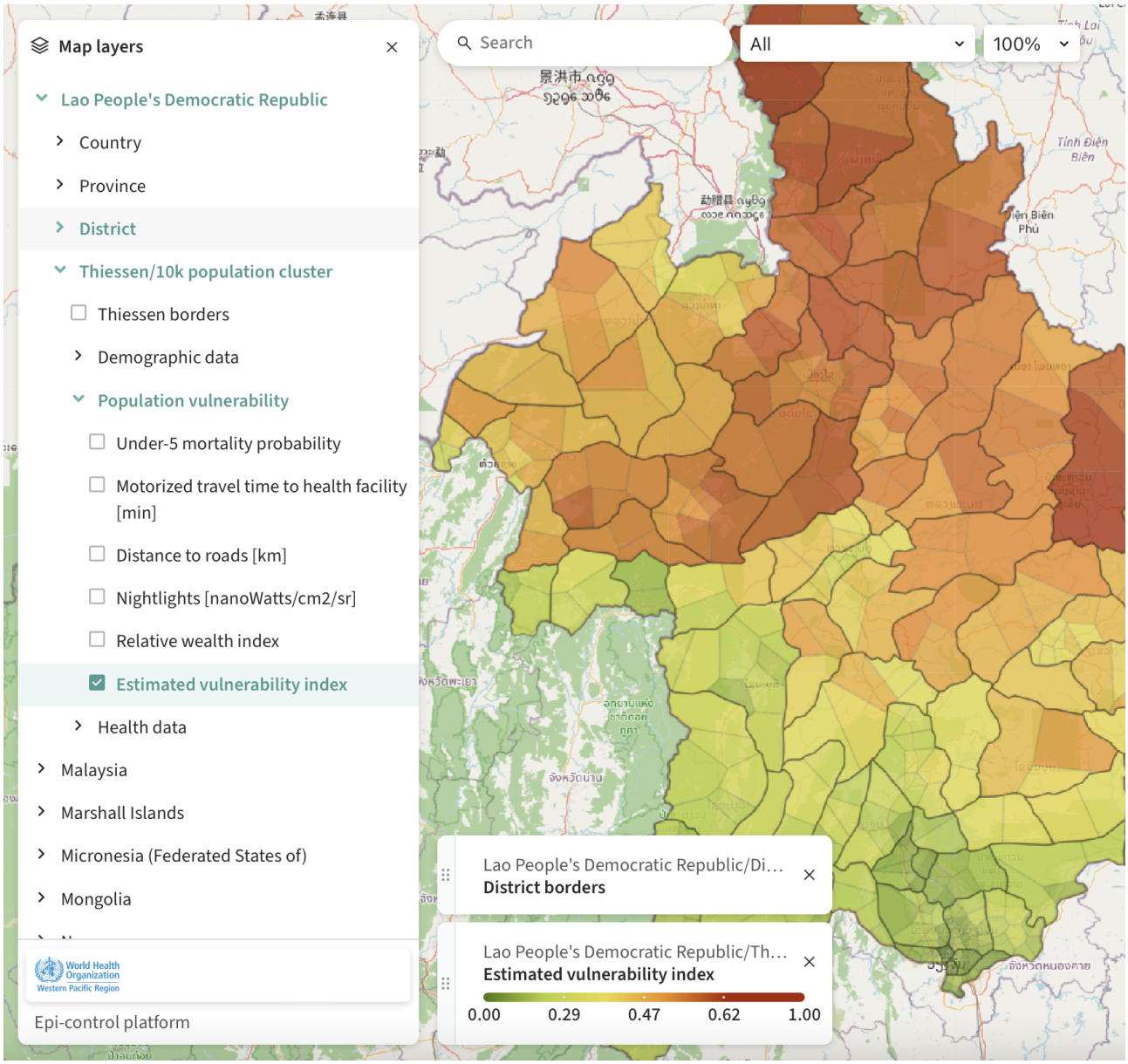
values within administrative units. The resulting index is a relative measure of estimated vulnerability, situating each administrative unit within the national distribution of vulnerability across all 3 domains simultaneously. The index is intended to support exploratory prioritization and dialogue with countries, not as a resource-allocation tool.

Figure 3. Thiessen-level (subdistrict-level) visualization of estimated vulnerability in the Lao People's Democratic Republic. The color scale ranges from green (low vulnerability) to orange red (high vulnerability).



XSL•FO
RenderX

Figure 4. Enlarged view of the Thiessen-level (subdistrict-level) visualization of estimated vulnerability in the Lao People's Democratic Republic. The color scale ranges from green (low vulnerability) to orange red (high vulnerability). District borders are shown in dark lines, and Thiessen borders are shown in lighter lines.



The mapping of data across all administrative levels ensures the platform's suitability for stakeholders across multiple professional cadres. Policy-level stakeholders may require a high-level overview, program implementers may need more granular information at the community level, and the scientific community may require both. Access levels and the scope of data can be customized based on local requirements, balancing comprehensiveness with simplicity.

Toward a More Equitable Data Future

The RTU map explorer is designed to support near real-time data updates when linked to a country's data platform or disease notification system, allowing new cases, interventions, or planned activities to be dynamically mapped on the platform. This integration is facilitated on a case-by-case basis according to country needs and data infrastructure. This also supports monitoring and evaluation efforts and contributes to pandemic preparedness strategies. As new survey data or modeled estimates become available, the data sources can be refreshed. For instance, the platform is currently being used in other countries, such as Nigeria [15], Pakistan [14], Uganda [19], and the Philippines, for planning community-based TB active case-finding interventions. Program stakeholders are able to track progress and plan interventions based on their access level.

Although the RTU map explorer platform has yet to achieve full uptake in the region, it has been integrated in other countries with mobile apps, diagnostic tools [20], X-ray equipment, and District Health Information Software 2 (DHIS2) platforms to access and map health-related data in real-time or to transmit notifications and data to other dashboards or devices.

Although various open-source platforms make subnational high-resolution data publicly available, this map explorer is unique in integrating data from different domains into a single platform, thereby maximizing its potential for scientific,

XSL•FO
RenderX

policy-level, and programmatic decision-making. Mapping and granular visualizations illuminate details that would remain obscured within aggregated numbers in tables and graphs, which is particularly crucial when attempting to reach underserved, remote, and vulnerable population groups [21]. The mapping of the local context enhances data intuitiveness and fosters opportunities for new ideas, approaches, and discussions [22].

In its current form, the platform is beneficial for government agencies, funding organizations, and local implementers. Funding agencies can use it to compare vulnerable regions both across and within countries to identify resource needs. Program planners and ministries can use it to identify high-priority regions within their respective countries. Local organizations can pinpoint communities or neighborhoods for targeted interventions, informed by the backdrop of local contextual data.

However, maintaining contextual and demographic data sources on such a platform comes with certain challenges. Reliable estimates and granular data may not be available for all countries and regions. These estimates are not updated on fixed timelines; therefore, the most recent granular estimates can be ≥5 years old for some countries. Second, most granular estimates are modeled datasets, which can inherently have large uncertainties. The platform in its current state is unable to visualize the associated CIs or uncertainties for these modeled estimates. The platform is not currently open source and comes with a licensing fee. The governance framework for platform adoption, including country-specific data-sharing, access, and stewardship arrangements, will be an important consideration when nonpublicly available data are incorporated into the platform. For a more tailored experience and to facilitate joint efforts in programmatic decision-making, the subsequent phase would involve the creation of individual mapping portals for specific countries. There is also a need for field validation and evaluation of the feasibility and uptake of the platform in the Western Pacific Region.

Data can generate greater impact when they empower those closest to the problem. Public health systems need to look beyond documenting inequality; there is a need for active tools that help make decisions. The progress of public health intelligence will depend not only on accumulating more data but also on transforming existing data into meaningful, local, relevant, and equitable insights. The RTU map explorer exemplifies technology that supports decision-making that is not just more intelligent, but also fairer.

Acknowledgments

No generative AI or AI-assisted technologies were used in the drafting, preparation, or analysis of this manuscript.

Funding

The World Health Organization Regional Office for the Western Pacific contracted EPCON for the implementation of the platform in support of its mandate to reach the most underserved populations in the region. EPCON did not receive any funding for the preparation of this manuscript.

Disclaimer

The authors alone are responsible for the views expressed in this publication and they do not necessarily represent the decisions, policy, or views of the World Health Organization. The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

Conflicts of Interest

SH, ALB, and JSS are employees of EPCON. VM is a cofounder and shareholder of EPCON. EPCON provided the analytical platform used in this study. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
List of data sources used to set up the platform.
[PDF File, 123 KB - publichealth_v12i1e93324_app1.pdf]

References

1. Nykiforuk CI, Flaman LM. Geographic information systems (GIS) for health promotion and public health: a review. Health Promot Pract 2011 Jan;12(1):63-73. [doi: 10.1177/1524839909334624] [Medline: 19546198]
2. Khashoggi BF, Murad A. Issues of healthcare planning and GIS: a review. ISPRS Int J Geo Inf 2020;9(6):352. [doi: 10.3390/ijgi9060352]
3. Katapally TR, Ibrahim ST. Digital health dashboards for decision-making to enable rapid responses during public health crises: replicable and scalable methodology. JMIR Res Protoc 2023 Jun 30;12:e46810. [doi: 10.2196/46810] [Medline: 37389905]

XSL•FO
RenderX

4. Laituri M, Kalra Y, Yang C. The disappearance of COVID-19 data dashboards: the case of ephemeral data. COVID 2025;5(1):12. [doi: 10.3390/covid5010012]
5. Riley TG, Mouat B, Shucksmith R. Real world data for real world problems: importance of appropriate spatial resolution modelling to inform decision makers in marine management. Ecol Model 2024 Dec;498:110864. [doi: 10.1016/j.ecolmodel.2024.110864]
6. Macharia PM, Beňová L, Ray N, et al. Putting health facilities on the map: a renewed call to create geolocated, comprehensive, updated, openly licensed dataset of health facilities in sub-Saharan African countries. BMC Med 2025 Apr 7;23(1):211. [doi: 10.1186/s12916-025-04023-z] [Medline: 40197252]
7. Rabiei R, Bastani P, Ahmadi H, Dehghan S, Almasi S. Developing public health surveillance dashboards: a scoping review on the design principles. BMC Public Health 2024 Feb 6;24(1):392. [doi: 10.1186/s12889-024-17841-2] [Medline: 38321469]
8. Reaching the unreached in the Western Pacific. World Health Organization. URL: https://www.who.int/westernpacific/newsroom/feature-stories/item/reaching-the-unreached-in-the-western-pacific [accessed 2025-12-18]
9. Sekarrini CE, Sumarmi, Bachri S, Taryana D, Giofandi EA. The application of geographic information system for dengue epidemic in Southeast Asia: a review on trends and opportunity. J Public Health Res 2022 Jul;11(3):22799036221104170. [doi: 10.1177/22799036221104170] [Medline: 35911430]
10. Pu Q, Yoo EH, Rothstein D, Cairo S, Malemo L. Improving the spatial accessibility of healthcare in North Kivu, Democratic Republic of Congo. Appl Geogr 2020 Aug;121(4):102262. [doi: 10.1016/j.apgeog.2020.102262]
11. Dulin MF, Ludden TM, Tapp H, et al. Using geographic information systems (GIS) to understand a community's primary care needs. J Am Board Fam Med 2010;23(1):13-21. [doi: 10.3122/jabfm.2010.01.090135] [Medline: 20051538]
12. Brakema EA, van der Kleij RM, Poot CC, et al. Mapping low-resource contexts to prepare for lung health interventions in four countries (FRESH AIR): a mixed-method study. Lancet Glob Health 2022 Jan;10(1):e63-e76. [doi: 10.1016/S2214-109X(21)00456-3] [Medline: 34919858]
13. Bayesian network approach. EPCON. URL: https://www.epcon.ai/bayesiannetworkapproach [accessed 2026-08-06]
14. Alege A, Hashmi S, Eneogu R, et al. Effectiveness of using AI-driven hotspot mapping for active case finding of tuberculosis in Southwestern Nigeria. Trop Med Infect Dis 2024 Apr 29;9(5):99. [doi: 10.3390/tropicalmed9050099] [Medline: 38787032]
15. Electronic case based surveillance: using predictive modelling and real-time data to plan active TB case finding in Pakistan. KIT. URL: https://www.kit.nl/institute/project/electronic-case-based-surveillance-using-predictive-modelling-and-real-time-data-to-plan-active-tb-case-finding-in-pakistan/ [accessed 2025-12-18]
16. MacQueen JB. Some methods for classification and analysis of multivariate observations. In: Le Cam LM, Neyman J, editors. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability: University of California Press; 1967:281-297.
17. Colston JM, Burgert CR. Using geospatial analysis to inform decision making in targeting health facility-based programs: a guidance document. : MEASURE Evaluation; 2014 URL: https://www.measureevaluation.org/resources/publications/ms-14-88.html [accessed 2026-08-06]
18. Social determinants of health. World Health Organization. URL: https://www.who.int/health-topics/social-determinants-of-health [accessed 2025-12-16]
19. Amanya G, Hashmi S, Stow JS, et al. AI-driven tuberculosis hotspot mapping to optimize active case-finding: implementing the Epi-control platform in Uganda. Trop Med Infect Dis 2026;11(2):36. [doi: 10.3390/tropicalmed11020036] [Medline: 41746006]
20. Scott LE, Shapiro AN, Da Silva MP, et al. Integrating molecular diagnostics and GIS mapping: a multidisciplinary approach to understanding tuberculosis disease dynamics in South Africa using Xpert MTB/RIF. Diagnostics (Basel) 2023 Oct 10;13(20):3163. [doi: 10.3390/diagnostics13203163] [Medline: 37891984]
21. Evans MV, Andréambeloson T, Randriamihaja M, et al. Geographic barriers to care persist at the community healthcare level: evidence from rural Madagascar. PLOS Glob Public Health 2022;2(12):e0001028. [doi: 10.1371/journal.pgph.0001028] [Medline: 36962826]
22. Brakema EA, van der Kleij RM, Poot CC, et al. A systematic approach to context-mapping to prepare for health interventions: development and validation of the SETTING-tool in four countries. BMJ Glob Health 2021 Jan;6(1):e003221. [doi: 10.1136/bmjgh-2020-003221] [Medline: 33436446]

Abbreviations

DHIS2: District Health Information Software 2
DHS: Demographic and Health Surveys
GIS: geographic information system
HDX: Humanitarian Data Exchange
IHME: Institute for Health Metrics and Evaluation
MCH: maternal and child health
NGO: nongovernmental organization

XSL•FO
RenderX

RTU: Reaching the Unreached
TB: tuberculosis
WASH: water, sanitation, and hygiene
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 13.Feb.2026; peer-reviewed by CM Ramirez, JK Kipterer; revised version received 20.May.2026; accepted 30.May.2026; published 26.Aug.2026.

Please cite as:
Hashmi S, Budts AL, Meurrens V, Stow JS, Morishita F, Yadav RP
Reimagining Data for Equity: Reaching the Unreached Map Explorer for Localized Public Health Intelligence in the Western Pacific Region
JMIR Public Health Surveill 2026;12:e93324
URL: https://publichealth.jmir.org/2026/1/e93324
doi:10.2196/93324

© Sumbul Hashmi, Anne-Laure Budts, Vincent Meurrens, Jessica Sarah Stow, Fukushi Morishita, Rajendra-Prasad Yadav. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 26.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Beyond Training: Systems Framework for Sustainable Health Informatics Investment in Africa

Eric Nzirakaindi Ikoona[1], MPH, MMED, MD, PhD; Lucy Namulemo[2,3], PhD; Ronald Kaluya[2], PhD; Rebecca Ikoona[4], DPharm; Foday Sahr[1], MD, PhD

[1]National Public Health Agency, 42A Main Motor Road, Wilberforce, Freetown, Western Area, Sierra Leone

[2]Uganda Counselling and Support Services, Kampala, Uganda

[3]Foothills Community-Based Interventions, Monticello, KY, United States

[4]Makerere University–Johns Hopkins Research Collaboration, Kampala, Uganda

Corresponding Author:
Eric Nzirakaindi Ikoona, MPH, MMED, MD, PhD
National Public Health Agency, 42A Main Motor Road, Wilberforce, Freetown, Western Area, Sierra Leone

Abstract

Across Africa, substantial investment has built national health information systems (HISs), including the surveillance platforms, reporting tools, and digital infrastructure through which health data flow. However, the health informatics capacity needed to sustain those systems remains fragile: trained health informaticians leave for better-resourced organizations, platforms fall into disrepair when donor funding ends, and data systems multiply without connecting. Health informatics is the discipline that designs, governs, and sustains HISs; the two are inseparable, and investment that builds one while neglecting the other cannot produce durable results. This viewpoint, grounded in the authors' direct implementation experience across sub-Saharan Africa and informed by published literature, argues that the cause is a structurally misaligned investment logic: resources concentrate on training health informaticians while the institutional, governance, and infrastructure conditions that determine whether those informaticians can perform are chronically underfunded. We propose that sustainable health informatics capacity requires a functioning HIS ecosystem and that this ecosystem rests on four interdependent pillars: (1) workforce development beyond training; (2) institutional strengthening; (3) governance and data standards; and (4) interoperable infrastructure, including the national HIS platforms, such as the District Health Information Software 2, through which health informatics practice operates. The pillars are interdependent: weakness in any one undermines the others in specific, predictable ways. The framework makes 3 contributions beyond existing World Health Organization (WHO) and Africa Centers for Disease Control and Prevention frameworks: it shows how weakness in one pillar actively undermines the others; it focuses specifically on health informatics and national HISs as the primary investment domain rather than as components of a broader digital health agenda; and it treats the enabling environment, specifically civil service structures, domestic financing, data sovereignty, and artificial intelligence governance, as core investment requirements rather than background conditions. Drawing on illustrative experiences from Ethiopia, Kenya, and Sierra Leone, we show how health informatics capacity succeeds or fails depending on the strength of all 4 conditions. A practical readiness checklist and audience-specific policy recommendations are provided for national health ministries, national public health institutes, regional bodies, and development partners.

(JMIR Public Health Surveill 2026;12:e89482) doi:10.2196/89482

KEYWORDS

digital health; health informatics; health information systems; District Health Information Software 2; DHIS2; workforce development; health systems strengthening; Africa; public health surveillance; artificial intelligence

Introduction

Over the past decade, African governments and development partners have invested substantially in digital health. More than 40 nations reportedly use the District Health Information Software 2 (DHIS2) for routine health information management [1,2], and health workers—from rural facilities to national disease control centers—now use digital platforms such as DHIS2-based surveillance modules, electronic disease notification systems, and mobile reporting tools to track outbreaks, monitor immunization coverage, and coordinate emergency responses [3,4]. The infrastructure investment is real and significant, but the capacity to sustain and use it effectively has not kept pace. This gap, between platforms built and platforms functioning, is the problem this paper addresses.

Understanding why requires a precise distinction between 2 related but different concepts. A health information system (HIS) is the operational infrastructure of a health system—the surveillance platforms, electronic health records, reporting tools, registries, and dataflows through which health data are

XSL•FO
RenderX

generated, managed, and used. DHIS2, the most widely deployed HIS platform in Africa, is an example. Health informatics is the discipline that designs, governs, and continuously improves those systems—the trained workforce, institutional arrangements, data standards, interoperability frameworks, and governance rules that determine whether an HIS functions well or poorly. An HIS is the system. Health informatics is what makes the system work. The relationship between them is the paper's central concern: health informatics capacity must be embedded within a functioning HIS ecosystem, and a functioning HIS ecosystem requires health informatics capacity to sustain it. Investment that treats them as separate problems produces neither.

However, the dominant investment logic does not reflect this relationship. Across Africa, health informatics investment concentrates on training: initiatives such as Africa Centers for Disease Control and Prevention's (CDC's) Africa Epidemic Services-Public Health Informatics (AES-PHI) fellowship produce competency-trained graduates and directly address workforce shortages [5,6]. This is necessary, but training produces individuals, not systems. A trained informatician placed in an institution without a funded post, a governance mandate, or functional HIS infrastructure cannot perform. The informatician and the system each need the other, and investment that builds one while neglecting the other cannot produce sustainable health informatics capacity. Existing frameworks recognize that workforce, governance, and infrastructure are all important: the World Health Organization (WHO), the World Bank, and Africa CDC have consistently identified them as interdependent requirements for sustainable health system performance [2,5,7-9], but they treat these elements as parallel components rather than as interdependent conditions whose misalignment is itself the cause of failure. What has been missing is a framework that explains, specifically for health informatics, how misalignment across these conditions limits the translation of training investment into functioning national HISs.

We therefore propose a 4-pillar systems framework for sustainable health informatics investment in Africa, one that goes beyond training to address the full set of conditions a health informatician needs to perform. The 4 pillars are workforce development beyond training, institutional strengthening, governance and data standards, and interoperable infrastructure. The framework is designed for 4 audiences: national health ministries responsible for health informatics investment decisions, national public health institutes (NPHIs) that serve as institutional anchors for surveillance and health informatics practice, regional bodies such as Africa CDC that set standards and coordinate capacity, and development partners whose financing decisions shape what gets built and what gets sustained.

From Paper to Platform: The African Digital Health Journey

Overview

For decades, African health systems ran on paper, with data arriving late, incomplete, and largely unused. The 2014 to 2016 Ebola outbreak exposed these gaps and accelerated investment in HISs across the region [1,2]. DHIS2, developed by the Health Information Systems Program at the University of Oslo, became the platform of choice: by 2020, over 40 African countries reportedly had national DHIS2 platforms supporting routine health information, surveillance, or both [1,10,11].

The country experiences mentioned under the *Ethiopia*, *Kenya*, and *Sierra Leone* sections are illustrative, not a systematic review. Their purpose is to show how the relative strength or weakness of specific enabling conditions shapes health informatics capacity. Those conditions map to the 4 pillars of the framework presented in the section that follows.

Ethiopia

Ethiopia's nationwide DHIS2 rollout, combined with the electronic community HIS, formed part of a sustained federal investment in digital health embedded in successive national health sector transformation plans [12]. An interrupted time series study subsequently reported improvements in maternal and child health indicators, documenting gains in data completeness and reporting timeliness [13]. In the authors' assessment, Ethiopia's trajectory illustrates an uneven pillar profile. Pillars 1 and 4 were relatively strong: the country built substantial informatics workforce capacity and invested in interoperable infrastructure. Pillars 2 and 3 lagged: harmonizing parallel vertical systems required cross-ministry mandates and consolidated data standards that took much longer to establish. The result was a system where skills and infrastructure existed but institutional coordination and governance frameworks were absent. This constrained comprehensive integration. It is a direct illustration of why pillar 4 investment alone cannot compensate for weak pillars 2 and 3.

Kenya

Kenya offers a contrasting trajectory. The 2011 to 2012 national DHIS2 rollout documented investment across all 4 enabling conditions: informatics workforce with defined institutional roles (pillar 1), coordination structures linking district and national health system levels (pillar 2), frameworks governing data access and use (pillar 3), and the technical infrastructure to connect them (pillar 4) [14]. The result was a more integrated national platform than comparator settings where investment was concentrated on a single domain [14,15]. The contrast with Ethiopia is not about resources, it is about alignment: investing across all 4 conditions simultaneously produced better integration than concentrating on 1 or 2 conditions.

Sierra Leone: Crisis-Driven Innovation and Systems Learning

Sierra Leone did not plan its way to building health informatics capacity; crisis forced it. The 2014 to 2016 Ebola epidemic exposed acute gaps in surveillance and coordination. It drove

XSL•FO
RenderX

rapid investment from the Ministry of Health and Sanitation, the US CDC, WHO, and partners [16,17]. In 2019, Sierra Leone became the first country in the WHO African region to achieve nationwide electronic health facility reporting of surveillance data [16].

The electronic Integrated Disease Surveillance and Response system, built on DHIS2, now operates across all 16 districts and more than 1500 health facilities [16,17]. Completeness of weekly IDSR reports rose from 68% in 2015 to above 98% in 2023, according to national DHIS2 data in the peer-reviewed literature [18]; these figures reflect reported national routine data and should be interpreted accordingly. Four enabling conditions drove this improvement: a trained surveillance workforce, a strengthened National Public Health Agency with a statutory mandate, standardized data definitions and reporting protocols, and a hybrid infrastructure combining SMS text messaging, mobile data entry, and web interfaces to accommodate varied connectivity [16,19].

When the COVID-19 pandemic arrived, Sierra Leone did not build new systems. National teams adapted the existing DHIS2-based platform for surveillance, contact tracing, and laboratory modules [11]. Precrisis investment in platforms and workforce enabled a faster, more coordinated response than would otherwise have been possible.

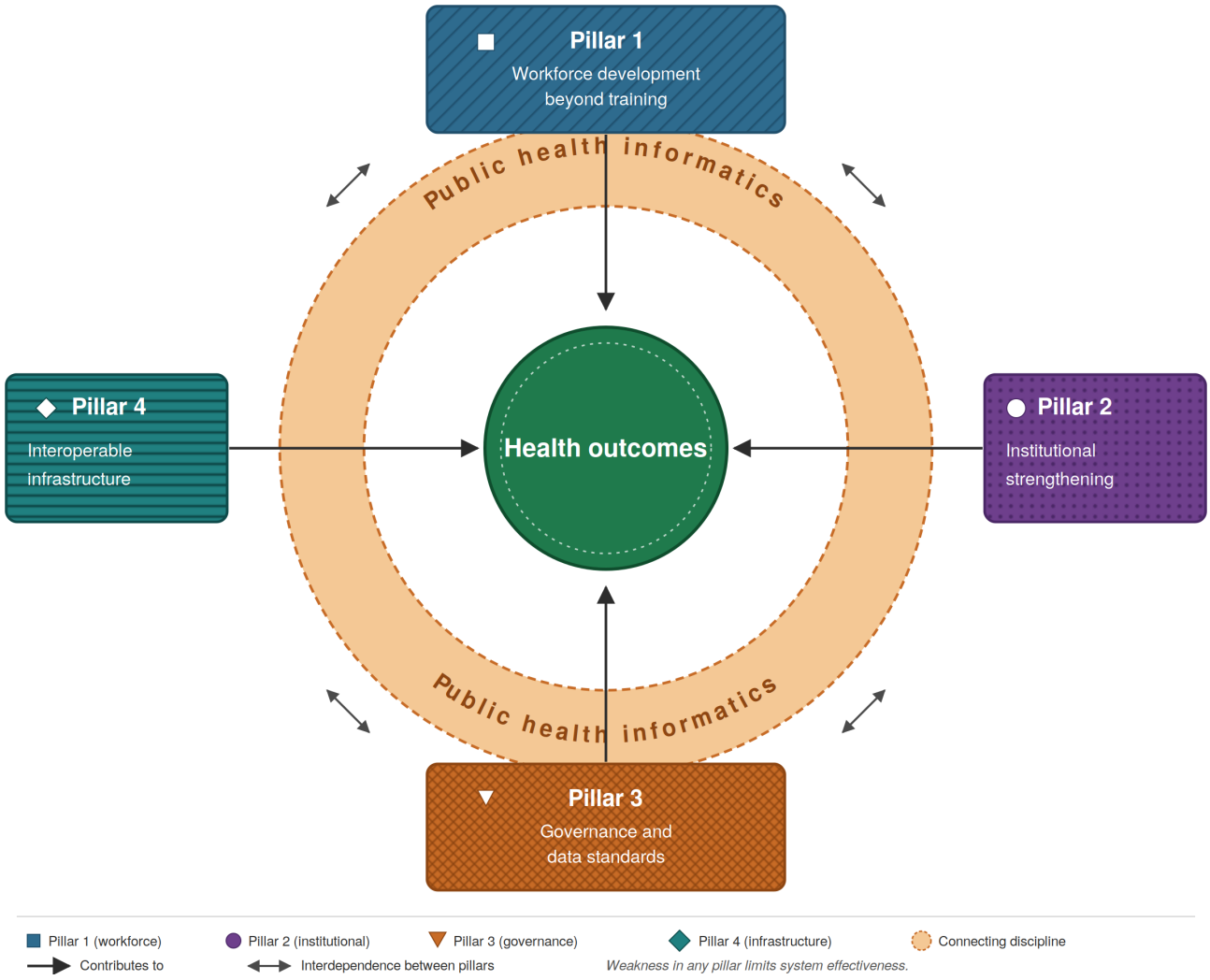
Two subsequent emergencies further tested this infrastructure. A synthetic cannabinoid crisis (locally called Kush) and concurrent mpox outbreaks [19,20] required rapid multisectoral coordination for which no dedicated parallel system existed. The National Public Health Agency coordinated responses using real-time DHIS2 data [19,20]—no new system was built. This is the clearest demonstration the Sierra Leone experience offers: when all 4 pillars are in place, routine HIS infrastructure can absorb novel threats without emergency workarounds.

A Systems Framework for Sustainable Health Informatics Investment

Overview

The country experiences above illustrate what happens when some of these conditions are in place and others are not. The framework shown in Figure 1 makes those conditions explicit and maps their interdependencies. Each pillar is addressed in the subsections below; the failure mechanisms linking them are addressed in the following section.

Figure 1. A systems framework for sustainable public health informatics investment in Africa. Four interdependent pillars (workforce development, institutional strengthening, governance and data standards, and interoperable infrastructure) converge on health outcomes. Public health informatics (orange ring) is the connecting discipline. Bidirectional arrows denote interdependence; weakness in any pillar limits system effectiveness.



XSL•FO
RenderX

Pillar 1: Workforce Development Beyond Training

Health informatics cannot function without trained practitioners. Africa CDC's workforce framework identifies health informatics as a priority domain [5], and the AES-PHI fellowship is one of the most important mechanisms for building this cadre [6]. However, training is the starting point, not the destination.

A stable workforce requires more than a training certificate. It requires formal civil service career pathways within ministries of health and NPHIs, salary structures that reduce attrition to the private sector and international organizations, structured mentoring, continuing professional development programs, and cross-border learning networks [21].

When these conditions are absent, fellowship graduates can migrate to better-resourced organizations, leaving the institutions that trained them without the human capital required to execute their mandates. This dynamic directly undermines pillar 2: institutions lose their most capable staff before those staff can build the routines and practices that make the institution function.

Pillar 2: Institutional Strengthening

Trained individuals cannot act without institutions that give them authority, resources, and a mandate. The gap between institutional presence and institutional capacity is well documented. Eighty percent of African countries have established a public health emergency operations center (PHEOC). However, fewer than one-third meet most core functionality criteria [22]. A structure on paper and a functioning institution are not the same thing.

The WHO–Africa CDC PHEOC framework establishes that functional operations centers require coherent policies, defined data systems, and stable skilled staffing working together—not physical infrastructure alone [23]. Clear mandates, standard operating procedures, and stable funding are the conditions under which data and skills can translate into coordinated action. Without them, neither does.

NPHIs are the natural institutional anchors for health informatics practice. They combine a statutory public health mandate with technical independence from line ministries, and can host surveillance systems, PHEOCs, and analytic units under a single roof [21,24]. In many settings, the binding constraint is not talent but the absence of a dedicated domestic budget line. Health informatics work that depends on donor funding cycles rarely aligns with national planning horizons. When grants end, the capacity they supported is at risk.

When institutional strengthening is weak, informaticians trained under pillar 1 can lack the resource authority, legal mandate, and budget to act. Their technical skills risk being stranded. Without an institutional anchor, governance frameworks and data-sharing agreements may remain aspirational: rules without the authority to enforce them.

Pillar 3: Governance and Data Standards

Data without governance are a liability, not an asset. Across Africa, many countries lack enforceable rules for cross-program data access, data-sharing agreements, and privacy protections [25]. Fragmented vertical systems, each funded by a separate donor project, complicate data harmonization and increase reporting burden on health workers. The WHO Regional Office for Africa identifies governance as a core HIS building block, and these gaps remain pervasive across the region [7]. When standards are absent and governance is unenforced, pillar 4 infrastructure defaults to program-specific specifications. Those specifications cannot be shared across systems, reproducing exactly the fragmentation that coordinated investment was intended to resolve.

The governance imperative extends to artificial intelligence (AI). AI tools for outbreak detection, clinical decision support, and resource allocation are being deployed across African health systems—in many cases ahead of the governance frameworks needed to guide them [26,27]. Without proactive national policies, the risks are concrete: AI data stewardship roles go undefined, consent requirements remain unspecified, and parameters for cross-border data sharing are left to external actors. These tools can then encode existing biases and concentrate analytical control outside the continent. Governance is not a constraint on informatics development. It is what makes that development safe, equitable, and nationally accountable.

Pillar 4: Interoperable Infrastructure

Infrastructure is not a background condition. It is an active determinant of whether the other three pillars can function. The Open Health Information Exchange architecture provides a widely adopted blueprint, defining components for client registries, facility registries, terminology services, and interoperability layers [28]. Sierra Leone's Integrated Disease Surveillance and Response experience illustrates how hybrid infrastructure design can accommodate real-world connectivity constraints while feeding a single national platform, combining SMS text messaging, offline data entry, and web interfaces.

Infrastructure investments must address reliable power and connectivity for rural facilities, hardware procurement and maintenance plans, interoperability layers connecting DHIS2 with laboratory and logistics systems, and cybersecurity protections. When infrastructure is unreliable or noninteroperable, health workers can revert to paper, data gaps accumulate, and rural facilities may effectively drop out of national surveillance and resource allocation. The professional confidence of the informatics workforce can erode, reinforcing the attrition that pillar 1 works to prevent.

Relation to Existing Frameworks and New Contribution

This framework builds on an established conceptual lineage. The WHO health systems building blocks [29] identified six interdependent foundations of functioning health systems: health workforce, health information, financing, medical products, leadership and governance, and service delivery. It has shaped health systems discourse for nearly two decades [29]. The Africa CDC workforce framework [5] and the WHO framework for implementing the global strategy on digital health [7] apply that logic to workforce development and digital health ecosystems respectively. These frameworks establish the consensus that no

XSL•FO
RenderX

single investment domain is sufficient. The contribution of this paper is not to identify new components of digital health systems. It is to explain how misalignment across these components limits the translation of health informatics investment into functioning national HISs. None of the existing frameworks specifies how weakness in one domain actively undermines investment in another or which direction that failure travels. None is scoped specifically to health informatics as the primary investment domain, and none treats the enabling environment, specifically civil service structures, domestic funding models, data sovereignty, and AI governance, as core investment requirements rather than supplementary considerations.

Against those three gaps, this framework makes three specific advances. First, it frames health informatics as the primary investment domain and treats HIS platforms as the core infrastructure through which that discipline operates, rather than as one component among many in a broader digital health agenda. Second, it explains how weakness in one pillar undermines the others: an informatician placed in an institution without a mandate, budget, or functional infrastructure cannot perform, regardless of the quality of their training; failure pathways run in both directions and explain why single-pillar investments so frequently underperform. Third, it treats the enabling environment, specifically civil service structures, domestic financing, data sovereignty, and AI governance, as core investment requirements rather than background conditions. Existing frameworks name these as context. This framework treats them as investment objects without which the other pillars cannot be sustained. Table 1 maps these gaps and contributions systematically.

Table . Comparison of this framework with existing Africa Centers for Disease Control and Prevention (CDC) and World Health Organization (WHO) frameworks.

Feature	WHO building blocks [29]; Africa CDC and WHO digital health frameworks [5,7]	This framework
Core domains identified	Six building blocks: health workforce, health information, financing, medical products, leadership or governance, and service delivery [29]; adapted to digital health as workforce, governance, infrastructure, and human resources [5,7]	Same 4 domains, explicitly named as pillars with defined interdependencies
Failure mechanisms	Building blocks note interdependence but do not map failure pathways; digital health frameworks list domains without interaction logic or sequencing guidance	Bidirectional failure pathways mapped (eg, weak pillar 2 strands pillar 1 investment, and absent pillar 3 fragments pillar 4)
Focus object	Broad digital health ecosystem	HIS[a] and public health informatics as primary investment objects; interoperability as shared national asset
Political economy	Not addressed	Civil service career pathways, domestic budget lines, and donor-dependency cycles
Data sovereignty and AI[b] governance	Mentioned generically or absent	Treated as core governance requirements with operational implications for platform hosting and policy review cycles
Equity dimensions	Not systematically addressed	Urban-rural connectivity gaps, gender-stratified access, and frontline worker inclusion in governance
Investment sequencing guidance	Not provided	Readiness checklist enables pillar-by-pillar gap assessment and prioritization

[a]HIS: health information system.

[b]AI: artificial intelligence.

How the pillars undermine each other follows predictable patterns. Weak institutional mandates (pillar 2) strand trained informaticians (pillar 1): without a funded post or decision-making authority, a fellowship graduate has no mechanism through which to act. Fragmented data governance (pillar 3) compounds this: absent standards mean workforce competencies trained for one program cannot transfer across systems. Absent governance also fragments infrastructure (pillar 4): platforms built without enforced interoperability standards cannot be connected, reproducing the vertical silos investment was meant to resolve. Moreover, as infrastructure scales, governance must keep pace: AI tools and cross-border dataflows that go unaddressed convert new technical capacity into new institutional risk. The failure runs in reverse too: unreliable infrastructure erodes the working environment of the informatics workforce, reinforcing the attrition that pillar 1 investment works to prevent. These are the patterns through which single-pillar investments so frequently underperform.

As the pillars depend on each other, the order in which they are strengthened matters. When a country trains health informaticians before establishing the institutional structures that give them a post and a mandate, the graduates leave. When it builds HIS infrastructure before putting governance and data standards in place, platforms become program-specific silos that cannot communicate. In low-baseline settings, pillar 2 must

XSL•FO
RenderX

come first: without a mandated institution with a budget and defined governance roles, neither workforce retention nor data governance enforcement is achievable. Pillar 3 should precede major pillar 4 investment for the same reason. In higher-baseline settings where pillars 2 and 3 are functional, pillars 1 and 4 can advance together. No sequence is universal: the readiness checklist in Table 2 provides a tool for identifying which pillar is the binding constraint in a given country context.

Table . Minimum enabling conditions, example indicators, and sequencing priority by pillar. Sequencing priority indicates whether a pillar is a prerequisite for others (must be established first), concurrent (can be developed in parallel), or dependent (requires other pillars to be functional first).

Pillar	Minimum enabling conditions	Example indicators	Sequencing priority
Pillar 1: workforce development	Civil service posts for informaticians; retention incentives; CPD[a] programs; structured mentoring	Vacancy rate; turnover rate; CPD completion rate; fellowship-to-post conversion rate	Concurrent in high-baseline settings; dependent on pillar 2 in low-baseline settings (no post=no retention)
Pillar 2: institutional strengthening	NPHI[b] or PHEOC[c] with statutory mandate; dedicated domestic budget line; clear HIS[d] governance roles	PHEOC functionality score; budget share for digital health; SOP[e] coverage for core functions	Prerequisite in most settings. Without mandate and budget, pillars 1, 3, and 4 lack an authorizing home
Pillar 3: governance and data standards	Approved national digital health policy; DSAs[f]; AI[g] governance principles; interoperability standards	Number of signed DSAs; data breach response time; AI use case registry; proportion of systems using national standards	Prerequisite before major pillar 4 investment. Infrastructure without governance defaults to vertical silos
Pillar 4: interoperable infrastructure	National HIS platform; interoperability layer connecting DHIS2, laboratory, and logistics systems; backup and uptime plan; cybersecurity protections	System uptime (%); API[h] transactions per month; facility connectivity rate; backup test success rate	Dependent on pillars 2 and 3. In high-baseline settings, simultaneous investment with pillar 1 retention pathways is optimal

[a]CPD: continuing professional development.
[b]NPHI: national public health institute.
[c]PHEOC: public health emergency operations center.
[d]HIS: health information system.
[e]SOP: standard operating procedure.
[f]DSA: data-sharing agreement.
[g]AI: artificial intelligence.
[h]API: application programming interface.

Practical Checklist: Minimum Enabling Conditions by Pillar

Table 2 provides indicative minimum enabling conditions, example indicators, and sequencing priority for each pillar. These are intended as a starting point for government and partner readiness assessments, not as a prescriptive checklist.

Implementation Challenges, Counterarguments, and Scope

This framework is analytically coherent. Implementing it is another matter. Politically, health ministries and development partners favor visible, countable outputs: training certificates and installed hardware. Systemic investments such as civil service reform and interoperability standards are harder to measure and harder to fund. Financially, sustaining recurring platform costs remains a persistent challenge: the funding landscape for disease surveillance digitalization in Africa is fragmented and predominantly externally driven [30]. Logistically, coordinating across ministries, regulatory bodies, and donor programs demands sustained political will and a named institutional owner with authority to enforce standards.

None of these obstacles is insurmountable. Ministries can build the political case for systemic investments by linking them to outcomes that are already valued: demonstrating that interoperable platforms reduced outbreak response times, for instance, makes the case for infrastructure investment in terms that existing accountability frameworks can register. Governments can pool regional resources to distribute platform maintenance costs and establish domestic budget lines to cofinance core infrastructure, reducing donor dependency over time.

The framework carries inherent limitations. It is grounded in published literature and the authors' direct implementation experience in a limited set of East and West African countries. Evidence for specific interventions is uneven, implementation contexts vary substantially across the continent, and the framework has not been prospectively validated. Country examples are illustrative; causal claims should not be drawn from them. Future empirical work could test and refine the framework through comparative implementation studies, workforce retention analyses linking civil service post availability to fellowship graduate outcomes, and validation of the pillar-based readiness indicators in Table 2.

A genuine counterargument deserves acknowledgment. Technology-first approaches have achieved real gains: rapid DHIS2 deployment during the Ebola response created

XSL•FO
RenderX

surveillance infrastructure that proved adaptable to subsequent crises [11,16,17]. There are settings where building the platform first was more pragmatic than waiting for all enabling conditions to be in place. This framework does not argue against emergency or opportunistic investment. It argues that such investment must be followed, as rapidly as possible, by the institutional and governance conditions that prevent the gains from remaining fragile. Sierra Leone's experience is illustrative precisely because the crisis-driven infrastructure investment was followed by deliberate strengthening of all 4 pillars.

Equity and Political-Economy Considerations

Digital health transformation does not occur in a neutral space. Within countries, urban facilities are typically connected before rural and remote clinics, compounding existing inequities in health system access [30,31]. Language barriers and gender norms shape who is recruited into informatics training programs and who benefits from digital tools [2,31]. Without deliberate corrective investment, digital health rollout can widen rather than narrow health system inequity. A system optimized purely for efficiency, connecting well-resourced facilities and training the most accessible graduates, will generate impressive aggregate indicators while leaving the most vulnerable populations underserved.

At the global level, data ownership and sovereignty are actively contested. External partners frequently host data or analytic platforms in servers outside the country, limiting national autonomy and generating well-founded concerns about data extraction [30,31]. Funding patterns that favor short-term pilots over sustained system investment perpetuate fragmentation and external dependency. Treating data sovereignty as a core pillar 3 governance requirement means mandating local hosting, national data stewardship roles, and African-led AI oversight. Development partners committed to African health security should align their platform-hosting and financing decisions accordingly.

Regional Integration and Cross-Border Collaboration

Infectious diseases do not respect national borders, and neither can HISs. Africa CDC's coordinating role during the COVID-19 pandemic demonstrated the value of shared platforms and regional institutional capacity [24], and the agency continues to support PHEOC development, cross-border surveillance networks, and common standards for laboratory and surveillance systems [5,21]. Each pillar has a cross-border dimension: regionally coordinated workforce development reduces brain-drain from smaller countries, harmonized governance and data standards lower the costs of cross-border data exchange, and common infrastructure standards allow national platforms to communicate during outbreaks without bespoke interoperability work for each event. These regional dimensions are not optional extensions of the framework. They are the conditions under which the framework can operate at the scale that continental health security requires.

Policy Implications and Recommendations

Translating the framework into action requires coordinated steps from governments, regional bodies, and development partners.

National governments have 4 priority actions. First, establish civil service career tracks for health informaticians with defined job descriptions, promotion pathways, and competitive salaries, as the primary mechanism for reducing postfellowship attrition [5,6,21]. Second, strengthen NPHIs and PHEOCs as institutional anchors for surveillance and digital health coordination, with investment in mandate, budget, and staffing capacity, not structure alone [22]. Third, establish dedicated domestic budget lines for national HIS platforms and infrastructure maintenance, treating interoperability as a shared public good [30]. Fourth, develop or update national digital health strategies to integrate HIS governance, AI oversight, equity indicators, and data sovereignty provisions, reviewed at least every 3 to 5 years [7,8,25].

Regional bodies have three priority actions. First, expand training programs in public health informatics with strong placement linkages to NPHIs and PHEOCs, building on the AES-PHI model to ensure graduates enter funded posts rather than unfilled institutional vacancies [5,6,21]. Second, promote harmonized standards for cross-border data exchange, privacy protection, and AI ethics, building on WHO and Open Health Information Exchange guidance [7,25,28]. Third, create peer-learning networks for technical platform maintenance, allowing countries to share infrastructure expertise and avoid duplicating the fixed costs of building national platform capacity independently.

Development partners have 4 priority actions. First, shift financing from parallel, program-specific information systems toward long-term support for national platforms and the interoperability standards that connect them [1,11,30]. Second, fund civil service post creation and workforce retention incentives alongside fellowship training. The return on training investment depends entirely on the institutional conditions that allow graduates to remain and act [5,6]. Third, condition infrastructure grants on the prior existence of a governance framework and a named institutional owner, so that pillar 4 investment does not precede the pillar 2 and pillar 3 conditions required to sustain it [23,25]. Fourth, adopt multipillar readiness assessments before designing new digital health programs; the checklist in Table 2 provides a starting framework for identifying which enabling conditions are weakest and sequencing investments accordingly.

Conclusions

Africa has built something real. Post-Ebola and post–COVID-19 pandemic investments have created a foundation of surveillance systems, PHEOCs, and digital platforms that did not exist a decade ago. The gains are documented and significant, but they are also fragile. Sustaining and expanding them requires a shift not in effort but in logic: the goal is not to produce trained

XSL•FO
RenderX

individuals but to build the enabling ecosystems in which those individuals can act.

What makes the framework useful is not the list of pillars but the logic connecting them. Each pillar creates the conditions the next one requires: workforce development produces the capacity, institutional strengthening gives it authority and resources, governance and data standards direct how that capacity is used, and interoperable infrastructure ensures it operates within a connected, nationally owned system. When any one is missing, the system does not merely underperform: it generates a category of failure that investment in the remaining pillars cannot compensate for. That is the core analytical insight this paper advances.

Empirical validation is the essential next step. Longitudinal studies tracking the retention and system-level impact of health informatics fellowship graduates would test the core pillar 1–pillar 2 interdependency. Multicountry comparisons of divergent pillar-investment trajectories would assess whether coordinated multipillar investment produces better outcomes than single-pillar approaches. Implementation research should examine whether the readiness checklist in Table 2 predicts health informatics system performance across diverse country contexts.

Public health informatics in Africa is not merely a technical enterprise; it is the information infrastructure on which evidence-based health governance depends. African institutions that own their data, have the workforce to analyze them, and operate within governance frameworks that protect and direct their use are better positioned to set their own health priorities and hold their own systems accountable. The 4-pillar framework is offered as a practical tool for building that capacity systematically, equitably, and durably.

Acknowledgments

All authors declared that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee (APF) support for the publication of this article. Generative artificial intelligence was not used in any portion of the drafting or generation of this manuscript.

Funding

No funding was received for this viewpoint paper.

Data Availability

No datasets were generated or analyzed for this viewpoint paper.

Authors' Contributions

ENI conceptualized the manuscript, wrote the original draft, and led revisions. LN, RK, and RI provided critical review and contributed to drafting sections. FS contributed to conceptualization, critical review, and contextualization for Sierra Leone and African regional structures. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Byrne E, Sæbø JI. Routine use of DHIS2 data: a scoping review. BMC Health Serv Res 2022 Oct 6;22(1):1234. [doi: 10.1186/s12913-022-08598-8] [Medline: 36203141]
2. Victor AA, Frank LJ, Makubalo LE, Kalu AA, Impouma B. Digital health in the African region should be integral to the health system's strengthening. Mayo Clin Proc Digit Health 2023 Sep;1(3):425-434. [doi: 10.1016/j.mcpdig.2023.06.003] [Medline: 37744952]
3. Adegoke OJ, Rachlin A, Porter AM, et al. Migration from Epi Info to District Health Information Software 2 for vaccine-preventable disease surveillance - World Health Organization African Region, 2019-2023. MMWR Morb Mortal Wkly Rep 2024 Jun 13;73(23):529-533. [doi: 10.15585/mmwr.mm7323a2] [Medline: 38870469]
4. Silenou BC, Nyirenda JL, Zaghloul A, et al. Availability and suitability of digital health tools in Africa for pandemic control: scoping review and cluster analysis. JMIR Public Health Surveill 2021 Dec 23;7(12):e30106. [doi: 10.2196/30106] [Medline: 34941551]
5. Framework for public health workforce development, 2020-2025. Africa Centres for Disease Control and Prevention. URL: https://africacdc.org/download/framework-for-public-health-workforce-development-2020-2025/ [accessed 2026-05-14]
6. The Africa Epidemic Services – Public Health Informatics track. Africa Centres for Disease Control and Prevention. URL: https://africacdc.org/africa-epidemic-services-public-health-informatics-track/ [accessed 2026-05-14]
7. Framework for implementing the global strategy on digital health in the WHO African Region. World Health Organization Regional Office for Africa. 2021. URL: https://www.afro.who.int/sites/default/files/2021-07/

XSL•FO
RenderX

AFR-RC71-10%20Framework%20for%20implementing%20the%20Global%20strategy%20on%20digital%20health%20in%20the%20WHO%20African%20Region.pdf [accessed 2026-05-14]
8. Global strategy on digital health 2020-2025. World Health Organization. 2021. URL: https://www.who.int/docs/default-source/documents/gs4dhdaa2a9f352b0445bafbc79ca799dce4d.pdf [accessed 2026-05-14]
9. Digital-in-health: unlocking the value for everyone. : World Bank Group; 2023 URL: https://www.worldbank.org/en/topic/health/publication/digital-in-health-unlocking-the-value-for-everyone [accessed 2026-05-14]
10. DHIS2 software releases. DHIS2. URL: https://dhis2.org/overview/ [accessed 2026-05-14]
11. Kinkade C, Russpatrick S, Potter R, et al. Extending and strengthening routine DHIS2 surveillance systems for COVID-19 responses in Sierra Leone, Sri Lanka, and Uganda. Emerg Infect Dis 2022 Dec;28(13):S42-S48. [doi: 10.3201/eid2813.220711] [Medline: 36502427]
12. Health sector transformation plan II (HSTP II) 2020/21-2024/25. Ministry of Health, Ethiopia. 2021. URL: https://extranet.who.int/countryplanningcycles/sites/default/files/public_file_rep/ETH_Ethiopia_Health-Sector-Transformation-Plan-II_2021-2026.pdf [accessed 2026-05-14]
13. Zerfu TA, Asressie M, Tareke AA, et al. Contributions of District Health Information Software 2 (DHIS2) to maternal and child health service performance in Ethiopia: an interrupted time series mixed-methods study. Arch Public Health 2025 Jul 2;83(1):173. [doi: 10.1186/s13690-025-01641-0] [Medline: 40605114]
14. Manya A, Braa J, Titlestad OH, Øverland LH, Mumo J. National roll out of District Health Information Software (DHIS 2) in Kenya, 2011 – central server and cloud based Infrastructure. In: IST-Africa 2012 Conference Proceedings: IST-Africa; 2016. URL: https://www.semanticscholar.org/paper/National-roll-out-of-District-Health-Information-2)-Manya-Braa/45e26ed166310cfbb6835a6a310340f2b474045b#citing-papers [accessed 2026-05-27]
15. Karuri J, Waiganjo P, Orwa D, Manya A. DHIS2: the tool to improve health data demand and use in Kenya. J Health Informatics Dev Ctries 2014;8(1) [FREE Full text]
16. Sierra Leone leads the way in Africa with fully functional electronic disease surveillance system. World Health Organization. 2019. URL: https://www.afro.who.int/news/sierra-leone-leads-way-africa-fully-functional-electronic-disease-surveillance-system [accessed 2026-05-14]
17. Martin DW, Sloan ML, Gleason BL, et al. Implementing nationwide facility-based electronic disease surveillance in Sierra Leone: lessons learned. Health Secur 2020 Jan;18(S1):S72-S80. [doi: 10.1089/hs.2019.0081] [Medline: 32004124]
18. Magoba B, Kanu JS, Squire JS, et al. Assessing the implementation of an electronic integrated disease surveillance and response system using DHIS2, 2024: Sierra Leone´s experience. J Interv Epidemiol Public Health 2025;8(2):5. [doi: 10.37432/jieph.supp.2025.8.1.13.5]
19. Bridget M, Gebru GN, Odongo GS, et al. Digitalizing disease surveillance: experience from Sierra Leone. Health Policy Plan 2025 Jan 11;40(1):85-96. [doi: 10.1093/heapol/czae039] [Medline: 38813658]
20. Mpox - Democratic Republic of Congo. World Health Organization. 2024. URL: https://www.who.int/emergencies/disease-outbreak-news/item/2024-DON522 [accessed 2026-05-14]
21. Baye T, Taame H, Raji T. Charting a healthier future: Africa CDC's strategic approach towards strengthening the health workforce of the African continent. J Public Health Afr 2023 Oct 31;14(10):2864. [doi: 10.4081/jphia.2023.2864] [Medline: 38020269]
22. Fekadu ST, Gebrewahid AL, Mankoula W, et al. Public health emergency operations centres in Africa: a cross-sectional study assessing the implementation status of core components and areas for improvement, December 2021. BMJ Open 2023 Jun 20;13(6):e068934. [doi: 10.1136/bmjopen-2022-068934] [Medline: 37339838]
23. Handbook for public health emergency operations center operations and management. World Health Organization African Region. URL: https://www.afro.who.int/publications/handbook-public-health-emergency-operations-center-operations-and-management [accessed 2026-05-14]
24. Fallah MP, Raji T, Ngongo AN, et al. The role of Africa Centres for Disease Control and Prevention during response to COVID-19 pandemic in Africa: lessons learnt for future pandemics preparedness, prevention, and response. BMJ Glob Health 2024 Feb 27;9(2):e014872. [doi: 10.1136/bmjgh-2023-014872] [Medline: 38413103]
25. Framework for integrating country and regional health data in the African Region: regional health data hub 2024-2030. World Health Organization Regional Office for Africa. 2024. URL: https://www.afro.who.int/sites/default/files/2024-07/AFR-RC74-7%20Framework%20for%20integrating%20country%20and%20regional%20health%20data%20in%20the%20African%20Region.pdf [accessed 2026-05-14]
26. Ethics and governance of artificial intelligence for health: WHO guidance. World Health Organization. 2021. URL: https://www.who.int/publications/i/item/9789240029200 [accessed 2026-05-14]
27. Adekola AP. Digital health innovations: transforming healthcare access in sub-Saharan Africa. JLSD 2025;12:1-22. [doi: 10.25159/2520-9515/18989]
28. OpenHIE architecture specification. Open Health Information Exchange. 2020. URL: https://ohie.org/wp-content/uploads/2020/12/OpenHIE-Specification-Release-3.0.pdf [accessed 2026-05-14]
29. Everybody's business: strengthening health systems to improve health outcomes: WHO's framework for action. World Health Organization. 2007. URL: https://www.who.int/publications/i/item/everybody-s-business----strengthening-health-systems-to-improve-health-outcomes [accessed 2026-05-14]

XSL•FO
RenderX

30. Kaburi BB, Harries M, Hauri AM, et al. Availability of published evidence on coverage, cost components, and funding support for digitalisation of infectious disease surveillance in Africa, 2003-2022: a systematic review. BMC Public Health 2024 Jun 28;24(1):1731. [doi: 10.1186/s12889-024-19205-2] [Medline: 38943132]
31. Latest documents. African Union. URL: https://au.int/en/documents/20220329/african-union-data-policy-framework [accessed 2026-05-14]

Abbreviations

AES-PHI: Africa Epidemic Services-Public Health Informatics
AI: artificial intelligence
CDC: Centers for Disease Control and Prevention
DHIS2: District Health Information Software 2
HIS: health information system
NPHI: national public health institute
PHEOC: public health emergency operations center
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 12.Dec.2025; peer-reviewed by BB Kaburi, M Fornazin, R Ruparel; revised version received 11.Mar.2026; accepted 23.Mar.2026; published 01.Jun.2026.

Please cite as:
Ikoona EN, Namulemo L, Kaluya R, Ikoona R, Sahr F
Beyond Training: Systems Framework for Sustainable Health Informatics Investment in Africa
JMIR Public Health Surveill 2026;12:e89482
URL: https://publichealth.jmir.org/2026/1/e89482
doi:10.2196/89482

© Eric Nzirakaindi Ikoona, Lucy Namulemo, Ronald Kaluya, Rebecca Ikoona, Foday Sahr. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 1.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Integrating AI Into Governmental Public Health Decision Making: Challenges, Considerations, and a Path Forward

Elizabeth Campbell, MS, MScPH, PhD; Oluremilekun Oyefolu, MD, MPH; Sarah Gillani, MPH; Hannah Goodtree, MPH; Alison Kelly, MPIA; Caitlin Rivers, MPH, PhD; Crystal Watson, MPH, DrPH

Johns Hopkins Center for Outbreak Response Innovation, 700 East Pratt Street, Suite 900, Baltimore, MD, United States

Corresponding Author:
Elizabeth Campbell, MS, MScPH, PhD
Johns Hopkins Center for Outbreak Response Innovation, 700 East Pratt Street, Suite 900, Baltimore, MD, United States

Abstract

Public health emergencies such as pandemics, natural disasters, and epidemics may require rapid, high-stakes decisions often made by elected officials with limited public health training. Artificial intelligence (AI) holds significant promise to enhance the quality, transparency, and timeliness of governmental decision-making during such crises. This paper examines the potential of AI as a decision-support tool for elected officials while identifying key technical, logistical, ethical, and policy challenges. Technical considerations include model accuracy, data representativeness, and privacy protection, while ethical imperatives center on fairness, transparency, and accountability to prevent amplification of existing health disparities. The paper further explores workforce development needs, emphasizing AI literacy and cross-sector collaboration to enable informed use of AI insights. This viewpoint presents a novel AI Decision Support Lifecycle framework specifically designed for governmental public health emergency response, mapping six phases from problem definition through post-emergency evaluation. We provide stakeholder-specific recommendations for model developers, health agencies, and elected officials, and illustrate practical application through a detailed case example and use cases. Drawing on empirical evidence regarding digital health technologies and AI governance, we emphasize that technology deployment alone is insufficient. Successful implementation requires complementary investments in organizational capacity, data infrastructure, workforce training, community engagement, and continuous evaluation. AI integration also requires robust governance frameworks, continuous model evaluation, and alignment with existing crisis management structures. Policy recommendations highlight the importance of ethical AI frameworks, risk assessments, and public engagement to foster trust. Ultimately, AI can strengthen public health decision-making if developed and implemented responsibly within transparent and equitable systems.

(JMIR Public Health Surveill 2026;12:e88470) doi:10.2196/88470

KEYWORDS

public health; elected officials; governance; AI; artificial intelligence; decision making; public health emergency response

Introduction

Public health emergencies, such as the COVID-19 pandemic, the 2014‑2016 West Africa Ebola epidemic, and natural disasters such as floods or hurricanes, pose serious and pressing individual and population health threats. Depending on the size and scope of these emergencies, they can require urgent and harmonized responses from governmental decision makers and public health systems at the local/tribal, state/territorial, and even national levels [1,2]. Sometimes, when an emergency is of significant size and scope, or is highly visible or consequential, decisions that are typically made by public health or emergency management experts shift to, or at least involve, elected officials, such as state governors [3]. This shift might be made because elected officials are politically responsible for the outcome of a response, or the shift might happen because only elected officials can exercise the legal powers necessary to respond. For example, in all states and territories, governors can issue emergency declarations to enact measures such as accessing and reallocating funds or suspending laws and regulations that may impede response efforts [4-6].

Regardless of the inciting event, when elected officials or other leaders are thrust into a public health decision-making role, often without formal public health or crisis response training, they may need additional support to ensure that decisions are well-considered. While public health and emergency management experts can, and should, play a significant supporting role in decision-making, leaders without good decision-making processes in place can easily fall victim to so-called decision-derailers–common biases, traps, and heuristics that can make decisions less-than sound [7].

Challenges to Decision Making During Public Health Emergencies

A hallmark of public health emergencies is an initial severe information deficit, while many competing voices clamor for

XSL•FO
RenderX

urgent action. There is also inevitably a tug-of-war between health and safety priorities and other societal imperatives such as economic stability. Further, the broad categories of "health," "economic," and "societal" impacts must themselves be unpacked; choices about how to respond, what measures and recommendations should be put in place, must consider and involve diverse populations who are affected in very different ways and to different degrees.

So, what can leaders do to ensure that their decision-making process has systematically identified and considered a full range of actions that could effectively address fundamental goals; carefully considered and balanced all the values at stake; incorporated the information available while hedging against the remaining uncertainties; and has built public support through its transparency and thoroughness?

Operationalizing Effective Decision Making During Public Health Emergencies

First, policymakers should clearly define, justify, and articulate the strategic goals that they aim to achieve and then share this information with the public. Next, it is critical to identify what principles leaders want to uphold in an emergency response. Protecting the health of the public is a general principle that most leaders identify with. Are there other important principles that should not be forgotten, like ensuring the consequences of interventions are fair across populations? As decisions are not made in a vacuum, the voices of diverse experts help to consider a variety of angles to a problem.

Once the goals, principles, and structure for decision-making are in place, an assessment process for identifying and evaluating options for action is needed. Evaluating the possible downstream impacts of options (both positive and negative), comparing options against one another, and assessing the logistical feasibility of the range of options. Finally, in addition to a continual updating as new evidence and information emerges, well-documented communication practices highlight the importance of telling the public that information may change and foreshadowing how new information can change plans.

In a crisis, these structures and processes are very difficult to implement, especially under time pressure and without prior planning. Developing plans ahead of a crisis is optimal, but limited time and resources make this rare, and leaders will still face an information deficit difficult to overcome with legacy approaches to data gathering.

AI for Public Health Decision-Making

Artificial intelligence holds considerable promise to help decision-makers gather data to inform decisions and implement decision-making processes to systematically consider options and undertake tradeoff analyses. Ultimately, AI has the potential to greatly enhance the quality of public health decision-making in these emergencies. However, while the potential for decision support is significant, numerous challenges persist in developing and operationalizing AI for public health crisis decision making. While digital health technologies and information and communication technology (ICT) infrastructure have expanded rapidly, empirical evidence regarding their impact on population health outcomes remains mixed [8-12]. Studies show that ICT adoption does not automatically translate into measurable health improvements, particularly when institutional capacity, governance quality, workforce skills, and implementation fidelity are insufficient. Previous digital health initiatives often showed weak or context-dependent effects due to inadequate organizational readiness, data quality issues, and a lack of complementary investments in regulatory frameworks and human capital.

The following viewpoint considers technical, logistical, and ethical challenges in developing AI tools to support elected officials in public health decision making and describes important considerations when addressing AI applications to optimize system use in the public health domain.

Aim and Scope

This viewpoint article aims to: (1) examine the potential role of AI as a decision-support tool for elected officials during public health emergencies and (2) identify key technical, ethical, and governance challenges that must be addressed.

This viewpoint focuses specifically on US federal, state, and local governmental public health decision-making contexts. While many principles may apply internationally, our discussion of policy frameworks, legal requirements, and organizational structures reflects the US governance system. Jurisdictions with different regulatory environments, data infrastructure capacity, or political structures may face distinct challenges requiring adaptation of these recommendations.

Target Audience

This viewpoint is written for three primary audiences: (1) state and local public health leaders and emergency management officials who must integrate AI tools into decision-making workflows; (2) AI developers and data scientists working with governmental agencies on public health applications; and (3) elected officials and policymakers who oversee public health emergency response and must understand both the potential benefits and risks of AI-supported decision making.

Methodological Approach

This is a viewpoint article that synthesizes insights from the emerging interdisciplinary literature on AI in public health, governmental AI governance, and algorithmic fairness. Literature was identified through targeted searches of key databases (PubMed, Web of Science, OAISTER) and policy sources (federal and state government websites, Office of Management and Budget [OMB] guidance, state legislative databases), focusing on AI applications in governmental public health decision-making, with particular attention to recent publications (2020 - 2025) addressing COVID-19 response and other public health emergencies.

While existing AI governance frameworks (eg, the National Institute of Standards and Technology (NIST) AI Risk Management Framework, Organisation for Economic Co-operation and Development (OECD) AI Principles) offer

XSL•FO
RenderX

domain-general guidance, and crisis management models address emergency decision structures, neither adequately addresses the intersection of AI-specific technical challenges and the distinctive governance realities of public health emergency response [1,13]. This includes the shift of decision authority to elected officials with limited technical training, political accountability pressures, and the need for real-time model adaptation. The AI Decision Support Lifecycle framework presented here is designed to fill this gap by providing phase-specific, stakeholder-differentiated guidance tailored to this context.

Technical Challenges in Model Development

Machine learning (ML) models that support public health decision-making are subject to technical challenges in their development pipeline prior to being scaled for public health response efforts. Underfitting and overfitting are two of the most common challenges in the model development process, which limit a model's utility. Models are sensitive to quality and representativeness in the data used to train them.

Overfitting occurs when a model becomes too specialized in its training data and fails to perform well on new, unseen data. Overfitting limits a model's generalizability on test or real-world data and leads to high variance in model performance based on changes to input data. For example, a forecasting model trained on historical influenza data that captures every nuance of past flu seasons may perform poorly during a novel pandemic like COVID-19, where transmission dynamics, public behavior, and health care utilization patterns differ substantially from historical precedent. Overfitting may be addressed by collecting more data to create a more robust training set, reducing the model parameters or undertaking feature selection, and using cross-validation. Underfitting occurs when a model is too simplistic and fails to capture underlying patterns in data. This leads to missed patterns and inaccurate predictions, even for training data. In emergency contexts, this might manifest as a basic model that predicts hospital admissions using only a single variable (eg, previous d's case count) while ignoring critical factors like vaccination rates, demographic composition, or seasonal patterns. Such models provide unreliable guidance and may lead decision-makers to underestimate resource needs. Increasing the number of model parameters or features and reducing regularization may help to reduce underfitting.

Ethical Considerations for AI in Public Health

Bias and Fairness in AI Algorithms for Public Health Decision Making

Ensuring algorithmic fairness and minimizing potential algorithmic biases is a key goal for successfully implementing AI into public health decision-making processes. Data used for developing AI models for public health practice may come from a variety of sources, including electronic health records, governmental data on demographic, socioeconomic, and environmental indicators, social media, and genomic data. These models are trained on vast amounts of data, which may inadvertently contain bias. Without actively and adequately addressing these biases, AI models for public health decision making may amplify and perpetuate existing health disparities.

Bias in public health AI models may come from a variety of sources. If training data is not representative of all subpopulations (eg, age, race/ethnicity, gender), this leads to sampling bias and may cause inaccurate model predictions for members of underrepresented groups [14-16]. Furthermore, AI models trained on past health care data may learn historical biases embedded in these data [17]. For example, Obermeyer et al identified bias in model performance against Black patients for a commercial algorithm used to predict population health risk. The model used health care costs as a proxy for health (ie, lower health care costs equating to better health). However, in the US health care system, less money is spent caring for Black patients compared to White patients, and Black patients are less likely to seek out and obtain medical care [18]. Health care data is also subject to measurement bias, which occurs when there are systemic errors in how health outcomes are measured and recorded (eg, misdiagnosis or underreporting) [19]. Finally, bias may be introduced algorithmically (eg, feature selection bias) or in model design processes [20-22]. AI algorithms also often work as black boxes, which further complicates their deployment in public health decision making efforts since it is difficult for both government officials and constituents to understand how a decision was reached. This lack of transparency and explainability can reduce trust in AI-driven policies, cause concerns about accountability, and create complex ethical and legal challenges should AI-driven recommendations lead to harm [23]. Should AI solutions produce incorrect, harmful, or misleading recommendations, this leads to public mistrust and decreased likelihood of buy-in for public health measures that require community-wide support (eg, social distancing).

Equity in Practice

To mitigate some of these potential risks, organizations responsible for developing and scaling public health AI models and governmental decision makers may undertake several actions. First, model designers should carefully curate large, representative datasets that reflect the diverse populations that public health decisions will impact. When models are developed, it is also critical that designers assess models for performance disparities, identify fairness metrics, and undertake careful documentation of model development processes to ensure transparency and accountability. Furthermore, it is essential that decision makers receive thorough training so that they are able to understand and interpret AI system recommendations to promote trust, accountability, and transparency. To operationalize equity in public health AI systems, organizations must implement concrete safeguards throughout the AI lifecycle, including: establishing compensated community advisory boards with decision-making authority from historically marginalized populations; engaging diverse stakeholders through participatory design processes to identify potential harms and validate outputs; conducting formal equity impact assessments before and during deployment with transparent public reporting; evaluating and reporting model performance metrics separately for demographic

XSL•FO
RenderX

and geographic subgroups (race/ethnicity, age, gender, socioeconomic status, disability, language, geography) with minimum performance thresholds for all groups rather than population averages alone; developing culturally appropriate communication strategies that provide plain language, multilingual explanations of AI use through trusted community channels with clear feedback and recourse mechanisms; and implementing continuous bias auditing through both internal reviews and independent third-party assessments with public documentation of findings and remediation steps. Finally, model development teams and government organizations should work together to regularly monitor and evaluate AI models to identify and correct potential biases, update and retrain models as new data becomes available, and ensure that models continue to adhere to existing legal and ethical standards [24].

Privacy and Data Security

AI and ML programs and utilization can both enhance data security or generate issues related to privacy, confidentiality breaches, and data security [25]. There are opportunities for both malicious and accidental incidents that can result in confidence reduction, misclassification of data, or misclassification of outputs, all with the potential for severe downstream effects [26]. AI and ML models require the collection, storage, and analysis of vast quantities of data, including health data, personal identifying information, and confidential information. Security risks are a huge concern for these models and should be a priority when developing and using AI or ML tools. Opportunities for data breaches or loss of security could occur at an individual level, including someone with authorized access using data maliciously, it could occur accidentally through non-secure networks, or it could occur through organized and intentional breaches into private systems [27].

The vast number of systems that may interact with an individual's personal data also increases the risk of privacy concerns, since more systems translates to more opportunities for risk [28]. It is essential that developers, researchers, and institutions develop, test, and continue to improve and invest in safeguards for data access, use, and control. These safeguards can include privacy risk mitigation strategies and privacy-enhancing technologies [29], including de-identifying data, safeguarding physical access to systems, and strong encryption and protection of cloud-based systems [28]. There are also ethical considerations for inadequate data privacy protections, including data breaches or unauthorized access. Individuals who have personal or confidential data used in these models should have provided fully informed consent [30], be updated in the event of a security compromise, and have the option to withdraw their data from the model. There are ongoing discussions on data ownership and profit from individual data as well [31]. Finally, regulatory bodies should exist that can enforce privacy protections and provide recourse when entities fail to protect sensitive data or violate consent for data use [30].

Implementation: Organizational Readiness

Integrating AI Into Education and Government Workforce Development

Successful AI deployment in the public sector hinges not only on technological capabilities but equally on organizational readiness and workforce expertise. Evidence from previous digital health initiatives demonstrates that technology alone is insufficient; effective implementation requires complementary investments in organizational capacity, data infrastructure, and workforce skills [8-12].

The G7 Toolkit for Artificial Intelligence in the Public Sector identifies workforce digital skills as a foundational requirement for effective and responsible AI deployment in government contexts. A multi-layered approach is essential to building this capacity. Foundational AI literacy programs can equip public officials with an understanding of both AI's potential and its inherent limitations. Expert workshops utilizing scenario-based exercises represent another valuable resource that could help decision-makers navigate the complex trade-offs between privacy protection, equity considerations, and operational efficiency. Historical analysis also offers great learning opportunities. By examining case studies demonstrating predictive disease surveillance and resource allocation during past emergencies, officials can visualize concrete applications of AI and develop strategies to manage implementation challenges as they arise.

Considerations for Implementation Into Workflows, Policy Formulation, Program Planning, and Resource Allocation

A successful deployment of artificial intelligence during public health emergencies extends well beyond technological considerations, requiring robust institutional frameworks and operational adaptations. Historically, expert consensus functioned as a governance mechanism through which policy alternatives were prioritized. The absence of a standardized governance framework for AI implementation may create significant hesitation among elected officials to support AI-derived recommendations, particularly in situations when actions are under public scrutiny [32].

AI integration into public health decision-making would also require a seamless alignment with existing processes and adaptive governance structures. The fragmentation of data across multiple stakeholders and information systems could result in delayed or incompatible information transfer, compromising model accuracy when rapid analytical insights are most critical [33]. At the operational level, AI outputs must be integrated directly into existing crisis-management dashboards and consoles so that public health officials can obtain timely and actionable insights without redundant data entry or additional interfaces [34].

The implementation of AI tools during a public health emergency involves significant resource considerations beyond software licensing and hardware acquisition. Investments in

XSL•FO
RenderX

data management infrastructure, cybersecurity protections, and dedicated personnel are also essential. Additionally, ongoing model retraining and impact assessments are required to maintain relevance as conditions evolve [35]. The resource allocation challenge is further complicated in emergency settings, where bandwidth and computational resources face competition from numerous concurrent priorities such as other critical emergency response functions.

Artificial intelligence systems can function as decision support mechanisms that enhance rather than supplant the judgment of elected officials and their technical advisors when designed with appropriate architectural safeguards. The safeguard features could include human validation functions that allow flagging specific output for expert review at key decision points [36], an iterative feedback loop function to allow officials to log disagreements with model outputs and feed these back into model retraining [33], an ensemble function that can combine model predictions with expert systems to ensure known constraints are reinforced alongside data-driven insights [37].

Case Study: To illustrate how these implementation considerations manifest in practice, we present a detailed case example:

Hospital Surge Prediction During Influenza Season

A state health department seeks to deploy an AI forecasting model to predict hospital bed needs 2‑4 weeks ahead during the influenza season, allowing proactive resource mobilization.

Problem Definition: The health department, in consultation with hospital associations and emergency management, identifies specific decision needs: estimating intensive care unit (ICU) bed demand, identifying which geographic regions will face capacity constraints, and determining timing for activating surge capacity plans.

Model Development: Data scientists develop an ensemble forecasting model using historical hospitalization data, influenza surveillance data, and demographic information. The model undergoes rigorous validation using holdout data from previous flu seasons, with performance assessed overall and for relevant subgroups (age categories, urban vs rural regions).

Integration: The model is integrated into the state's existing emergency operations center dashboard, providing daily updated forecasts with clear uncertainty ranges. Training is provided to epidemiologists and emergency coordinators on interpreting forecasts and understanding uncertainty.

Deployment: During flu season, forecasts inform daily situation reports to the governor's office. When the model predicts ICU capacity concerns in the jurisdiction's northern region, officials activate surge plans, coordinate patient transfers, and deploy mobile medical units. Importantly, epidemiologists review each forecast before distribution, flagging any predictions that seem inconsistent with other data sources.

Failure Mode: Mid-season, a new influenza strain emerges with higher severity. The model, trained on previous seasons' severity patterns, initially underestimates hospitalization demand. Human validation catches this discrepancy within 48 hours, prompting manual adjustments and communication to decision-makers about model limitations.

Mitigation: The incident illustrates the importance of human oversight, ensemble modeling (the team supplements the AI forecast with expert judgment and alternative models), and clear uncertainty communication (decision-makers understand that forecasts carry substantial uncertainty during novel epidemiological situations).

Evaluation and Continuous Learning

Systematic evaluation is essential for responsible AI deployment in public health emergencies. Agencies should conduct pre-deployment testing through prospective validation using historical emergency data, simulation exercises, and tabletop drills to assess accuracy, identify failure modes, and evaluate integration with decision workflows before operational use. Performance benchmarking should establish baseline accuracy metrics, track performance over time, and, where possible compare AI-supported versus non-AI-supported decision outcomes to assess actual impact on decision quality. Real-time monitoring requires dashboards that track model performance during deployment with automated alerts for degradation, supplemented by regular expert review to identify issues not captured by automated metrics. Agencies must establish adverse event reporting mechanisms similar to health care systems, with clear protocols for investigating suspected AI-related errors or harms and implementing corrective actions. Post-emergency after-action reviews should systematically assess whether AI tools were useful, what worked well, what failed, how tools should be modified, and whether unintended consequences or equity impacts occurred. Finally, agencies should define clear decommissioning criteria and transition plans for retiring models when they exhibit poor performance, conditions change, or better alternatives become available.

Policy Implications of AI in Public Health Decision Making

State-level legislative activity around governmental AI use has accelerated rapidly, with more than 150 bills introduced in 2024 and at least 30 states issuing formal guidance on AI use by state agencies. For example, Connecticut has implemented the AI Responsible Use Framework (Policy AI-01), which promotes ethical AI use, fairness, privacy, and transparency across all state agencies. Maryland enacted the Artificial Intelligence Governance Act of 2024, requiring each unit of state government to conduct inventories and assessments of systems that employ high-risk AI, and providing for the Department of Information Technology to develop policies and procedures concerning the development, procurement, deployment, use, and assessment of such systems [38]. Vermont established the Division of Artificial Intelligence within the Agency of Digital Services to review all aspects of AI systems developed, employed, or procured in state government [38].

At the federal level, several policies regulate the use of AI in governmental decision-making. The AI in Government Act of 2020 facilitates the adoption of AI technologies in the federal

XSL•FO
RenderX

government, aiming to improve cohesion and competency in the adoption and use of AI within federal agencies. Executive Order 13960, titled "Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government," establishes the policy of the United States to promote the innovation and use of AI to improve government operations and services in a manner that fosters public trust and confidence while protecting privacy, civil rights, civil liberties, and American values. Additionally, the Office of Management and Budget (OMB) issued Memorandum M-25‑21, "Accelerating Federal Use of AI through Innovation, Governance, and Public Trust," which provides guidance for federal agencies to enhance their AI governance and risk management practices [39-41]. More recently, Executive Order 14,319 (July 2025), titled "Preventing Woke AI in the Federal Government," required federal agencies to procure only AI models that prioritize truth-seeking and ideological neutrality, restricting the incorporation of diversity, equity, and inclusion frameworks into AI outputs [42]. This policy places constraints on public health AI decision-making by limiting the use of equity-informed analytic approaches that are commonly used to identify health disparities and guide targeted interventions. In December 2025, Executive Order 14365, titled "Ensuring a National Policy Framework for Artificial Intelligence," established a uniform national AI framework intended to preempt state regulations deemed burdensome to innovation. By authorizing federal challenges to inconsistent state laws and allowing the withholding of federal funding from non-compliant states, the order reshapes public health AI governance by centralizing oversight, potentially reducing states' ability to tailor AI safeguards, transparency standards, and equity protections to local public health needs [43].

[38]To strengthen AI regulations in government, several policies should be implemented. Governments should establish ethical AI frameworks that operationalize the fairness and bias mitigation principles discussed earlier, with continuous assessments to identify biases, ensure privacy protection, and evaluate AI system reliability. More states should implement policies like Connecticut, Maryland, and Vermont to regulate AI use within state agencies to ensure the ethical use of AI systems in government operations [38]. Additionally, mandatory impact assessments should be conducted before deploying AI systems, ensuring that risks, including bias and discrimination, are mitigated and disclosed publicly. Governments should also require AI systems to be explainable, allowing the public to understand the decision-making logic behind AI outcomes. For instance, New York has required state agencies to publicly disclose detailed information about their automated decision systems [44]. Public oversight should be ensured through independent governance structures, such as Chief AI Officers, who would be responsible for ensuring AI deployment aligns with ethical standards and transparency.

Building public trust in AI use requires active engagement with stakeholders, including community leaders, ethicists, and public health experts, to ensure AI applications meet the needs and values of the community. Governments must prioritize clear and accessible communication about how AI is used in decision-making, including detailed reports on AI system performance, challenges, and successes. Regular audits should be conducted to address potential biases, ensuring AI systems are fair and inclusive. Furthermore, ethical AI governance structures should be implemented, including ethics boards and external reviews, to monitor and guide responsible AI use. Creating avenues for public feedback, such as town halls or online platforms, will allow communities to voice concerns and provide input on improving AI systems. Strong privacy protections should be enacted to safeguard citizens' data, ensuring AI systems are compliant with privacy standards. Lastly, ongoing public education and training on AI literacy will help demystify AI and foster confidence in its use, ultimately increasing public trust in governmental AI decision-making.

However, even with emerging governance frameworks, empirical evidence suggests that formal governance structures do not automatically translate into improved outcomes [45-48]. Cross-country studies find that standard governance indicators (government effectiveness, rule of law, transparency) show weak or inconsistent relationships with population health outcomes once other factors are controlled. Similarly, AI governance arrangements (such as ethical guidelines, transparency principles, and oversight bodies) do not necessarily translate into effective implementation at the operational level. This disconnect occurs because high-level institutional indicators may fail to capture sector-specific capacities essential for public health emergency response: data quality, frontline health system integration, decision-maker competence, organizational culture, and implementation fidelity. Successful AI governance requires attention to these operational realities, not just policy documents.

The AI Decision Support Lifecycle for Public Health Emergencies

Drawing on the technical, ethical, organizational, and policy considerations discussed above, we propose an AI Decision Support Lifecycle framework specifically designed for governmental public health emergency response (Figure 1).

XSL•FO
RenderX

Figure 1. AI decision supported lifecycle in public health emergencies.

Phase 1 consists of defining the problem at hand and conducting a needs assessment. Stakeholders must clarify the decisions that will be made, the information that is required to help decision makers, the end users of AI outputs, and the consequences of wrong decisions. This phase involves engaging elected officials, public health leaders, emergency managers, and community representatives to ensure AI development targets genuine decision needs.

Phase 2 consists of equity-driven, technically robust model development and validation. Key details and steps include using representative data, evaluating performance across demographic and geographic subgroups, and documenting model limitations and uncertainty. Additional safeguards include independent expert review prior to operational deployment.

Phase 3 includes integrating AI models into workflows and pre-deployment testing. AI tools must be integrated into existing workflows and information systems. Steps include but are not limited to developing user interfaces compatible with incident command structures, training end users on interpretation and appropriate use, and establishing clear protocols for human validation and override of AI recommendations.

In *Phase 4***,** a model is ready for deployment during an emergency. During an emergency, AI tools provide decision support while human officials retain ultimate authority. When AI is being deployed during an emergency, model performance must be monitored in real time, validating outputs using expert evaluation before high-stakes decisions, and clear communication of uncertainty to decision makers.

Monitoring and adaptation comprise *Phase 5*, which includes tracking accuracy metrics, comparing predictions to observed outcomes, and adjusting models as new data becomes available.

Finally, in *Phase 6*, models are evaluated after emergency deployment, and lessons learned are gleaned. Evaluation includes understanding the model's usefulness for decision making, identifying what worked well and what didn't as well as any unintended consequences, and understanding how the tool may be used or modified for future emergencies.

A number of critical principles apply throughout the phases of the AI Decision Support Lifecycle for Public Health Emergencies. First, it is vital that there are robust privacy and confidentiality protections in place and that there is limited data use for limited and specific purposes. Next, equity principles should be positioned as a key priority in model development and implementation. Similarly, there must be clear documentation for model development and use, explainability for model outputs, and public accountability for the use of AI in public health decision-making. Finally, human authority must remain the centerpiece of public health emergency decision-making. AI can only serve as decision support, never an autonomous decision-maker.

XSL•FO
RenderX

Table 1 organizes and synthesizes technical, ethical, organizational, and policy challenges that occur throughout the lifecycle, with specific actionable recommendations for model developers, health agencies, and elected officials.

Table . Key challenges and stakeholder actions for AI integration in public health emergency decision-making.

Challenge area	Model developers	Health agencies & officials	Policymakers & elected officials
Technical challenges			
Model Development (underfitting, overfitting, data quality)	Ensure models are sensitive to data quality and representativeness; avoid over-specialization to training data	Work with modelers to validate models before deployment; establish validation processes for AI outputs	Understand model limitations; ensure AI enhances rather than supplants judgment
Ethical challenges			
Bias & Fairness	Curate representative datasets; assess for performance disparities across subgroups; document model development transparently	Continuously evaluate models for bias; implement continuous monitoring and regular auditing of AI systems	Ensure decisions account for diverse communities; conduct regular audits for fairness and inclusiveness
Privacy & Data Security	Design models with privacy protections; maintain confidentiality of training and testing data	Implement robust data management infrastructure and cybersecurity protections	Enforce transparent risk management and privacy protection policies; ensure civil rights protections
Organizational challenges			
Workforce Development & AI Literacy	Design user-friendly interfaces; provide training resources for end users	Implement foundational AI literacy programs; provide expert workshops using scenario-based training	Understand AI's potential and limitations; receive training to interpret and implement AI outputs
Workflow Integration & Implementation	Include human validation functions, feedback loops, and ensemble functions combining model predictions with expert systems	Integrate AI outputs into crisis-management dashboards; ensure seamless alignment with existing processes	Establish governance frameworks to reduce hesitation in supporting AI-derived recommendations
Resource Allocation & Infrastructure	Plan for ongoing maintenance and model updates beyond initial development	Invest in data management infrastructure, cybersecurity, and dedicated personnel	Allocate resources beyond software licensing and hardware acquisition
Policy & governance challenges			
Governance Frameworks & Standards	Follow ethical AI principles prioritizing fairness, transparency, and accountability	Develop adaptive governance structures; implement continuous risk assessments	Establish ethical AI frameworks at federal and state levels; implement mandatory impact assessments
Public Trust & Stakeholder Engagement	Design transparent systems that enable public understanding of AI decision-making processes	Engage with community leaders, ethicists, and public health experts; provide detailed performance reports	Ensure AI applications meet community needs and values; prioritize clear communication about AI use

Conclusion

Artificial intelligence holds substantial promise for supporting governmental public health emergency response, but responsible implementation requires adherence to core ethical principles rather than simply deploying technology. First, AI must function strictly as decision-support tools while preserving human authority. Elected officials and public health leaders retain ultimate accountability for decisions, with clear protocols for validating and overriding AI recommendations. Second, equity must be designed into AI systems from inception, not treated as an afterthought, through representative data, subgroup performance evaluation, community advisory boards, and transparent equity impact reporting. Third, successful AI deployment demands comprehensive organizational readiness, including workforce capacity (data scientists working collaboratively with epidemiologists and communication specialists), robust data infrastructure, governance structures, and sustained funding beyond initial acquisition costs.

Fourth, transparency and accountability mechanisms are non-negotiable, requiring public disclosure of when AI informs decisions, comprehensive model documentation, explainable outputs, and clear processes for challenging AI-informed decisions. Finally, continuous monitoring, evaluation, and adaptation throughout the AI lifecycle ensures systems remain reliable as conditions change, with pre-deployment testing, real-time performance monitoring, adverse event reporting, and willingness to decommission underperforming models. Achieving AI's potential requires genuine collaboration between technical developers, public health agencies, and elected officials, with each stakeholder group taking specific, complementary actions to ensure models meet legal and ethical standards while meaningfully supporting evidence-based decision-making that reduces rather than exacerbates health inequities.

XSL•FO
RenderX

Funding

The Center for Outbreak Response Innovation (CORI) is supported through Cooperative Agreement NU38FT000004 between CDC's Center for Forecasting and Outbreak Analytics and Johns Hopkins University's Bloomberg School of Public Health.

Conflicts of Interest

None declared.

References

1. US centers for disease control and prevention. Public Health Emergency Preparedness and Response Capabilities.: State Local Readiness; 2025. URL: https://www.cdc.gov/readiness/php/capabilities/index.html [accessed 2026-01-30]
2. Nelson C, Lurie N, Wasserman J, Zakowski S. Conceptualizing and defining public health emergency preparedness. Am J Public Health 2007 Apr;97 Suppl 1(Suppl 1):S9-11. [doi: 10.2105/AJPH.2007.114496] [Medline: 17413078]
3. Hodge JG, Dunning LT, Piatt JL. State public health emergency powers in response to COVID-19. Am J Public Health 2023 Mar;113(3):275-279. [doi: 10.2105/AJPH.2022.307158] [Medline: 36521072]
4. Orenstein DG. When law is not law: setting aside legal provisions during declared emergencies. J Law Med Ethics 2013 Mar;41 Suppl 1(S1):73-76. [doi: 10.1111/jlme.12044] [Medline: 23590746]
5. Hodge JG, Anderson ED. Principles and practice of legal triage during public health emergencies [Abstract].: Social Science Research Network; 2009. URL: https://papers.ssrn.com/abstract=1335342 [accessed 2025-07-10]
6. Sunshine G, Barrera N, Corcoran AJ, Penn M. Emergency declarations for public health issues: expanding our definition of emergency. J Law Med Ethics 2019 Jun;47(2_suppl):95-99. [doi: 10.1177/1073110519857328] [Medline: 31298138]
7. Higgins G, Freedman J. Improving decision making in crisis. JBCEP 2013;7(1):65. [doi: 10.69554/AYCK9571]
8. Agarwal R, Gao G(, DesRoches C, Jha AK. Research Commentary —The digital transformation of healthcare: current status and the road ahead. Information Systems Research 2010 Dec;21(4):796-809. [doi: 10.1287/isre.1100.0327]
9. Cline GB, Luiz JM. Information technology systems in public sector health facilities in developing countries: the case of South Africa. BMC Med Inform Decis Mak 2013 Jan 24;13:13. [doi: 10.1186/1472-6947-13-13] [Medline: 23347433]
10. Sheikh A, Sood HS, Bates DW. Leveraging health information technology to achieve the “triple aim” of healthcare reform. J Am Med Inform Assoc 2015 Jul;22(4):849-856. [doi: 10.1093/jamia/ocv022] [Medline: 25882032]
11. Fichman RG, Kohli R, Krishnan R. Editorial Overview —The role of information systems in healthcare: current research and future trends. Information Systems Research 2011 Sep;22(3):419-428. [doi: 10.1287/isre.1110.0382]
12. Scott RE, Mars M. Principles and framework for eHealth strategy development. J Med Internet Res 2013 Jul 30;15(7):e155. [doi: 10.2196/jmir.2250] [Medline: 23900066]
13. OECD, UNESCO, G7 Italia. G7 toolkit for artificial intelligence in the public sector. 2024 URL: https://www.oecd.org/en/publications/g7-toolkit-for-artificial-intelligence-in-the-public-sector_421c1244-en.html [accessed 2026-01-30]
14. Chen Z, Zhang JM, Sarro F, Harman M. A comprehensive empirical study of bias mitigation methods for machine learning classifiers. ACM Transactions on Software Engineering and Methodology 2023;32(4):1-30. [doi: 10.1145/3583561]
15. Benjamin R. Assessing risk, automating racism. Science 2019 Oct 25;366(6464):421-422. [doi: 10.1126/science.aaz3873] [Medline: 31649182]
16. Mhasawade V, Zhao Y, Chunara R. Machine learning and algorithmic fairness in public and population health. Nat Mach Intell 2021;3(8):659-666. [doi: 10.1038/s42256-021-00373-4]
17. McCradden MD, Joshi S, Anderson JA, Mazwi M, Goldenberg A, Zlotnik Shaul R. Patient safety and quality improvement: Ethical principles for a regulatory approach to bias in healthcare machine learning. J Am Med Inform Assoc 2020 Dec 9;27(12):2024-2027. [doi: 10.1093/jamia/ocaa085] [Medline: 32585698]
18. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. 2019. URL: https://www.science.org/doi/10.1126/science.aax2342 [accessed 2025-04-23]
19. Young JC, Conover MM, Funk MJ. Measurement error and misclassification in electronic medical records: methods to mitigate bias. Curr Epidemiol Rep 2018 Dec;5(4):343-356. [doi: 10.1007/s40471-018-0164-x] [Medline: 35633879]
20. Campbell EA, Bose S, Masino AJ. Conceptualizing bias in EHR data: A case study in performance disparities by demographic subgroups for a pediatric obesity incidence classifier. PLOS Digit Health 2024 Oct;3(10):e0000642. [doi: 10.1371/journal.pdig.0000642] [Medline: 39441784]
21. Andaur Navarro CL, Damen JAA, Takada T, et al. Risk of bias in studies on prediction models developed using supervised machine learning techniques: systematic review. BMJ 2021 Oct 20;375:n2281. [doi: 10.1136/bmj.n2281] [Medline: 34670780]
22. Pathak YV, Saikia S, Pathak S, Patel JK, Prajapati JB. Ethical Issues in AI for Bioinformatics and Chemoinformatics: CRC Press; 2023. [doi: 10.1201/9781003353751]
23. London AJ. Artificial intelligence and black-box medical decisions: accuracy versus explainability. Hastings Cent Rep 2019 Jan;49(1):15-21. [doi: 10.1002/hast.973] [Medline: 30790315]
24. Gerke S, Minssen T, Cohen G. Ethical and legal challenges of artificial intelligence-driven healthcare. Artif Intell Healthc 2020:295-336. [doi: 10.1016/B978-0-12-818438-7.00012-5]

XSL•FO
RenderX

25. Martin KD, Zimmermann J. Artificial intelligence and its implications for data privacy. Curr Opin Psychol 2024 Aug;58:101829. [doi: 10.1016/j.copsyc.2024.101829] [Medline: 38954851]
26. Oseni A, Moustafa N, Janicke H, Liu P, Tari Z, Vasilakos A. Security and privacy for artificial intelligence: opportunities and challenges. arXiv. Preprint posted online on 2021. [doi: 10.48550/arXiv.2102.04661]
27. Bak M, Madai VI, Fritzsche MC, Mayrhofer MT, McLennan S. You can't have ai both ways: balancing health data privacy and access fairly. Front Genet 2022;13:929453. [doi: 10.3389/fgene.2022.929453] [Medline: 35769991]
28. Yadav N, Pandey S, Gupta A, Dudani P, Gupta S, Rangarajan K. Data privacy in healthcare: in the era of artificial intelligence. Indian Dermatol Online J 2023;14(6):788-792. [doi: 10.4103/idoj.idoj_543_23] [Medline: 38099022]
29. Curzon J, Kosa TA, Akalu R, El-Khatib K. Privacy and artificial intelligence. IEEE Trans Artif Intell 2021;2(2):96-108. [doi: 10.1109/TAI.2021.3088084]
30. Andreotta AJ, Kirkham N, Rizzi M. AI, big data, and the future of consent. AI Soc 2022;37(4):1715-1728. [doi: 10.1007/s00146-021-01262-5] [Medline: 34483498]
31. Adapa VR. Navigating the privacy paradox: balancing ai advancement and data protection in the digital age. Int J Sci Res Comput Sci Eng Inf Technol 2024;10(6):99-110. [doi: 10.32628/CSEIT24106158]
32. Stanford V, Gresh L, Toledo J, Méndez J, Aldighieri S, Reveiz L. Evidence in decision-making in the context of COVID-19 in Latin America. Lancet Reg Health Am 2022 Oct;14:100322. [doi: 10.1016/j.lana.2022.100322] [Medline: 35879980]
33. Hu Y, Jacob J, Parker GJM, Hawkes DJ, Hurst JR, Stoyanov D. The challenges of deploying artificial intelligence models in a rapidly evolving pandemic. Nat Mach Intell 2020;2(6):298-300. [doi: 10.1038/s42256-020-0185-2]
34. Medaglia R, Zheng L. Mapping government social media research and moving it forward: a framework and a research agenda. Gov Inf Q 2017 Sep;34(3):496-510. [doi: 10.1016/j.giq.2017.06.001]
35. Ding X, Shang B, Xie C, Xin J, Yu F. Artificial intelligence in the COVID-19 pandemic: balancing benefits and ethical challenges in China's response. Humanit Soc Sci Commun 2025;12(1):1-19. [doi: 10.1057/s41599-025-04564-x]
36. McAndrew T, Gibson GC, Braun D, Srivastava A, Brown K. Chimeric Forecasting: an experiment to leverage human judgment to improve forecasts of infectious disease using simulated surveillance data. Epidemics 2024 Jun;47:100756. [doi: 10.1016/j.epidem.2024.100756] [Medline: 38452456]
37. Roy K, Zhang Q, Gaur M, Sheth A. Knowledge infused policy gradients for adaptive pandemic control. arXiv. Preprint posted online on 2021. [doi: 10.48550/arXiv.2102.06245]
38. Hooshidary S, Canada C, Clark W. Artificial intelligence in government: the federal and state landscape. NCSL. URL: https://www.ncsl.org/technology-and-communication/artificial-intelligence-in-government-the-federal-and-state-landscape [accessed 2025-04-24]
39. Campbell EA, Holl F, Marwah HK, Fraser HS, Craig SS. The impact of climate change on vulnerable populations in pediatrics: opportunities for AI, digital health, and beyond—a scoping review and selected case studies. Pediatr Res 2025 Oct;98(4):1250-1256. [doi: 10.1038/s41390-024-03719-x]
40. United States White House. Removing barriers to American leadership in artificial intelligence.: White House; 2025. URL: https://www.whitehouse.gov/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence [accessed 2025-04-24]
41. United States White House. White House releases new policies on federal agency AI use and procurement.: White House URL: https://www.whitehouse.gov/articles/2025/04/white-house-releases-new-policies-on-federal-agency-ai-use-and-procurement/ [accessed 2025-04-24]
42. United States White House. Preventing Woke AI in the Federal Government.: White House; 2025. URL: https://www.whitehouse.gov/presidential-actions/2025/07/preventing-woke-ai-in-the-federal-government [accessed 2026-02-09]
43. United States White House. Ensuring a National Policy Framework for Artificial Intelligence.: White House; 2025. URL: https://www.whitehouse.gov/presidential-actions/2025/12/eliminating-state-law-obstruction-of-national-artificial-intelligence-policy [accessed 2026-02-09]
44. Erence of state legislatures. Artificial Intelligence 2025 Legislation. 2025. URL: https://www.ncsl.org/technology-and-communication/artificial-intelligence-2025-legislation [accessed 2026-02-09]
45. Deaton A. The Great Escape: Health, wealth, and the origins of inequality.: Princeton University Press; 2013. URL: https://press.princeton.edu/books/hardcover/9780691153544/the-great-escape?srsltid=AfmBOooNBA9EvNKVTtDSApNys3_pCqq1V_CWoZvBVv_EmEdpuCfBMXtu [accessed 2026-02-06]
46. Filmer D, Pritchett L. The impact of public spending on health: does money matter? Soc Sci Med 1999 Nov;49(10):1309-1323. [doi: 10.1016/S0277-9536(99)00150-1]
47. Wagstaff A. Poverty and health sector inequalities. Bull World Health Organ 2002;80(2):97-105. [Medline: 11953787]
48. Holmberg S, Rothstein B. Dying of corruption. Health Econ Policy Law 2011 Oct;6(4):529-547. [doi: 10.1017/S174413311000023X] [Medline: 20809992]

Abbreviations

AI: artificial intelligence

XSL•FO
RenderX

ICT: information and communication technology
ML: machine learning
OMB: Office of Management and Budget

Edited by A Mavragani, T Sanchez; submitted 25.Nov.2025; peer-reviewed by M Akyuz, S Kakraba; revised version received 16.Mar.2026; accepted 20.Mar.2026; published 27.Apr.2026.

Please cite as:
Campbell E, Oyefolu O, Gillani S, Goodtree H, Kelly A, Rivers C, Watson C
Integrating AI Into Governmental Public Health Decision Making: Challenges, Considerations, and a Path Forward
JMIR Public Health Surveill 2026;12:e88470
URL: https://publichealth.jmir.org/2026/1/e88470
doi:10.2196/88470

© Elizabeth Campbell, Oluremilekun Oyefolu, Sarah Gillani, Hannah Goodtree, Alison Kelly, Caitlin Rivers, Crystal Watson. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 27.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Bias-Mitigated AI as a Foundation for Resilient and Effective Health Systems

Jarbas Barbosa da Silva Jr[1*], PhD; Maureen Birminghamm[1*], PhD; Ana Rivière Cinnamond[1*], PhD; Eldonna Boisson[1*], MPH; Mary Lou Valdez[1*], MPH; Sebastian Garcia Saiso[1*], MPH; Joao Paulo Souza[1*], MSc; Myrna C Marti[1*], MPH; Leah-Marie Richards[1*], MPH; Javier Guzman[2*], MPH; Jennifer Nelson[2*], MPH; Karina Pesce[3*], PDH; Ana Estela Haddad[4*], MPH; James Fitzgerald[1*], MPH; Ernesto Bascolo[1*], MPH; Marcelo D'Agostino[1], MSc

[1]World Health Organization Regional Office for the Americas, 525 23rd St NW, Washington, DC, United States

[2]Inter-American Development Bank, Washington, DC, United States

[3]Hospital Italiano de Buenos Aires, Buenos Aires, Argentina

[4]Secretariat of Information and Digital Health, Ministry of Health (Brazil), Brasília, Federal District, Brazil

[*]these authors contributed equally

Corresponding Author:
Marcelo D'Agostino, MSc
World Health Organization Regional Office for the Americas, 525 23rd St NW, Washington, DC, United States

Abstract

Artificial intelligence (AI) is rapidly reshaping the landscape of health care, from clinical diagnostics and disease surveillance to the prediction of individual health risks. Yet, its immense promise will only materialize if the tools we deploy work for everyone. Algorithms trained on incomplete or biased datasets risk embedding historical health disparities and can replicate patterns of uneven data representation, thereby limiting accuracy and generalizability across population groups. Addressing algorithmic bias should be treated as a core health quality standard, comparable in importance to safety and efficacy evaluations, to ensure consistent performance across all segments of the population. This paper aims to frame algorithmic bias in health-related AI as a quality, safety, and governance challenge for health systems rather than solely a technical problem for developers. It aims to inform policymakers, regulators, health system leaders, and developers by translating existing scientific evidence and regulatory guidance into operational governance considerations, with particular attention to the realities of low- and middle-income settings in the region of the Americas. This paper synthesizes existing knowledge and institutional experience into a practical, regionally grounded policy perspective. To operationalize this perspective, this paper first outlines the main forms of algorithmic bias relevant to health systems—including representation, measurement, aggregation, and deployment biases—and illustrates how each can emerge across the AI lifecycle. It then situates these technical challenges within the broader digital health context, where structural, commercial, and social dynamics may amplify inequities. This paper discusses the implications of biased data for emerging areas such as precision medicine before proposing a governance-oriented framework for bias mitigation that spans design, validation, deployment, and postmarket monitoring. It concludes with priority governance actions for policymakers, regulators, and health system leaders to embed fairness as a measurable component of health system performance.

(JMIR Public Health Surveill 2026;12:e88457) doi:10.2196/88457

KEYWORDS

primary health care; artificial intelligence; digital health; bias in artificial intelligence; AI

When Data Exclude, Health Follows: The Manifestations of Algorithmic Bias

The problem of bias is not one of malicious intent but of deep-seated scientific and structural oversight. While some of the examples discussed below originate in data collection practices or medical device calibration rather than in artificial intelligence (AI) models themselves, these upstream measurement biases are critical to understanding algorithmic bias in health. In practice, biased measurements and nonrepresentative clinical data are often used as training inputs for AI systems, allowing inequities embedded in devices or data standards to be inherited, amplified, and scaled through algorithmic decision-making. When data are nonrepresentative, bias often follows. Algorithms are only as sound as the data they are trained on, and when datasets fail to represent the full diversity of our populations across ancestry, gender, geography, and socioeconomic status, the resulting tools may perform well on average but systematically fail for underrepresented groups. This challenge has been documented in clinical settings and, in some cases, addressed, underscoring the need for improved model generalizability and data representation. The sources of bias are varied and can emerge at any stage of the AI lifecycle

XSL•FO
RenderX

[1,2]. To strengthen the analytical structure, bias can be grouped into four main categories: representation bias, measurement bias, aggregation bias, and deployment bias. This taxonomy helps identify specific interventions at each stage.

A clear example is representation bias, where training data does not reflect the target population. An AI system designed to detect skin cancer, for instance, achieved high accuracy when trained predominantly on images of light-skinned individuals. However, its diagnostic accuracy was nearly halved when applied to lesions on Black patients, reducing diagnostic performance in populations where early detection is already more difficult due to historical underrepresentation in medical datasets [3]. The result is an algorithm that may systematically discount reported symptoms or predict poor adherence, automating and scaling existing disparities, including in areas such as pain management and diagnostic triage decisions [1,4]. These examples illustrate that even well-established clinical language can reflect historical assumptions, an important consideration for developers of data-driven tools. Similar lifecycle-based taxonomies have been proposed in the broader machine learning literature, which consistently identify data collection and representation as primary sources of algorithmic unfairness, alongside proxy feature selection, model training, and deployment-related feedback effects [5].

Measurement bias can be embedded in the very tools we use to collect data. The pulse oximeter, a ubiquitous medical device, has been shown to overestimate blood oxygen levels in individuals with darker skin tones, a flaw rooted in the device's initial calibration on insufficiently diverse populations [1,6]. Although pulse oximetry bias is not, in itself, an AI-specific phenomenon, its relevance to this discussion lies in the fact that oxygen saturation data are frequently used in clinical datasets and predictive models, including AI-based risk stratification and decision support tools. When such measurements are systematically skewed for certain population groups, downstream algorithms trained on these data may reproduce or intensify these inaccuracies at scale. During the COVID-19 pandemic, this technical inaccuracy had profound implications, potentially delaying the recognition of severe hypoxemia and the delivery of critical care to already hard-hit communities [7]. Similar biases are found in clinical tools like spirometry equations that apply race-based "corrections" that may not be scientifically sound [6]. Similar considerations apply to race-based spirometry equations, where historically embedded assumptions in measurement standards can shape clinical datasets and, if left unexamined, be incorporated into AI-enabled models that rely on pulmonary function data for prediction or triage. In the Latin American and Caribbean region, where most medical devices are imported and not locally calibrated, this type of bias may be amplified, affecting vulnerable and remote populations that already face different barriers of access to specialized care.

Aggregation and deployment biases can emerge even with good data. Aggregation bias occurs when a single model is applied to a diverse population, ignoring crucial differences between subgroups. For years, cardiovascular risk models built predominantly on male data underdiagnosed heart attacks in women, whose symptoms can differ significantly [8]. Deployment bias arises when a tool developed in one context is applied to another without adaptation. Deployment bias can also appear in the prioritization of AI applications, where tools developed predominantly for specialized or hospital-based clinical contexts may inadvertently overshadow the needs of primary care and essential public health services, areas where data availability and research output are more limited. An algorithm trained in a high-resource tertiary hospital may perform poorly in a rural clinic with different patient demographics and resources [9]. This was also evident in the global response to COVID-19, where digital contact-tracing apps requiring modern smartphones and reliable internet inadvertently excluded older, poorer, and rural populations, the very groups most vulnerable to the pandemic's impact [10,11].

It is important to note that many of the well-documented empirical examples of algorithmic and measurement bias cited in the literature originate from high-income settings, reflecting where AI evaluation and postmarket surveillance mechanisms are most established. This does not imply that such biases are less relevant in Latin America and the Caribbean. On the contrary, health systems in the region frequently rely on imported medical devices, digital tools, and AI-enabled applications developed and validated elsewhere, often without systematic local calibration or performance assessment across diverse populations. The relative scarcity of region-specific evidence therefore represents a structural evidence gap rather than an absence of risk, underscoring the need for stronger regional capacity in evaluation, governance, and postdeployment monitoring of AI-enabled health technologies.

The Digital Context of Health

These technical biases do not exist in a vacuum. They are magnified by the emerging digital landscape, where structural forces, such as commercial and social ones, can shape and/or exacerbate existing disparities in health outcomes [12]. The same platforms that could deliver critical health information can also be vectors for dangerous misinformation or be used to market unhealthy products that may be particularly detrimental to vulnerable populations. Moreover, the vast datasets collected by technology companies, often without transparent consent, may be used in ways that deepen existing patterns of exclusion and social stratification. Addressing algorithmic bias in clinical tools is a critical first step, while also encouraging a digital ecosystem that balances innovation with transparent, responsible use of data in support of health objectives. Building institutional public-sector capacity for AI governance, including regulatory oversight, audit mechanisms, and digital literacy among health professionals, is essential to address health population needs, especially for low- and middle-income countries that face financial and human resource constraints.

Precision Medicine Demands Precision in Data

These biases extend to the most advanced frontiers of medicine. Precision medicine, which aims to tailor care to an individual's unique genetic, environmental, and lifestyle profile, is severely

XSL•FO
RenderX

hampered by a profound lack of diversity in the foundational data. An estimated 80% of participants in genome-wide association studies—the research that underpins many genetic tests—have been of European descent, a group that comprises just 16% of the world's population [13]. The consequences are alarming. The resulting genetic risk scores, used to predict the likelihood of developing diseases like cancer or heart disease, are most accurate for this single, well-studied group. For individuals of African, Asian, or Indigenous ancestry, these scores can be uninformative or, worse, misleading, limiting their clinical utility and the clinical relevance of advanced medical tools for diverse populations and reducing the full potential of precision medicine [12,14]. In Latin America, where more than half of the population identifies as Indigenous, Afro-descendant, or of mixed ancestry, underrepresentation in genomic research limits the accuracy and fairness of emerging AI-based diagnostic tools.

Expanding ancestral diversity in datasets is a critical prerequisite for making precision medicine more accurate, equitable, and clinically relevant across populations. Furthermore, investing in robust, representative data is a catalyst for discovery. Calls for increased diversity in precision medicine research often conflate multiple meanings of representation, spanning genetic ancestry, social identity, and inclusion objectives, which can lead to ambiguity in how datasets are constructed and interpreted [15]. Collecting pharmacogenomic data across diverse populations can prevent one-size-fits-all dosing errors and reveal novel population-specific genetic variants that could inspire new, targeted treatments [16,17]. For companies developing clinical AI tools and medical devices, ensuring reliable performance across all population groups reduces liability risks and facilitates regulatory approval, factors that support broader adoption and market sustainability. Establishing regional genomic research networks and public–private partnerships can help ensure that Latin America and the Caribbean are included in this new frontier of data-driven medicine.

A Framework for Fairness: Bias Mitigation Across the Algorithm Lifecycle

Achieving bias-mitigated AI requires vigilance and intentionality at every stage. A proactive, sociotechnical approach is needed. First, the process must begin with inclusive design and data collection. This means moving beyond convenience sampling to deliberately curate datasets that represent the full spectrum of the target population. Health-related institutions and developers should partner with minority-serving clinics and community organizations to fill critical data gaps. Where data are scarce, techniques like synthetic data augmentation may be carefully explored as a temporary bridge, but they are no substitute for real-world representation [18]. In the region of the Americas, countries can leverage the Information Systems for Health Framework and the Pan American Highway for Digital Health Initiative to establish secure, interoperable, and bias-aware data ecosystems that enable such inclusivity.

Second, it is critical to implement rigorous bias checks and fairness metrics during model building and validation. It is not enough to measure overall accuracy. Performance must be audited by subgroup to identify any significant disparities in error rates. A model that is 95% accurate for one group but only 80% for another is not a sign of robustness; it highlights the need for improved performance across diverse population segments to ensure reliability and impact in health. These audits should be a standard requirement before any algorithm is deployed in a clinical setting [19].

Strengthening algorithm performance across population subgroups also supports domestic preparedness by cultivating a digitally competent, bias-aware health workforce. These efforts align with voluntary standards such as the AI Risk Management Framework developed by the National Institute of Standards and Technology and the Good Machine Learning Practice principles issued by the US Food and Drug Administration, ensuring that AI in health care is both innovative and reliable. Collaborative partnerships between industry, regulators, and health care providers are essential to ensure that these tools are not only clinically sound but widely accessible. Expanding population representation in health datasets is a prerequisite for scientific accuracy, improved generalizability, and reduced liability across AI-enabled medical systems. To operationalize fairness, governments could adopt measurable indicators such as subgroup parity rates, demographic calibration metrics, and public reporting of bias audit results as part of AI system certification or procurement. Fairness metrics are inherently context-dependent and may involve trade-offs between competing objectives such as sensitivity, specificity, and equity across population groups. No single metric or threshold is universally applicable across all clinical or public health use cases. As a result, the selection and interpretation of fairness measures should be informed by clinical context, population characteristics, and local health system priorities and should be complemented by human oversight and local validation.

Third, regulatory agencies should take the lead in advancing transparency and explainability in health AI systems. While voluntary frameworks, such as the Good Machine Learning Practices and the AI Risk Management Framework, are important early steps, they can lay the groundwork for future formal regulations with enforceable requirements and mandatory mechanisms. Developers should be required to provide clear documentation on data sources, demographic representation, and known limitations, often referred to as "datasheets for datasets." This transparency will allow clinical institutions and oversight bodies to better evaluate algorithmic recommendations, ensuring these tools support, rather than replace, sound medical judgment. These actions align with global regulatory trends, including the EU AI Act, which mandates oversight for high-risk AI systems [20]. Regional adaptation of these principles through national AI oversight committees would promote harmonized, context-appropriate governance across the Americas.

Fourth, diverse multidisciplinary teams, including clinicians, data scientists, bioethicists, and representatives of affected communities, are essential to uncovering blind spots that homogeneous teams might miss. Engaging patients and community members can reveal contextual factors and potential harm that are not apparent from a purely technical perspective. Including gender experts, disability advocates, and Indigenous

XSL•FO
RenderX

community representatives in these teams ensures an intersectional approach to fairness.

Finally, strengthening health authority stewardship with clear and strong governance capacities, supported by accountability mechanisms, is essential. Regulatory agencies should define enforceable requirements for algorithmic performance, transparency, and postmarket monitoring. In parallel, institutional review bodies, such as algorithm bias–resistant committees, can play a complementary role within health systems or clinical institutions by independently evaluating tools for bias, transparency, and contextual fit before implementation. Once deployed, ongoing monitoring through real-world bias audits is necessary to detect performance drift and emerging disparities over time. To maintain trust, there must also be well-defined pathways for redress when AI tools are found to cause harm. Governments can also establish national registries of AI systems in health, linking performance data, fairness audits, and postmarket surveillance to ensure accountability and transparency.

To move from principles to practice, governance-oriented actions are needed that can be realistically implemented across diverse health system contexts. Table 1 summarizes a set of priority governance actions that health authorities and regulators can use to operationalize bias mitigation across the AI lifecycle. These actions are intended to be a flexible policy toolbox that can be adapted to different levels of regulatory maturity and resource availability, with particular relevance for low- and middle-income settings. The table summarizes governance-oriented actions proposed in this policy perspective to support bias mitigation in AI-enabled health tools across design, validation, deployment, and postmarket monitoring stages. The actions are intended for health policymakers, regulators, and health system leaders, with particular relevance for low- and middle-income settings in the region of the Americas.

Table . Priority governance actions to mitigate algorithmic bias across the artificial intelligence (AI) lifecycle in health systems.

Governance lever	Priority action	Primary actors	Policy rationale
Transparency and disclosure	Require population-representativeness disclosures for AI-enabled health tools, including data sources, demographic coverage, and known performance limitations.	Regulators, ministries of health.	Enables informed decision-making, reduces hidden bias risks, and supports accountability in procurement and deployment.
Performance assessment	Mandate subgroup-level performance reporting for high-risk AI applications rather than reliance on aggregate accuracy metrics.	Regulators, health technology assessment bodies.	Identifies differential impacts across population groups and prevents masking of inequities.
Predeployment review	Establish multidisciplinary review mechanisms within health authorities to assess bias, transparency, and contextual fit prior to deployment.	Ministries of health, regulatory agencies, ethics committees.	Leverages existing institutional structures to identify risks before scale-up.
Local validation	Promote context-specific validation of AI tools, particularly when systems are imported or adapted from other countries.	Health institutions, regulators, implementers.	Reduces deployment bias and ensures relevance to local populations and care pathways.
Postmarket monitoring	Integrate bias considerations into postmarket surveillance using routine health data and real-world evidence.	Regulators, health system operators.	Detects performance drift and emerging disparities after implementation.
Data governance	Strengthen data governance and interoperability frameworks to improve data quality, disaggregation, and representativeness.	Governments, digital health authorities.	Creates the foundational conditions for equitable AI development over time.
Workforce capacity	Invest in digital literacy and AI governance competencies for regulators, clinicians, and public health professionals.	Ministries of health, academic institutions.	Enables informed oversight and responsible adoption of AI tools.
Procurement and financing	Link fairness, transparency, and accountability criteria to public procurement and funding decisions.	Governments, public purchasers, donors.	Institutionalizes equity as a measurable component of health system quality.

From Ethical Mandate to a Health Standard

Addressing algorithmic bias should not be seen as a barrier to innovation. On the contrary, addressing such bias should be considered a catalyst for better science. An algorithm rigorously tested and validated to perform equally across diverse populations is inherently more robust, reliable, and effective. Confronting and correcting these biases allows AI to serve a dual purpose: improving health outcomes and exposing the

XSL•FO
RenderX

hidden disparities within health systems, using technology as a mirror to reveal where actions, applications, and policies must improve. Embedding bias detection into national quality standards and linking fairness to funding eligibility can institutionalize fairness as a measurable component of health system performance.

Conclusion and Call to Action

As the region of the Americas advances its digital transformation, the Pan American Health Organization urges countries, developers, and regulators to treat algorithmic fairness not as a technical afterthought but as a foundational element of resilient and inclusive health systems. This imperative aligns directly with the Pan American Health Organization's *Eight Guiding Principles of the Digital Transformation of the Health Sector*, notably Principle 2: Equity and Inclusion, which emphasizes the need to ensure that digital health benefits reach all population groups, and Principle 5: Ethical Use of Technology, which calls for the responsible and transparent design of digital tools [21]. Addressing algorithmic bias, from data collection to real-world deployment, is an ethical and health mandate that underpins scientific reliability, system trustworthiness, and equitable health outcomes.

Ensuring fairness also means that AI strengthens primary health care and essential public health functions, not only specialized clinical services, and future work should continue to explore how AI can be developed and deployed in ways that reinforce these levels of care, where evidence generation and data availability have historically been more limited.

AI holds extraordinary potential to enhance early detection, optimize resources, and extend care to underserved populations. But if left unchecked, it risks reinforcing the same data-driven health disparities that public health aims to dismantle. Ensuring that AI systems are designed to recognize and counter bias from the start will help create health technologies that are not only more accurate and reliable, but also truly reflective of the populations they aim to serve.

Acknowledgments

Disclosure of delegation to generative artificial intelligence (AI): The authors declare the use of generative AI tools during the research and manuscript preparation process, under full human supervision. The original concept, analytical framing, and initial draft of this editorial were developed by the authors, informed by institutional experience and discussions grounded in real-world public health cases. Generative AI–enabled tools, including NotebookLM and Gemini Pro (Deep Research), were used to support the identification, retrieval, and systematization of relevant peer-reviewed literature, as well as to cross-check cited evidence against published sources. Subsequently, a generative language model (ChatGPT, version 5.2) was used in a limited manner to support editorial language refinement and clarity. These tools were not used to generate scientific content, interpret data, or formulate conclusions. Responsibility for the final manuscript rests entirely with the authors.

Funding

No external funding was received for the preparation of this manuscript. The views expressed are those of the authors and do not necessarily reflect the official positions of their affiliated institutions.

Conflicts of Interest

None declared.

References

1. Norori N, Hu Q, Aellen FM, Faraci FD, Tzovara A. Addressing bias in big data and AI for health care: a call for open science. Patterns (N Y) 2021 Oct 8;2(10):100347. [doi: 10.1016/j.patter.2021.100347] [Medline: 34693373]
2. Franklin G, Stephens R, Piracha M, et al. The sociodemographic biases in machine learning algorithms: a biomedical informatics perspective. Life (Basel) 2024 May 21;14(6):652. [doi: 10.3390/life14060652] [Medline: 38929638]
3. Castillo A, Rivera-Hernandez M, Moody KA. A digital divide in the COVID-19 pandemic: information exchange among older Medicare beneficiaries and stakeholders during the COVID-19 pandemic. BMC Geriatr 2023 Jan 12;23(1):23. [doi: 10.1186/s12877-022-03674-4] [Medline: 36635684]
4. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science 2019 Oct 25;366(6464):447-453. [doi: 10.1126/science.aax2342] [Medline: 31649194]
5. Maripova T. Mitigating algorithmic bias in predictive models. Am J Eng Technol 2025;07(5):192-201. [doi: 10.37547/tajet/Volume07Issue05-19]
6. Shi C, Goodall M, Dumville J, et al. The accuracy of pulse oximetry in measuring oxygen saturation by levels of skin pigmentation: a systematic review and meta-analysis. BMC Med 2022 Aug 16;20(1):267. [doi: 10.1186/s12916-022-02452-8] [Medline: 35971142]
7. Department of Health and Social Care, UK. New action to tackle ethnic and other biases in medical devices. 2024. URL: https://www.gov.uk/government/news/new-action-to-tackle-ethnic-and-other-biases-in-medical-devices [accessed 2026-02-17]

XSL•FO
RenderX

8. Straw I, Rees G, Nachev P. Sex-based performance disparities in machine learning algorithms for cardiac disease prediction: exploratory study. J Med Internet Res 2024 Aug 26;26(1):e46936. [doi: 10.2196/46936] [Medline: 39186324]
9. Zech JR, Badgeley MA, Liu M, Costa AB, Titano JJ, Oermann EK. Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: a cross-sectional study. PLOS Med 2018 Nov;15(11):e1002683. [doi: 10.1371/journal.pmed.1002683] [Medline: 30399157]
10. Kodali PB, Hense S, Kopparty S, Kalapala GR, Haloi B. How Indians responded to the Arogya Setu app? Indian J Public Health 2020 Jun;64(Supplement):S228-S230. [doi: 10.4103/ijph.IJPH_499_20] [Medline: 32496261]
11. National Academies of Sciences, Engineering, and Medicine. Increasing ancestral diversity to enable precision medicine for neurological and psychiatric disorders. In: Stroud C, Posey Norris SM, Bain L, editors. From Molecular Insights to Patient Stratification for Neurological and Psychiatric Disorders: Proceedings of a Workshop: National Academies Press (US); 2022.
12. Vicuña L. Genetic associations with disease in populations with Indigenous American ancestries. Genet Mol Biol 2024;47Suppl 1(Suppl 1):e20230024. [doi: 10.1590/1678-4685-GMB-2023-0024] [Medline: 39254840]
13. Zhang H, De T, Zhong Y, Perera MA. The advantages and challenges of diversity in pharmacogenomics: can minority populations bring us closer to implementation? Clin Pharma and Therapeutics 2019 Aug;106(2):338-349. [doi: 10.1002/cpt.1491]
14. Bulik-Sullivan B, Finucane HK, Anttila V, et al. An atlas of genetic correlations across human diseases and traits. Nat Genet 2015 Nov;47(11):1236-1241. [doi: 10.1038/ng.3406] [Medline: 26414676]
15. Lee SSJ, Fullerton SM, McMahon CE, et al. Targeting representation: interpreting calls for diversity in precision medicine research. Yale J Biol Med 2022 Sep;95(3):317-326. [Medline: 36187415]
16. Claw KG, Dorr CR, Woodahl EL. Implementing community-engaged pharmacogenomics in Indigenous communities. Nat Commun 2024 Jan 31;15(1):920. [doi: 10.1038/s41467-024-45032-5] [Medline: 38296967]
17. Joseph J. Algorithmic bias in public health AI: a silent threat to equity in low-resource settings. Front Public Health 2025;13(13):1643180. [doi: 10.3389/fpubh.2025.1643180] [Medline: 40771228]
18. Zenone M, Kenworthy N, Maani N. The social media industry as a commercial determinant of health. Int J Health Policy Manag 2023;12:6840. [doi: 10.34172/ijhpm.2022.6840] [Medline: 35490262]
19. Silva J, Marti M, Báscolo E, et al. Digital transformation to overcoming access barriers. Rev Panam Salud Publica 2025 Nov 12;49:1. [doi: 10.26633/RPSP.2025.112]
20. Cross JL, Choma MA, Onofrey JA. Bias in medical AI: implications for clinical decision-making. PLOS Digit Health 2024 Nov;3(11):e0000651. [doi: 10.1371/journal.pdig.0000651] [Medline: 39509461]
21. Eight Guiding Principles of Digital Transformation of the Health Sector. A Call to Pan American Action. Pan American Health Organization. 2021. URL: https://iris.paho.org/handle/10665.2/54256 [accessed 2026-02-11]

Abbreviations

AI: artificial intelligence

Edited by A Mavragani, T Sanchez; submitted 25.Nov.2025; peer-reviewed by B Patel, K Khashu, W Qin; revised version received 26.Dec.2025; accepted 29.Dec.2025; published 23.Feb.2026.

Please cite as:
Barbosa da Silva Jr J, Birminghamm M, Rivière Cinnamond A, Boisson E, Valdez ML, Garcia Saiso S, Souza JP, Marti MC, Richards LM, Guzman J, Nelson J, Pesce K, Haddad AE, Fitzgerald J, Bascolo E, D'Agostino M
Bias-Mitigated AI as a Foundation for Resilient and Effective Health Systems
JMIR Public Health Surveill 2026;12:e88457
URL: https://publichealth.jmir.org/2026/1/e88457
doi:10.2196/88457

© Jarbas Barbosa da Silva Jr, Maureen Birminghamm, Ana Rivière Cinnamond, Eldonna Boisson, Mary Lou Valdez, Sebastian Garcia Saiso, Joao Paulo Souza, Myrna Marti, Leah-Marie Richards, Javier Guzman, Jennifer Nelson, Karina Pesce, Ana Estela Haddad, James Fitzgerald, Ernesto Bascolo, Marcelo Dagostino. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 23.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Viewpoint

The 100-Miles for QVOICES (Queer Voices in Oncology Igniting Change Through Empowered Stories of Survivorship) Campaign: Overview of the Benefits and Costs of Crowdfunding as a Funding Mechanism for Research

Viktor Clark[1,2], BA, HC-I, MS, PhD; Milena E Insalaco[1,2], MS; Charles K Kamen[1,2], MPH, PhD; Lauren V Ghazal[2,3], BA, BSN, MSN, PhD

[1]Division of Supportive Care in Cancer, School of Medicine, University of Rochester, Rochester, NY, United States
[2]Wilmot Cancer Center, University of Rochester, Rochester, NY, United States
[3]School of Nursing, University of Rochester, Rochester, NY, United States

Corresponding Author:
Viktor Clark, BA, HC-I, MS, PhD
Division of Supportive Care in Cancer
School of Medicine
University of Rochester
265 Crittenden Blvd
Rochester, NY, 14642
United States
Phone: 1 (585) 275 9959
Email: viktor_clark@urmc.rochester.edu

Abstract

In today's unstable funding climate, alternative research funding mechanisms are essential. Investigators are facing novel barriers to acquiring funding for their research. Early-stage investigators, in particular those who are beginning their careers in research, are especially vulnerable to these funding disruptions. Currently, finding alternative funding is especially relevant for scientists studying historically and intentionally excluded communities, such as lesbian, gay, bisexual, transgender, queer, or other sexual and/or gender expansive (LGBTQ+) populations, as many investigators have had their LGBTQ+ grants revoked by the Trump administration. Crowdfunding for research studies is a potential funding avenue that has grown in popularity, with more than US $12 million raised since 2012. This Viewpoint highlights crowdfunding as a potential funding model for LGBTQ+ cancer research, specifically to support the collection of preliminary data for predoctoral students and early-stage investigators. We describe the benefits and challenges of using crowdfunding as the sole funding mechanism for a mixed methods observational study among LGBTQ+ survivors of cancer. Despite challenges, crowdfunding can fund rigorous research on health disparities, notably in historically and intentionally excluded communities.

(JMIR Public Health Surveill 2026;12:e80079) doi:10.2196/80079

KEYWORDS

crowdfunding; early-stage investigators; LGBTQ+; cancer; survivorship

Background

Predoctoral Students and Early-Stage Investigators

Predoctoral students are currently enrolled in a doctoral program and eligible to conduct research [1], whereas early-stage investigators are defined as researchers in the early years (typically within 10 years) of their terminal degree who have not yet received a substantial independent research award (eg, R00 and R01) [2]. As of 2022, there were more than 250,000 full-time doctoral students in the fields of science, engineering, and health [3]; more than 55,000 doctoral degrees were awarded, more than 34,000 medical students were matched, and more than 62,000 postdoctoral trainees began their careers in the United States, representing a combined generation of early-stage investigators and predoctoral students [4]. As required by many independent research awards, predoctoral students and early-stage investigators must acquire preliminary data to achieve a successful funding application [5]. However, there are limited government-supported funding mechanisms that

XSL•FO
RenderX

provide resources to develop such preliminary data (eg, small R03s, R21s, and R15s) [6], and these have recently become even more restricted for research investigating historically and intentionally excluded communities [7]. Determining innovative strategies for collecting preliminary data among historically and intentionally excluded communities is necessary to ensure that researchers focusing on these disparities can progress in their careers and that the health disparities faced by these communities continue to be studied to work toward amelioration.

Crowdfunding

Crowdfunding (ie, the raising of money, capital, or other resources from a large number of people to support a project or venture) [8] is an innovative strategy that has been used to assist in funding for-profit business ventures, nonprofit social entrepreneurship projects, and even to offset medical expenses [9]. Since 2012, crowdfunding has grown, even within the scientific community, especially on platforms such as Experiment.com, which has been used to help fund more than 3000 research projects and raise more than US $12 million [10]. Crowdfunding has emerged as a notable phenomenon that researchers are examining to identify which types of projects are best supported by this financing model [11]. The growth of crowdfunding has been particularly notable among students and early-stage investigators who may require alternative funding pathways beyond traditional government-based models [12]. Crowdfunding may be particularly salient in light of the current funding crisis as a result of numerous executive orders signed by President Donald Trump [13].

Current Status of Research for Lesbian, Gay, Bisexual, Transgender, Queer, or Other Sexual and/or Gender Expansive Communities

Even before the executive orders implemented by President Trump, funding for research focused on some specific minoritized communities has been bleak, specifically among lesbian, gay, bisexual, transgender, queer, or other sexual and/or gender expansive (LGBTQ+) research efforts. Between 1989 and 2014, the National Institutes of Health (NIH) allocated 0.1% of its total budget toward LGBTQ+ research that was not related to HIV [14]. In 2020, this increased to 0.8%, when NIH funded 500 non–HIV-related projects pertaining to LGBTQ+ individuals (US $284,051,683) [15]. In 2015, NIH developed the Office of Sexual and Gender Minority Research, which helped support leading efforts to collect sexual orientation and gender identity data to determine the extent of health disparities between LGBTQ+ and cisgender heterosexual individuals [16]. With the support of the Office of Sexual and Gender Minority Research, longitudinal studies have been conducted that have revealed substantive disparities in health care for LGBTQ+ individuals. Examples of these disparities include higher rates of drug, alcohol, and tobacco use; greater mental health concerns, especially regarding aging; and a higher risk of several cancers (eg, skin, cervical, and anal) among LGBTQ+ individuals compared with cisgender heterosexual individuals [17,18]. In early 2025, the Office of Sexual and Gender Minority Research was dismantled through executive orders issued by President Trump [19]. Furthermore, in these executive orders, LGBTQ+ research has been deemed superfluous, leading to significant gaps in governmental funding [20]. As of 2025, nearly 10% (1 in 10) of the US adult population identifies as LGBTQ+ [21]. This proportion is even higher among younger generations, with nearly 30% (1 in 3) identifying as LGBTQ+ [22]. Consequently, health care systems are now tasked with providing care to LGBTQ+ individuals, yet many remain ill-equipped to do so [23]. However, disparities faced by LGBTQ+ people and gaps in knowledge regarding their experiences, particularly across the cancer control continuum, persist [24].

Viewpoint Aim

As these disparities persist and government funding becomes more restricted, the need for alternative funding for LGBTQ+ cancer research remains a high priority. In this Viewpoint, we highlight crowdfunding as a potential funding model for LGBTQ+ cancer research as an alternative to government-funded research for the collection of preliminary data, particularly for predoctoral students and early-stage investigators. We describe how crowdfunding was used for the QVOICES (Queer Voices in Oncology Igniting Change through Empowered stories of Survivorship) study, a mixed methods observational study among LGBTQ+ survivors of cancer.

Study Overview

Crowdfunding for QVOICES

In 2022, the QVOICES study was launched [25]. QVOICES was a mixed methods dissertation led by the first author (VC) and used crowdfunding to provide livable wage incentives for participants in the study. The QVOICES crowdfunding campaign was titled “100 Miles for QVOICES.” The first author (VC) engaged in a 100-mile (160-km) bike ride in support of this campaign. Donors to the crowdfunding campaign were encouraged to make donations ranging from US $1 to US $100. The campaign was shared across social media platforms (ie, Instagram [Meta Platforms, Inc] and Facebook [Meta Platforms, Inc]) and through email. Both social media ads and emails provided links that directed potential donors to a main crowdfunding webpage hosted on BetterWorld, a fundraising and crowdfunding platform focused mainly on nonprofit, charitable, and social impact initiatives [26]. BetterWorld was chosen due to its elimination of hosting fees, unlike other crowdfunding platforms such as GoFundMe or Kickstarter [26]. The main crowdfunding webpage had details about the QVOICES project and a self-created video that gave an overview of how donors’ funds would be used (Figure 1).

Funds were split into 3 intentions. First, funds were allocated to participant incentives. Second, funds were applied to offset study costs. Third, any remaining funds were donated to the National LGBT Cancer Network [27].

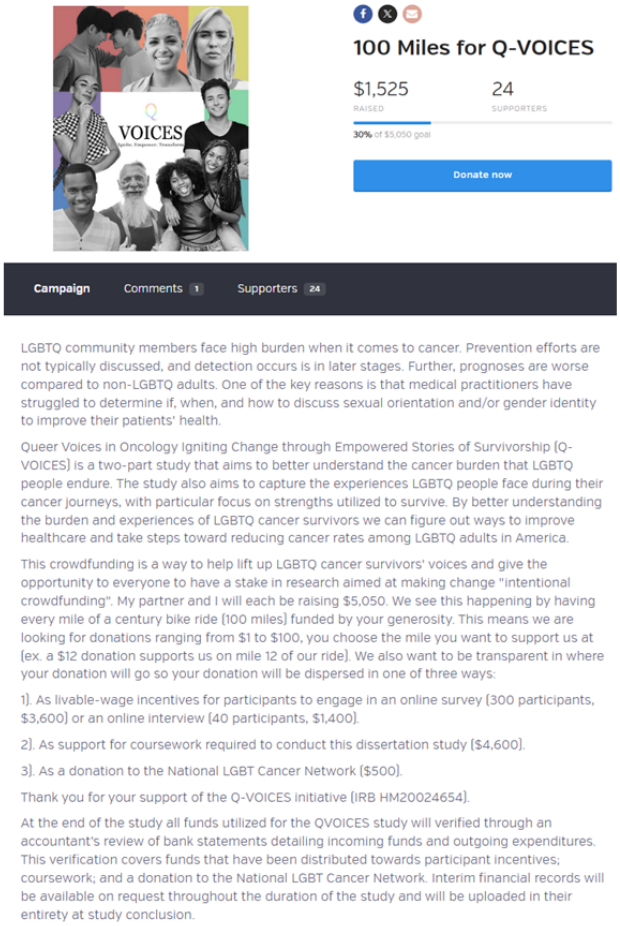
To reach key stakeholders, crowdfunding campaign materials were distributed in several ways. First, texts and emails were sent to the social networks of the first author (VC), including friends and family (Figure 2). Second, emails with an electronic flyer (Figure 3) were sent to (1) diversity, equity, and inclusion offices; (2) faculty engaged in LGBTQ+ research efforts, identified through faculty biographies, publication histories,

XSL•FO
RenderX

and abstracts; (3) National Cancer Institute–designated cancer centers determined to be LGBTQ+ inclusive, identified using the National LGBT Cancer Network "LGBT Welcoming Treatment Providers" tool [28]; and (4) LGBTQ+ affinity groups at various universities. At each National Cancer Institute–designated cancer center, faculty listed as LGBT-friendly providers by the National LGBT Cancer Network tool were specifically contacted. Within the LGBTQ+ affinity groups, the president or vice president was contacted if their information was available; if not, the general information email was used. Finally, the campaign was also posted in LGBTQ+ Facebook groups and on Instagram using Instagram carousels (Figure 4), with numerous hashtags (eg, #QVOICES, #crowdfunding, #LGBTQ, #cancer, #research, #donate, #ignite, #empower, and #transform).

The 100 Miles for QVOICES crowdfunding campaign ran from March 1, 2022, to June 30, 2022. Approximately 500 emails and texts were distributed, and Facebook ads were run daily during this period. An estimated US $150 was spent on Facebook ads. For marketing efforts, Instagram stories were also created every other day on a QVOICES-specific account between March 1, 2022, and June 30, 2022. QVOICES was approved by the Virginia Commonwealth University institutional review board (IRB) as an exempt study for all participant engagement. Funding initiatives through crowdsourcing were outside the scope of IRB jurisdiction.

Figure 1. QVOICES (Queer Voices in Oncology Igniting Change through Empowered stories of Survivorship) website.



XSL•FO
RenderX

Figure 2. 100-Miles for QVOICES (Queer Voices in Oncology Igniting Change through Empowered stories of Survivorship) texts.

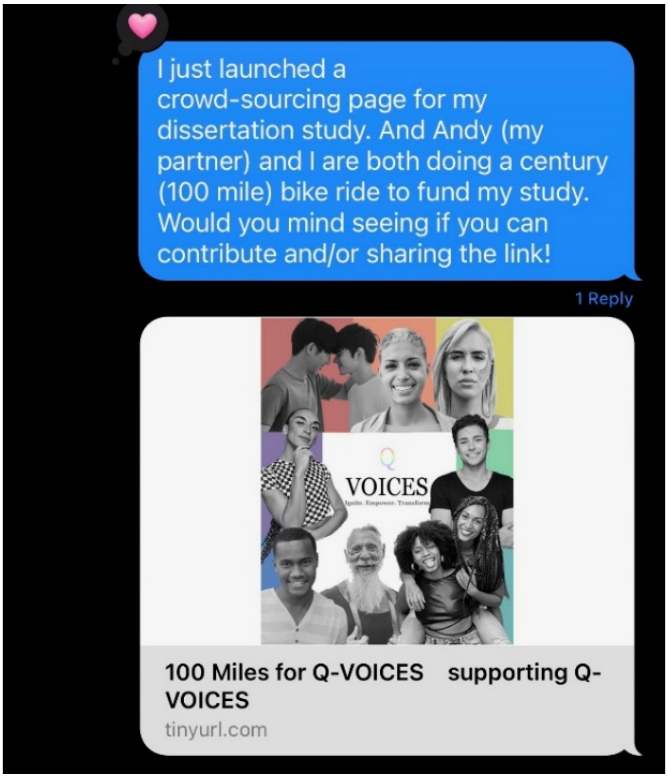
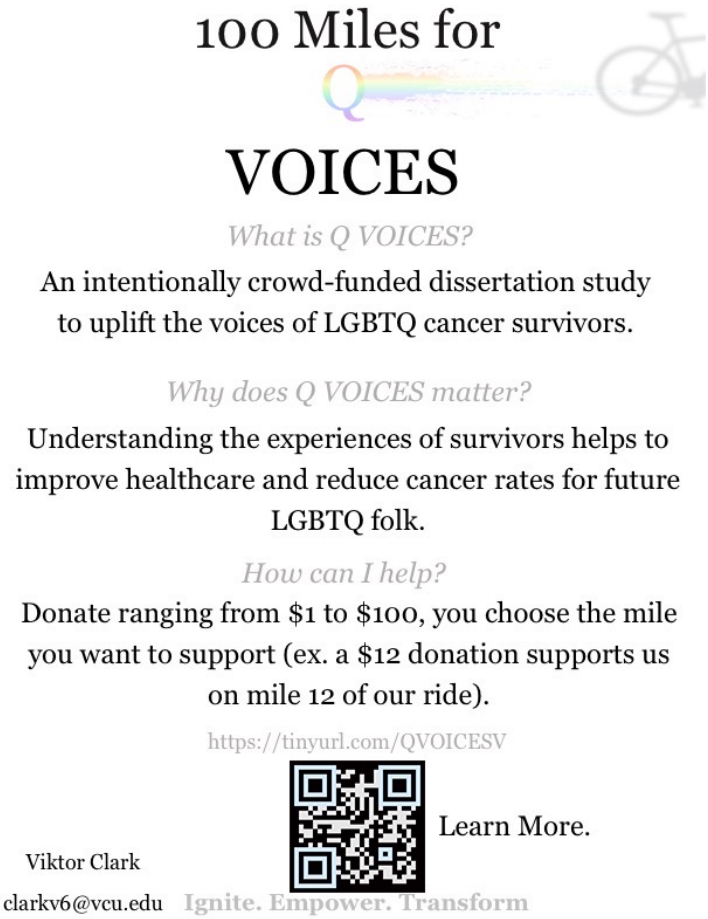
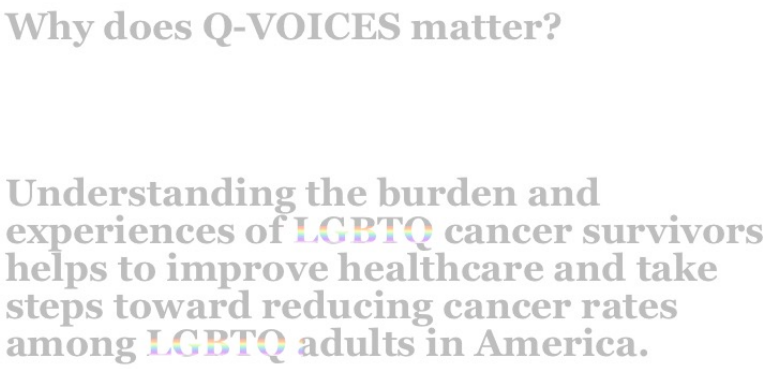


Figure 3. 100-Miles for QVOICES (Queer Voices in Oncology Igniting Change through Empowered stories of Survivorship) electronic flyer.



XSL•FO
RenderX

Figure 4. 100-Miles for QVOICES (Queer Voices in Oncology Igniting Change through Empowered stories of Survivorship) Instagram carousel.



Crowdfunding Results and Benefits

Overview

Crowdfunding efforts resulted in a total of US $2218 from 26 donors. Of these donations, the majority (n=24, 92%) were acquired through text messages from friends and family or through word-of-mouth referrals from friends and family. Two donations were received from faculty at institutions engaged in LGBTQ+ research.

Benefits

A crowdfunding approach to the QVOICES project led to several benefits. First, researchers were able to engage with community stakeholders and other scientists [29]. Second, branding around QVOICES established a means for all project-related materials to be interconnected [30]. Third, there was validation that LGBTQ+ focused cancer research was a priority, as evidenced by funding support [31]. Finally, maintaining autonomy in all funding decisions allowed for the use of innovative and attractive incentives (eg, participants were able to select their own gift cards [1 of 75 different types]) as compensation for engagement in the study [32].

Lessons Learned

Engaging With Community Stakeholders and Scientists

Throughout the distribution of crowdfunding materials, community stakeholders and scientists came together to help spread the word about the 100 Miles for QVOICES campaign. LGBTQ+ community members and scientists were often not direct sources of donations but instead served as champions by sharing information with allies who ultimately contributed. This was an interesting phenomenon and may be explained by LGBTQ+ community members historically being tasked with raising funds for their own initiatives and a new mentality shift toward collaborating with those in privileged positions to make change for marginalized communities [33].

Regarding engaging with scientists, past research has argued that those with postgraduate degrees have been found to be one of the lowest groups to contribute to crowdfunding efforts [29]. However, although their contributions to crowdfunding have been low, academic scientists over the past decade have been more accepting of using crowdfunding as a practical way to offset funding shortages and continue research initiatives [34].

Ethics of Crowdfunding

As noted, QVOICES being funded through crowdfunding efforts meant that the materials we disseminated to donors were not subject to IRB oversight. However, we carefully maneuvered how to engage with our donors and demonstrate tangible results from their contributions. On our donation page, we had a disclaimer stating that donating to the QVOICES project did not provide donors with rights to individual participant data; rather, the data would be reported as summative data in accordance with the IRB.

It was also important for us to be as transparent as possible with the use of funds that we were raising; therefore, we had detailed information on fund allocation laid out in a three-item financial budget: (1) livable wage participant incentives (US $8 per 15-minute survey and US $40 per hour interview), (2) research costs (ie, transcription software), and (3) a donation to the National LGBT Cancer Network if goal of US $5000 was met. We also informed donors that QVOICES had specific branding and that, in accordance with this branding, donors could search for and find any publications or conference proceedings connected to the QVOICES study. Donors' names, should they want them to appear in any disseminated materials, would be included on all manuscripts and conference proceedings, similar to that of a funding agency. All these efforts were to enhance the ethics of keeping the robustness and high quality of data collection at a peak, to keep the confidentiality of our participants safe, and to have our donors stay engaged and recognize that we were devoted to using their funds in a meaningful way.

Branding

Due to the increasing number of crowdfunding efforts, branding has become even more essential. Branding is defined as a means to differentiate oneself from competitors [35] and can include logos, messaging, and calls to action [36]. Successful branding has been determined to be instrumental in crowdfunding decision-making [37]. During the process of developing the 100-Miles QVOICES campaign, branding was an essential target in our development.

This was achieved through several means. First, QVOICES developed a distinct logo that was applied to all materials distributed for the campaign and the study (Figure 5). The pastel rainbow coloring of the "Q" in the logo intentionally established the LGBTQ+ community as the target audience. We also used

XSL•FO
RenderX

a simple 3-word tagline, "Ignite. Empower. Transform," which summarized our intention across all QVOICES efforts to use LGBTQ+ voices to *ignite* change, *empower* individuals, and *transform* cancer outcomes for these communities. Now, at the dissemination phase of QVOICES, this logo is placed on all posters, oral presentations, and manuscripts to recognize the crowdfunding donors and research participants who similarly supported the brand and research efforts of QVOICES.

Figure 5. QVOICES logo.

Validation and Prioritization of LGBTQ+ Cancer Research Among Donors

One of the most important aspects of crowdfunding is that financial support validates a crowd's belief in an effort [38]. We were able to achieve somewhat of a local validation. Our initial goal was US $5000, and we were successful in funding US $2218 (44.4%). We also had a <5% of response rate, resulting in 26 individual donations. Among these donors, LGBTQ+ cancer research was seen as a priority. This was reflected in our study, requesting donations ranging from US $1 to $100 to represent the 100 miles that would be biked throughout the campaign and achieving an average of US $88 (SD US $68.68) per donation.

Financial Autonomy

Finally, due to our funding resulting from crowdfunding, there were many aspects of the government-required financial process that we were able to avoid. One such process was the collection of social security numbers. Due to our community of interest (LGBTQ+ individuals), many participants were concerned about providing their social security numbers; therefore, reducing this barrier to participation was essential. In addition, often incentives need to be distributed at the conclusion of interviews or a few days later due to the processing of funds through systems. However, due to having autonomy in our funds, we were able to distribute our interview incentives at the beginning of the interview, allowing participants the autonomy to discontinue their participation at any point without penalty. We believe this approach led to more engagement, retention, and transparency throughout the interviews.

Barriers to Crowdfunding

Although there were many benefits to crowdfunding, there were also some costs that came with using crowdfunding, both in our marketing stream of the 100-Miles for QVOICES campaign and in using crowdfunding as our sole source of funding.

Mistrust or Stigma of the LGBTQ+ Topic Area

Because our campaign was run through the site BetterWorld, which is a newer platform and not as popular as Kickstarter or GoFundMe, this may have impacted donors' trust in the campaign. Additionally, because the majority of emails sent were "cold" emails (ie, initial contact made with the individual by our study team) and social media posts were cast to a large network, acquiring trust that we were going to use donations toward the promised efforts may have been difficult to obtain. Furthermore, because the emails were sent cold, many administrators, faculty, affinity officers, social media users, and text recipients were unfamiliar with our principal investigator (VC), which may have also caused some hesitation in donating toward the 100-Miles for QVOICES efforts.

Additionally, a complex stigma has been established for LGBTQ+ academicians regarding whether to engage in or even endorse LGBTQ+ research [39]. This complexity has only grown due to the politically hostile environment created by various executive orders under the Trump presidency [13]. QVOICES data were collected from June 2020 to January 2021 (during the first Trump presidency). Within this time, more than 150 anti-LGBTQ+ bills were introduced into state legislation [40]. This politically hostile environment may have further dissuaded people from donating due to concerns regarding their academic standings, even if they were interested in supporting the findings of the QVOICES dissertation work. Future studies should examine how trust in the campaign, trust in the sponsor of the campaign, and trust in the project topic impact intent to donate and conversion to actual donation from key stakeholders.

Lack of Sense of Urgency

In our campaign materials distributed through emails, texts, Facebook ads, and Instagram posts, we included (1) a blurb with a hyperlink to the BetterWorld crowdfunding site; (2) a video describing both the QVOICES study and the 100-mile campaign, with the website to the BetterWorld crowdfunding site included at the end; and (3) a PDF flyer for the 100-mile campaign with a call-to-action for how donors could get involved with a clickable hyperlink to the BetterWorld crowdfunding site. We felt that providing all 3 mediums maximized the potential for donations.

However, a more iterative and interactive modality may have resulted in better crowdfunding donation reactions [37]. This may have included developing Instagram posts with open comments, sending materials out in iterative waves (eg, first with blurbs about the study, then a flyer, and then a video) to

XSL•FO
RenderX

keep the crowdfunding materials accessible to potential donors, and using daily updates and discussions through the "BetterWorld" campaign page. Overall, while we had many well-developed marketing materials, the implementation of these materials could have been improved.

Additionally, providing an option for participants to opt out of iterative marketing materials, with a reason (eg, not interested in the project or have received too many funding requests about the topic area), may help to establish if saturation among key stakeholders on the topic area has been reached and whether an expansion of stakeholders is necessary.

Nuance

Crowdfunding has typically been used to fund business ventures and nonprofit social entrepreneurship projects and even to offset medical expenses [9]. Due to the nuance of crowdfunding being used within the research sector, this may have created some hesitation to engage in donation. According to the theory of Diffusion of Innovations [41], often with innovations, only 2.5% of individuals are seen as innovators or the first individuals to adopt an innovation, which is consistent with our low rate of return, as indicated by individuals who actually donated to our campaign.

If crowdfunding is to remain a permanent means of funding research, this rate should grow to approximately 13.4% of individuals, representing early adopters [41], followed by another 34% as the early majority, another 34% as the late majority [41], and finally 16% representing individuals who are the last to adopt crowdfunding as a reasonable means to fund research [41]. In alignment with the theory of Diffusion of Innovations, our goal of US $5000 was not met, and we had a <5% return rate. However, we see the 100-Miles for QVOICES campaign as the cutting edge of this innovation curve and anticipate crowdfunding to become more widely adopted as funding options shift.

Future Directions

From the 100-Miles for QVOICES campaign and the QVOICES study, we have learned many fundamental lessons to carry out other successful crowdfunding campaigns for research in the future. One of the biggest takeaways from this campaign was that when given the opportunity, donors—if interested and if barriers to donation are overcome—will rise to support the efforts of researchers, when they trust the effort, know that the research will be carried out, and know that their funds will be used in an effective and transparent way.

Our biggest lesson learned is that having a brand is only one step of the solution toward a successful campaign. Additionally, ensuring that the brand is backed by a strong marketing strategy, that there are large audiences to pivot toward, and that techniques exist to overcome barriers such as stigma are all essential to reaching the overarching goal of a campaign—particularly when conducting LGBTQ+ research. Future directions should apply these strategies to (1) various funding initiatives to enable comparative metrics to determine if topic area matters and (2) expand funding initiatives beyond solely participant incentives (eg, study materials, personnel, and salary support).

Conclusions

Crowdfunding can be a functional method for funding research that may need minimal support from nontraditional avenues (eg, outside of government funding or foundational grants), such as that of preliminary data collection among early-stage investigators. As government funding becomes more constrained and foundational grants become more saturated, crowdfunding provides an additional alternative for financial support [42]. Through the 100-Miles of QVOICES campaign, we have seen a glimpse of success that a crowdfunding campaign can bring in ensuring that high-quality research can be conducted and disseminated. Although crowdfunding can come with various challenges, with thoughtful planning, branding, marketing, and support, the possibilities of using crowdfunding as an innovative strategy to conduct research addressing health disparities is achievable.

Funding

No external funding was provided for the development of this project; all funding was provided through generous donations collected via the 100-Miles for QVOICES crowdfunding campaign.

Conflicts of Interest

None declared.

References

1. Pre-doctoral students. National Institutes of Health. URL: https://ors.od.nih.gov/pes/dis/AdministrativeStaff/Pages/predoctoralstudents.aspx [accessed 2025-11-30]
2. New and early-stage investigators. National Cancer Institute. URL: https://cancercontrol.cancer.gov/brp/funding/new-and-early-stage-investigators [accessed 2025-11-30]
3. Heuer R, Einaudi P, Kang K. Research doctorate conferrals rebound, leading to record number of U.S. doctorate recipients in 2022. National Center for Science and Engineering Statistics, National Science Foundation. 2022. URL: https://ncses.nsf.gov/pubs/nsf23353 [accessed 2025-11-30]

XSL•FO
RenderX

4. Smith B, Arbeit CA, Thompson H, Yamaner MI. Graduate enrollment and postdoctoral appointments in science, engineering, and health rise, driven largely by increase in number of women and temporary visa holders. National Center for Science and Engineering Statistics, National Science Foundation. URL: https://ncses.nsf.gov/pubs/nsf25316 [accessed 2025-11-30]
5. Highlight preliminary data in your next application. National Institute of Allergy and Infectious Diseases, National Institutes of Health. URL: https://www.niaid.nih.gov/grants-contracts/highlight-preliminary-data-your-next-application [accessed 2025-12-02]
6. Supported programs and mechanisms. Office of Research Infrastructure Programs, National Institutes of Health. URL: https://orip.nih.gov/funding/supported-programs-mechanisms [accessed 2025-12-02]
7. NIH terminates research grants for LGBTQ+, gender identity, and DEI studies. Columbia Law School. URL: https://climate.law.columbia.edu/content/nih-terminates-research-grants-lgbtq-gender-identity-and-dei-studies [accessed 2025-12-01]
8. Calic G. Crowdfunding. In: Warf B, editor. The SAGE Encyclopedia of the Internet. Thousand Oaks, CA: Sage Publications; 2025.
9. Smith T. Crowdfunding: what it is, how is works, and popular websites. P2PMarketData. URL: https://p2pmarketdata.com/articles/crowdfunding/ [accessed 2025-05-25]
10. Community statistics. Experiment. URL: https://experiment.com/stats [accessed 2025-04-23]
11. Mollick E. The Wharton crowdfunding study: understanding the impact of crowdfunding on creators, backers, and society. The Wharton School. URL: https://crowdfunding.wharton.upenn.edu/ [accessed 2025-11-25]
12. Sauermann H, Franzoni C, Shafi K. Crowdfunding scientific research: descriptive insights and correlates of funding success. PLoS One 2019;14(1):e0208384 [FREE Full text] [doi: 10.1371/journal.pone.0208384] [Medline: 30608943]
13. Kozlov M, Ryan C. How Trump 2.0 is slashing NIH-backed research - in charts. Nature 2025 Apr 10;640(8060):863-865. [doi: 10.1038/d41586-025-01099-8] [Medline: 40211085]
14. Coulter RW, Kenst KS, Bowen DJ, Scout C. Research funded by the National Institutes of Health on the health of lesbian, gay, bisexual, and transgender populations. Am J Public Health 2014 Feb;104(2):e105-e112. [doi: 10.2105/AJPH.2013.301501] [Medline: 24328665]
15. National Institutes of Health. 2020. URL: https://web.archive.org/web/20250104070652/https://dpcpsi.nih.gov/sites/default/files/2023-09/SGMRO-Overview-Snapshot-2020-508.pdf [accessed 2026-02-11]
16. Sexual and Gender Minority Research Office. URL: https://sgmresearch.com/ [accessed 2025-05-20]
17. Sell RL, Holliday ML. Sexual orientation data collection policy in the United States: public health malpractice. Am J Public Health 2014 Jun;104(6):967-969. [doi: 10.2105/AJPH.2013.301814] [Medline: 24825192]
18. Hogan J. Understanding cancer disparities among the LGBTQ+ community. Rutgers Cancer Institute. 2023. URL: https://cinj.org/understanding-cancer-disparities-among-lgbtq-community [accessed 2025-11-25]
19. Gaffney T. NIH puts former sexual and gender minority office employees on administrative leave. STAT. 2025. URL: https://www.statnews.com/2025/03/06/nih-puts-former-sexual-and-gender-minority-office-employees-on-administrative-leave/ [accessed 2025-04-28]
20. Nemec M. Trump anti-LGBTQ+ executive order litigation tracker. The National LGBTQ+ Bar Association. URL: https://lgbtqbar.org/programs/advocacy-resources/trump-executive-order-tracker/ [accessed 2025-12-03]
21. Jones JM. LGBTQ+ identification in U.S. rises to 9.3%. Gallup, Inc. 2025. URL: https://news.gallup.com/poll/656708/lgbtq-identification-rises.aspx [accessed 2025-04-28]
22. A political and cultural glimpse into America's future: generation z's views on generational change and the challenges and opportunities ahead. Public Religion Research Institute. URL: https://prri.org/research/generation-zs-views-on-generational-change-and-the-challenges-and-opportunities-ahead-a-political-and-cultural-glimpse-into-americas-future/ [accessed 2025-12-03]
23. Montero A, Hamel L, Artiga S, Dawson L. LGBT adults' experiences with discrimination and health care disparities: findings from the KFF survey of racism, discrimination, and health. KFF. 2024. URL: https://tinyurl.com/76dmatve [accessed 2025-05-25]
24. Beaton S, McKenzie-Johnson T. Barriers to cancer care in the LGBTQ+ community. J Surg Oncol 2024 Dec 11;130(7):1490-1495. [doi: 10.1002/jso.27980] [Medline: 39523924]
25. Clark V, Brown RF, Gada U, Kamen CS. Resilience related to anxiety, self-esteem, and internalized lesbian, gay, bisexual, and/or transgender negativity: a stepwise regression model among lesbian, gay, bisexual, transgender, queer, and/or another gender and/or sexually expansive identity cancer survivors. Transl Iss Psychol Sci 2025 Mar;11(1):107-118. [doi: 10.1037/tps0000451]
26. Online crowdfunding platform for nonprofits. BetterWorld. URL: https://betterworld.org/fundraising-software/crowdfunding/ [accessed 2025-04-25]
27. The National LGBT Cancer Network. 2025. URL: https://www.cancer-network.org/ [accessed 2025-04-28]
28. The National LGBT Cancer Network. 2025. URL: https://www.cancer-network.org/cancer-treatment-screening-providers/ [accessed 2026-02-09]
29. Byrnes JE, Ranganathan J, Walker BL, Faulkes Z. To crowdfund research, scientists must build an audience for their work. PLoS One 2014;9(12):e110329 [FREE Full text] [doi: 10.1371/journal.pone.0110329] [Medline: 25494306]

XSL•FO
RenderX

30. Newman TP, Beets B. Exploring the brand of science: implications for science communication research and practice. J Sci Commun 2023;22(02):A05. [doi: 10.22323/2.22020205]
31. Mollick E. The dynamics of crowdfunding: an exploratory study. J Bus Ventur 2014 Jan;29(1):1-16. [doi: 10.1016/j.jbusvent.2013.06.005]
32. Best-in-class gift cards platform for business rewards. BHN Rewards. URL: https://www.bhnrewards.com/rewards-gifts/ [accessed 2025-04-25]
33. Power R, Ussher JM, Perz J, Allison K, Hawkey AJ. "Surviving discrimination by pulling together": LGBTQI cancer patient and carer experiences of minority stress and social support. Front Oncol 2022;12:918016 [FREE Full text] [doi: 10.3389/fonc.2022.918016] [Medline: 35814403]
34. Lenart-Gansiniec R, Chen J. When will scientists say yes? Antecedents, consequences and limitations of crowdfunding in research. Stud High Educ 2023 Mar 18;48(8):1159-1171. [doi: 10.1080/03075079.2023.2186388]
35. Keller KL. Brand. In: Augier M, Teece DJ, editors. The Palgrave Encyclopedia of Strategic Management. London, United Kingdom: Palgrave Macmillan; 2016:123-128.
36. Gratton P. Brand identity: what it is and how to build one. Investopedia. 2025. URL: https://www.investopedia.com/terms/b/brand-identity.asp [accessed 2025-05-25]
37. Ellis J. Does branding matter for your crowdfunding campaign? Crowdfund Better. URL: https://www.crowdfundbetter.com/does-branding-matter-for-your-crowdfunding-campaign/ [accessed 2025-05-29]
38. Strohmaier D, Zeng J, Hafeez M. Trust, distrust, and crowdfunding: a study on perceptions of institutional mechanisms. Telemat Inform 2019 Oct;43:101252. [doi: 10.1016/j.tele.2019.101252]
39. Veldhuis CB. Doubly marginalized: addressing the minority stressors experienced by LGBTQ+ researchers who do LGBTQ+ research. Health Educ Behav 2022 Dec;49(6):960-974 [FREE Full text] [doi: 10.1177/10901981221116795] [Medline: 35972197]
40. Past legislation affecting LGBT rights across the country 2020. American Civil Liberties Union. URL: https://www.aclu.org/documents/past-legislation-affecting-lgbt-rights-across-country-2020 [accessed 2025-12-05]
41. Rogers EM. Diffusion of Innovations. 5th edition. New York, NY: Free Press; 2003.
42. Pitts F. NIH research funding freeze brings disruption. Chartis. 2025. URL: https://www.chartis.com/insights/nih-research-funding-freeze-brings-disruption [accessed 2025-05-25]

Abbreviations

IRB: institutional review board
LGBTQ+: lesbian, gay, bisexual, transgender, queer, or other sexual and/or gender expansive
NIH: National Institutes of Health
QVOICES: Queer Voices in Oncology Igniting Change through Empowered stories of Survivorship

Edited by A Mavragani, T Sanchez; submitted 03.Jul.2025; peer-reviewed by E Warner, O Olowokere; comments to author 10.Nov.2025; revised version received 19.Dec.2025; accepted 19.Jan.2026; published 26.Feb.2026.

Please cite as:
Clark V, Insalaco ME, Kamen CK, Ghazal LV
The 100-Miles for QVOICES (Queer Voices in Oncology Igniting Change Through Empowered Stories of Survivorship) Campaign: Overview of the Benefits and Costs of Crowdfunding as a Funding Mechanism for Research
JMIR Public Health Surveill 2026;12:e80079
URL: https://publichealth.jmir.org/2026/1/e80079
doi:10.2196/80079
PMID:

©Viktor Clark, Milena E Insalaco, Charles K Kamen, Lauren V Ghazal. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 26.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Everyday Digital Technology Use and Youth Health: Scoping Review of Longitudinal Studies

Preetika Banerjee[1], BE, MSPH; Louise Holly[2], MSc

[1]Department of Pharmacy, CHOICE Institute, School of Pharmacy, University of Washington, Health Sciences Building, 1956 NE Pacific St H362, Seattle, WA, United States

[2]Digital Transformations for Health Lab, University of Geneva, Geneva, Switzerland

Corresponding Author:
Preetika Banerjee, BE, MSPH
Department of Pharmacy, CHOICE Institute, School of Pharmacy, University of Washington, Health Sciences Building, 1956 NE Pacific St H362, Seattle, WA, United States

Abstract

Background: Everyday digital technologies such as social media, gaming, and internet use are deeply integrated into the lives of children, adolescents, and young adults. While these platforms can foster connection, learning, and entertainment, concerns have grown about their potential to influence mental, physical, and social well-being. Research on this topic has expanded rapidly over the past decade, yet much of it remains cross-sectional, limiting insights into long-term outcomes. Longitudinal studies are essential to capture evolving patterns of digital engagement, identify causal relationships, and guide effective policies and interventions that support youth in navigating digital environments. In particular, evidence is needed to distinguish between beneficial and harmful forms of digital engagement, such as social connection versus problematic use, and to understand how these impacts differ across diverse populations and contexts. The COVID-19 pandemic further accelerated young people's technology use, underscoring the urgency of examining both risks and opportunities. This review, therefore, synthesizes longitudinal research to map trends, identify knowledge gaps, and inform future directions.

Objective: The study aimed to systematically identify and map longitudinal studies examining associations between everyday digital technology use (eg, social media, gaming, and internet use) and the health and well-being of youth (25 years or younger) and to chart the types of evidence available by technology category, outcomes, and geographical setting in order to highlight key gaps for future research.

Methods: A systematic search of PubMed, Embase, and PsycArticles (2014‑2024) was conducted and reported in accordance with PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews). Data extraction covered demographics, digital technology categories, and health outcomes. Studies were grouped into 6 key themes: social media use and mental health, digital addiction and behavioral outcomes, physical activity and digital technology, digital health technologies and cognitive development, parental influence and digital technology, and digital well-being and risk behaviors.

Results: Of the 456 studies identified, 267 were longitudinal studies relevant to our research aims. Internet use (n=201 studies), social media (n=140 studies), and gaming (n=83 studies) dominated the themes. Mental health was the most frequently assessed outcome, with a focus on anxiety and depression. Geographically, 15% (40/267) of studies originated from low- and middle-income countries, with the majority from high-income settings such as the United States (n=76 studies) and Australia (n=15 studies). Nearly half (131/267, 49%) were published post 2020, reflecting heightened interest during the COVID-19 pandemic.

Conclusions: Longitudinal evidence on everyday digital technology use and youth health is growing but remains concentrated in mental health outcomes and high-income settings, with notable gaps in physical health, educational outcomes, and equity-focused research. These findings highlight the need for more diverse, methodologically robust longitudinal studies to inform context-sensitive policies and interventions that balance the risks and benefits of digital engagement for young people.

(JMIR Public Health Surveill 2026;12:e85094) doi:10.2196/85094

KEYWORDS

digital technologies; digital health interventions; youth health; adolescent health; longitudinal studies; scoping review

XSL•FO
RenderX

Introduction

Digital Technology Use and Youth Health

The rapid expansion of digital technologies has profoundly shaped young people's lives, introducing both opportunities and risks for their health and well-being [1]. Over the past decade, everyday digital technologies such as smartphones, social media platforms, and online games have become nearly ubiquitous in the lives of children and adolescents, with many young people spending several hours per day online across multiple devices [2]. Recent national and international reports indicate that a large proportion of adolescents exceed recommended daily screen-time limits and engage with multiple platforms concurrently, raising concerns about potential impacts on sleep, physical activity, and mental health [3]. At the same time, digital environments also provide opportunities for social connection, health information seeking, and access to mental health resources, underscoring the need to understand both the potential benefits and harms of digital engagement for youth well-being [4]

A growing body of empirical research has examined associations between social media use and adolescent mental health, with mixed but increasingly nuanced findings [5]. Several recent studies suggest that intensive or heavy use of social media is associated with higher levels of depressive symptoms, anxiety, and poor sleep, particularly among girls and young people with pre-existing vulnerabilities, although effect sizes are generally small to moderate [6,7]. Other work highlights that online social support, positive interactions, and access to digital mental health resources can mitigate some risks and, in some cases, promote improved well-being, illustrating the complex and bidirectional nature of these relationships [8,9].

Beyond social media, concerns about problematic or addictive patterns of digital technology use have intensified, with recent systematic reviews linking digital addiction, including problematic internet use, smartphone overuse, and gaming disorder, to poorer mental health, academic difficulties, and impaired social functioning in young people [10]. Emerging evidence also associates high levels of screen time and sedentary digital behaviors with higher risk of obesity, reduced physical activity, and disrupted sleep, although findings vary by age, sex, and type of digital activity [11]. At the same time, digital health technologies and serious games have shown promise for promoting physical activity, cognitive skills, and emotional regulation, suggesting that the health impacts of digital technologies depend on how, why, and in what contexts they are used [11].

Evidence Gaps and the Need for Longitudinal Research

Despite rapid growth in the literature, the current body of evidence is dominated by studies that are cross-sectional, focus on single platforms or outcomes, or are conducted in high-income countries, limiting understanding of causal pathways, developmental trajectories, and contextual factors such as socioeconomic position and digital literacy. Recent protocols and initiatives have called for more comprehensive approaches that examine diverse forms of digital technology use, multiple health domains, and inequities in access and risks, highlighting the need for syntheses that map this emerging evidence base [11]. Several reviews and expert statements emphasize that strong longitudinal evidence on the long-term health consequences of digital engagement in young people remains limited, particularly in low- and middle-income settings. In 2021, a joint commission of *The Lancet* and *Financial Times* called for more longitudinal studies to be conducted with young people, particularly in low- and middle-income countries (LMICs) to understand the impact of different digital determinants of young people's health [12]. Subsequent expert work has echoed these concerns, noting that current studies are often short term, rely heavily on self-reported screen use, and provide only fragmented evidence on mechanisms. Consequently, new longitudinal research agendas have been recommended [11,13]. Researchers, policymakers, and educators continue to grapple with understanding the effects of digital technology use on youth health, development, and education [13]. Longitudinal studies like Young Lives provide critical insights into how digital access, skills, and youth outcomes evolve over time, particularly in LMICs. Their findings underscore the importance of digital literacy and equitable access, highlighting how digital inclusion—or lack thereof—affects health and educational disparities [14]. Similarly, the Global Kids Online initiative explores how children's digital engagement influences their social, emotional, and cognitive development, emphasizing how unequal access to technology can deepen existing inequalities [15].

The Health Behavior in School-Aged Children (HBSC) study has also been instrumental in tracking the impact of digital technologies on youth well-being. By examining digital device usage and its correlation with mental and physical health, HBSC data reveal growing concerns over social media's influence on mental health, particularly in the wake of COVID-19, when digital communication became a crucial social lifeline [16]. These studies collectively illustrate the multifaceted relationship between digital engagement and youth outcomes, reinforcing the need for continued research to address the evolving challenges and opportunities presented by digital technologies.

While long-term research projects such as Young Lives, Global Kids Online, and HBSC offer important insights into how young people's use of digital technologies is evolving, significant evidence gaps exist regarding the long-term effects of digital technology use on different aspects of young people's health and well-being.

Objectives of This Scoping Review

The purpose of scoping reviews is to synthesize evidence and assess the scope of existing literature on a topic. The objective of this scoping review is to systematically identify and map the breadth of longitudinal evidence on the relationships between everyday digital technology use and the health and well-being of young people under the age of 25 years. For the purposes of this review, we define longitudinal studies as empirical studies that collected data from the same participants at 2 or more time points separated by at least 3 months, allowing examination of changes in digital technology exposure and health outcomes over time. Specifically, the review seeks to inform future research directions by (1) charting the characteristics of existing

XSL•FO
RenderX

longitudinal studies (including populations, types of digital technologies, outcomes, and settings), (2) summarizing how these studies conceptualize and measure digital engagement and health outcomes, and (3) identifying gaps and priorities for future longitudinal research.

Conceptually, this review is informed by emerging work on digital determinants of health, which examines how the various characteristics, features, uses, and governance of digital technologies and digital ecosystems influence health-related behaviors, health outcomes, and health care delivery [17]. For example, platform design, accessibility, data practices, and digital literacy have been identified as interacting with traditional social determinants to produce unequal risks and benefits of digital technologies across youth populations [18,19]. In addition, we draw on Bronfenbrenner ecological systems theory, which views young people's development as shaped by interacting influences at multiple levels, from family and peers to schools, communities, and broader sociotechnical and policy environments [20]. Everyday digital engagement can be understood as a proximal process embedded within these nested systems, with bidirectional influences whereby youth both shape and are shaped by their use of digital technologies.

Methods

Search Strategy

This scoping review was conducted and reported in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) [21]. A comprehensive literature search was conducted in PubMed, Embase, and PsycArticles to identify relevant studies published between 2014 and 2024. Search strategies were developed using the PICOTS (Patient Population, Intervention, Comparator, Outcomes, Timing, Setting) framework (Table 1). These 3 databases were chosen because they collectively cover a wide range of biomedical, psychological, and public health research on digital technologies and youth health and were fully accessible within our institutional subscriptions. Keywords related to the specified population and technologies were used to refine searches across databases. Boolean operators were employed to combine terms such as "youth," "adolescent," "digital technology," and "longitudinal study" to target relevant research. Database-specific search strings for PubMed, Embase, and PsycArticles, including all search terms, field tags, and limits, are provided in Multimedia Appendix 1.

Table . PICOTS (patient population, intervention, comparator, outcomes, timing, setting) framework used to structure the search strategy for this scoping review of longitudinal studies examining associations between everyday digital technology use and health and well-being outcomes among young people aged 25 years or younger (2014 - 2024).

Question	Keywords
P (patient or population)	Youth, adolescent, child, childhood
I (intervention)	Digital technology, social media, video games, internet use
C (comparator)	Not applicable for this study
O (outcomes)	No keywords applied to enable broad inclusion of different health and well-being outcomes
T (timing)	2014 - 2024
S (study type)	Longitudinal

The reporting of the search strategy follows key elements of the PRISMA-ScR checklist (Checklist 1). Search results from all databases were imported into Rayyan for reference management; Rayyan's deduplication function was used to identify duplicate records, which were then reviewed and confirmed by the first author before screening [22].

Given the exploratory nature of this scoping review, formal preregistration was not performed; however, a documented protocol was established and shared to enhance rigor and transparency. Quality appraisal was not conducted as the objective of the review was to map the literature and not assess quality of evidence. For similar reasons, no attempt was made to estimate pooled effect sizes or formally assess risk of bias.

Search strategies were developed using the PICOTS framework as a guide and tailored to each database, combining controlled vocabulary (eg, MeSH, Emtree) and free-text terms for youth (child, adolescent, young adult), digital technologies (digital technology, internet use, social media, video games, digital health, mobile apps, telemedicine, wearable electronic devices), and longitudinal study designs [23, 24]. Detailed search strings for each database and aim are provided in Multimedia Appendix 1 to enable replication.

Selection of Sources of Evidence

A total of 456 studies were initially identified in the screening process. After applying the inclusion and exclusion criteria detailed in Table 2, 267 studies were deemed relevant to the research question, focusing on longitudinal studies that examine the impact of digital technologies on youth well-being. The remaining 189 studies were excluded due to irrelevance, as they did not meet the criteria for exploring the relationship between everyday digital technology use and youth health outcomes in a longitudinal framework. This rigorous screening process ensured that only studies directly aligned with the research objectives were included for synthesis and analysis.

XSL•FO
RenderX

Table . Inclusion and exclusion criteria applied during screening of longitudinal empirical studies examining associations between everyday digital technology use and health or well-being outcomes among children, adolescents, and young adults (≤25 y) published between 2014 and 2024.

Category	Inclusion criteria	Exclusion criteria
Article type	Longitudinal empirical studies (cohort, panel, or repeated-measures designs) examining associations between everyday digital technology use and health or well-being outcomes in youth	Cross-sectional studies, case reports, qualitative-only studies, reviews, commentaries, editorials, conference abstracts without full data, and nonempirical papers
Population	Children, adolescents, and young adults aged ≤25 years (or samples with a mean age ≤25, SD 5 y) from any country or setting	Studies focusing exclusively on adults >25 years or mixed-age samples where data for participants ≤25 years cannot be disaggregated
Digital technology exposure	Everyday digital technologies such as internet use, social media, online gaming, smartphones, and other general digital media exposures	Studies focusing solely on highly specialized clinical technologies (eg, neurosurgical devices) or nondigital interventions; technology not related to personal digital media use
Outcomes	Health and well-being outcomes, including mental, physical, behavioral, social, or cognitive outcomes assessed at follow-up	Studies without health or well-being outcomes (eg, usability only) or outcomes unrelated to youth health (eg, purely technical performance metrics)
Study design and timeframe	Longitudinal design with at least 2 time points; published between 2014 and 2024 in peer-reviewed journals	Single time point (cross-sectional) designs; studies outside the 2014 - 2024 window or not peer reviewed
Language	Articles published in English	Articles published in languages other than English

Titles and abstracts were screened by the first author using predefined inclusion and exclusion criteria developed together with the second author. Any uncertainties were discussed with the second author before making final eligibility decisions. Full texts of potentially eligible articles were then assessed against the same criteria.

Data Extraction

Data from all included studies were extracted by the first author, with methodological guidance and spot checks provided by the second author to ensure accuracy and consistency. The form was piloted on a sample of studies and refined prior to full data charting. Data charting was conducted using a standardized Microsoft Excel form developed specifically for this review. The first author initially charted data from a small sample of eligible studies to pilot the form, after which the structure and definitions of each field were refined in consultation with the second author to improve clarity and consistency. The first author then charted data from all included studies, and the second author performed targeted checks of randomly selected entries to verify accuracy and resolve any uncertainties.

For each study, we extracted key data to organize and analyze the findings effectively. The following variables were recorded:

1. *Title:* The full title of the study for accurate identification
2. *Abstract:* A summary outlining the study's aims, methods, and results
3. *Journal:* The journal in which the study was published
4. *Volume, issue, pages:* Journal volume, issue number, and page range for citation purposes
5. *Authors:* The research team behind the study
6. *DOI:* The Digital Object Identifier (DOI) for easy access to study
7. *Theme:* The main research focus (eg, social media, video games, mobile applications)
8. *Subtheme:* A more specific focus within the theme (eg, for social media—"Impact on mental health")
9. *Age group:* The age category of the participants (eg, child, adolescent, young adult)
10. *Study type:* The design of the study (eg, longitudinal, cross-sectional)
11. *Outcome type:* The health or well-being outcome being assessed (eg, mental health, cognitive development)
12. *Gender:* The gender breakdown of the study population (eg, male, female, mixed)
13. *Geographical focus:* The region or country where the study took place
14. *Study name:* The official name of the study, if available
15. *Leading institution:* The institution responsible for the research

Themes Identified

Data charted from individual studies were synthesized using an inductive-deductive thematic approach appropriate for scoping reviews. Study aims and prior conceptual work informed an initial set of themes, which was then refined as new themes and insights emerged from the data. After familiarization with the charted data, the first author grouped studies according to similarities in digital technology type, health or well-being outcomes, and study aims. These preliminary groupings and definitions were then reviewed with the second author, who suggested refinements until consensus was reached on 6 overarching themes. Each included study was assigned to 1 primary theme based on its main focus; when a study spanned multiple topics, the dominant exposure-outcome relationship was used to guide classification. Within each theme, findings were synthesized descriptively, focusing on patterns in study designs, measures of digital engagement and outcomes, direction and consistency of reported associations, and notable gaps in populations or contexts. No meta-analysis or quantitative

XSL•FO
RenderX

pooling of effect estimates was performed, in line with the objectives of a scoping review.

The following 6 key themes were assigned to the studies based on their content and focus:

1. *Social media use and mental health:* Studies exploring the relationship between social media engagement and various mental health outcomes, including anxiety, depression, and social well-being
2. *Digital addiction and behavioral outcomes*: Research investigating the addictive potential of digital technologies and their effects on behavior, such as compulsive use of social media or video games
3. *Physical activity and digital technology:* Studies examining the interaction between physical activity levels and digital technology use, including how digital technologies may influence exercise behaviors
4. *Digital health technologies and cognitive development:* Research focused on the use of digital tools, such as mobile apps and telemedicine, to enhance cognitive development and support educational outcomes in young people
5. *Parental influence and digital technology:* Studies exploring the role of parents in shaping their children's use of digital technologies, including the impact of parental guidance on digital behaviors
6. *Digital well-being and risk behaviors:* Research addressing how digital technology use impacts risk behaviors such as substance abuse, risky sexual behavior, and online safety concerns

To facilitate reproducibility, the full search strategies, screening decision rules, and data-extraction template are available in Multimedia Appendix 1 and from the corresponding author upon reasonable request. The completed PRISMA-ScR checklist is provided in Checklist 1.

Results

Study Selection and Overview of Included Studies

The scoping review encompassed a total of 267 studies, which provided insight into various aspects of digital technology use and its impact on the health and well-being of young people over time. Internet use (n=201 studies), social media (n=140 studies), and gaming (n=83 studies) dominated the themes. Mental health was the most frequently assessed outcome, with a focus on anxiety and depression. Figure 1 presents the PRISMA-ScR flow of records through identification, screening, eligibility assessment, and inclusion. Across databases, 913 records were retrieved (PubMed n=518, Embase n=370, and PsycArticles n=25), of which 457 duplicates were removed before screening, leaving 456 unique records for title and abstract review. Of these, 190 records were excluded (wrong article type n=50, wrong population n=47, wrong outcome n=63, outside of date range n=30), with 266 studies remaining. One additional report was identified and retreived outside the initial screening process, resulting in 267 studies being included in the review.

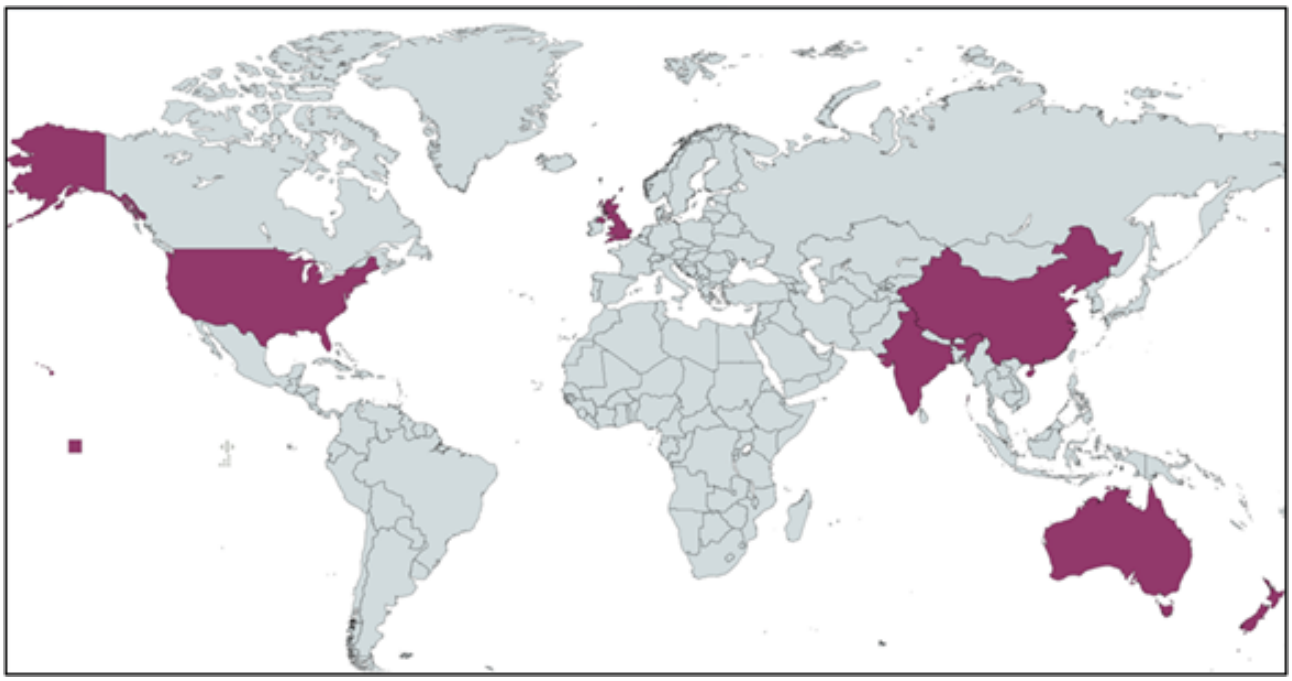
XSL•FO
RenderX

Figure 1. PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) flow diagram illustrating identification, screening, eligibility assessment, and inclusion of longitudinal studies examining everyday digital technology use and health or well-being outcomes among young people aged 25 years or younger (2014 - 2024).

XSL•FO
RenderX

Geographically, 15% (40/267) of studies originated from LMICs, with the majority coming from high-income settings such as the United States (n=76 studies) and Australia (n=15 studies; Figure 2). Nearly half (131/267, 49%) were published post 2020, reflecting heightened interest during the COVID-19 pandemic.

Figure 2. Geographic distribution of the 267 longitudinal studies included in the scoping review examining digital technology use and youth health outcomes, highlighting the predominance of studies conducted in high-income countries.



The analysis of these longitudinal studies revealed 6 key themes: social media use and mental health; digital addiction and behavioral outcomes; physical activity and digital technology; digital health technologies and cognitive development; parental influence and digital technology; and digital well-being and risk factors. Each theme highlights important findings related to the impact of digital technologies on health outcomes for youth. Table 3 summarizes the number of included studies by thematic category, country income group, and predominant age group, illustrating the concentration of research in high-income settings and among adolescents and young adults.

Table . Distribution of 267 longitudinal studies included in the scoping review examining digital technology use and youth health outcomes, categorized by thematic focus, country income group, and predominant participant age group.

Theme	Studies (N=267), n	Main country income groups[a]	Predominant age groups[b]
Social media use and mental health	69	Mostly high-income; some upper-middle-income	Adolescents, young adults
Digital addiction and behavioral outcomes	102	High-income and upper-middle-income	Adolescents, young adults
Physical activity and digital technology	18	Mostly high-income	Children, adolescents
Digital health technologies and cognitive development	33	Mostly high-income	Children, adolescents
Parental influence and digital technology	11	Mostly high-income	Children, adolescents
Digital well-being and risk behaviors	34	High-income and upper-middle-income	Adolescents, young adults

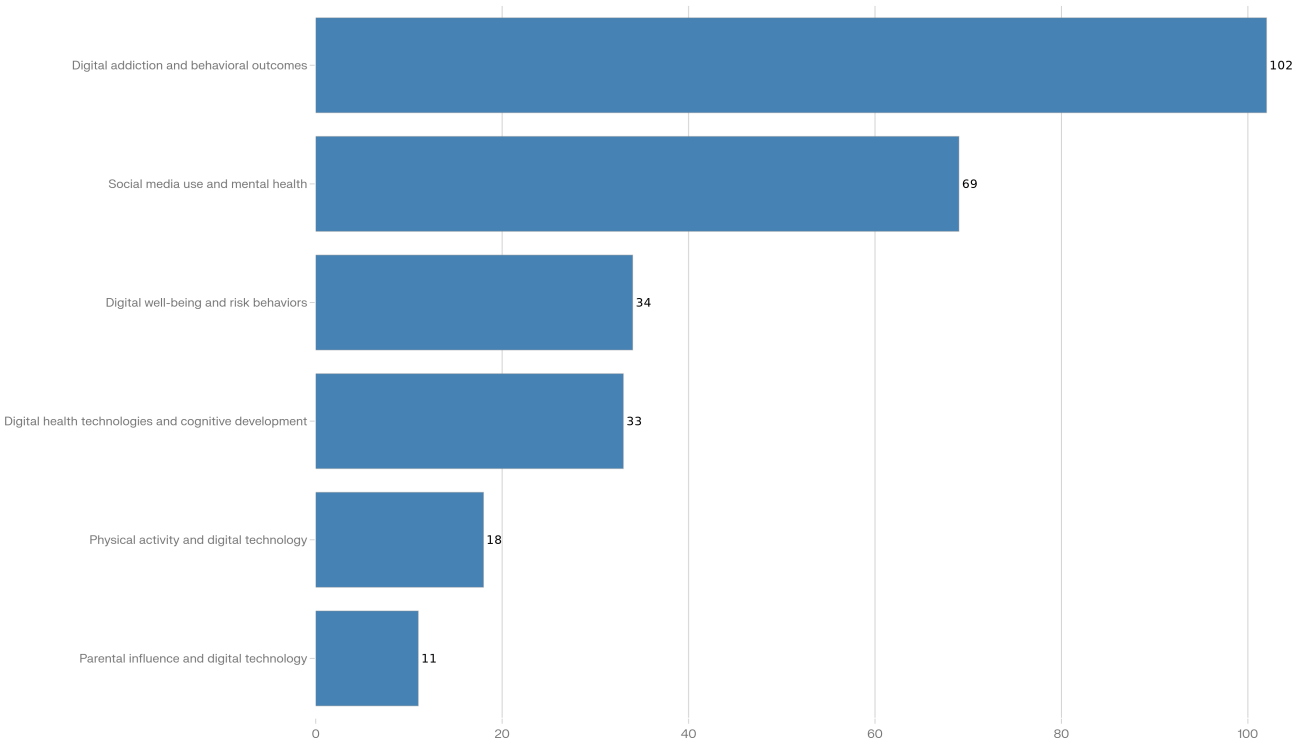
[a]Income groups classified according to the World Bank categories reported for study countries.

[b]Age groups as charted during data extraction: children (<12 y), adolescents (12 - 17 y), and young adults (18 - 25 y).

The distribution of longitudinal studies across the 6 thematic categories is presented in Figure 3. This visualization highlights the predominance of research on digital addiction and mental health outcomes compared to a relative scarcity of longitudinal studies on physical health and cognitive development impacts.

XSL•FO
RenderX

Figure 3. Number of longitudinal studies included in the scoping review by thematic category of digital technology use and youth health outcomes.



Theme 1: Social Media Use and Mental Health

A total of 69 longitudinal studies focused on the effects of social media on adolescent mental health. Of these, 45 studies examined depression and anxiety, showing a strong association between increased social media use and negative mental health outcomes, particularly depression and anxiety. Six studies explored the influence of social media on self-esteem and body image, indicating that social media can exacerbate body dissatisfaction, especially among adolescent girls [25-30]. Nine studies addressed the issue of cyberbullying, which was found to significantly affect mental health, with higher risks of depression, anxiety, and suicidal ideation [31-39]. Nine studies also highlighted positive aspects of social media use, such as online support groups and mental health resources [34,40-48]. Overall, the studies suggest that social media use has a complex, bidirectional relationship with mental health, with both positive and negative outcomes.

These studies reveal an important distinction between intensive and problematic social media use among adolescents. Intensive use, defined as high frequency or duration of engagement, was not consistently associated with adverse mental health outcomes, with several studies reporting no significant correlation between frequency of social media use and measures such as depression, anxiety, or stress [16,49]. In contrast, problematic social media use, characterized by addiction-like symptoms such as compulsive engagement, withdrawal, and loss of control, was moderately and consistently linked to increased symptoms of depression, anxiety, stress, and loneliness [16,50]. For example, meta-analytic and longitudinal studies reported that increases in problematic social media use are significantly associated with heightened depressive symptoms and loneliness over time, while intensive use alone does not show these associations [49,50]. These findings underscored the importance of differentiating between patterns of social media engagement, as problematic use appeared to pose a greater risk to adolescent mental health than high-frequency use alone.

Of the 69 studies in this theme, roughly two-thirds reported predominantly adverse associations between problematic or intensive social media use and later mental health outcomes, around one-quarter reported mixed or null findings, and a small minority emphasized mainly positive or buffering effects of online engagement. Taken together, the longitudinal evidence suggested that how adolescents use social media matters more than how much they use it, with addiction-like, emotionally dysregulated engagement aligning more closely with stress-generation and differential-susceptibility models than simple “screen-time” explanations. This theme highlighted the need for future studies to explicitly test mechanisms such as social comparison, sleep disruption, and online victimization as pathways linking problematic use to later depressive and anxiety symptoms.

Theme 2: Digital Addiction and Behavioral Outcomes

This theme addressed longitudinal studies of addictive-like behaviors fostered by everyday technology use (eg, internet, gaming, smartphone addiction) and their associations with mental health and behavioral functioning.

This review incorporated 102 studies on the impact of digital addiction. The majority of the studies (n=56 observational) demonstrated a link between excessive or compulsive digital technology use, particularly internet and gaming addiction, and behavioral issues like anxiety, depression, and poor social functioning. Notably, internet addiction (n=39 studies) and gaming addiction (n=38 studies) were strongly correlated with poor mental health and academic performance, while smartphone dependency (n=4 studies) and behavioral disorders

XSL•FO
RenderX

(n=20 studies) further exacerbated existing mental health issues [51-57]. Among 36 intervention studies, 54% (144/267) reported positive outcomes in terms of digital addiction treatment and management. However, a significant portion of studies revealed harmful effects of excessive digital use, which included disrupted sleep, increased mental health symptoms, and social isolation. Longitudinal research emphasized the long-term psychological consequences, such as increased depression and anxiety, associated with digital addiction [58-60].

Across the 102 digital addiction studies, most observational studies found that higher levels of internet or gaming addiction predicted worsening mood or functioning over time, whereas only a minority of intervention trials demonstrated sustained reductions in both addictive use and mental health symptoms. Overall, findings in this theme support conceptualizing digital addiction as part of a broader cluster of self-regulation and impulse-control difficulties, rather than as an isolated problem of device use. Several studies and reviews also highlighted how specific design features of social media platforms and games, such as variable and unpredictable rewards and always-on social notifications, were intended to maximize engagement and could reinforce compulsive checking and prolonged use, particularly among adolescents [61,62]. Longitudinal patterns linking persistent addiction-like use with worsening mood, school functioning, and social withdrawal pointed to the importance of interventions that targeted underlying regulatory skills and family or school contexts alongside usage reduction.

Theme 3: Physical Activity and Digital Technology

A total of 18 longitudinal studies were analyzed under this theme. A total of 13 studies explored the relationship between sedentary screen time and reduced physical activity. Two studies researched the effectiveness of digital tools, including fitness apps and activity trackers, in increasing physical activity levels. A further 3 studies focused on digital gamification or exergames designed to increase movement. Studies on sedentary screen time pointed to the negative effects of excessive screen time, which were associated with physically inactive lifestyles and an increased risk of obesity, particularly among girls. A bidirectional relationship was observed, with excessive screen time negatively impacting physical activity, while physical inactivity also exacerbated screen time use [63,64]. Only a small subset of intervention and gamification studies found sustained increases in physical activity linked to digital tools.

Across studies, digital technologies emerged as a double-edged tool for youth physical health: design features that scaffolded motivation and feedback could nudge activity, while unstructured screen use reinforced sedentary routines and obesogenic environments. These findings underscored the value of moving beyond total “screen time” metrics toward frameworks that differentiate health-promoting, deliberately designed movement-supporting tools from passive, displacement-oriented digital use. Future research should focus on long-term interventions to increase physical activity and reduce sedentary behavior across diverse youth populations.

Theme 4: Digital Health Technologies and Cognitive Development

A total of 33 longitudinal studies focused on the relationship between digital technology use and cognitive development. Fifteen studies evaluated deliberately designed digital health or educational tools to support cognitive outcomes (mobile health or training apps, n=10; telemedicine or telehealth for cognitive disorders, n=2; e-learning or serious game platforms, n=3), while 18 studies examined everyday digital media use (eg, overall screen time, general device or internet use) in relation to cognitive development. These studies generally indicated positive effects of digital health technologies on cognitive and behavioral outcomes, including attention improvement and executive function through exergames and mobile applications [65-76]. However, concerns regarding the long-term effectiveness of these technologies were prevalent, with many showing only short-term improvements or mixed results in sustained behavior change.

Eighteen studies explored the impact of everyday digital technology use on cognitive development. Across these studies of everyday digital technology use, higher overall screen time and background television exposure were frequently associated with weaker performance or slower gains in language, executive function, or school readiness, especially in younger children, although effect sizes were generally small. These patterns reinforced concerns about heavy, nonstructured screen use and support calls for more balanced, developmentally appropriate digital engagement

Of the 33 studies in this theme, around two-thirds documented short-term improvements in at least 1 cognitive or learning outcome following exposure to digital health or educational technologies (eg, training apps, exergames, e-learning platforms), whereas approximately one-third reported mixed, null, or unsustained effects, particularly at longer follow-up or in relation to more diffuse measures of everyday screen use.

Evidence in this theme suggested that digital health technologies are promising but not yet consistently transformative tools for youth cognitive development, with many benefits appearing short-term or context-dependent. Conceptually, this suggests that technology functions as an amplifier of existing learning environments and supports, rather than a stand-alone solution, and that sustained gains likely require integration with offline pedagogy, family engagement, and broader educational resources. At the same time, findings from observational studies in this theme underscored the importance of actively mitigating potential harms from heavy, nonstructured or background screen exposure, through limits on overall screen time, attention to content quality, and alignment of digital use with age-appropriate, offline learning and caregiving routines

Theme 5: Parental Influence and Digital Technology

A total of 11 longitudinal studies analyzed parental influence on children’s digital behaviors and online safety. The studies revealed that parental monitoring and control of screen time positively influenced youth behavior, with several studies showing improved outcomes in mental health and behavioral issues. However, some studies highlighted the challenges of

XSL•FO
RenderX

treatment adherence and the need for refinement in interventions aimed at parental involvement. Parental education about digital literacy and online safety was also shown to have a beneficial impact, although the effects were small and varied depending on the type of screen activity [76-80]. Future research should refine interventions to better support family dynamics and address treatment attrition.

Among the 11 studies examining parental influence, most found that some form of active monitoring, co-use, or guidance was linked to modest improvements in mental health, behavior, or online safety indicators, whereas purely restrictive strategies showed smaller and less consistent benefits and occasionally coincided with more conflictual digital use.

Collectively, these studies reinforce models of "co-regulation," in which parents shape not only how much time children spend online but also the meanings and risks attached to that time. Modest but consistent benefits of warm, structured guidance, compared with either lax or highly controlling approaches, highlighted the need for interventions that bolster parents' digital literacy and relational engagement rather than focusing solely on restrictive screen-time limits.

Theme 6: Digital Risk Factors

This theme focused on digital risk behaviors (eg, cyberbullying, sexting, exposure to harmful content) and how they co-occurred with and predict substance use, aggression, and other offline risks. This theme covered 34 longitudinal studies on risky digital behaviors—such as cyberbullying, sexting, and exposure to harmful content—and their relationship with other risk factors that influence young people's well-being, including substance abuse and aggressive behavior. Studies indicated that increased social media exposure and engagement in online activities such as cyberbullying, sexting, and online grooming correlated with higher risks of substance abuse and aggressive behaviors [30,81]. Furthermore, substance abuse was found to be linked to digital media use, particularly in adolescents, with frequent social media users reporting higher rates of alcohol, tobacco, and drug use compared to their peers [82-95]. The studies underlined the need for better regulation and monitoring of digital content to mitigate these risks [96]. Studies suggested that early media exposure could influence long-term risky behaviors, and intervention strategies focusing on media literacy and parental monitoring were highlighted as effective approaches [97-99].

Findings in this theme suggested that risky online behaviors often co-occurred with broader clusters of offline vulnerability, supporting syndemic and developmental-cascade perspectives in which digital and offline risks reinforced one another over time. Interventions that combine media literacy, skills for navigating online social pressures, and targeted support for substance use and aggression appeared more theoretically justified than approaches that address digital and offline risks in isolation.

Overall, the findings from these 6 themes underscored the complexity of digital technology's influence on youth health, with both positive and negative outcomes depending on usage patterns and the context of engagement. Across all 6 themes, the evidence depicted everyday digital technology use as a contextual, socially embedded determinant of youth health, with risks and benefits shaped by prior vulnerabilities, family and school environments, and the specific functions that digital tools serve in young people's lives.

Discussion

Principal Findings

The aim of this scoping review was to systematically identify and map the breadth of longitudinal evidence on the relationships between everyday digital technology use and the health and well-being of young people under the age of 25 years. The review sought to (1) chart the characteristics of existing longitudinal studies, (2) summarize how these studies conceptualize and measure digital engagement and health outcomes, and (3) identify gaps and priorities for future longitudinal research.

The review found that most longitudinal studies published between 2014 and 2024 focused on 3 aspects of technology use: internet browsing (n=201), social media (n=140), and gaming (n=83). Digital addiction and mental health, particularly anxiety and depression, were the dominant health outcomes studied. A major gap in the geographical scope of studies was identified: only 15% (40/267) of the included studies originated from LMICs.

Conceptual Implications

Taken together, the 6 dominant themes identified within the longitudinal studies support an ecological view of digital engagement and youth health, in which associations are rarely linear or uniform but instead depend on who is using which technologies, in what social and structural contexts, and for what purposes. For example, problematic, dysregulated use in the context of family conflict and offline adversity appears more strongly linked to subsequent internalizing symptoms and risky behaviors than similar levels of use in supportive environments, consistent with ecological and digital-determinants perspectives. Social media use was linked to increased loneliness and distress, especially during the COVID-19 pandemic. At the same time, positive uses, such as accessing health information or supportive communities, illustrate how digital spaces can also function as protective resources when embedded within enabling microsystems and macrosystems [18,20].

By combining thematic synthesis with simple frequency counts across technologies, outcomes, age groups, and regions, this review functions as an evidence map that highlights both densely studied areas (eg, social media and internalizing symptoms) and important gaps (eg, physical health, cognitive development outcomes, and research in low- and middle-income settings).

Cross-Cutting Patterns

Across themes, several cross-cutting patterns emerge from the longitudinal evidence. First, associations between digital engagement and youth health are generally small in magnitude, often contingent on baseline vulnerabilities, offline context, and the quality rather than the quantity of engagement, which aligns with differential susceptibility and person-by-context interaction

XSL•FO
RenderX

models. Second, studies that distinguish between specific modes of use (eg, passive vs active social media use, recreational vs problematic gaming) more consistently report directional relationships than those that rely on global screen-time measures, suggesting that conceptual models focused on "time displacement" alone are insufficient. Third, very few longitudinal studies explicitly test mediating pathways (eg, sleep disruption, social comparison, cyberbullying), limiting the ability to draw firm causal conclusions even when temporal ordering is clear. Digital addiction and digital risk behaviors represent related but distinct pathways through which online engagement influences youth health; these themes are therefore discussed separately but cross-referenced where studies span both domains.

Implications for Theory and Methods

Taken together, the mapped evidence supports a shift away from simple "screen-time" narratives toward frameworks that emphasize conditional and pathway-specific effects of digital engagement on youth health, for example, models in which digital experiences amplify existing mental health risks for some young people while offering compensatory social support or health information for others. Longitudinal designs that incorporate repeated measures of context, mediators, and offline supports are needed to move from description toward testing such conceptual models.

Implications for Public Health and Policy

From a public health perspective, these findings support a shift from individual responsibility narratives toward policies that address the digital environments in which young people spend time [100]. First, educators and health systems can use this evidence to develop age-appropriate digital literacy and mental health literacy programs that help young people recognize harmful content, manage problematic use, and leverage supportive online spaces, particularly for those already facing offline vulnerabilities [101]. Second, the concentration of harms around specific platform features, such as algorithmic amplification of extreme content and design elements that encourage compulsive engagement, underscores the need for regulation that embeds "safety-by-design" principles, transparency, and accountability for industry actors [102]. Third, inequities in who benefits from digital health technologies point to the importance of treating digital connectivity, skills, and protections as digital determinants of health and of integrating these into broader strategies to reduce mental health and well-being disparities among young people [103].

Our analysis reveals a marked increase in publications addressing digital technology and youth health since 2020. Most studies focus on mental health, particularly the effects of social media use, while fewer address physical health, educational outcomes, or equity considerations. Notably, there is a paucity of research from LMICs and limited longitudinal evidence on the long-term impacts of digital technology use. Given that 15% (40/267) of the included studies originated from LMICs, our conclusions about global patterns of digital engagement and health should be interpreted cautiously and viewed as more representative of high-income settings. Expanding research in these settings is crucial to understanding disparities in access and digital literacy. These gaps underscore the need for future research in underrepresented areas and populations. Additionally, the diversity of study designs reflects the evolving nature of this field, with a trend toward more sophisticated methodologies in recent years.

Strengths and Limitations

This review has several strengths, including the systematic mapping of longitudinal evidence across multiple domains of digital technology use and youth health outcomes. This scoping review synthesizes available longitudinal evidence on everyday digital technology use and youth well-being, highlighting predominant research foci and key gaps rather than providing a definitive assessment of effects. The findings point to priorities for future longitudinal research and for the careful design and evaluation of digital policies and interventions, particularly in underrepresented regions and populations.

This scoping review has several limitations that should be considered when interpreting the findings. First, the search was restricted to 3 databases (PubMed, Embase, and PsycArticles), to articles published in English, and to studies from 2014 to 2024, which means that relevant longitudinal research indexed elsewhere, published in other languages, or outside this time window may have been missed. The search was limited to PubMed, Embase, and PsycArticles; other multidisciplinary databases such as Scopus or Web of Science were not searched, which means that some relevant longitudinal studies indexed elsewhere may have been omitted.

Second, consistent with the aims and methodology of a scoping review, we did not conduct a formal meta-analysis, and the synthesis is descriptive rather than quantitative; therefore, the review cannot determine the direction or magnitude of causal effects. Third, most included studies were conducted in high-income countries, so the patterns summarized here may not be generalizable to low- and middle-income settings or to underrepresented youth populations.

Although a formal quality assessment or risk-of-bias appraisal was not undertaken in line with scoping review guidance, several recurrent methodological limitations should be noted at the level of the primary studies. Many relied on self-reported digital use and health outcomes, which are vulnerable to recall and social desirability biases, and follow-up periods were often relatively short, limiting inferences about long-term trajectories. Residual confounding is also likely, as few studies used repeated measures, time-varying covariates, or strong causal designs. Additionally, most samples were drawn from high-income countries and school- or clinic-based populations, which constrain generalizability. Taken together, these patterns suggest that both harmful and beneficial associations between digital engagement and youth health may be misestimated in the existing longitudinal evidence base and point to the need for more rigorous and diverse designs in future work.

Directions for Future Research

The findings from this study reinforce the call from the *Lancet* and *Financial Times* Commission for more longitudinal research on the digital determinants of young people's health and well-being. Future longitudinal research should examine the

XSL•FO
RenderX

long-term effects of everyday technology use and other forms of digital engagement on different dimensions of youth physical health, mental health, and well-being, including emerging technologies like virtual reality and artificial intelligence. Important objectives of these studies should be to increase our collective understanding of which features of digital technologies and platforms are most likely to promote or undermine health outcomes and why some young people are more vulnerable to the negative impacts of digital technology use than others. Greater effort is needed to generate evidence from LMICs to understand how sociocultural and economic contexts shape the impacts of technology use. Additionally, studies should examine the role of digital literacy in mitigating risks and tailoring interventions to diverse populations. Exploring digital interventions—such as literacy programs, online support communities, and mental health resources—will be essential in maximizing benefits while minimizing risks, ultimately fostering youth resilience and well-being.

Conclusions

This scoping review highlights the rapid expansion of longitudinal research examining digital technology use and youth health, while also revealing important conceptual and geographical gaps in the evidence base. Most existing studies focus on mental health outcomes in high-income settings, with far fewer investigating physical health, cognitive development, or youth populations in LMICs. Future research should prioritize more diverse populations, stronger longitudinal designs, and clearer conceptual models to better understand how different forms of digital engagement shape young people's health and well-being over time. Strengthening this evidence base will be essential for informing policies and interventions that maximize the benefits of digital technologies while mitigating potential harms.

Acknowledgments

The authors extend their sincere thanks to Aferdita Bytyqi for providing the opportunity to pursue this project and for her continued support throughout its development. Appreciation is further extended to the Digital Transformations for Health Lab for their support, which made the completion of this work possible. The authors declare the use of generative artificial intelligence (AI) tools during manuscript preparation. According to the GAIDeT (Generative AI Delegation Taxonomy; 2025), generative AI was used under full human supervision for proofreading and editing, summarizing text, and reformatting content, but not for data generation, analysis, or interpretation of results. The tools used were ChatGPT (GPT-4.5) and Perplexity. Responsibility for the integrity and accuracy of all content rests entirely with the human authors. Generative AI tools are not listed as authors and do not bear responsibility for the final manuscript.

Funding

This work was supported by the Digital Transformations for Health Lab as part of its research fellowship program. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. All work was undertaken as part of the authors' institutional roles.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request. The datasets generated and analyzed during this scoping review (including study-level extraction sheets and the coding framework for thematic classification) are available from the corresponding author upon reasonable request. All data are derived from previously published studies and do not include any individual-level participant information.

Authors' Contributions

PB served as the lead on data collection, analysis, and drafting of the manuscript. LH provided supervision, guidance on conceptualization, study design and methodology, and critical revisions of the manuscript. Both authors reviewed and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategies.
[DOCX File, 17 KB - publichealth_v12i1e85094_app1.docx]

Checklist 1
PRISMA checklist.
[PDF File, 185 KB - publichealth_v12i1e85094_app2.pdf]

XSL•FO
RenderX

References

1. Holly L, Wong BLH, Agrawal A, et al. Opportunities and threats for adolescent wellbeing provided by digital transformations. : Partnership for Maternal, Newborn & Child Health (PMNCH); 2021 URL: https://www.governinghealthfutures2030.org/wp-content/uploads/2021/10/Adolescent-Wellbeing-Background-Paper-digital.pdf?utm_source=chatgpt.com [accessed 2026-03-31]
2. Masri-Zada T, Martirosyan S, Abdou A, et al. The impact of social media & technology on child and adolescent mental health. J Psychiatry Psychiatr Disord 2025;9(2):111-130. [doi: 10.26502/jppd.2572-519X0242] [Medline: 40520349]
3. Zablotsky B, Ng AE, Black LI, et al. Associations between screen time use and health outcomes among US teenagers. Prev Chronic Dis 2025 Jul 10;22:E38. [doi: 10.5888/pcd22.240537] [Medline: 40638804]
4. Chen T, Ou J, Li G, Luo H. Promoting mental health in children and adolescents through digital technology: a systematic review and meta-analysis. Front Psychol 2024;15:1356554. [doi: 10.3389/fpsyg.2024.1356554] [Medline: 38533221]
5. Plackett R, Sheringham J, Dykxhoorn J. The longitudinal impact of social media use on adolescent mental health in the UK. Eur J Public Health 2022 Oct 21;32(Supplement_3). [doi: 10.1093/eurpub/ckac129.054]
6. Cabezas-Klinger H, Fernandez-Daza FF, Mina-Paz Y. Associations between social media use and mental disorders in adolescents and young adults: a systematic review and meta-analysis of recent evidence. Behav Sci (Basel) 2025 Oct 24;15(11):1450. [doi: 10.3390/bs15111450] [Medline: 41301253]
7. McAlister KL, Beatty CC, Smith-Caswell JE, Yourell JL, Huberty JL. Social media use in adolescents: bans, benefits, and emotion regulation behaviors. JMIR Ment Health 2024 Nov 4;11:e64626. [doi: 10.2196/64626] [Medline: 39496203]
8. Shiferaw BD, Tang J, Wang Y, et al. Impact of digital addiction on youth health: a systematic review and meta-analysis. J Behav Addict 2025 Sep 30;14(3):1129-1158. [doi: 10.1556/2006.2025.00081] [Medline: 40928886]
9. Khalaf AM, Alubied AA, Khalaf AM, Rifaey AA. The impact of social media on the mental health of adolescents and young adults: a systematic review. Cureus 2023 Aug;15(8):e42990. [doi: 10.7759/cureus.42990] [Medline: 37671234]
10. Cohen K, Rapoport A, Friis E, Hill S, Feldman S, Schleider J. The alongside digital wellness program for youth: longitudinal pre-post outcomes study. JMIR Form Res 2025 Oct 8;9:e73180. [doi: 10.2196/73180] [Medline: 41061253]
11. Cascio CN, Selkie E, Moreno MA. Effect of technology and digital media use on adolescent health and development: protocol for a multimethod longitudinal study. JMIR Res Protoc 2023 Sep 13;12:e50984. [doi: 10.2196/50984] [Medline: 37703071]
12. Kickbusch I, Piselli D, Agrawal A, et al. The Lancet and Financial Times Commission on governing health futures 2030: growing up in a digital world. Lancet 2021 Nov 6;398(10312):1727-1776. [doi: 10.1016/S0140-6736(21)01824-9] [Medline: 34706260]
13. Dienlin T, Johannes N. The impact of digital technology use on adolescent well-being. Dialogues Clin Neurosci 2020 Jun;22(2):135-142. [doi: 10.31887/DCNS.2020.22.2/tdienlin] [Medline: 32699513]
14. Ford K, Freund R. Young lives under pressure: protecting and promoting young people's mental health at a time of global crises. : Young Lives; 2022 URL: https://www.younglives.org.uk/sites/default/files/2025-06/YL-PolicyBrief-55-Jun25.pdf [accessed 2026-03-31]
15. Global Kids Online. URL: http://globalkidsonline.net/ [accessed 2025-05-19]
16. Boniel-Nissim M, Marino C, Galeotti T, et al. A focus on adolescent social media use and gaming in Europe, Central Asia and Canada: health behaviour in school-aged children international report from the 2021/2022 survey. : World Health Organization Regional Office for Europe; 2024 URL: https://iris.who.int/server/api/core/bitstreams/f787fae6-e564-477c-a0e6-d54f24b38f0b/content [accessed 2026-03-31]
17. Holly L, Naing SY, Pitt H, Thomas S, Kickbusch I. Health promotion and the digital determinants of health. Health Promot Int 2025 Mar 5;40(2):daaf033. [doi: 10.1093/heapro/daaf033] [Medline: 40139247]
18. Chidambaram S, Jain B, Jain U, et al. An introduction to digital determinants of health. PLOS Digit Health 2024 Jan;3(1):e0000346. [doi: 10.1371/journal.pdig.0000346] [Medline: 38175828]
19. Mooney P, Lam LTV, Todd AR, Partridge SR, Raeside R. The inclusion and impact of digital determinants of health in digital nutrition interventions for adolescents: a systematic review. Health Promot Int 2025 Sep 3;40(5):daaf154. [doi: 10.1093/heapro/daaf154] [Medline: 41020319]
20. McHale SM, Dotterer A, Kim JY. An ecological perspective on the media and youth development. Am Behav Sci 2009 Apr;52(8):1186-1203. [doi: 10.1177/0002764209331541] [Medline: 22247564]
21. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
22. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 5;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
23. Zrubka Z, Champion A, Holtorf AP, et al. The PICOTS-ComTeC framework for defining digital health interventions: an ISPOR special interest group report. Value Health 2024 Apr;27(4):383-396. [doi: 10.1016/j.jval.2024.01.009] [Medline: 38569772]
24. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372(71):n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]

XSL•FO
RenderX

25. Booker CL, Kelly YJ, Sacker A. Gender differences in the associations between age trends of social media interaction and well-being among 10-15 year olds in the UK. BMC Public Health 2018 Mar 20;18(1):321. [doi: 10.1186/s12889-018-5220-4] [Medline: 29554883]
26. Straatmann VS, Almquist YB, Oliveira AJ, Rostila M, Lopes CS. Cross-lagged structural equation models for the relationship between health-related state and behaviours and body bullying in adolescence: findings from longitudinal study ELANA. PLOS ONE 2018;13(1):e0191253. [doi: 10.1371/journal.pone.0191253] [Medline: 29342218]
27. Zimmer-Gembeck MJ, Rudolph J, Webb HJ, Henderson L, Hawes T. Face-to-face and cyber-victimization: a longitudinal study of offline appearance anxiety and online appearance preoccupation. J Youth Adolesc 2021 Dec;50(12):2311-2323. [doi: 10.1007/s10964-020-01367-y] [Medline: 33449288]
28. Vall-Roqué H, Andrés A, González-Pacheco H, Saldaña C. Women's body dissatisfaction, physical appearance comparisons, and Instagram use throughout the COVID-19 pandemic: a longitudinal study. Int J Eat Disord 2023 Jan;56(1):118-131. [doi: 10.1002/eat.23827] [Medline: 36268646]
29. Livet A, Boers E, Laroque F, Afzali MH, McVey G, Conrod PJ. Pathways from adolescent screen time to eating related symptoms: a multilevel longitudinal mediation analysis through self-esteem. Psychol Health 2024 Sep;39(9):1167-1182. [doi: 10.1080/08870446.2022.2141239] [Medline: 36345595]
30. Doornwaard SM, Bickham DS, Rich M, Vanwesenbeeck I, van den Eijnden R, ter Bogt TFM. Sex-related online behaviors and adolescents' body and sexual self-perceptions. Pediatrics 2014 Dec;134(6):1103-1110. [doi: 10.1542/peds.2014-0592] [Medline: 25404728]
31. Wang C, Gao T, Cheng X, Li B. Social media use, cyber victimization, and adjustment during COVID-19 virtual learning: a short-term longitudinal study among Chinese middle school students. Sch Psychol 2024 Mar;39(2):176-186. [doi: 10.1037/spq0000582] [Medline: 37883009]
32. Viner RM, Gireesh A, Stiglic N, et al. Roles of cyberbullying, sleep, and physical activity in mediating the effects of social media use on mental health and wellbeing among young people in England: a secondary analysis of longitudinal data. Lancet Child Adolesc Health 2019 Oct;3(10):685-696. [doi: 10.1016/S2352-4642(19)30186-5] [Medline: 31420213]
33. Pittet I, Berchtold A, Akré C, Michaud PA, Surís JC. Are adolescents with chronic conditions particularly at risk for bullying? Arch Dis Child 2010 Sep;95(9):711-716. [doi: 10.1136/adc.2008.146571] [Medline: 19307194]
34. Kerr B, Wilburn K, Lam E, Moreno M. Associations between problem alcohol use and active and passive social media posts. J Adolesc Health 2018 Feb;62(2):S135. [doi: 10.1016/j.jadohealth.2017.11.275]
35. Cole DA, Zelkowitz RL, Nick E, et al. Longitudinal and incremental relation of cybervictimization to negative self-cognitions and depressive symptoms in young adolescents. J Abnorm Child Psychol 2016 Oct;44(7):1321-1332. [doi: 10.1007/s10802-015-0123-7] [Medline: 26747449]
36. Meter DJ, Bauman S. When sharing is a bad idea: the effects of online social network engagement and sharing passwords with friends on cyberbullying involvement. Cyberpsychol Behav Soc Netw 2015 Aug;18(8):437-442. [doi: 10.1089/cyber.2015.0081] [Medline: 26252928]
37. Rose CA, Tynes BM. Longitudinal associations between cybervictimization and mental health among U.S. adolescents. J Adolesc Health 2015 Sep;57(3):305-312. [doi: 10.1016/j.jadohealth.2015.05.002] [Medline: 26115909]
38. Cross D, Lester L, Barnes A. A longitudinal study of the social and emotional predictors and consequences of cyber and traditional bullying victimisation. Int J Public Health 2015 Feb;60(2):207-217. [doi: 10.1007/s00038-015-0655-1] [Medline: 25645100]
39. Gradinger P, Yanagida T, Strohmeier D, Spiel C. Effectiveness and sustainability of the ViSC Social Competence Program to prevent cyberbullying and cyber-victimization: class and individual level moderators. Aggress Behav 2016;42(2):181-193. [doi: 10.1002/ab.21631] [Medline: 26879896]
40. Meisel SN, Nesi J, Janssen T, Jackson KM. Adolescent (mis)perceptions of peer alcohol posts on social media: prospective associations with alcohol attitudes and use. Alcohol Clin Exp Res 2022 Nov;46(11):2054-2067. [doi: 10.1111/acer.14935] [Medline: 36378079]
41. Fitzpatrick C, Boers E. Developmental associations between media use and adolescent prosocial behavior. Health Educ Behav 2022 Apr;49(2):265-271. [doi: 10.1177/10901981211035702] [Medline: 34605695]
42. Lim MSC, Cooper S, Lewis L, et al. Prospective mixed methods study of online and offline social networks and the development of sexual agency in adolescence: the Social Networks and Agency Project (SNAP) protocol. BMJ Open 2019 May 19;9(5):e024329. [doi: 10.1136/bmjopen-2018-024329] [Medline: 31110083]
43. Moreno MA, Kacvinsky L, Pumper M, Wachowski L, Whitehill JM. Associations between social media displays and event-specific alcohol consumption by college students. WMJ 2013 Dec;112(6):251-256. [Medline: 24511865]
44. Wartberg L, Lindenberg K. Predictors of spontaneous remission of problematic internet use in adolescence: a one-year follow-up study. Int J Environ Res Public Health 2020 Jan 9;17(2):448. [doi: 10.3390/ijerph17020448] [Medline: 31936677]
45. Teppers E, Luyckx K, Klimstra TA, Goossens L. Loneliness and Facebook motives in adolescence: a longitudinal inquiry into directionality of effect. J Adolesc 2014 Jul;37(5):691-699. [doi: 10.1016/j.adolescence.2013.11.003] [Medline: 24321573]
46. Alexandraki K, Stavropoulos V, Burleigh TL, King DL, Griffiths MD. Internet pornography viewing preference as a risk factor for adolescent Internet addiction: the moderating role of classroom personality factors. J Behav Addict 2018 Jun 1;7(2):423-432. [doi: 10.1556/2006.7.2018.34] [Medline: 29788747]

XSL•FO
RenderX

47. Coyne SM, Stockdale L. Growing up with grand theft auto: a 10-year study of longitudinal growth of violent video game play in adolescents. Cyberpsychol Behav Soc Netw 2021 Jan;24(1):11-16. [doi: 10.1089/cyber.2020.0049] [Medline: 33337262]
48. Rosenberg M, Houghton S, Hunter SC, et al. A latent growth curve model to estimate electronic screen use patterns amongst adolescents aged 10 to 17 years. BMC Public Health 2018 Mar 7;18(1):332. [doi: 10.1186/s12889-018-5240-0] [Medline: 29514633]
49. Shannon H, Bush K, Villeneuve PJ, Hellemans KG, Guimond S. Problematic social media use in adolescents and young adults: systematic review and meta-analysis. JMIR Ment Health 2022 Apr 14;9(4):e33450. [doi: 10.2196/33450] [Medline: 35436240]
50. Boer M, van den Eijnden R, Boniel-Nissim M, et al. Adolescents' intense and problematic social media use and their well-being in 29 countries. J Adolesc Health 2020 Jun;66(6S):S89-S99. [doi: 10.1016/j.jadohealth.2020.02.014] [Medline: 32446614]
51. Gentile DA, Choo H, Liau A, et al. Pathological video game use among youths: a two-year longitudinal study. Pediatrics 2011 Feb;127(2):e319-e329. [doi: 10.1542/peds.2010-1353] [Medline: 21242221]
52. Huang P, Zhou Y, Li D, et al. Developmental trajectories of adolescent internet addiction: interpersonal predictors and adjustment outcomes. Res Child Adolesc Psychopathol 2023 Mar;51(3):355-367. [doi: 10.1007/s10802-022-00987-1] [Medline: 36334225]
53. Yang X, Guo WJ, Tao YJ, et al. A bidirectional association between internet addiction and depression: a large-sample longitudinal study among Chinese university students. J Affect Disord 2022 Feb 15;299:416-424. [doi: 10.1016/j.jad.2021.12.013] [Medline: 34906641]
54. Pagani LS, Jalbert M, Derevensky JL. From one year to the next: video gaming life-style predicts subsequent psychosocial risk in adolescent boys and girls. J Paediatr Child Health 2022 Sep;58(9):1554-1560. [doi: 10.1111/jpc.16032] [Medline: 35657173]
55. Shang R, Pang H, Jiang J, et al. Internet addiction and depressive and anxious symptoms among Chinese rural left-behind adolescents: mediating roles of resilience and friendship quality. Child Care Health Dev 2024 Jan;50(1):e13160. [doi: 10.1111/cch.13160] [Medline: 37594198]
56. Li G, Hou G, Yang D, Jian H, Wang W. Relationship between anxiety, depression, sex, obesity, and internet addiction in Chinese adolescents: a short-term longitudinal study. Addict Behav 2019 Mar;90:421-427. [doi: 10.1016/j.addbeh.2018.12.009] [Medline: 30553156]
57. Chen IH, Lin YC, Lin CY, Wang WC, Gamble JH. The trajectory of psychological distress and problematic internet gaming among primary school boys: a longitudinal study across different periods of COVID-19 in China. J Mens Health 2022;18(3):1-14. [doi: 10.31083/j.jomh1803070]
58. Donald JN, Ciarrochi J, Parker PD, Sahdra BK. Compulsive internet use and the development of self-esteem and hope: a four-year longitudinal study. J Pers 2019 Oct;87(5):981-995. [doi: 10.1111/jopy.12450] [Medline: 30592047]
59. Geng X, Zhang J, Liu Y, et al. Problematic use of the internet among adolescents: a four-wave longitudinal study of trajectories, predictors and outcomes. J Behav Addict 2023 Jun 29;12(2):435-447. [doi: 10.1556/2006.2023.00021] [Medline: 37294627]
60. Lee SY, Lee HK, Choi JS, et al. The Matthew effect in recovery from smartphone addiction in a 6-month longitudinal study of children and adolescents. Int J Environ Res Public Health 2020 Jul 1;17(13):4751. [doi: 10.3390/ijerph17134751] [Medline: 32630338]
61. Wichstrøm L, Stenseng F, Belsky J, von Soest T, Hygen BW. Symptoms of internet gaming disorder in youth: predictors and comorbidity. J Abnorm Child Psychol 2019 Jan;47(1):71-83. [doi: 10.1007/s10802-018-0422-x] [Medline: 29623484]
62. Kim D, Nam JK, Keum C. Adolescent internet gaming addiction and personality characteristics by game genre. PLOS ONE 2022;17(2):e0263645. [doi: 10.1371/journal.pone.0263645] [Medline: 35143558]
63. Straatmann VS, Almquist YB, Oliveira AJ, Veiga GV, Rostila M, Lopes CS. Stability and bidirectional relationship between physical activity and sedentary behaviours in Brazilian adolescents: longitudinal findings from a school cohort study. PLoS ONE 2019;14(1):e0211470. [doi: 10.1371/journal.pone.0211470] [Medline: 30682158]
64. Cliff DP, Howard SJ, Radesky JS, McNeill J, Vella SA. Early childhood media exposure and self-regulation: bidirectional longitudinal associations. Acad Pediatr 2018;18(7):813-819. [doi: 10.1016/j.acap.2018.04.012] [Medline: 29704999]
65. Anguera JA, Rowe MA, Volponi JJ, et al. Enhancing attention in children using an integrated cognitive-physical videogame: a pilot study. NPJ Digit Med 2023 Apr 12;6(1):65. [doi: 10.1038/s41746-023-00812-z] [Medline: 37046040]
66. Opoku Asare K, Terhorst Y, Vega J, Peltonen E, Lagerspetz E, Ferreira D. Predicting depression from smartphone behavioral markers using machine learning methods, hyperparameter optimization, and feature importance analysis: exploratory study. JMIR Mhealth Uhealth 2021 Jul 12;9(7):e26540. [doi: 10.2196/26540] [Medline: 34255713]
67. Maclennan G, McDonald A, McPherson G, Avenell A, Gaynes L, Shuman N. Abstracts from the Society for Clinical Trials annual meeting, Boston, Massachusetts, USA - May 19-22, 2013. Clin Trials 2013 Aug;10(2_suppl):S1-S88. [doi: 10.1177/1740774513497438] [Medline: 24188963]

XSL•FO
RenderX

68. Woodruff SI, Conway TL, Edwards CC, Elliott SP, Crittenden J. Evaluation of an internet virtual world chat room for adolescent smoking cessation. Addict Behav 2007 Sep;32(9):1769-1786. [doi: 10.1016/j.addbeh.2006.12.008] [Medline: 17250972]
69. Hu YZ, Urakami J, Wei HT, Vomberg LH, Chignell M. Longitudinal analysis of sustained performance on gamified cognitive assessment tasks. Appl Neuropsychol Adult 2024;31(4):502-526. [doi: 10.1080/23279095.2022.2039931] [Medline: 35196467]
70. Schluter P, Lee M, Hamilton G, Coe G, Messer-Perkins H, Smith B. Keep on brushing: a longitudinal study of motivational text messaging in young adults aged 18-24 years receiving work and income support. J Public Health Dent 2015;75(2):118-125. [doi: 10.1111/jphd.12079] [Medline: 25409935]
71. Steinemann ST, Geelan BJ, Zaehringer S, et al. Potentials and pitfalls of increasing prosocial behavior and self-efficacy over time using an online personalized platform. PLoS ONE 2020;15(6):e0234422. [doi: 10.1371/journal.pone.0234422] [Medline: 32584828]
72. Bender K, Begun S, DePrince A, Haffejee B, Kaufmann S. Utilizing technology for longitudinal communication with homeless youth. Soc Work Health Care 2014 Oct;53(9):865-882. [doi: 10.1080/00981389.2014.925532] [Medline: 25321934]
73. Ye S, Pope ZC, Lee JE, Gao Z. Effects of school-based exergaming on urban children's physical activity and cardiorespiratory fitness: a quasi-experimental study. Int J Environ Res Public Health 2019 Oct 23;16(21):4080. [doi: 10.3390/ijerph16214080] [Medline: 31652857]
74. Cooray N, Ho C, Bestman A, et al. Exploring the potential of a behavior theory-informed digital intervention for infant fall prevention: mixed methods longitudinal study. JMIR Pediatr Parent 2024 Jan 3;7:e47361. [doi: 10.2196/47361] [Medline: 38170580]
75. Prabu Kumar A, Omprakash A, Chokkalingam Mani PK, et al. E-learning and E-modules in medical education-a SOAR analysis using perception of undergraduate students. PLoS ONE 2023;18(5):e0284882. [doi: 10.1371/journal.pone.0284882] [Medline: 37205679]
76. Cabanas-Sánchez V, García-Cervantes L, Esteban-Gonzalo L, Girela-Rejón MJ, Castro-Piñero J, Veiga Ó. Social correlates of sedentary behavior in young people: the UP&DOWN study. J Sport Health Sci 2020 Mar;9(2):189-196. [doi: 10.1016/j.jshs.2019.03.005] [Medline: 32099727]
77. Salway RE, Emm-Collison L, Sebire S, Thompson JL, Jago R. Associations between socioeconomic position and changes in children's screen-viewing between ages 6 and 9: a longitudinal study. BMJ Open 2019 Dec 10;9(12):e027481. [doi: 10.1136/bmjopen-2018-027481] [Medline: 31826887]
78. McDaniel BT, Radesky JS. Technoference: longitudinal associations between parent technology use, parenting stress, and child behavior problems. Pediatr Res 2018 Aug;84(2):210-218. [doi: 10.1038/s41390-018-0052-6] [Medline: 29895837]
79. Radesky JS, Silverstein M, Zuckerman B, Christakis DA. Infant self-regulation and early childhood media exposure. Pediatrics 2014 May;133(5):e1172-e1178. [doi: 10.1542/peds.2013-2367] [Medline: 24733868]
80. Radesky JS, Kaciroti N, Weeks HM, Schaller A, Miller AL. Longitudinal associations between use of mobile devices for calming and emotional reactivity and executive functioning in children aged 3 to 5 years. JAMA Pediatr 2023 Jan 1;177(1):62-70. [doi: 10.1001/jamapediatrics.2022.4793] [Medline: 36508199]
81. van Oosten JMF, Peter J, Boot I. Exploring associations between exposure to sexy online self-presentations and adolescents' sexual attitudes and behavior. J Youth Adolesc 2015 May;44(5):1078-1091. [doi: 10.1007/s10964-014-0194-8] [Medline: 25287000]
82. Trager BM, Morgan RM, Boyle SC, Montiel Ishino FA, LaBrie JW. Capturing the bigger picture: a gestalt of general and alcohol-specific social media usage during the transition to college as a predictor of first-year alcohol use and consequences. Addict Behav 2023 Jan;136:107472. [doi: 10.1016/j.addbeh.2022.107472] [Medline: 36067637]
83. Brunborg GS, Skogen JC, Andreas JB. Time spent on social media and alcohol use among adolescents: a longitudinal study. Addict Behav 2022 Jul;130:107294. [doi: 10.1016/j.addbeh.2022.107294] [Medline: 35231842]
84. Calhoun BH, Lee CM, Fairlie AM. Exposure to media messages portraying effects of alcohol use in a young adult sample. Subst Use Misuse 2022;57(8):1281-1286. [doi: 10.1080/10826084.2022.2076880] [Medline: 35621116]
85. Hendriks H, de Nooy W, Gebhardt WA, van den Putte B. Causal effects of alcohol-related Facebook posts on drinking behavior: longitudinal experimental study. J Med Internet Res 2021 Nov 11;23(11):e28237. [doi: 10.2196/28237] [Medline: 34762061]
86. Ng Fat L, Cable N, Kelly Y. Associations between social media usage and alcohol use among youths and young adults: findings from Understanding Society. Addiction 2021 Nov;116(11):2995-3005. [doi: 10.1111/add.15482] [Medline: 33886119]
87. Boers E, Afzali MH, Conrod P. A longitudinal study on the relationship between screen time and adolescent alcohol use: the mediating role of social norms. Prev Med 2020 Mar;132:105992. [doi: 10.1016/j.ypmed.2020.105992] [Medline: 31954144]
88. Moreno MA, Cox ED, Young HN, Haaland W. Underage college students' alcohol displays on Facebook and real-time alcohol behaviors. J Adolesc Health 2015 Jun;56(6):646-651. [doi: 10.1016/j.jadohealth.2015.02.020] [Medline: 26003580]
89. Hull JG, Brunelle TJ, Prescott AT, Sargent JD. A longitudinal study of risk-glorifying video games and behavioral deviance. J Pers Soc Psychol 2014 Aug;107(2):300-325. [doi: 10.1037/a0036058] [Medline: 25090130]

XSL•FO
RenderX

90. Rew L, Mackert M, Zhaohui S. Relationship of media location to adolescent health-risk behaviors: gender differences. J Adolesc Health 2014 Feb;54(2):S47. [doi: 10.1016/j.jadohealth.2013.10.107]
91. D'Angelo J, Moreno M. Facebook-induced friend shift and identity shift: a longitudinal study of Facebook posting and collegiate drinking. Cyberpsychol Behav Soc Netw 2019 Mar;22(3):186-191. [doi: 10.1089/cyber.2018.0246] [Medline: 30855991]
92. Tan ASL, Chen JT, Keen R, et al. Culturally tailored anti-smoking messages: a randomized trial with U.S. sexual minority young women. Am J Prev Med 2024 May;66(5):840-849. [doi: 10.1016/j.amepre.2023.12.001] [Medline: 38065403]
93. Loukas A, Paddock EM, Li X, Harrell MB, Pasch KE, Perry CL. Electronic nicotine delivery systems marketing and initiation among youth and young adults. Pediatrics 2019 Sep;144(3):e20183601. [doi: 10.1542/peds.2018-3601] [Medline: 31451608]
94. Zheng X, Li W, Wong SW, Lin HC. Social media and E-cigarette use among US youth: longitudinal evidence on the role of online advertisement exposure and risk perception. Addict Behav 2021 Aug;119:106916. [doi: 10.1016/j.addbeh.2021.106916] [Medline: 33798917]
95. Depue JB, Southwell BG, Betzner AE, Walsh BM. Encoded exposure to tobacco use in social media predicts subsequent smoking behavior. Am J Health Promot 2015;29(4):259-261. [doi: 10.4278/ajhp.130214-ARB-69] [Medline: 24670071]
96. Maes C, Vandenbosch L. Adolescents' use of sexually explicit internet material over the course of 2019-2020 in the context of the COVID-19 pandemic: a three-wave panel study. Arch Sex Behav 2022 Jan;51(1):105-121. [doi: 10.1007/s10508-021-02122-5] [Medline: 35001225]
97. Soutullo C, Huss M, Coghill D. Abstracts of the 16th International Congress of ESCAP European Society for Child and Adolescent Psychiatry. Eur Child Adolesc Psychiatry 2015 Jun;24(S1):1-303. [doi: 10.1007/s00787-015-0714-4]
98. Jones S, Chu YH, Burke MP, Teixeira A, Blake CE, Frongillo EA. A case for targeting marketing and availability in school food policy: adolescents' food purchases at school and exposure to television, internet, and video games. J Hunger Environ Nutr 2012 Jan;7(1):1-10. [doi: 10.1080/19320248.2012.651389]
99. Morgan PL, Wang Y, Woods AD. Risk and protective factors for frequent electronic device use of online technologies. Child Dev 2021 Mar;92(2):704-714. [doi: 10.1111/cdev.13532] [Medline: 33427312]
100. Chhabra J, Pilkington V, Benakovic R, Wilson MJ, La Sala L, Seidler Z. Social media and youth mental health: scoping review of platform and policy recommendations. J Med Internet Res 2025 Jun 20;27:e72061. [doi: 10.2196/72061] [Medline: 40540734]
101. Yeo G, Reich SM, Liaw NA, Chia EYM. The effect of digital mental health literacy interventions on mental health: systematic review and meta-analysis. J Med Internet Res 2024 Feb 29;26:e51268. [doi: 10.2196/51268] [Medline: 38421687]
102. Hamilton JL, Torous J, Szlyk HS, et al. Leveraging digital media to promote youth mental health: flipping the script on social media-related risk. Curr Treat Options Psychiatry 2024 Jun;11(2):67-75. [doi: 10.1007/s40501-024-00315-y] [Medline: 39258150]
103. Arias López MDP, Ong BA, Borrat Frigola X, et al. Digital literacy as a new determinant of health: a scoping review. PLOS Digit Health 2023 Oct;2(10):e0000279. [doi: 10.1371/journal.pdig.0000279] [Medline: 37824584]

Abbreviations

HBSC: Health Behavior in School-aged Children
LMICs: low- and middle-income countries
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews

Edited by A Mavragani, T Sanchez; submitted 30.Sep.2025; peer-reviewed by R Ruparel, T Yusuff; revised version received 20.Mar.2026; accepted 20.Mar.2026; published 27.Apr.2026.

Please cite as:
Banerjee P, Holly L
Everyday Digital Technology Use and Youth Health: Scoping Review of Longitudinal Studies
JMIR Public Health Surveill 2026;12:e85094
URL: https://publichealth.jmir.org/2026/1/e85094
doi:10.2196/85094

© Preetika Banerjee, Louise Holly. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 27.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic

XSL•FO
RenderX

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

XSL•FO
RenderX

Digital Tools' Effectiveness on Physical Activity Outcomes in Children and Adolescents: Umbrella Review

Garden Tabacchi[1], Prof Dr; Antonino Scardina[1], PhD; Antonella Amato[2], Prof Dr; Marta Giardina[2], PhD; Giulia Accardi[3], Prof Dr; Valentina Di Liberto[3], Prof Dr; Giuseppe Giglia[3], Prof Dr Med; Sonya Vasto[2], Prof Dr; Monica Frinchi[4], Prof Dr; Paolo Boffetta[5], Prof Dr Med; Walter Mazzucco[2], Prof Dr Med; Marianna Bellafiore[1], Prof Dr

[1]Department of Psychology, Educational Science and Human Movement, University of Palermo, Via Giovanni Pascoli 6, Palermo, PA, Italy

[2]Department of Biological, Chemical and Pharmaceutical Sciences and Technologies, University of Palermo, Palermo, PA, Italy

[3]Department of Biomedicine, Neuroscience and Advanced Diagnostics (BIND), University of Palermo, Palermo, PA, Italy

[4]Department of Medicine and Surgery, Università degli Studi di Enna Kore, Enna, Sicily, Italy

[5]Department of Medical and Surgical Sciences, University of Bologna, Bologna, Italy

Corresponding Author:
Garden Tabacchi, Prof Dr
Department of Psychology, Educational Science and Human Movement, University of Palermo, Via Giovanni Pascoli 6, Palermo, PA, Italy

Abstract

Background: A substantial proportion of children and adolescents worldwide do not meet current physical activity (PA) guidelines. So digital tools interventions have been implemented worldwide. However, evidence regarding the effectiveness of these remains inconsistent, underscoring the need for a comprehensive synthesis of the available literature.

Objective: This umbrella review aimed to summarize and critically evaluate the characteristics and effectiveness of digital interventions designed to increase PA in children and adolescents.

Methods: An overview of systematic reviews (SRs) and meta-analyses of randomized controlled trials (RCTs) was conducted in accordance with the PRIOR (Preferred Reporting Items for Overviews of Reviews) and the PRISMA-S (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Literature Search Extension) checklists. Reviews published between 2018 and 2025 were searched in SCOPUS, PubMed or MEDLINE, Web of Science, Cochrane Database of Systematic Reviews, and SPORTDiscus or EBSCO, using a combination of terms addressing the type of digital tool (eg, apps, wearables, etc) and device-based PA outcomes. Eligible SRs and meta-analyses focused on populations aged 6 - 17 years and evaluated digital interventions aimed at increasing PA. Methodological quality was assessed using AMSTAR 2 (A Measurement Tool to Assess Systematic Reviews) for SRs and meta-analyses, and Risk of Bias 2 for RCTs. Intervention characteristics and effectiveness outcomes were summarized as frequencies, and chi-square tests were applied to explore differences in effectiveness across study features.

Results: Forty-eight SRs or meta-analyses comprising 62 RCTs were included. The reviews' quality was moderate to poor, and 7/62 (11.3%) of RCTs were judged to have a high risk of bias. The mainly addressed PA outcomes were moderate-to-vigorous PA and step counts, commonly measured using accelerometers and pedometers. Overall, 45.2% of interventions demonstrated effectiveness for at least 1 PA outcome. Higher effectiveness rates were observed in RCTs that targeted PA as the sole primary outcome (19/32, 59.4%), used wearables as both the digital intervention component (7/11, 63.6%) and delivery device (8/12, 66.7%), and used pedometers for PA outcome measurement (13/16, 81.3%). Significant differences in effectiveness were found for the type of PA assessment device (P=.003) and for interventions targeting low-income populations, which showed lower effectiveness (P=.01). Additional trends were identified for geographic region (P=.06), intervention setting (P=.09), baseline activity level (P=.06), intervention focus (P=.09), and device brand (P=.09).

Conclusions: This novel umbrella review provides a comprehensive synthesis of digital PA interventions in youth, foreseeing potential factors that may influence their effectiveness, and highlighting methodological limitations. It offers evidence-based insights for practitioners, educators, and policymakers, helping to identify digital tools most likely to successfully increase PA in youth. Future research should prioritize stronger methodological rigor and more precise intervention designs. This has clear value for the public health practice to reduce long-term disease risk.

Trial Registration: PROSPERO CRD42024510602; https://www.crd.york.ac.uk/PROSPERO/view/CRD42024510602

(JMIR Public Health Surveill 2026;12:e75769) doi:10.2196/75769

XSL•FO
RenderX

KEYWORDS

physical activity; digital; effectiveness; children; adolescents; mobile phone

Introduction

Physical activity (PA) is an important determinant of health status. Nearly a third of the population in the world does not meet the PA recommended levels [1], and this can be observed already in young ages, with 80% of adolescents aged 11 - 17 years not meeting the current PA guidelines [2]. This low PA level in children and adolescents may lead to obesity, poor cardiometabolic health, low fitness levels, sleep disturbances, and other diseases in adulthood.

Continuous monitoring of PA is thus needed to identify trends and priorities in public health, and effective interventions are mandatory to improve PA in youngsters and prevent future health status impairment. Among the most common actions, increasing evidence outlines the use of digital interventions in youth to improve PA and reduce sedentary behaviors (SBs), beyond addressing other health-related attitudes such as eating, tobacco smoking, or alcohol drinking.

Recently, the Global Action Plan on Physical Activity 2018 - 2030 of the World Health Organization underlined the need to increase PA by 15% by 2030, through different policy actions that also include digital approaches such as apps, wearables, or virtual coaching [3,4].

Digital technologies are particularly appealing to children and adolescents, and this is supported by recent data indicating that among those aged 15 years, more than 50% spent 30+ hours/week using digital devices, and in some countries up to 43% spent 60+ hours/week [5]. The main used technology is the internet, with 1 in 3 internet users in the world being children and adolescents [6], and 98% of those aged 15 years having a smartphone with internet connection [5]. The main digital devices used are smartphones, with the rate of adoption having increased particularly in those of young ages; 95% of teens from the US and 86% of European adolescents use smartphones and social media every day [5,7]. Other digital devices used in this age can be PCs, tablets, smartwatches, activity trackers, virtual reality headsets, and other wearables, and digital components can include websites, apps, text messages, voice messages, gamification components, virtual coaching, virtual reality, etc.

As all these tools can offer effective ways to improve and promote PA in youth, they have been increasingly included in tailored interventions in recent years [8,9]. These interventions were conducted in different settings, including education contexts such as schools, although only 40% of primary schools, 50% of lower secondary schools, and 65% of upper secondary schools were connected to the internet for pedagogical purposes [10]. Several studies were found on behavior change techniques (ie, activity monitoring with feedback, goal setting, graded tasks, or social incentives) [11-13], which are supposed to be effective for initiating and maintaining active behaviors over time by enhancing motivation and accountability.

Nonetheless, still clearer information is needed on the effectiveness of digital approaches in modifying PA behaviors of young populations. Literature shows that interventions aimed at increasing PA have been performed in different ways, with different designs and digital tools [14]. Thus, the main objective of this umbrella review is to systematically summarize and critically evaluate the characteristics and effectiveness of digital interventions aimed at increasing PA in children and adolescents, with particular attention to intervention components, determinants of effectiveness, and methodological limitations.

Methods

Protocol and Registration

This umbrella review was registered in the PROSPERO (International Prospective Register of Systematic Reviews; CRD42024510602). The search strategy was amended from what was originally specified at registration and in the protocol. The search was performed by following the PRIOR (Preferred Reporting Items for Overviews of Reviews) [15] and the PRISMA-S (Preferred Reporting Items for Systematic Review and Meta-Analyses Literature Search Extension) [16] checklists (Checklists 1 and 2).

Study Design

The present review is part of a larger systematic review (SR) protocol, which includes studies identifying digital interventions aimed to improve different lifestyles, such as PA, dietary behaviors, migraine, or sleep patterns, within the DARE (Digital Lifelong Prevention) Project. The DARE Project is an initiative funded by the Italian Ministry of Universities and Research within the National Recovery and Resilience Plan, aimed at using digital technologies to create a data-driven health care ecosystem that can contribute to promoting healthy lifestyles and disease prevention and cure.

A systematically conducted overview of reviews and meta-analyses was chosen as the best design for the search and is hereafter referred to as an umbrella review. An umbrella review can be viewed as a way to obtain evidence from different existing reviews that address a broader question and can be useful for comparing interventions and developing guidelines, thus providing a very high level of evidence [17,18]. As a huge number of studies have been conducted and published in the literature on this topic, most of them lacking methodological rigor, the present search was limited to capturing SRs and meta-analyses that collected experimental studies, that is, randomized controlled trials (RCTs). SRs included by the authors followed a strict, peer-reviewed protocol to ensure all available evidence was considered impartially, involving a comprehensive literature search, critical appraisal of study quality, and statistical synthesis of results to provide a clear, evidence-based answer.

Search Strategy

The review was conducted on 5 databases: Scopus, PubMed or MEDLINE, Web of Science, the Cochrane Database of Systematic Reviews via the Cochrane Library, and SPORTDiscus via EBSCOhost. The search was conducted in

XSL•FO
RenderX

November 2025, with record publication dates restricted to the years 2018 - 2025, informed by evidence that research on second-generation technologies (such as smartphones and wearables) has risen sharply since 2013 [19]; a 2018 review would therefore be expected to encompass a larger set of previously used tools.

A manual search was conducted in parallel by browsing the reference lists of all included reviews, by searching for information on the websites, and from gray literature. For papers not freely available on the internet, the authors were contacted and asked to provide the full-text versions.

Keywords used for the search included a combination of terms addressing the type of digital tool and the main PA outcomes (Multimedia Appendix 1). In detail, the type of digital tools searched included apps, wearables, social media, messaging, exergames, digital assistants, and different devices such as pedometers or accelerometers (Multimedia Appendix 1). No language restrictions were applied to the search strategy. Filters used were the range publication year from 2018 to 2025, the study design being an SR or meta-analysis, and the age range between 6 and 65 years. With regard to age, a comprehensive search for the age ranges 6 - 65 years was performed first, as this was the target of the authors' pilot study within the DARE Project, and afterward, studies targeting only children and adolescents (6 - 17 y) were selected for the purposes of this study. The decision to conduct separate analyses for younger populations and adults was motivated by important age-related differences in the use of digital technologies. Young people tend to engage with digital tools that are distinct from those commonly used by adults. Moreover, these tools may influence behaviors in different ways across age groups. In particular, the impact of digital technologies on PA practice may vary according to developmental stage. For these reasons, age-specific analyses were considered necessary to ensure a more accurate interpretation of the findings.

References were managed and duplicates removed using Zotero (version 6.0.37, Corporation for Digital Scholarship). The selection process, with a first screening of titles and abstracts and a second screening of full texts, was conducted by 2 independent reviewers. Possible divergences were solved by discussion between the two reviewers or with the intervention of a third reviewer.

Eligibility Criteria

The PICOS (Population, Intervention, Comparison, Outcome, and Study Type) criteria were used to derive the eligibility of the studies.

For the population, we selected healthy human participants in the age range of 6 - 17 years. Healthy participants are those who are not diagnosed with any disease or who do not have disabilities that could affect PA outcome assessment. Reviews targeted to overweight or obese participants without another declared disease were retained. Moreover, populations aged 6 - 65 years were initially included in the review, based on the needs of the project; afterward, studies targeting only populations aged between 6 and 17 years were selected for the present review.

For intervention, we had interventions using digital approaches to increase PA. Such digital approaches should have included the use of one or a combination of the following: IT devices (eg, smartphone, tablet, and PC); wearable activity trackers (eg, accelerometers and pedometers); applications (eg, stand-alone apps or apps associated with wearable devices); social media tools (eg, chatbots or conversational agents, and social networks); messaging tools (eg, email, SMS text messages, and podcasts); web-based tools (eg, websites, and video chats or video sharing); exergames, active video games, or gamification approaches; and synchronous or nonsynchronous PA coaching or training.

For comparison, we use any other intervention or no intervention.

For the outcome, we used PA as the primary or secondary outcome.

For study type, we used SRs and meta-analyses including RCTs. The following reviews were excluded: reviews of reviews or meta-analyses; scoping or narrative reviews; those not reporting PA-specific outcomes (eg, those reporting only SBs, only dietary or nutrition outcomes, or only weight-related outcomes); those focusing only on wearables that cannot be applied to huge populations for economical or organizational reasons (eg, virtual reality); those focusing on individuals (telemedicine); studies where the intervention was not delivered by at least 1 digital tool; papers where the digital component was used only to "measure" PA and not as a means to increase PA; those reporting only nondigital self-reported tools (such as paper questionnaires) for the assessment of PA; interventions for secondary or tertiary prevention (eg, targeted to individuals with preclinical signs); studies targeted to specific groups, such as smokers or pregnant women.

Data Extraction

The data from the selected reviews were extracted by the same 2 independent reviewers that conducted the selection process, and the following information was reported: first author and year of publication; document type; design of studies included in the review; country; age group; special population; date range of the search; number of databases searched; digital device for intervention delivery; digital component for intervention delivery; focus of intervention or digital tool; outcomes; number of total studies included in the review; number of total RCTs in all reviews; number of RCTs of interest; and inclusion criteria, first author, and publication date of RCTs describing tools of interest.

For the primary studies identified within each review, the following information was extracted: first author and year of publication; country; age group; special population; school attendance; sample dimension and type; intervention duration, arms, or name; setting; focus task; theoretical foundation; digital component or components of intervention; nondigital component or components; digital device for intervention delivery; intervention group (IG) and control group (CG); device-based or self-reported PA measure; PA objective measurement tool; fitness tests assessed; other assessed measures; and effectiveness on PA or SB outcomes. Regarding outcome effectiveness, an

XSL•FO
RenderX

intervention was classified as partially effective when positive effects were observed for some, but not all, PA outcomes.

Data Synthesis Methods

Data regarding the characteristics of the different reviews and of the RCTs selected are presented as numbers and percentage frequencies. Effectiveness data were stratified by the type of device-based PA outcome and by the characteristics of the interventions. Differences between groups in relation to their effectiveness have been estimated through the Pearson chi-square test. All statistical estimates were conducted using STATA MP (version 12.1, StataCorp, 2011) software. As a meta-analysis was not feasible given the design of this review, heterogeneity between studies was explored using qualitative and descriptive methods combined with narrative synthesis. Sensitivity analyses were qualitatively performed as well to understand differences in populations, interventions, and outcomes that may explain variability in findings. Considered factors for sensitivity analyses were the quality of included reviews and RCTs, the study population characteristics, the intervention features, and the outcome measurement.

Degree of Overlap

The extent of overlap of RCTs across the included SRs and meta-analyses was examined using the corrected covered area (CCA) approach [20]. As STATA does not provide a dedicated command for computing the CCA, a citation matrix was constructed, with primary studies listed in rows and the different SRs or meta-analyses in columns. A custom STATA script was developed to automate the calculation according to the formula: $CCA = (N - r) / (r \times c - r)$, where N represents the total number of publications, r the number of unique studies, and c the number of reviews or meta-analyses. The resulting values were converted to percentages by multiplying by 100. A CCA of 0% indicates no overlap (each review or meta-analysis includes only unique RCT); 0% - 5% reflects slight overlap; 6% - 10% moderate overlap; 11% - 15% high overlap; and values above 15% indicate very high overlap. A value of 100% denotes complete overlap, meaning that all reviews or meta-analyses include the same primary studies.

Quality Assessment of the Reviews or Meta-Analyses

The methodological quality of included reviews and meta-analyses reporting randomized controlled interventions was assessed using the AMSTAR 2 (A Measurement Tool to Assess Systematic Reviews) tool [21]. This tool allows a quality assessment based on 7 critical and 9 noncritical domains, rather than a total score, generating an overall rating of “high” quality (those reviews or meta-analyses that meet 7/7 critical domains), “moderate” (6/7), “low” (those meeting all but one critical domain and missing few noncritical domains), and “critically low” (not meeting multiple critical domains).

Quality Assessment of the RCTs

The methodological quality of each included RCT was evaluated using the 5 domains of the revised Cochrane Risk of Bias 2 (RoB 2) tool for RCTs [22]. Overall judgments for individual studies were derived according to the RoB 2 algorithm, categorizing studies as having a low risk of bias, some concerns, or a high risk of bias. A total of 2 reviewers conducted the assessments independently, and any discrepancies were resolved through consultation with a third reviewer.

Results

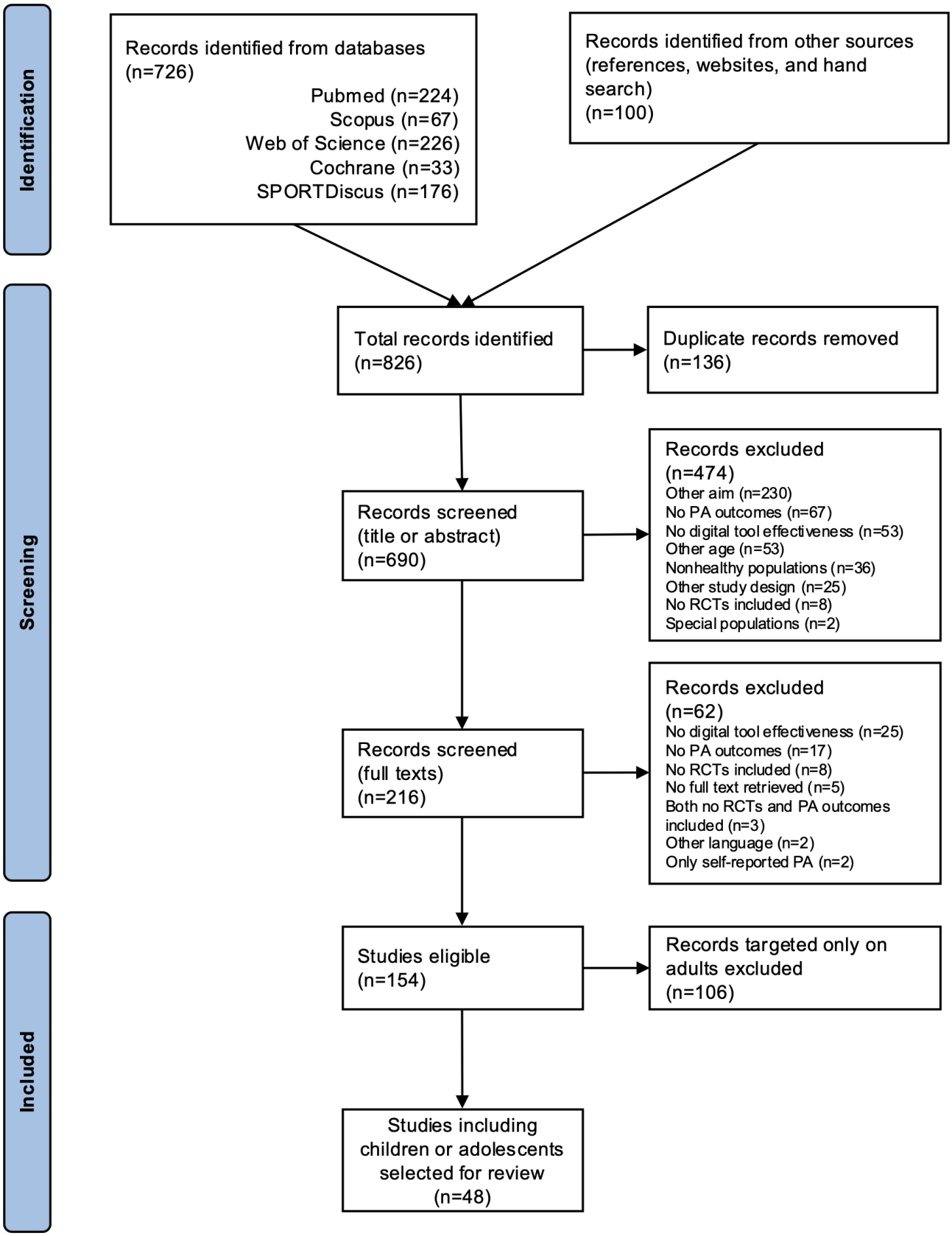
Search Flow Results

Figure 1 shows the flow diagram of the search.

The database search included a total of 726 records, which were added to the 100 records identified by the manual search. After excluding 136 duplicates, a total of 690 reviews were screened by title and abstract, and subsequently, the full texts of the selected 216 records were examined. A total of 154 studies were eligible, but they included all-age populations. A further selection finally included a total of 48 SRs and meta-analyses targeted specifically to children and adolescents [23-70].

XSL•FO
RenderX

Figure 1. PRISMA flow diagram of the search. PA: physical activity; PRISMA: Preferred Reporting Items for Systematic Review and Meta-Analyses; RCT: randomized controlled trial.



XSL•FO
RenderX

Description of the SRs or Meta-Analyses

Characteristics of the SRs or Meta-Analyses

The characteristics of the included SRs or meta-analyses are shown in Multimedia Appendix 2, while Figure S1 Multimedia Appendix 3 provides a summary of the percentage frequencies of the different characteristics.

As shown in Figure S1 Multimedia Appendix 3, in 43.8% of the reviews, a meta-analysis was added to the SR [23,26,27,31,32,40,42,43,46,47,50,52,54,58,60,61,63,66-69]. The considered SRs or meta-analyses included any kind of study, both experimental and observational. The major part of SRs or meta-analyses was published by European research groups, mainly from the United Kingdom [28,29,32,34,39,41,45,59,61], Germany [23,24,33,47,50], Italy [25,38,55], and Spain [26,43,63]. The other European SRs or meta-analyses were published by research groups from Portugal [35,56], France [40], the Netherlands [44], and Denmark [65]. Less than one-third of the reviews were from research groups based in different Asian countries, such as China [31,49,53,54,60,66-69], Hong Kong [52], Korea [52], Malaysia [58,62], Iran [36], Singapore [37], and Qatar [41]. Most American SRs or meta-analyses were from the United States [30,47,48,57,71-75], and the others from Canada [51] and Brazil [42,64]. Only 2 reviews were published by Australian research groups [27,70].

Most reviews included studies using a miscellaneous assortment of device types [24,27,29,32-35,37,38,40-42,45,48,53,56,57,59,61,63,68-71] including mobiles, computers, tablets, and wearables, or mobiles and wearables, followed by only mobiles or smartphones [23,25,30,31,36,39,44,46,47,49-51,55,66] or only wearables [26,28,43,52,54,58,67], or consoles as devices for their gamification tools [60,62,64,65].

Parallelly, most SRs or meta-analyses searched for a miscellaneous assortment of different components for the intervention delivery [23,24,27,29-34,37,38,41,46,55-59,61,63,70], while a lower number searched only for gamification tools, including exergames, serious games, or games in devices [35,40,42,45,48,49,60,62,64,65,68,71], and in one case the exergame use was associated with coaching or telehealth [45]. Some reviews examined apps [25,36,44,47,50,51,66], and some investigated wearables [26,28,43,52,54,67]; 2 reviews were focused on chatbots [53,69] and only 1 on text messaging [39].

The focus task comprised exclusively of PAs in less than half of the reviews [24,28,30,31,33,37,40,41,44,45,47-50,52,54,57,58,60,61,63,76]; in the other reviews, the main outcomes were PAs associated with SB [23,26,38,39,43,67-69], with weight, obesity, or body composition [29,42,64,65], with diet [35] or physical fitness [66]; other reviews were focused on weight or obesity management [25,32,34,36,62,70], and on multiple health behaviors such as alcohol consumption, tobacco smoking, or diet and sleep, beyond PA and SB [27,46,51,53,55,56,59,71].

The selected reviews included an average of 19 (SD 12.2) experimental studies each for a total of 1023 studies. An average of 3.7 (SD 2.9) RCTs was present in each review. A total of 165 primary RCTs were identified across the 48 SRs or meta-analyses, according to the inclusion criteria (RCTs, healthy populations, PA outcomes present, ages between 6 and 17 years, published in the last 15 years, and effectiveness of digital tools assessed). The resulting CCA was 3.5%, indicating a slight degree of overlap. The 62 unique component studies identified were included in the final analysis [77-138].

Quality of the SRs or Meta-Analyses

The schematic results of the SRs' and meta-analyses' quality are shown in Multimedia Appendix 4. A high percentage of the reviews was rated as critically low (43.8%) [24-30,33-39,41,44,45,48,49,51,70], a rating that was often attributed to authors' failure to address one of the critical criteria, namely providing a list of excluded studies along with justifications for their exclusion.

Description of RCTs

General Characteristics

All extracted data from RCTs are shown in Multimedia Appendix 5. A total of 22 interventions did not have a specific name. The same intervention was described in 2 papers as follows: Active Teen Leaders Avoiding Screen-Time obesity prevention program [110,128], Dads and Daughters Exercising and Empowered intervention [115,116], Diab and Nano study [80,82], Girls on the Move intervention [120,124], MyMovez project [135,136], and the Raising Awareness of PA study [123,137]. The Nutrition and Enjoyable Activity for Teen Girls intervention was described in 3 papers [88,89,109].

Considerable heterogeneity was found for the sample dimension that was included between a minimum of 10 and a maximum of 1519 schoolchildren, and for intervention duration, which ranged between 2 weeks and 24 months.

In Figure S2 Multimedia Appendix 3, a summary of the main characteristics of the interventions, divided by general RCT aspects, is presented. The included RCTs were published between 2010 and 2025, mostly in America, and were primarily conducted in school settings, with others in home or community environments. Most interventions targeted middle school children, followed by primary school children and adolescents, with some focusing on overweight or obese students, and 11.3% conducted in low-income areas. The majority used a 2-arm design with IGs and CGs and mainly focused on PA alone, followed by PA combined with other behaviors or obesity-related outcomes.

A total of 38 interventions were based on different theories for behavioral changes. The most used theoretical models were the following: the self-determination theory by Deci and Ryan [139] and Ryan and Deci [140] was used in 16 RCTs [78,81,95-97,99,100,120,121,124,127,128,132,135,136,138]; the social cognitive theory by Bandura [141] was used in 15 RCTs [86,88,89,101,102,108,114,121,123,125,128,130,131,134,137]; 5 interventions [94,102,122,135,136] were based on the theory of planned behavior by Aizen [142]; the transtheoretical model

XSL•FO
RenderX

of behavior change that was suggested by Prochaska and Velicer [143] was applied in 2 interventions [86,107].

Digital Aspects

Figure S3 Multimedia Appendix 3 shows a summary of the main characteristics of the trials related to the digital aspects of the intervention delivery. Most interventions included a wearable as a digital component of the intervention; in some cases, the wearable was a unique, tested digital tool [77,79,92,103,105,106,108,113,114,117,118,129], in some other trials, it was included in a multicomponent intervention [84,86,88,89,99,100,102,107,109,110,114-116,126,128,132,135-137]. Several RCTs used gamification, apps, web-based technologies, and text messaging; few interventions used social media, live coaching, or intelligent personal assistants. The most used digital devices were mobile phones or smartphones or wearables; 19.4% of interventions used the console or arcade machine and the computer, laptop, or tablet.

Around half of the RCTs used nondigital components for the intervention delivery, such as goals, rewards, incentives, advice, diaries, school programs, face-to-face educational sessions, lessons, or booklets (Multimedia Appendix 5).

A total of 53.2% of interventions used device-based measures only to evaluate PA outcomes, while the rest used device-based measures together with self-report measures, such as data from questionnaires or step diaries (Figure S3 Multimedia Appendix 3).

Device-based measurement tools were mainly accelerometers or activity trackers, that is, ActiGraph (Ametris, LLC) [80-83,85,86,88-90,92,93,98,99,101,102,104,109-112,117,120,123,124,128,130-133,137], Fitbit Flex or Fitbit Zip (Fitbit, Inc) [121,123,131,134-136], GENEActiv (ActivInsights Ltd) accelerometers [78,84,87] and Polar Active (Polar Products Inc) sensors [103,107,126,144]. Around one-third of the interventions used pedometers for step count, mainly Yamax (Polygon Direct [UK] Ltd) [77,79,94,108,113-116,129], Tractivity (Tractivity Ltd) [95-97], and Omron (OMRON Healthcare Co, Ltd) [106,118]. In 5 RCTs, other device-based measurement tools were embedded, for example, in the game machines such as the PlayStation (Sony Interactive Entertainment LLC) memory cards [82,93,111,112,122].

Device-based measures consisted mostly of moderate-to-vigorous physical activity (MVPA) or other PA outcomes such as vigorous physical activity (VPA), moderate PA, light physical activity (LPA), and total PA [80-82,85-88,88,92,93,95-98,101,103,104,109-112,117,119,121,122,126-128,130,131,133,137]; step counts were evaluated in half of the interventions [77,79,84,91,92,94-99,102,106,108,113-116,118,121,123,125,127,129-132,134-136,138].

A total of 19 RCTs also assessed physical fitness abilities through fitness tests; for example, cardiorespiratory fitness through the 20 m shuttle test [84,104,111,120], the 1-mile walk or run test [90,113], or the grip strength through the dynamometer [84,103,107,128]. A total of 86.9% of the interventions collected other outcomes, including anthropometry or body composition, cognitive or psychosocial aspects, dietary outcomes, clinical biomarkers, smoking, alcohol, sleep, and media use.

Effectiveness

As reported in Table 1 and Multimedia Appendix 5, a total of 54.8% of the interventions were not effective on the considered device-based PA outcomes [78,80-82,85,88-90,92,93,98,99,102,108-113,120-126,128-130,134-138]. The other RCTs showed effectiveness (37.1%) [77,79,83,84,86,94-97,100,101,104-107,114-116,118,127,131-133] or partial effectiveness (8.1%) [87,91,103,117,119] on one or more PA outcomes, meaning that the digital intervention had positive effects on some PA outcomes and not on other PA outcomes; for example, in the intervention of Comeras-Chueca et al [87], the Active video game increased the LPA, but not moderate PA, VPA, or total PA; or in the RCT from Petrušič et al [119], the digital placebo video games contributed to increasing the level of LPA but not of the VPA.

XSL•FO
RenderX

Table . Frequencies (in number and percentage) of the RCTs[a] effectiveness overall and by specific device-based outcome.[b]

Device-based outcomes	Total	Effective	Ineffective
Total RCTs, n (%)	62 (100)	28[c] (45.2)	34 (54.8)
Total outcomes, n (%)	122 (100)	40 (32.8)	82 (67.2)
Steps, n (%)	31 (50)	16 (25.8)	15 (24.2)
MVPA[d], n (%)	38 (61.3)	12 (19.4)	26 (41.9)
LPA[e], n (%)	13 (21)	4 (6.5)	9 (14.5)
MPA[f], n (%)	6 (9.7)	1 (1.6)	5 (8.1)
VPA[g], n (%)	9 (14.5)	1 (1.6)	8 (12.9)
TPA[h], n (%)	7 (11.3)	1 (1.6)	6 (9.7)
LVPA[i], n (%)	1 (1.6)	0 (0)	1 (1.6)
CPM[j], n (%)	12 (19.4)	2 (3.2)	10 (16.1)
Active minutes, n (%)	5 (8.1)	3 (4.8)	2 (3.2)

[a]RCT: randomized controlled trial.

[b]Referred to the 62 randomized controlled trials.

[c]A total of 5 of 28 interventions were partially effective, that is, effectiveness was demonstrated for some physical activity outcomes of the intervention.

[d]MVPA: moderate-to-vigorous physical activity.

[e]LPA: light physical activity.

[f]MPA: moderate physical activity.

[g]VPA: vigorous physical activity.

[h]TPA: total physical activity.

[i]LVPA: light-to-vigorous physical activity.

[j]CPM: counts per minute.

Moreover, some interventions demonstrated stronger effects when implemented over shorter durations or when assessed at short-term follow-up, typically within 3 to 6 months [91,103,118].

When stratifying RCTs' effectiveness by the type of device-based outcomes, a total of 25.8% of the RCTs were effective in increasing the number of steps, while a similar percentage demonstrated no effectiveness (Table 1). The major part of the RCTs assessing MVPA was not effective in this increase; all the other device-based PA outcomes mostly did not increase after the digital intervention (Table 1). Totally, 67.2% of the PA outcomes were not positively affected by the intervention.

Stratified analyses of intervention characteristics (Table 2) showed higher proportions of effective interventions in trials conducted in Asian populations; those targeting PA as the sole primary outcome; interventions using wearables as the digital component or delivery device; trials measuring PA with pedometers; and interventions using specific devices (Tractivity, Omron, Polar Active, and Yamax) rather than ActiGraph or other accelerometers.

XSL•FO
RenderX

Table . Frequencies (in numbers and percentages) of RCTs'[a] effectiveness on the device-based outcomes, by their characteristics.

Group	Total participants (N=62), n	Effective or partially effective on device-based PA[b] outcomes		Ineffective on device-based PA outcomes		*P* value[c]
		Participants, n/N, (%)	Participants wrt[d] the total cohort (ie, n/62×100), %	Participants, n/N, (%)	Participants wrt the total cohort (ie, n/62×100), %	
Age (years)						.26
Younger people (6-13)	39	20/39 (51.3)	32.3	19/39 (48.7)	30.6	
Older people (14-17)	21	8/21 (38.1)	12.9	13/21 (61.9)	21	
Younger people and older people (6-17)	2	0/2 (0)	0	2/2 (100)	3.2	
Continent						.06
America	24	10/24 (41.7)	16.1	14/24 (58.3)	22.6	
Asia	5	5/5 (100)	8.1	0/5 (0)	0	
Europe	17	8/17 (47.1)	12.9	9/17 (52.9)	14.5	
Oceania	16	5/16 (31.3)	8.1	11/16 (68.8)	17.7	
Setting						.09
School or other community	42	22/42 (52.4)	35.5	20/42 (47.6)	32.3	
Home	19	5/19 (26.3)	8.1	14/19 (73.7)	22.6	
Any setting	1	1/1 (100)	1.6	0/1 (0)	0	
Publication date						.92
2010 - 2017	35	16/35 (45.7)	25.8	19/35 (54.3)	30.6	
2018 - 2025	27	12/27 (44.4)	19.4	15/27 (55.6)	24.2	
Special populations						
Weight status						.20
Normal weight	48	25/48 (52.1)	40.3	23/48 (47.9)	37.1	
At risk of obesity	2	0/2 (0)	0	2/2 (100)	3.2	
Normal, overweight, or obese	3	1/3 (33.3)	1.6	2/3 (66.7)	3.2	
Overweight or obese	9	2/9 (22.2)	3.2	7/9 (77.8)	11.3	
Income						.01
Any income	55	28/55 (50.9)	45.2	27/55 (49.1)	43.5	
Low income	7	0/7 (0)	0	7/7 (100)	11.3	
PA level						.06
Sufficiently active	58	28/58 (48.3)	45.2	30/58 (51.7)	48.4	
Insufficiently active	4	0/4 (0)	0	4/4 (100)	6.5	
Intervention duration						.16
3 months or less	36	19/36 (52.8)	30.6	17/36 (47.2)	27.4	
More than 3 months	21	6/21 (28.6)	9.7	15/21 (71.4)	24.2	
N/A[e]	5	3/5 (60)	4.8	2/5 (40)	3.2	

XSL•FO
RenderX

Group	Total participants (N=62), n	Effective or partially effective on device-based PA[b] outcomes		Ineffective on device-based PA outcomes		*P* value[c]
		Participants, n/N, (%)	Participants wrt[d] the total cohort (ie, n/62×100), %	Participants, n/N, (%)	Participants wrt the total cohort (ie, n/62×100), %	
Intervention arms						.54
2-arm	50	23/50 (46)	37.1	27/50 (54)	43.5	
3-arm	10	4/10 (40)	6.5	6/10 (60)	9.7	
4-arm	1	1/1 (100)	1.6	0/1 (0)	0	
N/A	1	0/1 (0)	0	1/1 (100)	1.6	
Theoretical foundation						.16
Yes	37	14/37 (37.8)	22.6	23/37 (62.2)	37.1	
N/A	25	14/25 (56)	22.6	11/25 (44)	17.7	
Focus task						.09
PA only	32	19/32 (59.4)	30.6	13/32 (40.6)	21	
PA and other	17	4/17 (23.5)	6.5	13/17 (76.5)	21	
Obesity, weight, body composition, or energy balance	10	3/10 (30)	4.8	7/10 (70)	11.3	
Other (health or sleep)	3	2/3 (66.7)	3.2	1/3 (33.3)	1.6	
Digital component						.68
App	3	1/3 (33.3)	1.6	2/3 (66.7)	3.2	
Gamification	18	8/18 (44.4)	12.9	10/18 (55.6)	16.1	
Miscellaneous	23	9/23 (39.1)	14.5	14/23 (60.9)	22.6	
Text messaging	2	1/2 (50)	1.6	1/2 (50)	1.6	
Wearable	11	7/11 (63.6)	11.3	4/11 (36.4)	6.5	
Web-based	4	2/4 (50)	3.2	2/4 (50)	3.2	
Nondigital component added						.42
Yes	32	16/32 (50)	25.8	16/32 (50)	25.8	
No	30	12/30 (40)	19.4	18/30 (60)	29	
Digital device for intervention delivery						.37
Computer, laptop, tablet, or iPod touch (Apple Inc)	8	4/8 (50)	6.5	4/8 (50)	6.5	
Console or arcade	12	5/12 (41.7)	8.1	7/12 (58.3)	11.3	
Mobile phone or smartphone	14	7/14 (50)	11.3	7/14 (50)	11.3	
Miscellaneous	10	3/10 (30)	4.8	7/10 (70)	11.3	
Wearable	12	8/12 (66.7)	12.9	4/12 (33.3)	6.5	
N/A	6	1/6 (16.7)	1.6	5/6 (83.3)	8.1	
Device-based and self-report measure						.68
Device-based only	33	17/33 (51.5)	27.4	16/33 (48.5)	25.8	
Device-based and self-reported	29	11/29 (37.9)	17.7	18/29 (62.1)	29	

XSL•FO
RenderX

Group	Total participants (N=62), n	Effective or partially effective on device-based PA[b] outcomes		Ineffective on device-based PA outcomes		*P* value[c]
		Participants, n/N, (%)	Participants wrt[d] the total cohort (ie, n/62×100), %	Participants, n/N, (%)	Participants wrt the total cohort (ie, n/62×100), %	
Device for outcome measurement						.003
Accelerometer only	36	14/36 (38.9)	22.6	22/36 (61.1)	35.5	
Accelerometer and calorimeter	1	1/1 (100)	1.6	0/1 (0)	0	
Accelerometer and exergame in-built device	4	0/4 (0)	0	4/4 (100)	6.5	
Accelerometer and pedometer	4	0/4 (0)	0	4/4 (100)	6.5	
Pedometer only	16	13/16 (81.3)	21	3/16 (18.8)	4.8	
Exergame in-built device	1	0/1 (0)	0	1/1 (100)	1.6	
Brand of device for outcome measurement						.09
ActiGraph accelerometer	26	8/26 (30.8)	12.9	18/26 (69.2)	29	
Polar Active accelerometer	3	2/3 (66.7)	3.2	1/3 (33.3)	1.6	
New Lifestyles (NEW LIFESTYLES, Inc) accelerometer	2	1/2 (50)	1.6	1/2 (50)	1.6	
Fitbit Flex accelerometer	4	0/4 (0)	0	4/4 (100)	6.5	
Tractivity accelerometer	3	3/3 (100)	4.8	0/3 (0)	0	
GENEActiv accelerometer	2	1/2 (50)	1.6	1/2 (50)	1.6	
Yamax pedometer	9	6/9 (66.7)	9.7	3/9 (33.3)	4.8	
Omron pedometer	2	2/2 (100)	3.2	0/2 (0)	0	
Two different devices[f]	4	1/4 (25)	1.6	3/4 (75)	4.8	
Other[g]	7	4/7 (57.1)	6.5	3/7 (42.9)	4.8	
Device-based PA outcome						.21
Steps	31	16/31 (51.6)	25.8	15/31 (48.4)	24.2	
MVPA[h]	38	12/38 (31.6)	19.4	26/38 (68.4)	41.9	
Other PA outcomes[i]	31	11/31 (35.5)	17.7	20/31 (64.5)	32.3	
Risk of bias						.25
Low	25	12/25 (48)	19.4	13/25 (52)	21	
Some concern	30	16/30 (53.3)	25.8	14/30 (46.7)	22.6	

XSL•FO
RenderX

Group	Total participants (N=62), n	Effective or partially effective on device-based PA[b] outcomes		Ineffective on device-based PA outcomes		*P* value[c]
		Participants, n/N, (%)	Participants wrt[d] the total cohort (ie, n/62×100), %	Participants, n/N, (%)	Participants wrt the total cohort (ie, n/62×100), %	
High	7	6/7 (85.7)	9.7	1/7 (14.3)	1.6	

[a]RCT: randomized controlled trial.

[b]PA: physical activity.

[c]Difference between groups estimated through the Pearson chi-square test.

[d]wrt: with respect to.

[e]N/A: not applicable.

[f]GENEActiv accelerometer and Misfit Ray (Fossil Group) accelerometer, ActiGraph accelerometer and New Lifestyles accelerometer (New Lifestyles, Inc), ActiGraph accelerometer and Sony Playstation 2 memory cards, and ActiGraph accelerometer and Omron pedometer.

[g]Walk4Life (Active Together) pedometer, Zamzee accelerometer (Firefly Design LLC), accelerometer (brand not available), MOX (Maastricht Instruments BV) accelerometer, Hoggan Health (Hoggan Scientific, LLC) exergame in-built device, Samsung mobile app Health pedometer (Samsung Electronics Co, Ltd), and movisens (movisens GmbH) accelerometer.

[h]MVPA: moderate-to-vigorous physical activity.

[i]CPM: counts per minute; LPA: light physical activity; LVPA: light-to-vigorous physical activity; VPA: vigorous physical activity; active minutes.

Examining in detail, the results on the digital components for intervention delivery, the gamification tools that were efficacious or partially efficacious on the device-based PA outcomes were the MobileKids Monster Manor installed on iPhone (Apple Inc); Xbox 360 (Microsoft Corp); games of the Nintendo Wii, such as Wii Sports-Tennis, Sonic and Mario at the Olympics or Celebrity Sports Showdown-Horse Racing, Just Dance 3; games from Kinect (Microsoft Corp), such as Kinect Your Shape: Fitness Evolved 2012, Disneyland Adventures and Kinect Sports Season 2, and Kinetic Adventures! Some games from Xbox 360 were efficacious in 2 interventions [104,131], and nonefficacious in 1 intervention from Staiano et al [130]. The game Dance Dance Revolution was efficacious when it was used in the intervention from Maloney et al [112], while in the intervention from Errickson et al [93], it did not show any effect on PA outcomes. The ineffective games were Diab and Nano, as well as active video games such as PlayStation or Sports and Dance Factory, Sport Champions, Move Fitness, Start the Party and Medieval Moves, Dance Star Party and Sorcery, Camp Conquer, a FitQuest smartphone exergame, and Dance Central.

Effective apps in increasing schoolchildren's PA were the Dads and Daughters Exercising and Empowered app [115] and the mobile app MapMyFitness [127]; they both were developed quite recently and targeted to primary and secondary schoolchildren, respectively. The smartphone app "Eat Wisely and Move Happily" used in the Diet, Exercise and Cardiovascular Health–Children [145] was partially effective, while 3 apps were ineffective in increasing PA, such as the immersive app "Zombies Run! 5K training" and the nonimmersive app "Get Running-Couch to 5K" from the AIMFIT (Apps for Improving Fitness) [90], the app from the Active Teen Leaders Avoiding Screen-Time obesity prevention program [110], and from MyMovez project [135,136].

Conversely, higher proportions of ineffective interventions were observed in RCTs conducted in Australia or New Zealand; home-based trials; interventions targeting individuals at risk for obesity, overweight or obese participants, or mixed populations including overweight youth; and trials focused on low-income or insufficiently active children and adolescents. Interventions lasting longer than 3 months and those targeting PA alongside other behaviors or obesity-related outcomes were mostly ineffective (Table 2). App-based and mixed digital interventions also showed limited effectiveness. Trials using multiple delivery devices were largely ineffective, whereas combining accelerometers with exergame or pedometer devices, or using exergame-only devices, was associated with higher effectiveness.

Nearly 70% of the interventions using ActiGraph accelerometers were not effective, all interventions using Fitbit Flex were ineffective, and ineffectiveness increased to 75% when 2 devices were used. Overall, 68.4% of interventions did not increase MVPA, and 64.5% did not improve other PA outcomes. Chi-square analyses showed significant associations for low-income populations (*P*=.01) and outcome measurement device (*P*=.003), with additional trends for country of origin, setting, baseline PA level, intervention focus, and sensor brand (Table 2).

Quality of the RCTs

A total of 11.3% of the RCTs showed a high risk of bias, while the 48.4% had some concern (Multimedia Appendix 6). One of the critical points regards the "deviations from the intended interventions" (33.9% of the interventions) and "measurement of the outcome" (19.4%).

Discussion

Principal Findings

The present umbrella review provides an overview of SRs and meta-analyses that collected RCTs aimed at evaluating the effectiveness of digital approaches to increase PA in healthy children and adolescents. It summarizes and critically evaluates the digital intervention components, as well as the possible determinants and methodological limitations of their effectiveness.

XSL•FO
RenderX

Slightly less than half of the interventions reported either total or partial effectiveness of the digital tools in increasing device-based PA outcomes. A wide variation in study designs, target age groups, intervention objectives, digital devices, and intervention components was identified across the reviews and trials.

Effectiveness over the population target of the present review appears to be influenced by multiple factors, including the focus task of the intervention, the type of digital component used to deliver the intervention and its delivery device, the device used for the objective measurement of PA, some specific population target, and other characteristics such as the geographic region, the intervention setting, and the device brand.

Country

Overview

A first observation concerns the underrepresentation of certain countries in the implementation of the included trials, particularly when considered in relation to the population size of countries with advanced economies across continents. Most interventions were conducted in North America, predominantly in the United States; however, this concentration does not adequately reflect the size and diversity of the North American population, because when compared with the approximately 380 million inhabitants of the United States and Canada [146], the number of interventions remains relatively limited. Around one quarter of the included RCTs were conducted in Europe, mainly in the United Kingdom, while another quarter originated from Australia and New Zealand; the latter proportion appears comparatively high given the much smaller population size of these countries (approximately 31 million) compared with Europe (approximately 744 million) [146]. In contrast, Asian countries with advanced economies were notably underrepresented despite their large combined population (approximately 206 million) [146].

Moreover, the influence of the country on the effectiveness of digital tools is confirmed by the trend observed after the nonparametric association analyses. Several factors may plausibly explain why digital tools appear to be more effective in increasing PA among children in Asian populations. Higher levels of digital literacy and early exposure to technology in many Asian countries (especially in the economically advanced regions) may facilitate greater engagement with apps, wearables, and gamified platforms. In several Asian contexts, structured educational systems, strong acceptance of technology, and high parental support may enhance adherence and implementation fidelity. Additionally, school curricula may better integrate digital health tools into daily routines, increasing consistency of exposure. Urban environments, limited outdoor spaces, and academic pressures may further increase the appeal of digital interventions, helping to explain regional differences in their effectiveness for promoting PA in children. Future research should aim to address this imbalance by promoting RCTs in underrepresented regions and deeply investigating the differences in the effectiveness between the different countries. Expanding the geographical diversity of intervention settings would improve the generalizability of findings and provide a more comprehensive understanding of the effectiveness of digital interventions across different cultural and socioeconomic contexts.

Age and Special Population Groups

This review identified a tendency for RCTs to collect fewer data on high school students compared with those in primary and middle school, suggesting a need for increased research focused on this age group. Furthermore, specific population subgroups, such as students who are overweight or obese, or those from low socioeconomic backgrounds, were rarely included in the reviewed trials. This finding is consistent with the observations reported by Kracht et al [14]. All the interventions conducted to low-income populations were significantly ineffective, which suggests that it would be urgent to plan digital interventions that can positively affect PA also in children and adolescents from low-income families. Similar results were observed for subpopulations of insufficiently active children and adolescents, for whom all digital interventions did not result in effectiveness. Accordingly, future intervention studies should place greater emphasis on the inclusion of these underrepresented subgroups when designing and evaluating digital interventions.

Setting

With regard to the intervention setting, community-based environments, including schools, were substantially more prevalent than home-based settings, and a significant trend of difference in effectiveness for increasing PA was observed between these 2 contexts. As demonstrated in previous studies, social support, together with parental involvement and responsibility in implementing interventions with their children, plays a crucial role in facilitating behavior change [147,148]. Moreover, the ecological model of health and disease highlights how the onset of chronic conditions is influenced by environmental factors, beginning within the family context and extending to schools and the broader community [149]. Accordingly, both school and home environments are essential for the development of interventions aimed at modifying lifestyle behaviors [150,151]. Future research should therefore take greater consideration in conducting RCTs within home-based settings.

Focus Task

PA was a primary focus in most reviews, followed by the weight or obesity management tasks, and this was also reflected in the RCTs extracted by the reviews. Those interventions that had only PA as an outcome were apparently more effective than those focusing primarily on weight or obesity or PA together with other variables. Even though the explanation for this effect is not clear, it could be hypothesized that closer and more specific attention to a topic could improve the intervention procedure to then obtain more effective results. From a theoretical perspective, behavior change frameworks such as the self-determination theory [139] (which has been used in several RCTs of the present overview) suggest that interventions targeting clearly defined mechanisms (eg, motivation, capability, and opportunity) are more likely to be implemented with fidelity and evaluated appropriately. Narrowly focused interventions allow for clearer operationalization of outcomes and improved

XSL•FO
RenderX

alignment between intervention components and measurement strategies. Empirically, previous SRs of digital PA interventions in children and adolescents have shown that interventions with clearly specified objectives and well-defined outcomes tend to report more consistent and interpretable effects compared with broadly defined or multicomponent interventions [66,152,153]. Moreover, methodological guidance for complex interventions (eg, the Medical Research Council framework [154]) emphasizes that precise intervention focus facilitates better study design, implementation, and evaluation.

Theoretical Foundation

This review revealed that studies including a theoretical foundation in the trial procedure seemed to be less effective than those that did not report any theory. This result is not in line with the literature reporting that the behavior change techniques and theoretical frameworks represent the core mechanisms through which interventions exert their effects. Although the most widely adopted standardized taxonomy of behavior change techniques was introduced only in 2013 [14], potentially constraining its use in earlier primary studies and reviews, the RCTs included in the present review that incorporated theoretical frameworks were predominantly conducted and published after this date. This advocates the need for reviewing and specifically addressing target populations with all the behavior change techniques and theories on the basis of digital tools interventions.

Digital Components for the Intervention Delivery

Gamification

The gamification was a digital component frequently investigated in the reviews and a commonly used tool in the retrieved RCTs. This is expected when we consider the target age of the populations under study; gamification tools such as exergames, serious games, or games embedded in devices are commonly used in this age as they are more appealing than other digital intervention tools. Gamification is intended as the use of suggestive digital games and game elements such as points or reward systems in nongame contexts, to influence users' behavior and motivation [155]; gamification targeting PA is specifically designed to involve children in movement-based gameplay, encouraging them to be more active, as many common exergames (also called active video games) do. Recently, the concept of serious games has been of great interest, as they are designed for a primary purpose other than pure entertainment, thus having mainly an educational aim. However, there is often misleading information in the literature about the use of the term "serious game" because some authors define a serious game as any game used for an intervention study; that is, Liu et al [145] in their meta-analysis included different games that showed a significant effect of serious games on the increase of PA in children. In our work, there was no intervention reporting the term serious games, but only "exergames," "role-playing videogames," "active video games," or "arcade machines," which could be considered as serious games because they were designed or used for the purpose of the intervention. The term "serious game" should then be clarified in future publications to shed light on an agreement between the results of the different interventions.

In summary, it was expected that interventions using gamification and apps as digital components could be more effective than other tools in increasing the PA of young populations, but this was not found out in this umbrella review. The explanation is not so easy because different factors can influence the effectiveness of digital interventions, but it would be useful to explore in detail the single intervention, which is not the purpose of this review. Our recommendation, therefore, is that future studies should plan interventions using accurate gamification tools and apps that could be more effective for schoolchildren's PA.

Wearables

The major part of interventions testing a wearable as a unique delivery tool and component reported effective results on increasing PA. These results are in line with the literature stating that wearable technologies have emerged as important digital tools for promoting PA among children and adolescents [156-158]. By providing real-time feedback, goal setting, and self-monitoring features, wearables can enhance awareness of movement behaviors and motivate users to increase daily activity levels. In younger populations, these devices support the development of self-regulation skills and reinforce positive behavior through immediate reinforcement and gamified elements. Moreover, wearable-based interventions can be integrated with mobile apps and social features, facilitating parental involvement, peer support, and structured activity challenges. Importantly, wearables also enable objective assessment of activity patterns, improving the accuracy of outcome measurement in intervention studies. Collectively, these features position wearable technologies as effective and scalable tools for supporting PA promotion in pediatric populations. It is suggested that these devices should be prioritized in future RCTs as intervention delivery tools.

Web-Based Tools, Apps, and Text Messages

Very few interventions examined the effects of web-based tools, mobile apps, and text messaging as unique components of the interventions, as these tools were generally included in a multicomponent intervention. Similar proportions of effective and ineffective outcomes between these 3 tools were observed in this review. In particular, RCTs that used web-based components were mainly addressed to older children attending middle or high schools, in contrast with gamification approaches that were targeted principally to schoolchildren from elementary schools, which is consistent with findings from the international literature [159-162]. Overall, these digital approaches demonstrate potential but yield inconsistent effects on PA promotion in schoolchildren [163].

Nondigital Components

Nondigital components were used in slightly more than half of the interventions, and they were mostly goals, rewards, and incentives. Previous works have shown that multicomponent interventions, including nondigital components such as counseling for children and parents, class seminars, and health promotion curricula, could have a positive impact on PA outcomes [159]. Therefore, different communication channels should be combined with digital tools because they are based on the possible behavior change [164].

XSL•FO
RenderX

Digital Device for Intervention Delivery and for PA Outcome Assessment

Most SRs or meta-analyses and RCTs included studies that used smartphones as devices for the intervention delivery, and this is consistent with the fact that they are the most common devices used to access the internet games, apps, or other digital components [165]. Effectiveness was not likely to be affected by the specific device. The same could be stated for the use of devices such as console or arcade machines, computers, tablets, or iPod touch in relation to the increase in PA outcomes. The use of a miscellaneous array of devices, instead, revealed a major proportion of ineffective interventions.

Other frequently used devices were the wearables, in most cases overlapping with the wearable that was used as a component of the digital intervention delivered. The main used wearables were accelerometers and pedometers, such as ActiGraph, Polar Active, or Tractivity, all tools that have been widely validated and broadly used [166-168]. Additionally, Yamax or Omron pedometers were mainly used for the step count measurement. The same type of device was used within the intervention for the assessment of PA outcomes. Greater significant effectiveness was found for the use of pedometers, compared for example to the accelerometers. Pedometers may be more effective than accelerometers in increasing PA among children and adolescents because they provide simple, easily interpretable feedback, such as step counts, that is readily understood by young users. This immediacy can enhance motivation and goal setting, particularly when daily step targets are used. Pedometers are also less complex and often more user-friendly, which may promote consistent wear and engagement. In contrast, accelerometers are typically designed primarily for measurement rather than behavior change and often provide limited or delayed feedback to participants [67,77,169]. The visibility of step counts can encourage self-monitoring and reinforce active behaviors throughout the day. As a result, pedometers may better support behavior change processes in youth populations.

Device-Based PA Outcomes

Some considerations have to be made about the specific outcomes that were assessed through the devices. Step count was positively affected by the digital intervention in half of the considered interventions, while MVPA or other PA outcomes, such as light-to-vigorous PA, LPA, MVPA, VPA, counts per minute, and active minutes, did not increase in the majority of the interventions. The failure in increasing step or other PA outcomes could be explained by supposing that engagement with the technology declined over time, reducing sustained behavior change. Some interventions relied primarily on monitoring without incorporating sufficient behavior change techniques such as goal setting, feedback, or rewards.

Limited integration of social support from parents, peers, or schools may also have weakened intervention effects [170]. In addition, interventions of short duration may have been insufficient to produce measurable increases in habitual PA, or those that were too long may have led to a loss of interest [171]. Technical issues, poor usability, or lack of age-appropriate design could have further limited children's and adolescents' motivation to increase daily steps and other PA outcomes [44]. Finally, measurement challenges and high baseline MVPA levels in some samples may have limited the capacity to observe significant increases [172].

Overlapping

The observed slight degree of overlap indicates that most SRs and meta-analyses include different primary studies. This statement supports more robust and balanced conclusions in this umbrella review, as the risk of double-counting evidence is reduced, and the independence of findings across reviews is strengthened. Although a slight overlap enhances the credibility of the overall synthesis by ensuring that conclusions are not driven by a small number of repeated trials, it may also increase heterogeneity, as findings are drawn from diverse interventions and populations [173]. Interpretation, therefore, requires careful consideration of contextual differences.

Quality of the Studies

The critically low rating of the major part of the SRs and meta-analyses was frequently due to the omission of a key methodological requirement, specifically the reporting of excluded reviews and the reasons for their exclusion. The lack of transparency regarding excluded reviews could limit the completeness of the evidence base and the potential for selection bias.

With regard to the quality assessment of the RCTs, the domain of the RoB 2 tool "deviations from the intended interventions" was particularly critical. Digital interventions strongly depend on user engagement, adherence, and correct use of the technology, which are often inconsistent in younger populations. Children and adolescents may discontinue use, use the tools incorrectly, or engage at varying intensities. When participants deviate from the assigned intervention, the estimated intervention effect may be biased, especially if these deviations differ systematically between IG and CGs [174]. When not properly addressed analytically, they may lead to overestimation or underestimation of intervention effects or heterogeneity in reported outcomes.

Limitations

A limitation of the present umbrella review is that several included SRs and meta-analyses were of critically low methodological quality; however, excluding all such reviews would have resulted in the omission of a substantial proportion of relevant evidence. This implies a potential and substantial influence on the interpretation of evidence of the digital tools' effectiveness on increasing PA in children and adolescents. For example, methodological weaknesses, such as unclear inclusion criteria, incomplete search strategies, poor transparency in data extraction and synthesis, may reduce reproducibility and confidence in findings; inadequate assessment of risk of bias can result in overestimation or underestimation of intervention effects; insufficiently explored heterogeneity may mask important differences between intervention types or populations; inconsistent outcome definitions and measurement methods further limit comparability across reviews [175]. Therefore, the findings of this umbrella review should be interpreted with caution.

XSL•FO
RenderX

Similarly, RCTs assessed as having a high risk of bias constituted a substantial portion of the evidence base; therefore, excluding them would have been detrimental to the review. This could have influenced the results of this umbrella review, in the sense of overestimating the effectiveness of interventions, masking real differences between types of digital tools, such as mobile apps, wearable devices, or exergames. Common sources of bias, including inadequate randomization, lack of blinding, and high attrition, can systematically skew outcomes. Overall, the presence of high-risk trials weakens the ability to grade the certainty of evidence and may ultimately lead to inappropriate or poorly informed recommendations for practice and policy. Moreover, the lack of certainty of assessment (eg, through the GRADE [Grading of Recommendations, Assessment, Development, and Evaluation] tool) was a deliberate choice; it is a common and currently recognized practice in the field that a significant portion of published umbrella reviews do not assess the certainty of evidence, particularly when authors focus instead on the methodological quality of the included reviews using the AMSTAR 2 tool [176,177].

Another limit of this review is that it did not quantitatively synthesize through a meta-analytical approach (meta-meta-analyses or meta-analysis of the RCTs) the overall effectiveness, the heterogeneity, and the effects obtained after stratification and sensitivity analyses. Digital tools such as apps, wearables, exergames, and gamified platforms often differ in content, intensity, duration, and implementation context, making direct comparisons challenging. This variability limits the ability to synthesize findings quantitatively and reduces confidence in pooled effect estimates. In this overview, heterogeneity was explored through qualitative methods combined with narrative synthesis, and this approach helped in the identification of patterns and potential moderators even when quantitative pooling is not possible. Heterogeneity may obscure true intervention effects, as positive outcomes observed in some interventions may not be generalizable across different settings or age groups [178]. It also complicates the identification of which specific digital components or delivery strategies are most effective for schoolchildren. Consequently, conclusions must be framed more descriptively, and recommendations for practice and policy should emphasize context-specific effectiveness rather than universal applicability. Therefore, the authors envisage that the next step is to conduct a meta-analysis of RCTs to more robustly assess the effects of digital approaches on increases in PA. Such analyses should be complemented by sensitivity analyses, for example, excluding interventions of lower methodological quality, to enhance the robustness and interpretability of the findings.

As some of the included RCTs were conducted during the COVID-19 pandemic (2020‑2022) or in the postpandemic period, changes in PA patterns and digital tool use during these years may have influenced the observed outcomes [179], and this should be considered for the interpretation of results.

Strengths

This umbrella review was based on rigorous and systematic methodology, relying on reviews that collected RCTs and that used objective device-based measures of PA. These measures improve the reliability of the review by providing more accurate and unbiased outcome assessment than self-reported data. Moreover, they reduce recall and social desirability bias, enhance comparability across studies, and allow more precise detection of changes in activity levels. As a result, they strengthen the quality of the evidence base, reduce heterogeneity related to outcome measurement, and increase confidence in conclusions regarding the effectiveness of digital interventions [180,181].

Risk of bias was evaluated at both the review and primary study levels, given that only a small subset of interventions from each review met the inclusion criteria for inclusion in this umbrella review. As a result, assessing risk of bias exclusively at the review level may have led to misleading conclusions, because the original reviews incorporated a broader evidence base.

Another strength of this review is its innovative character, as it is among the first umbrella reviews to comprehensively synthesize evidence from recent SRs and meta-analyses on digital PA interventions in children and adolescents. It provides novel comparative insights into methodological features and population subgroups that may drive or limit the success of digital approaches.

The findings of this work have several important real-world implications and contributions. They provide evidence-based guidance for practitioners, schools, and policymakers on which digital tools are more likely to effectively increase PA in children and adolescents, supporting more informed decision-making. The results highlight the importance of selecting appropriate devices, outcome measures, and intervention designs to maximize impact in real-world settings. This work also identifies gaps in equity and population coverage, informing future interventions aimed at reducing disparities in PA. Finally, the findings contribute to the optimization and scaling of digital health interventions by clarifying which approaches are most promising for practical implementation.

Recommendations for Future Research and Implications

While the authors envisage quantitatively analyzing the effect of the interventions on the specific PA outcomes in children and adolescents in a future study through meta-analytic procedures, they wish to guide the strengthening of rigor for future SRs and interventions and support greater engagement in PA during childhood and adolescence.

First of all, the selection of device-based measures, which are able to provide reliable and comparable data, should be stressed. A clarification on the way to use gamification should be provided, together with a clearer explanation of the term “serious game” when it is used within the interventions. The importance of incorporating gamification tools, wearables, and/or apps as integral components of interventions should be highlighted, as they may be particularly effective for school-aged children [74,163].

Future interventions should place greater emphasis on conducting RCTs within home-based settings [150]. PA management should be the primary aim of the intervention; limiting the focus of the RCT to PA, without going too far into

XSL•FO
RenderX

other outcomes that may cause a loss of the focus, could help in obtaining accurate data and effective outcomes [182]. There is a need to intensify RCTs targeted at high school populations, also considering that this age range is particularly engaged with digital technologies. Sustaining engagement over time remains a key challenge [183], highlighting the importance of developing and evaluating strategies that support long-term participation and adherence to digital interventions.

Better planning of interventions targeted to low-income and to insufficiently active subpopulations of children and adolescents represents a key aspect [184]. Future research should aim to promote and support RCTs in underrepresented regions, particularly in Asian and other non-Western high-income countries.

A greater emphasis on methodological transparency throughout the review and intervention selection process is suggested; in particular, clear and explicit criteria for study inclusion and exclusion should be reported. Transparency is needed in quality assessment procedures, including the rationale for quality scoring and study weighting; in addition, future reviews should clearly describe how methodological quality influences the interpretation of findings. The need for systematic evaluation and deliberate selection of behavior change techniques and theoretical models that are explicitly tailored to specific target populations within digital intervention contexts is underscored [185].

The results of this umbrella review can suggest practical actions both to schools and policymakers. Schools should avoid embedding PA as a minor component within broad well-being or educational programs; can integrate wearable-based programs (eg, step challenges and activity feedback) into physical education or daily routines; should favor pedometers or basic wearables over self-report tools to monitor student activity; and should accurately choose the type of device when implementing digital PA programs. Schools in disadvantaged areas may require additional support, such as teacher training, parental engagement strategies, or longer intervention duration.

Policymakers should prioritize funding and endorsement of interventions where PA improvement is explicitly defined, measured, and prioritized; can justify investments in wearable technologies as scalable and effective tools for youth PA promotion; can recommend objective PA monitoring standards in school-based programs to improve evaluation quality and accountability; can develop guidelines specifying preferred assessment tools to improve program effectiveness and comparability; should avoid one-size-fits-all approaches and allocate additional resources to adapt digital interventions to low-income contexts.

Conclusions

This innovative study represents one of the first umbrella reviews describing digital PA interventions in youth, determinants of effectiveness, and methodological gaps. It provides evidence-based guidance for practitioners, schools, and policymakers on which features of the interventions are more likely to effectively increase PA in children and adolescents, supporting more informed decision-making. Interventions appear to be most effective when they prioritize focused PA objectives, incorporate wearables as both the core digital intervention component and the delivery device, use pedometers to measure PA outcomes, and choose accurately the device to assess PA as a quality criterion in program selection.

All these adaptations in future studies could be of great importance to generating high-quality primary RCTs and provide consistent conclusions about the usefulness of digital approaches to increase PA. The findings of this umbrella review have important implications for public health practice and policy. Digital PA interventions represent a scalable, cost-effective approach to addressing low PA levels among children and adolescents, a key modifiable risk factor for noncommunicable diseases across the life course. By identifying which types of digital tools show the greatest potential and which design and contextual factors influence effectiveness, this review supports more strategic investment in evidence-based digital solutions within schools, communities, and health care systems. The results also underscore the need for public health stakeholders to move beyond one-size-fits-all approaches and to ensure that digital interventions are developmentally appropriate, engaging, and equitably accessible, particularly for underserved populations. Strengthening methodological rigor in future research will improve the reliability of evidence used to guide policy decisions and large-scale implementation.

Collectively, these insights can help inform integrated public health strategies that leverage digital innovation to promote lifelong PA habits, reduce health inequalities among youth, and the burden of chronic diseases in adulthood.

Acknowledgments

The authors declare the use of generative artificial intelligence (GenAI) in the research and writing process. According to the GAIDeT (Generative Artificial Intelligence Delegation Taxonomy, 2025), the following tasks were delegated to GenAI tools under full human supervision: (1) proofreading and editing, (2) summarizing text, and (3) translation. The GenAI tool used was ChatGPT-3.5 (OpenAI). The responsibility for the final paper lies entirely with the authors. GenAI tools are not listed as authors and do not bear responsibility for the outcomes. This declaration was submitted by collective responsibility.

Funding

This research was cofunded by the Italian Complementary National Plan PNC-I.1 "Research Initiatives for Innovative Technologies and Pathways in the Health and Welfare Sector" Decreto Direttoriale 931 of 06/06/2022, "DARE - Digital Lifelong Prevention"

XSL•FO
RenderX

initiative (PNC0000002, CUP: B53C22006450001). The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

The following data will be available per request: data analysis forms of systematic reviews or meta-analyses, and data analysis forms of randomized controlled trials.

Authors' Contributions

GT made substantial contributions to conception, design, data acquisition, analysis, and interpretation; AS, MG, and MF supported the data acquisition; AA, GA, GG, SV, and VDL contributed to the methodological improvement of this paper; PB, WM, and MB supervised this paper; and WM secured financial support.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strings.
[DOCX File, 17 KB - publichealth_v12i1e75769_app1.docx]

Multimedia Appendix 2
Characteristics of the retrieved systematic reviews or meta-analyses.
[DOCX File, 38 KB - publichealth_v12i1e75769_app2.docx]

Multimedia Appendix 3
Figures in percentage frequencies.
[DOCX File, 548 KB - publichealth_v12i1e75769_app3.docx]

Multimedia Appendix 4
AMSTAR 2 quality rating table of the selected systematic reviews and meta-analyses. AMSTAR 2: A Measurement Tool to Assess Systematic Reviews.
[DOCX File, 29 KB - publichealth_v12i1e75769_app4.docx]

Multimedia Appendix 5
Characteristics of the included randomized controlled trials.
[DOCX File, 60 KB - publichealth_v12i1e75769_app5.docx]

Multimedia Appendix 6
Risk of bias of the selected randomized controlled trials.
[PDF File, 480 KB - publichealth_v12i1e75769_app6.pdf]

Checklist 1
PRISMA checklist.
[PDF File, 138 KB - publichealth_v12i1e75769_app7.pdf]

Checklist 2
PRIOR checklist.
[PDF File, 1068 KB - publichealth_v12i1e75769_app8.pdf]

References

1. Strain T, Flaxman S, Guthold R, et al. National, regional, and global trends in insufficient physical activity among adults from 2000 to 2022: a pooled analysis of 507 population-based surveys with 5·7 million participants. Lancet Glob Health 2024 Aug;12(8):e1232-e1243. [doi: 10.1016/S2214-109X(24)00150-5] [Medline: 38942042]
2. Guthold R, Stevens GA, Riley LM, Bull FC. Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1·6 million participants. Lancet Child Adolesc Health 2020 Jan;4(1):23-35. [doi: 10.1016/S2352-4642(19)30323-2] [Medline: 31761562]

XSL•FO
RenderX

3. Global action plan on physical activity 2018–2030: more active people for a healthier world. : World Health Organization; 2018 URL: https://iris.who.int/handle/10665/272722 [accessed 2026-02-25]
4. Wright M, Reitegger F, Cela H, Papst A, Gasteiger-Klicpera B. Interventions with digital tools for mental health promotion among 11-18 year olds: a systematic review and meta-analysis. J Youth Adolesc 2023 Apr;52(4):754-779. [doi: 10.1007/s10964-023-01735-4] [Medline: 36754917]
5. How's life for children in the digital age. : OECD Publishing; 2025 URL: https://www.oecd.org/en/publications/how-s-life-for-children-in-the-digital-age_0854b900-en.html [accessed 2026-02-25]
6. Children in a digital world. : UNICEF; 2017 URL: https://www.unicef.org/media/48581/file/SOWC_2017_ENG.pdf [accessed 2026-02-25]
7. European citizenship – report. : European Commission; 2021 URL: https://data.europa.eu/doi/10.2775/731315 [accessed 2026-02-26]
8. De Santis KK, Jahnel T, Matthias K, Mergenthal L, Al Khayyal H, Zeeb H. Evaluation of digital interventions for physical activity promotion: scoping review. JMIR Public Health Surveill 2022 May 23;8(5):e37820. [doi: 10.2196/37820] [Medline: 35604757]
9. McCallum C, Rooksby J, Gray CM. Evaluating the impact of physical activity apps and wearables: interdisciplinary review. JMIR mHealth uHealth 2018 Mar 23;6(3):e58. [doi: 10.2196/mhealth.9054] [Medline: 29572200]
10. Youth Report 2024: Technology in Education: A Tool on Our Terms!: United Nations Educational, Scientific and Cultural Organization (UNESCO); 2024. [doi: 10.54676/ELSQ4648]
11. Carraça E, Encantado J, Battista F, et al. Effective behavior change techniques to promote physical activity in adults with overweight or obesity: a systematic review and meta-analysis. Obes Rev 2021 Jul;22 Suppl 4(Suppl 4):e13258. [doi: 10.1111/obr.13258] [Medline: 33949778]
12. Domin A, Spruijt-Metz D, Theisen D, Ouzzahra Y, Vögele C. Smartphone-based interventions for physical activity promotion: scoping review of the evidence over the last 10 years. JMIR mHealth uHealth 2021 Jul 21;9(7):e24308. [doi: 10.2196/24308] [Medline: 34287209]
13. Ghanvatkar S, Kankanhalli A, Rajan V. User models for personalized physical activity interventions: scoping review. JMIR mHealth uHealth 2019 Jan 16;7(1):e11098. [doi: 10.2196/11098] [Medline: 30664474]
14. Kracht CL, Hutchesson M, Ahmed M, et al. E-&mHealth interventions targeting nutrition, physical activity, sedentary behavior, and/or obesity among children: a scoping review of systematic reviews and meta-analyses. Obes Rev 2021 Dec;22(12):e13331. [doi: 10.1111/obr.13331] [Medline: 34476890]
15. Gates M, Gates A, Pieper D, et al. Reporting guideline for overviews of reviews of healthcare interventions: development of the PRIOR statement. BMJ 2022 Aug 9;378:e070849. [doi: 10.1136/bmj-2022-070849] [Medline: 35944924]
16. Rethlefsen ML, Kirtley S, Waffenschmidt S, et al. PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. Syst Rev 2021 Jan 26;10(1):39. [doi: 10.1186/s13643-020-01542-z] [Medline: 33499930]
17. Biondi-Zoccai G, editor. Umbrella Reviews: Evidence Synthesis with Overviews of Reviews and Meta-Epidemiologic Studies, 1st edition: Springer International Publishing : Imprint: Springer; 2016. [doi: 10.1007/978-3-319-25655-9]
18. Fusar-Poli P, Radua J. Ten simple rules for conducting umbrella reviews. Evid Based Ment Health 2018 Aug;21(3):95-100. [doi: 10.1136/ebmental-2018-300014] [Medline: 30006442]
19. Müller AM, Maher CA, Vandelanotte C, et al. Physical activity, sedentary behavior, and diet-related eHealth and mHealth research: bibliometric analysis. J Med Internet Res 2018;20(4):e122. [doi: 10.2196/jmir.8954]
20. Pieper D, Antoine SL, Mathes T, Neugebauer EAM, Eikermann M. Systematic review finds overlapping reviews were not mentioned in every other overview. J Clin Epidemiol 2014 Apr;67(4):368-375. [doi: 10.1016/j.jclinepi.2013.11.007]
21. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 2017 Sep 21;358:j4008. [doi: 10.1136/bmj.j4008] [Medline: 28935701]
22. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug 28;366:l4898. [doi: 10.1136/bmj.l4898] [Medline: 31462531]
23. Baumann H, Fiedler J, Wunsch K, Woll A, Wollesen B. mHealth interventions to reduce physical inactivity and sedentary behavior in children and adolescents: systematic review and meta-analysis of randomized controlled trials. JMIR mHealth uHealth 2022 May 11;10(5):e35920. [doi: 10.2196/35920] [Medline: 35544294]
24. Böhm B, Karwiese SD, Böhm H, Oberhoffer R. Effects of mobile health including wearable activity trackers to increase physical activity outcomes among healthy children and adolescents: systematic review. JMIR mHealth uHealth 2019 Apr 30;7(4):e8298. [doi: 10.2196/mhealth.8298] [Medline: 31038460]
25. Bonvicini L, Pingani I, Venturelli F, et al. Effectiveness of mobile health interventions targeting parents to prevent and treat childhood obesity: systematic review. Prev Med Rep 2022 Oct;29:101940. [doi: 10.1016/j.pmedr.2022.101940] [Medline: 36161123]
26. Casado-Robles C, Viciana J, Guijarro-Romero S, Mayorga-Vega D. Effects of consumer-wearable activity tracker-based programs on objectively measured daily physical activity and sedentary behavior among school-aged children: a systematic review and meta-analysis. Sports Med Open 2022 Jan 31;8(1):18. [doi: 10.1186/s40798-021-00407-6] [Medline: 35099630]

XSL•FO
RenderX

27. Champion KE, Parmenter B, McGowan C, et al. Effectiveness of school-based eHealth interventions to prevent multiple lifestyle risk behaviours among adolescents: a systematic review and meta-analysis. Lancet Digit Health 2019 Sep;1(5):e206-e221. [doi: 10.1016/S2589-7500(19)30088-3] [Medline: 33323269]
28. Creaser AV, Clemes SA, Costa S, et al. The acceptability, feasibility, and effectiveness of wearable activity trackers for increasing physical activity in children and adolescents: a systematic review. Int J Environ Res Public Health 2021 Jun 8;18(12):6211. [doi: 10.3390/ijerph18126211] [Medline: 34201248]
29. Dobbie LJ, Tahrani A, Alam U, James J, Wilding J, Cuthbertson DJ. Exercise in obesity-the role of technology in health services: can this approach work? Curr Obes Rep 2022 Sep;11(3):93-106. [doi: 10.1007/s13679-021-00461-x] [Medline: 34791611]
30. Emberson MA, Lalande A, Wang D, McDonough DJ, Liu W, Gao Z. Effectiveness of smartphone-based physical activity interventions on individuals' health outcomes: a systematic review. In: Abreu M, editor. Biomed Res Int 2021;2021:6296896. [doi: 10.1155/2021/6296896] [Medline: 34409104]
31. He Z, Wu H, Yu F, et al. Effects of smartphone-based interventions on physical activity in children and adolescents: systematic review and meta-analysis. JMIR mHealth uHealth 2021 Feb 1;9(2):e22601. [doi: 10.2196/22601] [Medline: 33522980]
32. Jacob CM, Hardy-Johnson PL, Inskip HM, et al. A systematic review and meta-analysis of school-based interventions with health education to reduce body mass index in adolescents aged 10 to 19 years. Int J Behav Nutr Phys Act 2021 Jan 4;18(1):1. [doi: 10.1186/s12966-020-01065-9] [Medline: 33397403]
33. Klos L, Feil K, Eberhardt T, Jekauc D. Interventions to promote positive affect and physical activity in children, adolescents and young adults-a systematic review. Sports (Basel) 2020 Feb 20;8(2):26. [doi: 10.3390/sports8020026] [Medline: 32093347]
34. Lam C, Milne-Ives M, Harrington R, et al. Internet of things-enabled technologies as an intervention for childhood obesity: a systematic review. In: Mulvaney S, editor. PLOS Digit Health 2022 Apr;1(4):e0000024. [doi: 10.1371/journal.pdig.0000024] [Medline: 36812526]
35. Lamas S, Rebelo S, da Costa S, Sousa H, Zagalo N, Pinto E. The influence of serious games in the promotion of healthy diet and physical activity health: a systematic review. Nutrients 2023 Mar 14;15(6):1399. [doi: 10.3390/nu15061399] [Medline: 36986129]
36. Langarizadeh M, Sadeghi M, As'habi A, Rahmati P, Sheikhtaheri A. Mobile apps for weight management in children and adolescents; an updated systematic review. Patient Educ Couns 2021 Sep;104(9):2181-2188. [doi: 10.1016/j.pec.2021.01.035] [Medline: 33573915]
37. Lee YS, Chia M, Komar J. A systematic review of physical activity intervention programs in ASEAN countries: efficacy and future directions. Int J Environ Res Public Health 2022 Apr 28;19(9):5357. [doi: 10.3390/ijerph19095357] [Medline: 35564751]
38. Longobucco Y, Ricci M, Scrimaglia S, Camedda C, Dallolio L, Masini A. Effects of school nurse-led interventions in collaboration with kinesiologists in promoting physical activity and reducing sedentary behaviors in children and adolescents: a systematic review. Healthcare (Basel) 2023 May 26;11(11):1567. [doi: 10.3390/healthcare11111567] [Medline: 37297707]
39. Ludwig K, Arthur R, Sculthorpe N, Fountain H, Buchan DS. Text messaging interventions for improvement in physical activity and sedentary behavior in youth: systematic review. JMIR mHealth uHealth 2018 Sep 17;6(9):e10799. [doi: 10.2196/10799] [Medline: 30224335]
40. Mazeas A, Duclos M, Pereira B, Chalabaev A. Evaluating the effectiveness of gamification on physical activity: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2022 Jan 4;24(1):e26779. [doi: 10.2196/26779] [Medline: 34982715]
41. Nash EA, Critchley JA, Pearson F, et al. A systematic review of interventions to promote physical activity in six Gulf countries. In: Palazón-Bru A, editor. PLOS ONE 2021;16(10):e0259058. [doi: 10.1371/journal.pone.0259058]
42. Oliveira CB, Pinto RZ, Saraiva BTC, et al. Effects of active video games on children and adolescents: a systematic review with meta‐analysis. Scand Med Sci Sports 2020 Jan;30(1):4-12. [doi: 10.1111/sms.13539]
43. Rodrigo-Sanjoaquín J, Corral-Abós A, Aibar Solana A, Zaragoza Casterad J, Lhuisset L, Bois JE. Effectiveness of school-based interventions targeting physical activity and sedentary time among children: a systematic review and meta-analysis of accelerometer-assessed controlled trials. Public Health (Fairfax) 2022 Dec;213:147-156. [doi: 10.1016/j.puhe.2022.10.004]
44. Schwarz A, Winkens LHH, de Vet E, Ossendrijver D, Bouwsema K, Simons M. Design features associated with engagement in mobile health physical activity interventions among youth: systematic review of qualitative and quantitative studies. JMIR mHealth uHealth 2023;11:e40898. [doi: 10.2196/40898]
45. Seims AL, Hall J, Bingham DD, et al. Interventions targeting children and young people's physical activity behavior at home: a systematic review. In: Lee KY, editor. PLOS ONE 2023;18(8):e0289831. [doi: 10.1371/journal.pone.0289831] [Medline: 37556477]
46. Shin Y, Kim SK, Lee M. Mobile phone interventions to improve adolescents' physical health: a systematic review and meta-analysis. Public Health Nurs 2019 Nov;36(6):787-799. [doi: 10.1111/phn.12655] [Medline: 31456259]

XSL•FO
RenderX

47. Stecher C, Pfisterer B, Harden SM, et al. Assessing the pragmatic nature of mobile health interventions promoting physical activity: systematic review and meta-analysis. JMIR mHealth uHealth 2023 May 4;11:e43162. [doi: 10.2196/43162] [Medline: 37140972]
48. Williams WM, Ayres CG. Can active video games improve physical activity in adolescents? A review of RCT. Int J Environ Res Public Health 2020 Jan 20;17(2):669. [doi: 10.3390/ijerph17020669] [Medline: 31968642]
49. Xu L, Shi H, Shen M, et al. The effects of mHealth-based gamification interventions on participation in physical activity: systematic review. JMIR mHealth uHealth 2022;10(2):e27794. [doi: 10.2196/27794]
50. Yang Y, Hu H, Koenigstorfer J. Effects of gamified smartphone applications on physical activity: a systematic review and meta-analysis. Am J Prev Med 2022 Apr;62(4):602-613. [doi: 10.1016/j.amepre.2021.10.005]
51. Yau KW, Tang TS, Görges M, et al. Effectiveness of mobile apps in promoting healthy behavior changes and preventing obesity in children: systematic review. JMIR Pediatr Parent 2022 Mar 28;5(1):e34967. [doi: 10.2196/34967] [Medline: 35343908]
52. Au WW, Recchia F, Fong DY, et al. Effect of wearable activity trackers on physical activity in children and adolescents: a systematic review and meta-analysis. Lancet Digit Health 2024 Sep;6(9):e625-e639. [doi: 10.1016/S2589-7500(24)00139-0] [Medline: 39112110]
53. Chen J, Hu RZ, Zhuang YX, et al. Natural language processing chatbot-based interventions for improvement of diet, physical activity, and tobacco smoking behaviors: systematic review. JMIR mHealth uHealth 2025 Jun 11;13:e66403. [doi: 10.2196/66403] [Medline: 40503914]
54. Chen X, Wang F, Zhang H, Lin Y, Zhu S, Yang Y. Effectiveness of wearable activity trackers on physical activity among adolescents in school-based settings: a systematic review and meta-analysis. BMC Public Health 2025;25(1):1050. [doi: 10.1186/s12889-025-22170-z]
55. De Luca V, Virgolesi M, Vetrani C, et al. Digital interventions for weight control to prevent obesity in adolescents: a systematic review. Front Public Health 2025;13:1584595. [doi: 10.3389/fpubh.2025.1584595] [Medline: 40487516]
56. França C, Santos F, Martins F, et al. Digital health in schools: a systematic review. Sustainability 2022;14(21):13848. [doi: 10.3390/su142113848]
57. Ha T, Moon J, Yu H, Fan X, Paulson L. A systematic review of technology-infused physical activity interventions in K-12 school settings: effectiveness, roles, and implementation strategies. Int J Behav Nutr Phys Act 2025 Aug 23;22(1):113. [doi: 10.1186/s12966-025-01811-x] [Medline: 40849486]
58. Kassim PSJ, Muhammad NA, Rahman NFA, Sidik SM, Essau CA, Shah SA. Effectiveness of digital behaviour change interventions for promoting physical activity in adolescents with overweight and obesity: a systematic review and meta-analysis. Digit Health 2025 Jan;11:20552076251314904. [doi: 10.1177/20552076251314904]
59. Kemp BJ, Thompson DR, Watson CJ, McGuigan K, Woodside JV, Ski CF. Effectiveness of family-based eHealth interventions in cardiovascular disease risk reduction: a systematic review. Prev Med 2021 Aug;149:106608. [doi: 10.1016/j.ypmed.2021.106608] [Medline: 33984372]
60. Li SJ, Ma HM, Zhu LQ, et al. Exergame-based behavior change interventions for promoting physical activity: systematic review and meta-analysis of randomized controlled studies. J Med Internet Res 2025 Aug 8;27:e62906. [doi: 10.2196/62906] [Medline: 40779747]
61. Love R, Adams J, van Sluijs EMF. Are school-based physical activity interventions effective and equitable? A meta-analysis of cluster randomized controlled trials with accelerometer-assessed activity. Obes Rev 2019 Jun;20(6):859-870. [doi: 10.1111/obr.12823] [Medline: 30628172]
62. Na A, Na A. A systematic review on the intervention using exercise gaming to reduce childhood obesity. IMJM 2019;18(2). [doi: 10.31436/imjm.v18i2.78]
63. Sequí-Domínguez I, Cavero-Redondo I, Álvarez-Bueno C, López-Gil JF, Martínez-Vizcaíno V, Pascual-Morena C. Effectiveness of eHealth interventions promoting physical activity in children and adolescents: systematic review and meta-analysis. J Med Internet Res 2024 Feb 21;26:e41649. [doi: 10.2196/41649] [Medline: 38381490]
64. Figueroa VAS, Martin ARSS, Hernández PAF, et al. Exergames para mejorar la condición física y la composición corporal en la niñez y adolescencia: una revisión sistemática [Article in Spanish]. Retos 2025;67:917-928. [doi: 10.47197/retos.v67.110803]
65. Spring FDH, Lærkholm G, Jensen RB, Kloppenborg JT. The effect of exergaming on BMI and fitness in children and adolescents with obesity: a systematic review. Acta Paediatr 2025 Jul;114(7):1522-1537. [doi: 10.1111/apa.70048] [Medline: 40035397]
66. Wang JW, Zhu Z, Shuling Z, et al. Effectiveness of mHealth app-based interventions for increasing physical activity and improving physical fitness in children and adolescents: systematic review and meta-analysis. JMIR mHealth uHealth 2024 Apr 30;12:e51478. [doi: 10.2196/51478] [Medline: 38687568]
67. Wang W, Huang C, Shen Y, Cheng J, Wang L. Effectiveness of step-count monitoring interventions in increasing physical activity among children and adolescents: a systematic review and meta-analysis. Digit Health 2025;11:20552076251374249. [doi: 10.1177/20552076251374249] [Medline: 40918079]

XSL•FO
RenderX

68. Wang M, Xu J, Zhou X, Li X, Zheng Y. Effectiveness of gamification interventions to improve physical activity and sedentary behavior in children and adolescents: systematic review and meta-analysis. JMIR Serious Games 2025 Sep 18;13:e68151. [doi: 10.2196/68151] [Medline: 40966596]
69. Wang Q, Yang W, Xu K, Lam LT. The effect of chatbot-based exercise interventions on physical activity, exercise habits, and sedentary behavior: a systematic review and meta-analysis of randomized controlled trials. Digit Health 2025;11:20552076251353731. [doi: 10.1177/20552076251353731] [Medline: 40656859]
70. Chai LK, Farletti R, Fathi L, Littlewood R. A rapid review of the impact of family-based digital interventions for obesity prevention and treatment on obesity-related outcomes in primary school-aged children. Nutrients 2022 Nov 15;14(22):4837. [doi: 10.3390/nu14224837] [Medline: 36432522]
71. Rheu MM, Jang Y, Peng W. Enhancing healthy behaviors through virtual self: a systematic review of health interventions using avatars. Games Health J 2020 Apr;9(2):85-94. [doi: 10.1089/g4h.2018.0134] [Medline: 31724888]
72. Metzendorf MI, Wieland LS, Richter B. Mobile health (m-health) smartphone interventions for adolescents and adults with overweight or obesity. Cochrane Database Syst Rev 2024 Feb 20;2(2):CD013591. [doi: 10.1002/14651858.CD013591.pub2] [Medline: 38375882]
73. Pfledderer CD, Kwon S, Strehli I, Byun W, Burns RD. The effects of playground interventions on accelerometer-assessed physical activity in pediatric populations: a meta-analysis. IJERPH 2022;19(6):3445. [doi: 10.3390/ijerph19063445]
74. Rodríguez-González P, Hassan MA, Gao Z. Effects of family-based interventions using mobile apps on youth's physical activity: a systematic review. J Clin Med 2022 Aug 17;11(16):4798. [doi: 10.3390/jcm11164798] [Medline: 36013037]
75. Schembre SM, Liao Y, Robertson MC, et al. Just-in-time feedback in diet and physical activity interventions: systematic review and practical design framework. J Med Internet Res 2018 Mar 22;20(3):e106. [doi: 10.2196/jmir.8701] [Medline: 29567638]
76. Lee AM, Chavez S, Bian J, et al. Efficacy and effectiveness of mobile health technologies for facilitating physical activity in adolescents: scoping review. JMIR mHealth uHealth 2019 Feb 12;7(2):e11847. [doi: 10.2196/11847] [Medline: 30747716]
77. Allafi AR. Effects of rewards and pedometer-feedback on children's physical activity: a school-based intervention study. Prog Nutr 2020;22(1):122-126. [doi: 10.23751/pn.v22i1.8117]
78. Babic MJ, Smith JJ, Morgan PJ, et al. Intervention to reduce recreational screen-time in adolescents: outcomes and mediators from the "Switch-Off 4 Healthy Minds" (S4HM) cluster randomized controlled trial. Prev Med 2016 Oct;91:50-57. [doi: 10.1016/j.ypmed.2016.07.014] [Medline: 27471018]
79. Baldursdottir B, Taehtinen RE, Sigfusdottir ID, Krettek A, Valdimarsdottir HB. Impact of a physical activity intervention on adolescents' subjective sleep quality: a pilot study. Glob Health Promot 2017 Dec;24(4):14-22. [doi: 10.1177/1757975915626112] [Medline: 27173502]
80. Baranowski T, Baranowski J, Thompson D, et al. Video game play, child diet, and physical activity behavior change a randomized clinical trial. Am J Prev Med 2011 Jan;40(1):33-38. [doi: 10.1016/j.amepre.2010.09.029] [Medline: 21146765]
81. Baranowski T, Abdelsamad D, Baranowski J, et al. Impact of an active video game on healthy children's physical activity. Pediatrics 2012 Mar 1;129(3):e636-e642. [doi: 10.1542/peds.2011-2050]
82. Baranowski T, Baranowski J, Chen TA, et al. Videogames that encourage healthy behavior did not alter fasting insulin or other diabetes risks in children: randomized clinical trial. Games Health J 2019 Aug;8(4):257-264. [doi: 10.1089/g4h.2018.0097]
83. Brannon EE, Cushing CC, Walters RW, Crick C, Noser AE, Mullins LL. Goal feedback from whom? A physical activity intervention using an N-of-1 RCT. Psychol Health 2018 Jun;33(6):701-712. [doi: 10.1080/08870446.2017.1385783] [Medline: 28988493]
84. Caillaud C, Ledger S, Diaz C, Clerc G, Galy O, Yacef K. iEngage: a digital health education program designed to enhance physical activity in young adolescents. In: Jaafar Z, editor. PLOS ONE 2022;17(10):e0274644. [doi: 10.1371/journal.pone.0274644]
85. Carlin A, Logue C, Flynn J, Murphy MH, Gallagher AM. Development and feasibility of a family-based health behavior intervention using intelligent personal assistants: randomized controlled trial. JMIR Form Res 2021 Jan 28;5(1):e17501. [doi: 10.2196/17501] [Medline: 33507155]
86. Chen JL, Weiss S, Heyman MB, Cooper B, Lustig RH. The efficacy of the web-based childhood obesity prevention program in Chinese American adolescents (Web ABC study). J Adolesc Health 2011 Aug;49(2):148-154. [doi: 10.1016/j.jadohealth.2010.11.243] [Medline: 21783046]
87. Comeras-Chueca C, Villalba-Heredia L, Perez-Lasierra JL, et al. Active video games improve muscular fitness and motor skills in children with overweight or obesity. Int J Environ Res Public Health 2022 Feb 24;19(5):2642. [doi: 10.3390/ijerph19052642] [Medline: 35270330]
88. Dewar DL, Morgan PJ, Plotnikoff RC, et al. The Nutrition and Enjoyable Activity for Teen Girls Study. Am J Prev Med 2013 Sep;45(3):313-317. [doi: 10.1016/j.amepre.2013.04.014]
89. Dewar DL, Morgan PJ, Plotnikoff RC, Okely AD, Batterham M, Lubans DR. Exploring changes in physical activity, sedentary behaviors and hypothesized mediators in the NEAT girls group randomized controlled trial. J Sci Med Sport 2014 Jan;17(1):39-46. [doi: 10.1016/j.jsams.2013.02.003] [Medline: 23506657]

XSL•FO
RenderX

90. Direito A, Jiang Y, Whittaker R, Maddison R. Apps for Improving Fitness and increasing physical activity among young people: the AIMFIT pragmatic randomized controlled trial. J Med Internet Res 2015 Aug 27;17(8):e210. [doi: 10.2196/jmir.4568] [Medline: 26316499]
91. Duncan MJ, Staples V. The impact of a school-based active video game play intervention on children's physical activity during recess. Hum Mov 2010;11(1):95-99. [doi: 10.2478/v10038-009-0023-1]
92. Duncan MJ, Fitton Davies K, Okwose N, Harwood AE, Jakovljevic DG. Effects of a 10-week integrated curriculum intervention on physical activity, resting blood pressure, motor skills, and well-being in 6- to 7-year-olds. J Phys Act Health 2024;21(6):595-605. [doi: 10.1123/jpah.2023-0239]
93. Errickson SP, Maloney AE, Thorpe D, Giuliani C, Rosenberg AM. "Dance Dance Revolution" used by 7- and 8-year-olds to boost physical activity: is coaching necessary for adherence to an exercise prescription? Games Health J 2012 Feb;1(1):45-50. [doi: 10.1089/g4h.2011.0028] [Medline: 26196431]
94. Ezendam NPM. Evaluation of the web-based computer-tailored FATaintPHAT intervention to promote energy balance among adolescents. Arch Pediatr Adolesc Med 2012 Mar 5;166(3):248. [doi: 10.1001/archpediatrics.2011.204]
95. Garde A, Umedaly A, Abulnaga SM, et al. Assessment of a mobile game ("MobileKids Monster Manor") to promote physical activity among children. Games Health J 2015 Apr;4(2):149-158. [doi: 10.1089/g4h.2014.0095] [Medline: 26181809]
96. Garde A, Umedaly A, Abulnaga SM, et al. Evaluation of a novel mobile exergame in a school-based environment. Cyberpsychol Behav Soc Netw 2016 Mar;19(3):186-192. [doi: 10.1089/cyber.2015.0281]
97. Garde A, Chowdhury M, Rollinson AU, et al. A multi-week assessment of a mobile exergame intervention in an elementary school. Games Health J 2018 Feb;7(1):1-8. [doi: 10.1089/g4h.2017.0023] [Medline: 29394109]
98. Graves LEF, Ridgers ND, Atkinson G, Stratton G. The effect of active video gaming on children's physical activity, behavior preferences and body composition. Pediatr Exerc Sci 2010;22(4):535-546. [doi: 10.1123/pes.22.4.535]
99. Guagliano JM, Armitage SM, Brown HE, et al. A whole family-based physical activity promotion intervention: findings from the Families Reporting Every Step to Health (FRESH) pilot randomised controlled trial. Int J Behav Nutr Phys Act 2020 Sep 22;17(1):120. [doi: 10.1186/s12966-020-01025-3] [Medline: 32962724]
100. Guthrie N, Bradlyn A, Thompson SK, et al. Development of an accelerometer-linked online intervention system to promote physical activity in adolescents. In: Adams MA, editor. PLOS ONE 2015;10(5):e0128639. [doi: 10.1371/journal.pone.0128639] [Medline: 26010359]
101. Honas JJ, Szabo-Reed AN, Sherman J, et al. The intensity of 10-minute classroom physical activity breaks in elementary school children. IJPAH 2025;4(1):4. [doi: 10.18122/ijpah.4.1.4.boisestate]
102. Jake-Schoffman DE, Turner-McGrievy G, Wilcox S, Moore JB, Hussey JR, Kaczynski AT. The mFIT (Motivating Families With Interactive Technology) study: a randomized pilot to promote physical activity and healthy eating through mobile technology. J technol behav sci 2018 Sep;3(3):179-189. [doi: 10.1007/s41347-018-0052-8]
103. Jauho AM, Pyky R, Ahola R, et al. Effect of wrist-worn activity monitor feedback on physical activity behavior: a randomized controlled trial in Finnish young men. Prev Med Rep 2015;2:628-634. [doi: 10.1016/j.pmedr.2015.07.005] [Medline: 26844128]
104. Lau PWC, Wang JJ, Maddison R. A randomized-controlled trial of school-based active videogame intervention on Chinese children's aerobic fitness, physical activity level, and psychological correlates. Games Health J 2016 Dec;5(6):405-412. [doi: 10.1089/g4h.2016.0057] [Medline: 27855265]
105. Layne T, Simonton K, Irwin C. Effects of a sport education instructional model and heart rate monitor system on the physical activity and jump rope performance of fourth grade students. J Phys Educ Sport 2022;22(4):889-899. [doi: 10.7752/jpes.2022.04113]
106. Lee LL, Kuo YC, Fanaw D, Perng SJ, Juang IF. The effect of an intervention combining self-efficacy theory and pedometers on promoting physical activity among adolescents. J Clin Nurs 2012 Apr;21(7-8):914-922. [doi: 10.1111/j.1365-2702.2011.03881.x] [Medline: 22082301]
107. Leinonen AM, Pyky R, Ahola R, et al. Feasibility of gamified mobile service aimed at physical activation in young men: population-based randomized controlled study (MOPO). JMIR mHealth uHealth 2017 Oct 10;5(10):e146. [doi: 10.2196/mhealth.6675] [Medline: 29017991]
108. Lubans DR, Morgan PJ, Aguiar EJ, Callister R. Randomized controlled trial of the Physical Activity Leaders (PALs) program for adolescent boys from disadvantaged secondary schools. Prev Med 2011;52(3-4):239-246. [doi: 10.1016/j.ypmed.2011.01.009] [Medline: 21276812]
109. Lubans DR, Morgan PJ, Okely AD, et al. Preventing obesity among adolescent girls: one-year outcomes of the Nutrition and Enjoyable Activity for Teen Girls (NEAT Girls) cluster randomized controlled trial. Arch Pediatr Adolesc Med 2012 Sep 1;166(9):821-827. [doi: 10.1001/archpediatrics.2012.41] [Medline: 22566517]
110. Lubans DR, Smith JJ, Plotnikoff RC, et al. Assessing the sustained impact of a school-based obesity prevention program for adolescent boys: the ATLAS cluster randomized controlled trial. Int J Behav Nutr Phys Act 2016 Dec;13(1):92. [doi: 10.1186/s12966-016-0420-8]
111. Maddison R, Foley L, Ni Mhurchu C, et al. Effects of active video games on body composition: a randomized controlled trial. Am J Clin Nutr 2011 Jul;94(1):156-163. [doi: 10.3945/ajcn.110.009142] [Medline: 21562081]

XSL•FO
RenderX

112. Maloney AE, Threlkeld KA, Cook WL. Comparative effectiveness of a 12-week physical activity intervention for overweight and obese youth: exergaming with “Dance Dance Revolution”. Games Health J 2012 Apr;1(2):96-103. [doi: 10.1089/g4h.2011.0009]
113. Manley D, Cowan P, Graff C, et al. Self-efficacy, physical activity, and aerobic fitness in middle school children: examination of a pedometer intervention program. J Pediatr Nurs 2014;29(3):228-237. [doi: 10.1016/j.pedn.2013.10.011] [Medline: 24263251]
114. Morgan PJ, Collins CE, Plotnikoff RC, et al. The ‘Healthy Dads, Healthy Kids’ community randomized controlled trial: a community-based healthy lifestyle program for fathers and their children. Prev Med 2014 Apr;61:90-99. [doi: 10.1016/j.ypmed.2013.12.019]
115. Morgan PJ, Young MD, Barnes AT, Eather N, Pollock ER, Lubans DR. Engaging fathers to increase physical activity in girls: the “Dads and Daughters Exercising and Empowered” (DADEE) randomized controlled trial. Ann Behav Med 2019 Jan 1;53(1):39-52. [doi: 10.1093/abm/kay015] [Medline: 29648571]
116. Morgan PJ, Rayward AT, Young MD, et al. Establishing effectiveness of a community-based, physical activity program for fathers and daughters: a randomized controlled trial. Ann Behav Med 2022 Jul 12;56(7):698-711. [doi: 10.1093/abm/kaab056]
117. Morris JL, Daly-Smith A, Defeyter MA, et al. A pedometer-based physically active learning intervention: the importance of using preintervention physical activity categories to assess effectiveness. Pediatr Exerc Sci 2019 Aug 1;31(3):356-362. [doi: 10.1123/pes.2018-0128] [Medline: 30612529]
118. Ngo CS, Pan C, Finkelstein EA, et al. A cluster randomised controlled trial evaluating an incentive‐based outdoor physical activity programme to increase outdoor time and prevent myopia in children. Ophthalmic Physiologic Optic 2014 May;34(3):362-368. [doi: 10.1111/opo.12112]
119. Petrušič T, Bogataj Š, Štemberger V. Teaching gymnastics elements in elementary school with placebo effect digital games: cluster randomised controlled trial. KINSI 2022;28(3):5-15. [doi: 10.52165/kinsi.28.3.5-15]
120. Pfeiffer KA, Robbins LB, Ling J, et al. Effects of the Girls on the Move randomized trial on adiposity and aerobic performance (secondary outcomes) in low-income adolescent girls. Pediatr Obes 2019 Nov;14(11):e12559. [doi: 10.1111/ijpo.12559] [Medline: 31267695]
121. Pope L, Garnett B, Dibble M. Lessons learned through the implementation of an eHealth physical activity gaming intervention with high school youth. Games Health J 2018 Apr;7(2):136-142. [doi: 10.1089/g4h.2017.0164] [Medline: 29393679]
122. Rhodes RE, Blanchard CM, Bredin SSD, Beauchamp MR, Maddison R, Warburton DER. Stationary cycling exergame use among inactive children in the family home: a randomized trial. J Behav Med 2017 Dec;40(6):978-988. [doi: 10.1007/s10865-017-9866-7] [Medline: 28631101]
123. Ridgers ND, Timperio A, Ball K, et al. Effect of commercial wearables and digital behaviour change resources on the physical activity of adolescents attending schools in socio-economically disadvantaged areas: the RAW-PA cluster-randomised controlled trial. Int J Behav Nutr Phys Act 2021 Apr 12;18(1):52. [doi: 10.1186/s12966-021-01110-1] [Medline: 33845853]
124. Robbins LB, Ling J, Sharma DB, et al. Intervention effects of “Girls on the Move” on increasing physical activity: a group randomized trial. Ann Behav Med 2019 Mar 28;53(5):493-500. [doi: 10.1093/abm/kay054] [Medline: 29985968]
125. Robertson J, Macvean A, Fawkner S, Baker G, Jepson RG. Savouring our mistakes: learning from the FitQuest project. Int J Child Comput Interact 2018 Jun;16:55-67. [doi: 10.1016/j.ijcci.2017.12.003]
126. Ruotsalainen H, Kyngäs H, Tammelin T, Heikkinen H, Kääriäinen M. Effectiveness of Facebook-delivered lifestyle counselling and physical activity self-monitoring on physical activity and body mass index in overweight and obese adolescents: a randomized controlled trial. Nurs Res Pract 2015;2015:159205. [doi: 10.1155/2015/159205] [Medline: 26697218]
127. Seah MLC, Koh KT. The efficacy of using mobile applications in changing adolescent girls’ physical activity behaviour during weekends. Eur Phy Educ Rev 2021 Feb;27(1):113-131. [doi: 10.1177/1356336X20930741]
128. Smith JJ, Morgan PJ, Plotnikoff RC, et al. Smart-phone obesity prevention trial for adolescent boys in low-income communities: the ATLAS RCT. Pediatrics 2014 Sep;134(3):e723-e731. [doi: 10.1542/peds.2014-1012] [Medline: 25157000]
129. Neto AS, Corrêa RC, Farias JD, et al. Efeitos de uma intervenção com pedômetro sobre o risco metabólico em crianças obesas [Article in Portuguese]. Rev Bras Med Esporte 2016 Dec;22(6):476-479. [doi: 10.1590/1517-869220162206146692]
130. Staiano AE, Beyl RA, Hsia DS, Katzmarzyk PT, Newton RL. Twelve weeks of dance exergaming in overweight and obese adolescent girls: transfer effects on physical activity, screen time, and self-efficacy. J Sport Health Sci 2017 Mar;6(1):4-10. [doi: 10.1016/j.jshs.2016.11.005] [Medline: 28491483]
131. Staiano AE, Beyl RA, Guan W, Hendrick CA, Hsia DS, Newton RL. Home-based exergaming among children with overweight and obesity: a randomized clinical trial. Pediatr Obes 2018 Nov;13(11):724-733. [doi: 10.1111/ijpo.12438] [Medline: 30027607]
132. Thompson D, Cantu D, Ramirez B, et al. Texting to increase adolescent physical activity: feasibility assessment. Am J Health Behav 2016 Jul;40(4):472-483. [doi: 10.5993/AJHB.40.4.9] [Medline: 27338994]
133. Trost SG, Sundal D, Foster GD, Lent MR, Vojta D. Effects of a pediatric weight management program with and without active video games a randomized trial. JAMA Pediatr 2014 May;168(5):407-413. [doi: 10.1001/jamapediatrics.2013.3436] [Medline: 24589566]

XSL•FO
RenderX

134. Tugault-Lafleur CN, De-Jongh González O, Macdonald J, et al. Efficacy of the Aim2Be intervention in changing lifestyle behaviors among adolescents with overweight and obesity: randomized controlled trial. J Med Internet Res 2023 Apr 25;25:e38545. [doi: 10.2196/38545] [Medline: 37097726]
135. van Woudenberg TJ, Bevelander KE, Burk WJ, Smit CR, Buijs L, Buijzen M. A randomized controlled trial testing a social network intervention to promote physical activity among adolescents. BMC Public Health 2018 Dec;18(1):542. [doi: 10.1186/s12889-018-5451-4]
136. Van Woudenberg TJ, Bevelander KE, Burk WJ, Smit CR, Buijs L, Buijzen M. Testing a social network intervention using vlogs to promote physical activity among adolescents: a randomized controlled trial. Front Psychol 2020;10:2913. [doi: 10.3389/fpsyg.2019.02913]
137. Verswijveren S, Abbott G, Lai SK, et al. Mediators of effects on physical activity and sedentary time in an activity tracker and behavior change intervention for adolescents: secondary analysis of a cluster randomized controlled trial. JMIR mHealth uHealth 2022 Aug 16;10(8):e35261. [doi: 10.2196/35261] [Medline: 35972777]
138. Wunsch K, Fiedler J, Hubenschmid S, Reiterer H, Renner B, Woll A. An mHealth intervention promoting physical activity and healthy eating in a family setting (SMARTFAMILY): randomized controlled trial. JMIR mHealth uHealth 2024 Apr 26;12:e51201. [doi: 10.2196/51201] [Medline: 38669071]
139. Deci EL, Ryan RM. Self-determination theory. In: Handbook of Theories of Social Psychology: Volume 1: SAGE Publications Ltd; 2012:416-437. [doi: 10.4135/9781446249215.n21]
140. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [doi: 10.1037//0003-066x.55.1.68] [Medline: 11392867]
141. Bandura A. Social foundations of thought and action. In: The Health Psychology Reader: SAGE Publications Ltd; 2002:94-106. [doi: 10.4135/9781446221129.n6]
142. Ajzen I. The theory of planned behavior. Organ Behav Hum Decis Process 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
143. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997 Sep;12(1):38-48. [doi: 10.4278/0890-1171-12.1.38]
144. Pyky R, Koivumaa-Honkanen H, Leinonen AM, et al. Effect of tailored, gamified, mobile physical activity intervention on life satisfaction and self-rated health in young adolescent men: a population-based, randomized controlled trial (MOPO study). Comput Human Behav 2017 Jul;72:13-22. [doi: 10.1016/j.chb.2017.02.032]
145. Liu M, Guan X, Guo X, et al. Impact of serious games on body composition, physical activity, and dietary change in children and adolescents: a systematic review and meta-analysis of randomized controlled trials. Nutrients 2024 Apr 26;16(9):1290. [doi: 10.3390/nu16091290] [Medline: 38732536]
146. Current world population. World Population Review. URL: https://worldpopulationreview.com/ [accessed 2026-02-26]
147. Golley RK, Hendrie GA, Slater A, Corsini N. Interventions that involve parents to improve children's weight-related nutrition intake and activity patterns - what nutrition and activity targets and behaviour change techniques are associated with intervention effectiveness? Obes Rev 2011 Feb;12(2):114-130. [doi: 10.1111/j.1467-789X.2010.00745.x] [Medline: 20406416]
148. Anderson EL, Howe LD, Kipping RR, et al. Long-term effects of the Active for Life Year 5 (AFLY5) school-based cluster-randomised controlled trial. BMJ Open 2016 Nov;6(11):e010957. [doi: 10.1136/bmjopen-2015-010957]
149. Brofenbrenner U. Six theories of child development: revised formulations and current issues. In: Ecological Systems Theory: Jessica Kingsley Publishers; 1992:187-249.
150. Hendrie GA, Brindal E, Corsini N, Gardner C, Baird D, Golley RK. Combined home and school obesity prevention interventions for children: what behavior change strategies and intervention characteristics are associated with effectiveness? Health Educ Behav 2012 Apr;39(2):159-171. [doi: 10.1177/1090198111420286] [Medline: 21984691]
151. Vilardell-Dávila A, Martínez-Andrade G, Klünder-Klünder M, et al. A multi-component educational intervention for addressing levels of physical activity and sedentary behaviors of schoolchildren. Int J Environ Res Public Health 2023 Feb 9;20(4):3003. [doi: 10.3390/ijerph20043003] [Medline: 36833697]
152. Rose T, Barker M, Maria Jacob C, et al. A systematic review of digital interventions for improving the diet and physical activity behaviors of adolescents. J Adolesc Health 2017 Dec;61(6):669-677. [doi: 10.1016/j.jadohealth.2017.05.024] [Medline: 28822682]
153. Swindle T, Poosala AB, Zeng N, Børsheim E, Andres A, Bellows LL. Digital intervention strategies for increasing physical activity among preschoolers: systematic review. J Med Internet Res 2022;24(1):e28230. [doi: 10.2196/28230]
154. Shahsavari H, Matourypour P, Ghiyasvandian S, Nejad MG. Medical research council framework for development and evaluation of complex interventions: a comprehensive guidance. J Edu Health Promot 2020;9(1):88. [doi: 10.4103/jehp.jehp_649_19]
155. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining "gamification". Presented at: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments; Sep 28-30, 2011; Tampere, Finland p. 9-15. [doi: 10.1145/2181037.2181040]

XSL•FO
RenderX

156. Mateo-Orcajada A, Vaquero-Cristóbal R, Abenza-Cano L. Mobile application interventions to increase physical activity and their effect on kinanthropometrics, body composition and fitness variables in adolescent aged 12-16 years old: an umbrella review. Child Care Health Dev 2024 Jan;50(1):e13146. [doi: 10.1111/cch.13146] [Medline: 37387258]
157. Panicker RM, Chandrasekaran B. “Wearables on vogue”: a scoping review on wearables on physical activity and sedentary behavior during COVID-19 pandemic. Sport Sci Health 2022;18(3):641-657. [doi: 10.1007/s11332-021-00885-x] [Medline: 35018193]
158. Zhang W, Xiong K, Zhu C, Evans R, Zhou L, Podrini C. Promoting child and adolescent health through wearable technology: a systematic review. Digit Health 2024;10:20552076241260507. [doi: 10.1177/20552076241260507] [Medline: 38868368]
159. Nguyen P, Le LKD, Nguyen D, Gao L, Dunstan DW, Moodie M. The effectiveness of sedentary behaviour interventions on sitting time and screen time in children and adults: an umbrella review of systematic reviews. Int J Behav Nutr Phys Act 2020 Dec;17(1):117. [doi: 10.1186/s12966-020-01009-3]
160. Romero-Rodríguez JM, Martínez-Menéndez A, Alonso-García S, Victoria-Maldonado JJ. The reality of the gamification methodology in primary education: a systematic review. Int J Educ Res 2024;128:102481. [doi: 10.1016/j.ijer.2024.102481]
161. Zhu J, Chen M, Jiang W, et al. Gamified digital exercise interventions for children and adolescents: protocol for a systematic review and behavior change technique analysis. Digit Health 2025;11:20552076251365017. [doi: 10.1177/20552076251365017] [Medline: 40761775]
162. Alalawi A, Blank L, Goyder E. Umbrella review of international evidence for the effectiveness of school-based physical activity interventions. In: Tilga H, editor. PLOS ONE 2024;19(6):e0304513. [doi: 10.1371/journal.pone.0304513] [Medline: 38870155]
163. Singh B, Ahmed M, Staiano AE, et al. Lifestyle eHealth and mHealth interventions for children and adolescents: systematic umbrella review and meta-meta-analysis. J Med Internet Res 2025 Oct 17;27:e69065. [doi: 10.2196/69065] [Medline: 41106821]
164. Glanz K, Rimer BK, Viswanath K, editors. Health Behavior and Health Education: Theory, Research, and Practice, 4th edition: Jossey-Bass; 2008. URL: https://transformationalchange.pbworks.com/f/HealtBehavior-Education.pdf [accessed 2026-02-26]
165. Indagine Conoscitiva su Bullismo e Cyberbullismo [Article in Italian]. URL: https://www.istat.it/ [accessed 2026-02-26]
166. Lee JA, Williams SM, Brown DD, Laurson KR. Concurrent validation of the Actigraph gt3x+, Polar Active accelerometer, Omron HJ-720 and Yamax Digiwalker SW-701 pedometer step counts in lab-based and free-living settings. J Sports Sci 2015 Jun 15;33(10):991-1000. [doi: 10.1080/02640414.2014.981848]
167. Redenius N, Kim Y, Byun W. Concurrent validity of the Fitbit for assessing sedentary behavior and moderate-to-vigorous physical activity. BMC Med Res Methodol 2019 Dec;19(1):29. [doi: 10.1186/s12874-019-0668-1]
168. Santos-Lozano A, Santín-Medeiros F, Cardon G, et al. Actigraph GT3X: validation and determination of physical activity intensity cut points. Int J Sports Med 2013 Nov;34(11):975-982. [doi: 10.1055/s-0033-1337945] [Medline: 23700330]
169. Peng S, Othman AT, Khairani AZ, Zeng G, Xiaogang Z, Fang Y. Effectiveness of pedometer- and accelerometer-based interventions in improving physical activity and health-related outcomes among college students: a systematic review and meta-analysis. Digit Health 2023;9:20552076231188213. [doi: 10.1177/20552076231188213] [Medline: 37492032]
170. Laird Y, Fawkner S, Kelly P, McNamee L, Niven A. The role of social support on physical activity behaviour in adolescent girls: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2016 Dec;13(1):79. [doi: 10.1186/s12966-016-0405-7]
171. Xie C, Zhang Z, Zhang X, Li Y, Shi P, Wang S. Effects of interventions on physical activity behavior change in children and adolescents based on a trans-theoretical model: a systematic review. BMC Public Health 2025;25(1):657. [doi: 10.1186/s12889-025-21336-z]
172. Borde R, Smith JJ, Sutherland R, Nathan N, Lubans DR. Methodological considerations and impact of school-based interventions on objectively measured physical activity in adolescents: a systematic review and meta-analysis. Obes Rev 2017 Apr;18(4):476-490. [doi: 10.1111/obr.12517] [Medline: 28187241]
173. Lunny C, Pieper D, Thabet P, Kanji S. Managing overlap of primary study results across systematic reviews: practical considerations for authors of overviews of reviews. BMC Med Res Methodol 2021 Jul 7;21(1):140. [doi: 10.1186/s12874-021-01269-y] [Medline: 34233615]
174. Higgins JPT, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration’s tool for assessing risk of bias in randomised trials. BMJ 2011 Oct 18;343(oct18 2):d5928-d5928. [doi: 10.1136/bmj.d5928]
175. Fernandez RS, Sharifnia AM, Khalil H. Umbrella reviews: a methodological guide. Eur J Cardiovasc Nurs 2025 Sep 5;24(6):996-1002. [doi: 10.1093/eurjcn/zvaf012] [Medline: 39844649]
176. Sadoyu S, Tanni KA, Punrum N, et al. Methodological approaches for assessing certainty of the evidence in umbrella reviews: a scoping review. PLoS ONE 2022;17(6):e0269009. [doi: 10.1371/journal.pone.0269009] [Medline: 35675337]
177. Kuitunen I. Characterising umbrella reviews in paediatrics: a systematic analysis. Acta Paediatr 2025 Jan;114(1):51-55. [doi: 10.1111/apa.17438] [Medline: 39367681]
178. West SL, Gartlehner G, Mansfield AJ, et al. Comparative Effectiveness Review Methods: Clinical Heterogeneity: Agency for Healthcare Research and Quality (US); 2010. URL: http://www.ncbi.nlm.nih.gov/books/NBK53310 [accessed 2026-02-26]

XSL•FO
RenderX

179. Wyatt B, Forstmann N, Badier N, et al. Changes in physical activity, heart rate, and sleep measured by activity trackers during the COVID-19 pandemic across 34 countries: retrospective analysis. J Med Internet Res 2025 Apr 4;27:e68199. [doi: 10.2196/68199] [Medline: 40184182]
180. Gasana J, O'Keeffe T, Withers TM, Greaves CJ. A systematic review and meta-analysis of the long-term effects of physical activity interventions on objectively measured outcomes. BMC Public Health 2023 Sep 2;23(1):1697. [doi: 10.1186/s12889-023-16541-7] [Medline: 37660119]
181. Wahlich C, Chaudhry UAR, Fortescue R, et al. Effectiveness of adult community-based physical activity interventions with objective physical activity measurements and long-term follow-up: a systematic review and meta-analysis. BMJ Open 2020 Apr;10(5):e034541. [doi: 10.1136/bmjopen-2019-034541]
182. Lum M, Wolfenden L, Jones J, et al. Interventions to improve child physical activity in the early childhood education and care setting: an umbrella review. IJERPH 2022;19(4):1963. [doi: 10.3390/ijerph19041963]
183. Mierniczek M, Mierniczek M, Mierniczek A, Kaczmarska K, Rosołowska K, Dudek J. Digital health interventions for school-aged children: effectiveness, implementation, and impact on 24-hour movement and psychosocial well-being. J Educ Health Sport 2025;81:66752. [doi: 10.12775/JEHS.2025.81.66752]
184. König LM, Western MJ, Denton AH, Krukowski RA. Umbrella review of social inequality in digital interventions targeting dietary and physical activity behaviors. npj Digit Med 2025;8(1):11. [doi: 10.1038/s41746-024-01405-0]
185. Mair JL, Salamanca-Sanabria A, Augsburger M, et al. Effective behavior change techniques in digital health interventions for the prevention or management of noncommunicable diseases: an umbrella review. Ann Behav Med 2023 Sep 13;57(10):817-835. [doi: 10.1093/abm/kaad041]

Abbreviations

AIMFIT: Apps for Improving Fitness
AMSTAR 2: A Measurement Tool to Assess Systematic Reviews
CCA: corrected covered area
CG: control group
DARE: Digital Lifelong Prevention
GRADE: Grading of Recommendations, Assessment, Development, and Evaluation
IG: intervention group
LPA: light physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
PICOS: Population, Intervention, Comparison, Outcome, and Study Type
PRIOR: Preferred Reporting Items for Overviews of Reviews
PRISMA-S: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Literature Search Extension
PROSPERO: International Prospective Register of Systematic Reviews
RCT: randomized controlled trial
RoB 2: Risk of Bias 2
SB: sedentary behavior
SR: systematic review
VPA: vigorous physical activity

Edited by A Mavragani, S Brini; submitted 11.Apr.2025; peer-reviewed by DLIHK Peiris, I Prestanti, P Sandu, S Honorio; accepted 21.Jan.2026; published 24.Mar.2026.

Please cite as:
Tabacchi G, Scardina A, Amato A, Giardina M, Accardi G, Di Liberto V, Giglia G, Vasto S, Frinchi M, Boffetta P, Mazzucco W, Bellafiore M
Digital Tools' Effectiveness on Physical Activity Outcomes in Children and Adolescents: Umbrella Review
JMIR Public Health Surveill 2026;12:e75769
URL: https://publichealth.jmir.org/2026/1/e75769
doi:10.2196/75769

© Garden Tabacchi, Antonino Scardina, Antonella Amato, Marta Giardina, Giulia Accardi, Valentina Di Liberto, Giuseppe Giglia, Sonya Vasto, Monica Frinchi, Paolo Boffetta, Walter Mazzucco, Marianna Bellafiore. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 24.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted

XSL•FO
RenderX

use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

The Effectiveness of Contact Tracing to Reduce Transmission of Infectious Diseases During Epidemic or Pandemic Response: Rapid Systematic Review

Michael Nunns[1], MSc, PhD; Samantha Febrey[1], BSc; Kieran Becker[1], BSc; Morgan Weiland[1], BMBS, BMedSci, MSc; Jill Buckland[1], MSc; Rebecca Abbott[1,2], PhD; Rebecca Whear[1,2], MSc; Alison Bethel[1,2], BSc; Liz Shaw[1], MSc, DClinPsy; Kate Boddy[1,2], MA, MSc; Serena Carville[3], PhD; Tamsyn Harris[3], PhD; Jo Thompson Coon[1,2], PhD; G J Melendez-Torres[1,2], MPH, DPhil

[1]Isca Evidence, University of Exeter Medical School, Faculty of Life Sciences, University of Exeter, Exeter, United Kingdom

[2]University of Exeter Medical School, NIHR Applied Research Collaboration South West Peninsula (PenARC), University of Exeter, Exeter, United Kingdom

[3]UK Health Security Agency, London, United Kingdom

Corresponding Author:
Michael Nunns, MSc, PhD
Isca Evidence, University of Exeter Medical School
Faculty of Life Sciences
University of Exeter
SC 3.09, St Luke's Campus
Heavitree Road
Exeter, EX1 2LU
United Kingdom
Phone: 44 1392722399
Email: m.p.nunns@exeter.ac.uk

Abstract

Background: Contact tracing (CT), the process of identifying and managing contacts of infected cases, is one public health and social measure that may reduce the spread of infectious diseases. While previous systematic reviews of CT exist, a comprehensive review of both the effectiveness and potential unintended consequences has not been undertaken to our knowledge. Understanding effective CT strategies could help governments and health authorities prepare effectively for emergency epidemic or pandemic situations.

Objective: This study aims to systematically review the evidence on the effectiveness of CT across infectious diseases with epidemic or pandemic potential. Effectiveness is measured in terms of impacts on disease transmission, health care use, mortality, or unintended consequences.

Methods: We searched 6 bibliographic databases (MEDLINE, Embase, Global Health, CINAHL Ultimate, Cochrane, and Scopus) between November 29 and December 3, 2024, with supplementary citation searching. We sought human studies comparing CT with interventions with no CT or other forms of CT, delivered in the community, in prespecified diseases of epidemic or pandemic potential. We included studies with any measure of disease transmission, related health care use, or unintended consequences of CT and prioritized studies with concurrent comparators. Screening, data extraction, and critical appraisal were performed in duplicate. Due to substantial heterogeneity, a narrative synthesis was performed. This review was informed by meetings with a patient and public involvement and engagement group.

Results: After deduplication, a total of 12,816 titles and abstracts were screened, with 198 records assessed for eligibility at full text. Five additional studies were found through supplementary searching. Finally, 88 reports (of 86 studies) were included, of which 57 reports (of 55 studies) were prioritized. Two main routes of transmission were represented: respiratory (tuberculosis [TB], 15 studies; COVID-19, 5 studies) and blood-borne or sexually transmitted infections (STIs; 35 studies, of which 13 were in HIV, and 22 were bacterial or parasitic infections). No evidence was found on vector-borne, direct contact, or food- or water-borne routes of transmission. Evidence was highly heterogeneous, and more than half of the studies had notable methodological limitations. While there was no difference between CT and comparator interventions for most outcomes, there

XSL•FO
RenderX

was some evidence of reductions in disease prevalence in TB and for provider-initiated CT to be superior to patient-led approaches in STIs. Only 2 studies reported measures of unintended consequences.

Conclusions: We found inconsistent evidence for the effectiveness of CT, focused primarily on TB and on contrasts between provider-initiated CT and patient-led referral in STIs and HIV. High heterogeneity in study design precluded clear assertions regarding optimal strategies for CT, including with respect to relevant subgroups. Future work should consider generalizability of CT mechanisms across contexts, including by route of transmission and from the Global South, and a more thorough account of unintended consequences.

(JMIR Public Health Surveill 2026;12:e84805) doi:10.2196/84805

KEYWORDS

contact tracing; epidemic response; pandemic preparedness; partner notification; public health and social measures; systematic review

Introduction

Overview

The term Public Health and Social Measures (PHSMs) describes a range of nonpharmaceutical interventions implemented by individuals, communities, and governments to reduce the risk and scale of infection during a health emergency such as an epidemic or pandemic [1]. PHSMs play a role through all phases of such emergencies, especially while clinical countermeasures (vaccines and therapeutics) are being developed and deployed. It is therefore important to understand the effectiveness of PHSMs and their potential unintended consequences to allow governments and public health organizations to prepare for epidemics and pandemics.

Contact tracing (CT) is one important PHSM, defined by the World Health Organization (WHO) as the process of identifying, assessing, and managing people who have been exposed to someone who has been infected with the infection of interest [1]. CT may take many forms or be a component of multicomponent interventions and is an important element of organizational responses to any infectious disease. This review focused on CT, as well as innovations linked to CT that are generalizable across a range of infectious diseases.

One systematic review of CT in the control of infectious diseases from 2022 found that provider-initiated CT can be effective in managing infectious diseases but that evidence was relatively scarce outside of tuberculosis (TB), highly heterogeneous, and of mixed quality [2]. Similar reviews of CT have been published in relation to COVID-19 [3-8]. These reviews either included modeling or simulation studies, used broader definitions of CT that included human resource–intensive strategies such as active case finding, CT in combination with other measures, or included aspects of CT that are not generalizable to pathogens that spread by other routes of transmission (such as partner-delivered treatment for sexually transmitted infections [STIs]). Given the variability in the existing review evidence, and our research aim to have a comprehensive review of CT in the context of epidemic or pandemic response, a de novo systematic review examining the effectiveness of CT for different pathogens and routes of transmission is warranted.

Research Question

What is the effectiveness of CT in reducing transmission of infectious disease as part of an epidemic or pandemic response?

Methods

Overview

The protocol for this review was drafted in conjunction with key interest holders at the UK Health Security Agency and a patient and public involvement and engagement group (described below). It was prospectively registered on PROSPERO (International Prospective Register of Systematic Reviews) prior to commencement of the review (CRD42025628278). The review was reported following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [9] (see checklist in Multimedia Appendix 1). This was a rapid review conducted between November 2024 and April 2025. Full methodological details are available in the protocol and full report [10].

Identification of Studies

The search strategy was developed in MEDLINE by an information specialist (JB), peer reviewed by a second information specialist (AB), and translated for the other databases. Validated search filters for randomized controlled trials (RCTs) were obtained via the InterTASC Information Specialists' Sub-Group (ISSG) Filters Resource [11] for searching MEDLINE [12], Embase [12], and CINAHL Ultimate [13] and nonvalidated search filters for observational studies [14], evaluation studies, and nonrandomized trials or quasi-experimental studies [15].

The list of diseases was developed and validated with subject matter experts as a pragmatic list of representative diseases of each of the 5 routes of transmission (respiratory, vector-borne, direct contact, blood-borne, food-borne, and water-borne [ingestion]). The searches used a combination of relevant controlled vocabulary terms (eg, MeSH [Medical Subject]) and free-text terms.

A total of 6 bibliographic databases were searched between November 29 and December 3, 2024: MEDLINE (1946-current), Embase (1974-current), and Global Health (1973-current) via Ovid; CINAHL Ultimate (1937-current) via EBSCOhost; Cochrane Central via Cochrane; and Scopus (1788-current) via Elsevier. No date or language restrictions were used. Results

XSL•FO
RenderX

were downloaded into EndNote (version 20; Clarivate Analytics), which was used for finding and removing duplicates.

Supplementary searching involved forward and backward citation searching of included references using Scopus. After deduplication using EndNote, the remaining 1996 records were single screened. Search strategies are available in Multimedia Appendix 2.

Inclusion and Exclusion Criteria

The inclusion and exclusion criteria were defined in consultation with interest holders and are summarized below. A more detailed description is provided in Multimedia Appendix 3.

Participants or Population

Include: all humans in the context of an outbreak, epidemic, or pandemic of one of the prespecified infectious diseases of interest (full list provided in Multimedia Appendix 3), using WHO definitions for an outbreak, epidemic, or pandemic [16]. We focused on conditions with the potential to reach epidemic or pandemic status, as it was considered that this could help inform the use of CT to prevent an outbreak from escalating.

Intervention

The inclusion criteria were as follows: interventions that incorporate CT, defined as the process of starting from a confirmed case diagnosed with an infectious disease (index case), identifying who the index case may have come into contact with, and attempting to communicate with the contacts. CT may be a stand-alone intervention or part of a multicomponent intervention that may involve further management of contacts (eg, treatment or quarantine) or other activities to improve infection detection (eg, education and training of health care staff). Interventions should take place in a community setting (including CT that is led by a health care setting but not limited to that health care setting).

Exclusion criteria were as follows: any other form of PHSMs or intervention to reduce transmission that does not incorporate CT; studies in health care settings where the goal was control of nosocomial outbreaks.

Comparator

Different CT approaches compared with each other or compared with no CT were included.

Other PHSMs as comparators where these were not part of background measures in both arms were excluded.

Outcomes

Studies with at least 1 measure of effectiveness (ie, measures of transmission of the focal disease, health care use, or mortality) or unintended consequences were included, as detailed in Multimedia Appendix 3.

Study Design

Experimental studies with a control group, including RCTs, quasi-experimental studies, and before-and-after studies, and observational studies with a control group, including cross-sectional and case-control cohorts were included.

Studies without control groups, reviews of any type, crossover study designs, modeling studies, case series, case reports, and qualitative studies were excluded.

We included only peer-reviewed research published in English, with no restriction on date of publication or setting.

Selection of Evidence

After piloting, the title and abstract of each record retrieved by the search were screened by 2 independent reviewers (MN, SF, MW, K Becker, RA, JTC, GJMT, and JB). The full text of each remaining record was then screened by 2 independent reviewers (MN, SF, MW, and K Becker). Disagreements at each stage were resolved through discussion. Articles excluded at full text were coded to indicate the first reason for exclusion. Several additional clarifications to the inclusion criteria were needed as study selection progressed due to the complex nature of interventions:

- Testing interventions were included if testing (ie, partner-delivered testing kits) was part of the CT process.
- Treatment interventions were excluded if there was no difference in CT activity between arms.
- Community-led surveillance interventions (where “community hot spots” with high disease prevalence were actively targeted for screening) were excluded.
- Any paper in which only proxy measures for a relevant outcome were presented was excluded.
- Active case finding or intensified active case finding was included if the process began from a known index case, rather than indiscriminate case finding based on location, for example.

Prioritization

Because of the volume of evidence identified, we prioritized the most robust evidence (studies with concurrent control). This was a protocol deviation and was registered on PROSPERO. Top-level data were extracted from nonprioritized studies, tabulated, and briefly narratively described.

Critical Appraisal

After piloting, each included trial was critically appraised using a modified version of the Effective Public Health Practice Project Quality Assessment Tool for Quantitative Studies, a validated tool suitable for multiple study designs [17]. Each study was rated either “strong,” “moderate,” or “weak” in 6 domains, with the number of “weak” scores given per study used to determine its overall score: studies were rated as “strong” if they had no “weak” scores, “moderate” with 1 “weak” score, and “weak” if they had 2 or more “weak” scores. Modifications and interpretations established during critical appraisal are described in the full report [10].

Appraisal was independently conducted by 2 reviewers (K Becker, MW, SF, LS, and RW), with disagreements resolved through discussion. The findings of critical appraisal informed the interpretation of results and did not influence the inclusion of studies in the review or the synthesis.

XSL•FO
RenderX

Data Extraction

After piloting, key information was extracted from included trials by 1 reviewer and checked by a second (K Becker, MW, SF, and RA). Discrepancies were resolved through discussion. Data were extracted in relation to author and study details, sample characteristics, trial arm details, outcomes, outcome data, and inequalities addressed.

Synthesis

Extracted data were tabulated and narratively described. Included studies were grouped based on route of transmission and condition or infection. STIs were grouped within the blood-borne route to align with the approach of the Department of Health and Social Care pandemic preparedness program at the time of conducting the review [18]. Interventions and outcomes were then described within those groups, with studies presenting similar interventions and comparators discussed together where possible. Comparators were described as usual care or standard of care as described by study authors. The CT element of each intervention was categorized into 1 of 4 groups:

- Provider-initiated contact tracing (PICT): CT is instigated by a trained provider (often health care worker) who notifies the index case's contacts.
- Partner referral (PR): CT is instigated by the index case who notifies their contacts.
- Household investigation (HI): the home of the index case is visited by a health care worker to notify and assess household members for signs of infection.
- Technology-based contact tracing (TECH): CT facilitated using digital tools, programs, or applications.

Contract referral (where the index case is given a set amount of time to contact their partners before a health care worker notifies the contacts) was categorized as a mix of both PICT and PR.

Where study design, route of transmission, intervention, comparator, and outcome were similar in multiple studies, we considered meta-analysis. In the event, heterogeneity precluded this, and we opted for a narrative synthesis approach, guided by Synthesis Without Meta-Analysis [19].

Patient and Public Involvement and Engagement

This review benefited from discussions with PERSPEX (Public Engagement in Research for Health and Social Policy at Exeter), a group of 17 public collaborators who bring their carer, patient, or public perspective to the work of Isca Evidence. PERSPEX members meet monthly online, and membership is culturally, geographically, and demographically diverse. The review team discussed the project 3 times with PERSPEX, enabling the review team to gain insights about CT from a patient, carer, and public perspective.

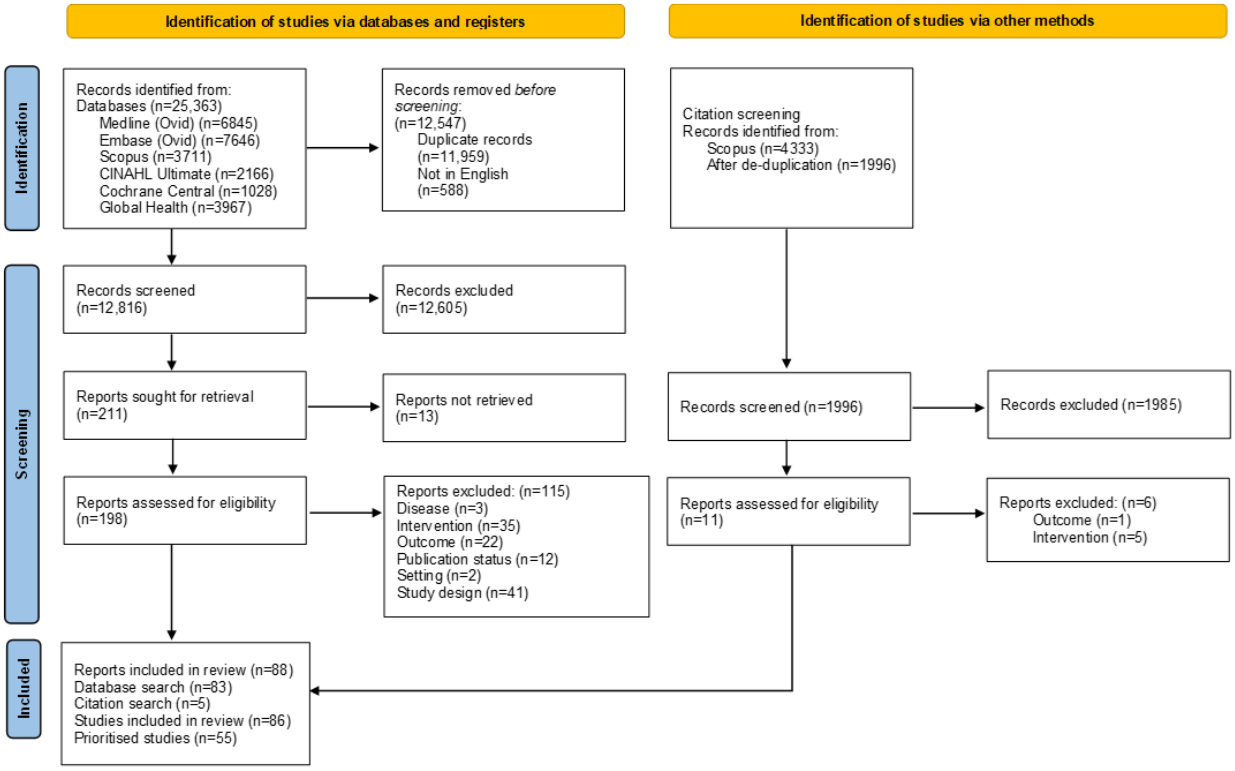
Results

Study Selection

The initial searches identified 25,363 records. After deduplication and removal of non–English-language references, a total of 12,816 titles and abstracts were screened for relevance. After excluding 12,605 records for clearly not meeting 1 or more inclusion criteria, a total of 211 full-text articles were sought. Thirteen were unavailable, resulting in 198 records assessed for eligibility at full text. A total of 5 additional studies were found through supplementary searching. The most common reasons for exclusion at the full-text stage were intervention (40 reports) and study design (41 reports). Finally, 88 reports (of 86 studies) were included in the review. Of these, 57 reports (of 55 studies) were prioritized. The PRISMA flow diagram shown in Figure 1 summarizes this process. Multimedia Appendix 4 contains a list of studies with reasons for exclusion at full-text. Multimedia Appendix 5 contains the search summary table, which documents the origin of included references [20].

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart summarizing the results of the literature search and screening for eligibility.



Sample Characteristics

Table S1 in Multimedia Appendix 6 displays the sample characteristics of the 55 prioritized studies, reported in 57 papers [21-77], including study design and location, condition of interest, study dates and sample sizes, settings and recruitment criteria, and sample demographic characteristics. A total of 36 studies were a form of RCT [21,22,24-31,35-39,42,43,45,51-60,65-67,69-71,74-77]. The 19 non-RCT studies included 9 observational studies with concurrent control [23,32,33,44,61-64,73], 4 controlled before-and-after studies [46-49], 3 interrupted time series [40,41,50], and 3 nonrandomized trials with concurrent controls [34,68,72]. Populations across 6 continents were sampled, with the majority taking place in Africa (n=18 [24,26,27,29,32,39,42,43,47,48,50,51,54,60,65,66,70,76]), North America (n=13 [44,52,53,56-58,61,63,68,71,72,74,77], and Europe (n=11 [21,22,34-38,40,41,59,64,69,73]. A total of 8 studies took place in Asia [23,28,45,46,49,55,62,67], 3 in South America [25,30,75], 1 in Oceania [33], and 1 across Asia and Oceania [31].

All evidence was related to either respiratory or sexual or blood-borne routes of transmission. Within the respiratory route, 5 studies related to COVID-19 [40,41,64,73,76] and 15 to TB [23,25,42,43,46-51,54,55,65,67,75]. Under the sexual or blood-borne route, 35 studies were related to STIs [21,22,24,26-39,44,45,52,53,56-63,66,68-72,74,77]. Of the STIs, 13 studies included HIV [24,26-29,31,32,44,45,58,60,62,63], 13 chlamydia [21,22,33-38,52,53,56,59,68,69,77], 7 gonorrhea [34,39,53,56,68,72,77], 3 nongonococcal urethritis or urethritis [34,52,56], 6 syphilis [30,39,44,61,70,71], and 3 trichomoniasis [39,57,74]. Several papers investigated more than 1 STI [34,39,44,52,53,56,68,77]. One paper included chancroid and lymphogranuloma venereum alongside syphilis, gonorrhea, and trichomoniasis [39], while another referred to STIs but did not specify which ones [66].

Where reported, sample size ranged from 55 [31] to 59,027 [41], with the total sample in this review being 125,635 participants. A total of 27 studies were set in a form of clinic or outpatient services [22,24,26-28,32-39,42,45,50-52,54,56,58,62,63,66,68,70,72,74,78]. Of these, 14 studies were in generalist services [32,33,35,36,39,42,45,50,51,54,58,68,70,72,74] and 13 in services specialized for the condition of interest, such as an STI clinic [22,24,26-28,34,37,38,52,56,62,63,66,77]. A total of 19 studies took place in the community [21,23,25,29,41,47-49,55,59,61,64,65,67,69,71,73,75,76] and 5 in both clinic and community settings [30,43,44,57,60]. One study took place in both a clinic and on the internet [53], 1 was set in a prison [31], 1 in a nationwide program [40], and 1 setting was not reported [46].

Critical Appraisal

Of the 55 appraised studies [21-77], a total of 9 received an overall rating of “strong” [28,42,44,57,64,65,70,76,77], 18 received an overall rating of “moderate” [21,26,30,39,40,45,49,51,52,55,59-61,67,69,71,74,75], and the remaining 28 were rated as “weak” [22-25,27,29,31-38,41,43,46-48,50,53,54,56,58,62,63,66,68,72,73]. Overall, the criteria on which studies scored most highly were study design and data collection. Studies scored most poorly on criteria related to researcher and participant blinding. Ratings are summarized in Table 1. A more detailed description of critical appraisal findings is available in the full report [10].

XSL•FO
RenderX

Table 1. Section and overall critical appraisal ratings for prioritized studies.

Study ID	Selection bias	Study design	Confounders	Blinding	Data collection	Withdrawals	Overall rating
Andersen et al [21]	MO[a]	ST[b]	WK[c]	MO	ST	ST	MO
Apoola and Beardsley [22]	ST	ST	WK	WK	ST	ST	WK
Bai et al [23]	MO	WK	WK	WK	ST	N/A[d]	WK
Brown et al [24]	MO	ST	ST	WK	ST	WK	WK
Cavalcante et al [25]	WK	ST	ST	WK	ST	WK	WK
Chen et al [26]	MO	ST	ST	WK	ST	ST	MO
Cherutich et al [27]	MO	ST	WK	WK	WK	ST	WK
Chiou et al [28]	MO	ST	ST	MO	ST	ST	ST
Choko et al [29]	ST	ST	WK	WK	ST	MO	WK
Clark et al [30]	MO	ST	ST	MO	WK	ST	MO
Culbert et al [31]	ST	ST	WK	WK	WK	ST	WK
Dibia et al [32]	MO	WK	WK	MO	ST	N/A	WK
England et al [33]	MO	MO	WK	WK	ST	N/A	WK
Estcourt et al [34]	WK	WK	WK	WK	WK	WK	WK
Estcourt et al [35,36]	MO	ST	ST	WK	WK	MO	WK
Estcourt et al [37,38]	WK	ST	ST	WK	ST	WK	WK
Faxelid et al [39]	MO	ST	ST	MO	WK	N/A	MO
Fetzer and Graeber [40]	MO	MO	WK	MO	ST	N/A	MO
Findlater et al [41]	MO	WK	WK	MO	ST	N/A	WK
Hanrahan et al [42]	MO	ST	ST	MO	ST	ST	ST
Hanrahan et al [43]	MO	ST	ST	WK	ST	WK	WK
Heumann et al [44]	MO	MO	ST	MO	ST	N/A	ST
Hu et al [45]	MO	ST	ST	MO	ST	WK	MO
Huang et al [46]	MO	MO	WK	WK	ST	WK	WK
Jerene et al [47]	MO	WK	WK	WK	ST	WK	WK
José et al [48]	MO	MO	WK	WK	ST	N/A	WK
Joshi et al [49]	MO	MO	WK	MO	ST	N/A	MO
Kagujje et al [50]	MO	MO	WK	MO	ST	WK	WK
Kaswaswa et al [51]	WK	ST	ST	MO	ST	ST	MO
Katz et al [52]	MO	ST	ST	MO	WK	N/A	MO
Kerani et al [53]	WK	ST	WK	MO	WK	MO	WK
Ketema et al [54]	MO	ST	WK	MO	ST	WK	WK
Khatana et al [55]	MO	ST	ST	MO	ST	WK	MO
Kissinger et al [56]	MO	ST	ST	WK	WK	MO	WK
Kissinger et al [57]	MO	ST	ST	MO	ST	ST	ST
Landis et al [58]	WK	ST	WK	WK	WK	N/A	WK
Low et al [59]	MO	ST	WK	MO	MO	MO	MO
Lugada et al [60]	MO	ST	ST	MO	ST	WK	MO
Lukac et al [61]	ST	MO	WK	MO	ST	MO	MO
Luo et al [62]	WK	WK	WK	WK	ST	N/A	WK
Malave et al [63]	ST	MO	WK	MO	WK	N/A	WK

XSL•FO
RenderX

Study ID	Selection bias	Study design	Confounders	Blinding	Data collection	Withdrawals	Overall rating
Malheiro et al [64]	MO	MO	ST	MO	ST	N/A	ST
Martinson et al [65]	MO	ST	ST	MO	ST	ST	ST
Mathews et al [66]	WK	ST	WK	MO	MO	ST	WK
Morishita et al [67]	ST	ST	WK	ST	ST	N/A	MO
Oh et al [68]	MO	MO	WK	MO	WK	N/A	WK
Østergaard et al [69]	WK	ST	ST	MO	ST	MO	MO
Parkes-Ratanshi et al [70]	MO	ST	ST	MO	ST	ST	ST
Peterman et al [71]	ST	ST	ST	WK	ST	ST	MO
Potterat and Rothenberg [72]	MO	ST	WK	WK	WK	N/A	WK
Raymenants et al [73]	MO	MO	WK	MO	WK	ST	WK
Schwebke and Desmond [74]	MO	ST	ST	MO	ST	WK	MO
Shah et al [75]	MO	ST	ST	MO	ST	WK	MO
Tchakounte et al [76]	MO	ST	ST	MO	ST	MO	ST
Wilson et al [77]	MO	ST	ST	MO	ST	ST	ST

[a]MO: moderate.

[b]ST: strong.

[c]WK: weak.

[d]N/A: not applicable.

Deprioritized Studies

Table S2 in Multimedia Appendix 6 displays the 31 studies that were deprioritized because of lack of a concurrent control. Like the prioritized evidence, the deprioritized evidence consisted largely of STI studies, with 9 studies on bacterial infections (eg, chlamydia) [78-86] and 8 studies on HIV [87-94]. In addition, there were 11 studies focused on TB [95-105], 2 on COVID-19 [106,107], and 1 on chronic hepatitis B [108]. None of the 31 deprioritized studies investigated a route of transmission that was not already covered by the prioritized evidence. Sample sizes ranged from 70 [78] to 2,022,127 participants [101], and the median was 919 (IQR 253-3192). Most studies compared an intervention with some variation of standard practice protocol (n=24).

Synthesis

Respiratory or Airborne Transmission

Studies of CT in respiratory transmission focused on TB and COVID-19. Table S1 in Multimedia Appendix 7 summarizes the characteristics, key components, and overview of outcomes for these studies. Further details about the characteristics of interventions evaluated within these studies are available in Table S2 in Multimedia Appendix 7, and detailed outcome data are available in Table S3 in Multimedia Appendix 7. More extensive interpretation is available in the full report [10].

Characteristics of CT Interventions in TB

Of the 15 trials focusing on TB, 6 were cluster RCTs [25,42,51,65,67,75], 3 RCTs [43,54,55], 4 controlled before-and-after designs [46-49], 1 interrupted time series [50], and 1 observational study [23]. The TB studies were conducted outside high-income countries, except for 1 study conducted in Taiwan [46]. Nearly half of the TB studies were rated as being of weak quality, with only 2 rated as strong [42,65].

The interventions in 7 of the 15 TB studies incorporated PICT methods [27,29,47,50,53,54,58], 4 were HI [51,55,65,79], 3 involved a combination of PICT and HI [46,52,71], and 1 involved PR [55]. The majority of TB interventions were community-based, and most aimed to counteract the disease by encouraging people to attend health care facilities for screening or by facilitating access to home-based screening; 11 of the 15 interventions involved home visits [23,42,47-50,55,65,67,75]. Home visits varied in nature, but tended to involve health care staff [23,55] or recruited community members [46-48,50,67] carrying out symptom screening at the home. Some studies also carried out testing in the visits [55,65]. A total of 3 studies involved community members [46,47] or research fieldworkers [65] following up patients' households for monitoring, while incentives were offered to contacts to present for screening in 4 studies [23,25,42,43]. Several studies included training components in which health care workers or key community figures were educated to help carry out the intervention (n=8) [23,46-48,50,54,55,67].

A total of 6 studies specifically focused on improving case finding in younger ages [51,52,54,55,59,75]. Three studies referred contacts aged 5 years or younger irrespective of symptom status [51,52,59]. Three focused their intervention on children aged 15 years or younger [75]. Some interventions were multicomponent, with activities such as attempts to increase awareness of TB [50] and screening children at school [49] being examples of additional components.

TB interventions were usually compared with a form of standard care. Often this was a national program, and in all but 1 comparator this involved passive case finding, whereby potential

XSL•FO
RenderX

cases self-report for testing; the 2023 study by Hanrahan et al [43] used household CT. In 4 studies, the control or baseline program was described as a form of directly observed treatment, short-course (DOTS) program that was adapted or enhanced to improve CT and case finding [23,25,46,49]. It is likely that baseline programs used in other studies resembled DOTS, but they instead described their comparators as a national standard of care [25,42,47,50,55]. One study incorporated CT and active monitoring vs CT and passive monitoring [75].

TB interventions tended to be longer or have longer follow-ups than COVID-19 and STI interventions. The median and most frequently reported intervention duration (or time until last follow-up) was 12 months (n=5) [23,47-49,54], with a range of 3 months [51] to 5 years [25].

While all TB studies involved treatment for contacts with a pharmaceutical in both the intervention and comparator groups, some interventions focused more heavily on improving this aspect compared with the comparator. In particular, 7 studies used various approaches to encourage treatment access or adherence [25,46-48,50,54,65]. These approaches included integration of indigenous case managers, community-led home visits, improving access to chest radiography for young children, and provision of prophylaxis.

Synthesis of Findings of CT in TB

Disease Incidence or Prevalence

Six of 7 studies reporting TB incidence or prevalence observed a decrease in this outcome with CT [23,25,48,50,67,75]. Effect estimates varied in nature; no between-group statistical analyses were performed in 2 studies [50,67], and 4 of these studies were rated as weak quality [23,25,48,50]. Therefore, findings should be interpreted with caution.

All 6 interventions associated with reduced disease incidence or prevalence used home visits, and 3 included active follow-up of contacts [23,25,48]. The 1 study showing no effect also included home visits, but the authors speculated that working in a high-burden area, where motivation for testing and treatment was already high, may explain their null findings [65].

Case Detection

A total of 4 studies reported case detection outcomes, with 2 studies showing greater numbers of cases detected with intervention [49,55] and 2 showing no difference [51,54]. In the controlled before-and-after study by Joshi et al [49] intensified childhood TB case finding yielded 33% more cases after the intervention compared with 16% more in regions without intensified childhood TB case finding. Intensified case finding included activities such as household contact screening and mobile chest health camps, while usual care was poorly described. However, standard DOTS took place separately for TB cases as usual in both intervention and control areas, as per national guidelines [49]. In the quasi-RCT by Khatana et al [55], their active case-finding strategy led to a nearly 4-fold increase in the odds of detecting TB cases (odds ratio [OR] 3.97, 95% CI 1.73-9.11).

Treatment Rates

A total of 6 TB studies reported treatment rates. The 3 studies reporting increased treatment rates were all of weak quality [46,47,54]. Only 1 of these outcomes related to successful completion of treatment [46], while all others recorded initiation of treatment. Two of the 3 studies reporting increased treatment rates involved interventions designed to be culturally relevant or sensitive to the populations of interest, with usual care being generic [46,47]. In both cases, this meant that intervention deliverers understood local customs and dialects and could access people living in remote areas [46,47].

Only 3 studies reported outcomes in more than 1 category [51,54,65], with mixed findings in all instances. No studies reported on the unintended consequences of CT.

Characteristics of CT Interventions in COVID-19

A total of 5 trials evaluating CT in COVID-19 were included [40,41,64,73,76], based in the United Kingdom, Belgium, Portugal, and Cameroon. Two studies were rated as strong quality [64,76], 1 as moderate quality [40], and 2 as weak quality [41,73]. Interventions in 3 studies were categorized as technology-based CT methods [44,45,80], and 2 incorporated PICT [68,77].

The cluster RCT by Tchakounte et al [76] investigated digitized CT using a smartphone app. This intervention involved individuals registering for an app, where upon receiving a positive test, their information was sent to the CT unit, a clinical interview conducted, and automated text messages sent to notify contacts of their exposure to COVID-19. A total of 2 studies reported on the same natural experiment which occurred due to an error in the UK CT system [40,41], when a technical issue caused around 16,000 positive cases to be missed by the national CT system, thereby delaying the tracing and self-isolation of missed cases. Finally, there were 2 cohort studies with concurrent control groups [64,73]. One study evaluated a standard COVID-19 CT program with contact-eliciting interviews and mandatory quarantine [64]. Symptomatic close contacts were then referred for evaluation and testing [64]. The other program tested different durations of tracing windows, with contacts traced if their last interaction with the index case was 2 days before onset or test for the standard window, and 3-7 days in the extended window group [73]. Comparators were standard care protocols including national CT programs [40,41], as well as manual CT [76] and no CT [64,73]. The median intervention period was 1 week, and 3 studies had a 1-week intervention arm [40,41,73]. Intervention durations ranged from 2 days [73] to 23 weeks [76]. Interventions were delivered by local public health authorities [64], government contact tracers [73], and health care workers [76].

Synthesis of Findings of CT in COVID-19

Both studies evaluating the impact of CT on disease prevalence found positive associations [40,73]. Fetzer et al [40], reported there was an association between delayed notifications to quarantine, as a result of a fault in the UK Test & Trace program, and increased disease prevalence and mortality per case traced with a 6-14 day delay (multiple outcomes in Table S3 in Multimedia Appendix 7) [40].

XSL•FO
RenderX

Extending the backward CT window from 2 to 3-7 days led to 42% more cases being detected among contacts [40,73]. However, 3 studies found no impact of CT on case detection rates [41,64,76].

Additional outcomes were reported in the 2 studies evaluating the effect of delayed notification of close contacts due to an error with the UK Test & Trace system [40,41]. In the study by Findlater et al [41], they observed that the delay in CT led to primary contacts of index cases being slightly more likely to be admitted to hospital within 14 days (OR 1.1, 95% CI 1.0-1.2), although there was no impact on patient mortality or hospital admission rates in their secondary contacts. Fetzer et al [40] reported higher mortality rates in the delayed CT group (additional 0.007 deaths per capita; SE 0.003; $P<.05$). No studies reported on the unintended consequences of CT.

Summary

A total of 15 evaluations focused on TB [23,25,42,43,46-51,54,55,65,67,75], and 5 on COVID-19 [40,41,64,73,76]. Studies on TB typically evaluated interventions with the broad aim of encouraging the identification, testing, and treatment of contacts via home visits, active follow-up, incentives, and other measures. There was just 1 study that incorporated CT and active monitoring vs CT and passive monitoring [75]. Findings were inconclusive across outcome domains and studies, but the greatest change associated with intervention was observed for reductions in disease incidence or prevalence. However, the nature of the evidence, which was highly heterogeneous in design and quality, means that only tentative conclusions may be drawn. Five studies of varying designs, quality and outcomes, provide limited evidence about the impact of CT interventions on outcomes in the context of COVID-19.

Sexual or Blood-Borne Transmission

Overview

A total of 35 studies focused on CT in sexually transmitted or blood-borne diseases. Table S1 in Multimedia Appendix 8 provides a broad overview of their key characteristics, trial arm components, and findings. Further detail about the characteristics of interventions evaluated within these studies is available in Table S2 in Multimedia Appendix 8, and detailed outcome data are available in Table S3 in Multimedia Appendix 8.

Characteristics of CT Interventions in STIs

Of the 35 prioritized studies falling into the sexually transmitted route of spread, 13 included chlamydia [21,22,33-38,52,53,56,59,68,69,77], 13 included HIV [24,26-29,31,32,44,45,58,60,62,63], 6 included syphilis [30,39,44,61,70,71], 6 included gonorrhea [34,53,56,68,72,77], 3 included nongonococcal urethritis [35,52,56], 3 included trichomoniasis [39,57,74], 1 included chancroid and lymphogranuloma venereum [39], and 1 did not specify which STIs were included [66]. A total of 7 trials evaluated 2 intervention arms [24,29,34-36,52,70,71], and 2 evaluated 3 intervention arms [30,62].

The duration of the interventions was recorded in 28 studies [22,24,26-31,33-39,45,53,56-58,62,63,66,68-72,74,77]. The duration ranged from as little as 45 minutes for the counseling session in the study by Mathews et al [66] to 13 months, the whole study period that personnel were introduced in the study by England et al [33], but 82% of the interventions were less than 6 months in duration. Most interventions involved a clinic setting (n=26) [22,24,26-30,33,34,37-39,44,45,52,53,56,57,62,63, 66,68,70-72,74,77] or took place within the community (n=6) [21,57,60,61,69,71]. One study was based in a prison with community follow-up of partners [31]. Two were primary care–based [35,36,59], 1 an unspecified health facility [32], and 1 a public health facility [58]. Interventions were delivered by a range of different health care professionals, and in all but 11 studies [24,26,34,52,57,59,60,63,68,74,77] this was the same for both intervention and control.

As the interventions were broadly similar across all conditions within this group, they have been summarized together.

One study involved household contact investigation where all household members of the index case were identified and tested, not solely sexual partners of the index case [60]. In all remaining studies, the CT method involved partner notification (PN), where the sexual partners of patients diagnosed with an STI were identified and informed of their exposure to an infection and the need to be tested and treated [21,22,24,26-39,44,45,52,53,56-59,61-63,66,68-72,74,77]. One study also included notification of social contacts [26]. The majority of the PN methods were patient referral notification, where the index patient was responsible for informing their partners of their risk of exposure and referring them to services for testing (n=17) [21,22,29,30,34-39,53,56,57,59,61,69,70,72,77]. Contact or referral slips were used in 11 of these studies [22,24,26,30,39,56,57,59,70,72,77], and 2 studies trialed the use of a web-based notification service [30,53]. Provider referral (where notification was instigated by a health care professional) was used in 9 studies [24,27,32,44,52,62,63,68,71], and 3 studies involved contract referral, whereby the index case was given limited time to notify partners before the provider stepped in [24,26,74]. In 5 studies, index cases had the choice of PN approach [28,31,33,45,58,66]. In 1 study, provider referral was used if the index case refused to notify their partners [39], and in another the index case could choose which method of PN to use (patient-initiated with nurse support or provider-initiated) [61].

In 4 studies, the nature of the partner elicitation interview with the index patient was varied [44,59,72,74]. A total of 12 studies incorporated reminders or active follow-up as part of their interventions [27,29-31,35-38,45,52,62,68,70,72].

A total of 12 studies adapted the testing process to encourage partners to engage with testing, thereby enhancing CT [21,22,27,29,31,34-38,45,62,69,71]. In 7 of these studies, index cases were provided with sampling kits to pass directly to partners [21,22,29,37,38,45,62,69]. In 2 studies urine sampling was evaluated rather than swab sampling [21,22]. Of the 7 home-based testing interventions, 3 studies provided the means to return samples to the laboratory [21,45,69], whereas partners or index cases were responsible for returning tests to a site in 3 studies [22,29,62]. The study by Choko et al [19] had an additional intervention arm incorporating a financial incentive

XSL•FO
RenderX

for positive partners that attended clinic [29]. Estcourt et al [34-38] conducted 3 studies that involved the provision of sampling kits and treatment either after partner consultation through a telephone hotline [35-38] or consultation with a pharmacist [34-36]. Sampling kits and treatment were provided to the index case for their partners in the 2022/2024 study [37,38], and this was also an option for some partners in the 2015/2016 study, while other partners collected the tests and treatment from clinic reception or from the pharmacist [34-36]. Luo et al [62] incorporated an additional intervention arm where testing was completed as a couple in a clinic setting [62]. Home or field testing was offered to partners in 4 studies [27,31,60,71].

Anonymity of the index case was a feature of the interventions in 8 studies [26,30-32,45,62,67,68], 5 of these were HIV studies [26,31,32,45,62].

New or specialist personnel were introduced to deliver PN services in 5 studies [33,52,59,60,63,74]. In 1 study this was the introduction of 2 public health officers to assist with CT [33], and in another a practice nurse–led PN strategy was evaluated [59]. In 3 studies disease intervention specialists were used [52,63,74]. The introduction of different staff allowed for some studies (n=3) to evaluate the effectiveness of delivering a field-based PN service [52,60,74].

A total of 5 studies incorporated additional counseling and education for the index case, in part providing extra time and support to plan for PN [28,33,39,66,77]. Index cases in the study by Chiou et al [28] also had phone access to a counselor during working hours and through an app or via email 24 hours/day [28].

Most of the study comparators (n=25) involved patient referral PN methods [21,22,24,26,27,29-31,34-38,45,52,53,56-58,60,62,68-70,74,77], and of these, 13 used contact or referral slips, cards, or vouchers as a means of facilitating notification [20,21,23,25,28,29,33,51,57,59,67,69,76]. A total of 3 studies involved provider referral comparators [44,61,63], 2 involved contract referral [71,72], 4 included various PN methods [28,32,33,59]. In 1 study, no detail was given for which method of PN the comparator used [66]. Two comparators involved no CT [27,39]. In the study by Cherutich et al [27], the comparator arm was delayed (by 6 weeks) provider-initiated PN (compared with immediate provider PN in the intervention arm), and outcomes were collected prior to this delayed PN occurring. The study by Faxelid et al [39] did not involve PN.

In 30 studies, the comparator was the usual care or standard of care [21,22,24,26,28-39,44,45,52,53,56-60,66,68-70,72,74,77]. In 11 studies, the intervention group received usual care or standard care plus an additional element in the intervention [28,30,34-38,45,53,66,70,74].

Information (written or online) provided to the index case was a feature of the comparator in 6 studies [30,34-38,68,74]. In 9 studies, the comparator involved a counseling session or interview with the index case, which was face-to-face in 8 studies [24,28,31,52,58,66,72,77] and telephone based in 1 study [44]. The comparator group in 9 studies involved some form of follow-up [28,30,31,37,38,45,58,62,68,70]. This predominantly involved follow-up of the index case (n=7) for feedback and experience on partner notification [28,31,37,38,45,62,68,70]. The comparator groups in 2 studies followed up partners directly [30,58].

Synthesis of Findings of CT Interventions in STIs

Case Detection

Case detection was the most frequently assessed outcome domain, being evaluated in 21 studies [21,24,26-29,31-33,44,45,58-63,69-72]. Seven of these studies observed that interventions were associated with a favorable outcome [26,27,29,31,44,60,69], while the rest observed no difference between intervention arms. In addition, the study by Luo et al [62] provided case detection rates but did not analyze between-group differences. All but 1 of the positive outcomes came from studies of HIV, including in the study by Heumann et al [44], which focused on both HIV and syphilis and only found improved detection of the former. The study by Lugada et al [60] may be considered an outlier in that it used household contact investigation to test for HIV rather than relying on PN. Compared with clinic-based detection, this led to a near 3-fold increase in the odds of case detection (OR 2.76, 95% CI 1.97-3.86; $P<.001$). The study by Ostergaard et al [69] found that home sampling, rather than office sampling, was associated with more infected partners identified per index case (mean difference 1.63 cases; 95% CI 1.1-2.3).

The remaining 5 studies observing positive case detection outcomes used a wide variety of intervention components, and other than the use of active follow-up in 3 studies [26,27,31], they rarely shared common components. Comparators were more passive forms of CT, including passive “self-tell” PN [26,31], telephone rather than in-person contact elicitation interviews [44], delayed notification [27], and referral slips alone [29]. Given that the components used in the successful studies were also used across several other interventions showing no effect, it is unclear whether specific components are associated with success. For example, the studies led by Brown et al [24], Hu et al [45], and Chen et al [26] all used a form of provider-led PN in their interventions, all in HIV samples, but while Chen et al [26] saw an improvement in case detection rates, Brown et al [24] and Hu et al [45] observed no differences.

The magnitude of effect varied widely, with Chen et al [26] finding the ratio of persons referred per index case (ratio 1.9, 95% CI 1.2-3.1) and new cases per index case (ratio 2.0, 95% CI 1.2-3.2) to be nearly doubled with intervention, and Cherutich et al [27] finding a nearly 5-fold greater likelihood of detecting new cases per index (incidence rate ratio [IRR] 5.0, 95% CI 3.2-7.9). Choko et al [29] detected 4.3% more new HIV-positive contacts with HIV self-testing kits and 6.9% more with the addition of an incentive to HIV self-testing [29], while Culbert et al [31] identified 5 new HIV-positive partners compared with 0 with assisted PN vs self-tell alone (no statistics provided) [31]. Finally, Heumann et al [44] found that the number of partners diagnosed with HIV was more than twice as high in the in-person interview arm compared with the telephone interview arm (risk ratio [RR] 2.17, 95% CI 1.04-4.50) [44].

XSL•FO
RenderX

Treatment Rates

A total of 15 studies looked at treatment rates, with 7 finding that interventions were associated with improvements [27,34,39,44,52,56,61]. Two of these studies included people living with HIV [27,44]; the rest were interested in multiple STIs [34,39,52,56,61]. Similarities between successful studies were minimal, with interventions having different approaches and components, all of which were also present in unsuccessful studies. In the study by Cherutich et al [27], treatment rates were more than 4 times higher at 6 weeks in the group that was immediately assisted with PN services compared with the group that received no services (IRR 4.4, 95% CI 2.6-7.4) [27]. The addition of a consultation hotline (24% greater) or consultation with a pharmacist (23% greater) led to greater treatment rates than routine PN in the study by Estcourt et al [34]. In-person interviews, rather than telephone interviews, led to a 19% greater likelihood of partners being treated for syphilis in the study by Heumann et al [44] (RR 1.19, 95% CI 1.03-1.37). Katz et al [52], in their study from 1988, found that active follow-up was associated with increased treatment rates among contacts compared with both an interview (72% vs 22% treatment rates; $P<.001$) and standard nursing referral (72% vs 18%; $P<.001$), while Kissinger et al [56] found that the provision of booklets led to greater odds of contacts of patients with STIs being treated (OR 1.66, 95% CI 1.22-2.27). Faxelid et al [39] observed that individual counseling and assisted PN resulted in a greater proportion of partners of male index patients attending treatments than with standard care (no CT; 85% vs 55%); however, this was not observed in partners of female index patients (65% vs 56%). In the study by Lukac et al [61], patient-initiated PN, with the support of public health nurses, led to better treatment rates than provider-initiated PN (65% vs 59%, no statistics provided).

Disease Incidence or Prevalence

A total of 7 studies evaluated disease incidence or prevalence, with 2 detecting a reduction in incidence or prevalence associated with intervention [56,77], and the rest finding no difference between intervention and comparator arms [35,37,57,66,74]. In the studies observing improved outcomes, Kissinger et al [56] reported that the odds of reinfection among index cases were lower (OR 0.22, 95% CI 0.11-0.44), while Wilson et al [77] observed a 5% drop in reinfection rates of index cases ($P=.02$).

Unintended Consequences

A total of 2 studies assessed unintended consequences. Both were RCTs, 1 from South Africa and 1 from Kenya, and both found that the incidence of intimate partner violence was not affected by the intervention [27,66]. However, Mathews et al [66] observed that abandonment was slightly but significantly more likely following the intervention (risk difference 1.7%, 95% CI 0.2%-3.3%; $P=.02$).

Summary

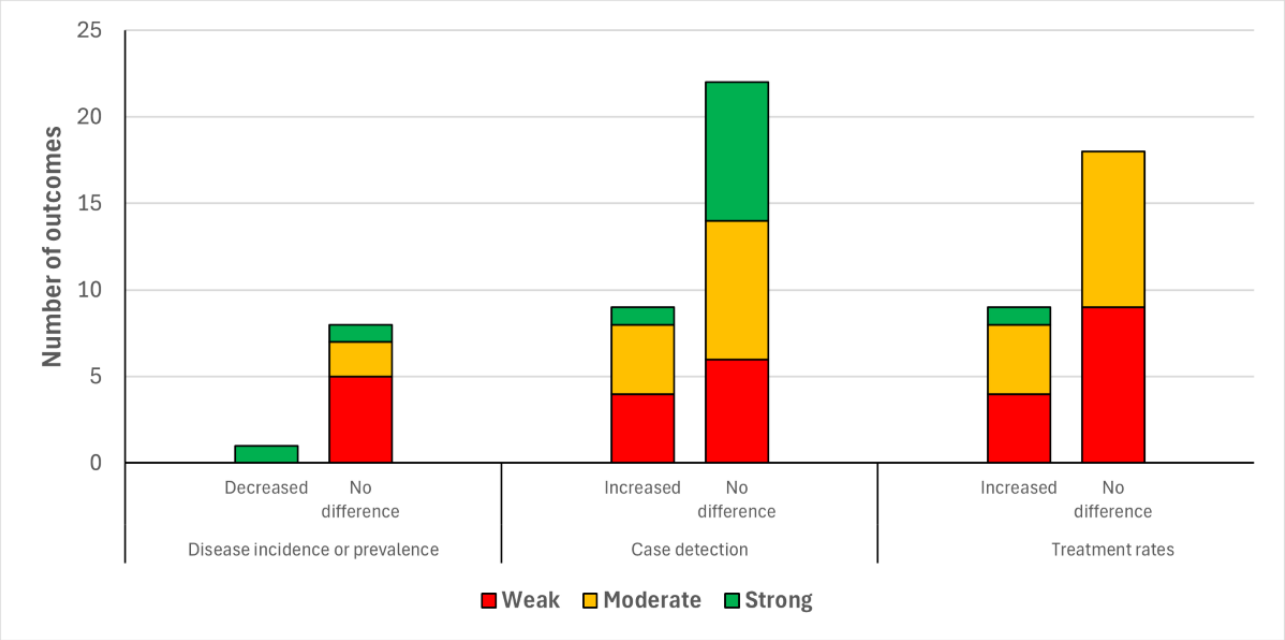
This review included 35 prioritized studies focusing on CT interventions in STIs, including 13 in HIV. In all but 1 study, interventions were a form of PN, which was either provider- or patient-led. Such interventions typically incorporated the use of referral slips, CT, active follow-up, and self-testing kits.

Despite the sizeable body of evidence, there was considerable heterogeneity across study design, intervention components, and outcomes evaluated, and there were only 5 studies rated as being of “strong” quality. Across all effectiveness outcome domains (case detection, disease incidence or prevalence, treatment rates, and other outcomes), it was most often observed that interventions had no effect on outcomes, though this was not exclusively the case, with several instances of improved outcomes, as depicted in Figure 2. Case detection, the most frequently assessed outcome domain, was typical of the pattern of results, featuring in 21 studies, with 7 interventions associated with greater case detection rates and the rest showing no differences compared with their comparators. There were only 2 instances in which outcomes related to unintended consequences were reported: 1 showing no difference between intervention and comparator for the risk of intimate partner violence, and 1 instance in which the intervention was reportedly associated with an unfavorable outcome—an increased likelihood of abandonment. Overall, there were 19 instances of improved outcomes across the 69 outcomes reported, with 1 negative outcome. Positive findings were more likely to be observed when provider referral was compared with patient referral.

There was little consistency in findings or patterns of intervention components that may be more likely to be associated with successful outcomes. The heterogeneity of the evidence, in terms of design (both at the study and intervention level), quality, and outcomes, precludes robust conclusions. However, the overall picture for CT in sexually transmitted or blood-borne diseases is one in which interventions either do no harm or may improve outcomes.

XSL•FO
RenderX

Figure 2. Distribution of effectiveness outcomes across the domains of disease incidence or prevalence, case detection, and treatment rates, for contact tracing (CT) in sexually transmitted or blood-borne infections. Color represents the quality rating of studies providing the outcome measure.



Public and Patient Involvement and Engagement

The review was discussed 3 times with PERSPEX, our standing patient and public involvement and engagement group. This enabled the team to gain insights about CT from patient, carer, and public perspectives. In December 2024, the review team introduced the concept of CT and discussed its use as a public health intervention in general terms. Members drew on their experiences of COVID-19 CT. A majority opinion shared by members was that reviewers should not underestimate the concerns about, and lack of trust in, government agencies that are prevalent in some communities, which could affect the success of CT initiatives. Members were also concerned that some communities may experience additional barriers affecting the outcome of CT. This emphasized the importance of including health equity data in the review protocol. This was taken forward into the review and is available in the full report [10].

In February 2025, PERSPEX received a brief update on the progress of the review. In March 2025, the team reported the emergent main discussion items from the review. PERSPEX members raised a number of matters, which included a query about the role of technology in CT; a view that people respond in different ways to CT and not everyone sees the need to take action; and, for members of PERSPEX whose country of origin was outside the United Kingdom, discussion about the cost of medications even if they are traced and how this could affect the success of CT. Members also spoke about the taboos in many families and cultures around some illnesses and how this could hinder CT and subsequent management within the home. This discussion about culture fed into our thinking about the generalizability of interventions.

Some members shared their experiences of TB management requirements and visa entry for their relatives and visitors from South Asia and Africa. They also discussed UK maternity ward visitor restrictions aimed at preventing the spread of TB. These were shared as positive preventive measures, understood within the community and therefore indicative that CT may be positively received if supported by simultaneous patient education and delivered by community-based teams.

Discussion

Statement of Principal Findings

Our systematic review of the best available evidence relating to the effectiveness of CT identified 57 reports of 55 studies including randomized or prospective nonrandomized comparisons, interrupted time-series designs, or retrospective studies with concurrent controls. A sizeable evidence base drawing on uncontrolled pre-post comparisons was deprioritized given the extent of the higher-quality evidence identified. Included studies related to bacterial or parasitic STIs, HIV, TB, and COVID-19, covering only 2 relevant routes of transmission. Most studies compared different forms of CT against each other. More than half of the included studies had notable methodological limitations, raising an important question about the certainty of our conclusions, though no one critical appraisal domain was identified as driving this. In brief, included studies were inconclusive with respect to the impacts of CT in COVID-19 but provided some tentative evidence for reductions in disease incidence and prevalence in TB attributable to CT. Included studies focusing on STIs and HIV were similarly inconclusive, but some signal of improved effectiveness emerged in the use of provider-led referral vs patient-led referral. Interventions were broadly categorized as including technological methods, HI methods, PR, and PICT, with clear splits based on route of transmission. The heterogeneity of interventions, outcome definitions, and comparators precludes a clear assertion of effectiveness of any one CT strategy.

Comparison With Previous Literature

Although there are some differences in the pool of evidence drawn upon, our findings are broadly in keeping with recent

XSL•FO
RenderX

reviews that covered multiple conditions, led by Hossain et al [2] and Guy et al [3]. Hossain et al [2] focused on PICT as the intervention, with slight differences in definitions of CT and outcomes of interest, while Guy et al [3] included descriptive studies, leading to inclusion of nearly 380 studies, had a broad definition of CT, and included a focus on sociocultural factors influencing uptake and success of CT strategies. Ours was also the only review to consider the unintended consequences of CT. Despite these differences, all 3 reviews concluded that there was some promise for PICT interventions but that evidence was highly heterogeneous and of mixed quality. The varying foci of reviews provide complementary evidence in the quest to develop successful CT strategies suitable for a variety of infectious diseases, such as identifying CT strategies that may be successful across a variety of infectious diseases and learning about the important sociocultural factors influencing their uptake.

Interpretation and Implications

Background Disease Context

We included conditions with epidemic or pandemic potential. In the case of COVID-19, all comparisons drew from a pandemic in progress. Evidence relating to TB generally drew from explicitly endemic contexts. However, a challenge in the evidence related to characterization of the "background" disease context; for example, whether STI CT was implemented in the context of a live outbreak. As a general point, it was nearly always unclear from non–COVID-19 studies how background disease surveillance specifically influenced the design of innovations in CT. Most studies resorted to appeals to health care system efficiency or general disease burden to justify why CT was evaluated.

Intervention Generalizability

Related to the above, a key challenge in the evidence was identifying how different approaches to CT might generalize to "true" pandemic contexts, or across conditions and routes of transmission. The substantial literature relating to bacterial STIs and HIV focused on PN, which implies a specific, and easily identifiable, form of "contact" that may be less relevant in a pandemic driven by airborne transmission, for example. Similarly, evidence relating to TB focused primarily on household contacts. Another facet of generalizability related to intervention context. Our review was global in nature, meaning we included substantial amounts of evidence from the Global South. A challenge, but also an opportunity, for public health in well-resourced settings will be to identify how innovations successful in low-income or middle-income country contexts might be relevant in the Global North [109]. Finally, generalizability may be impacted by the human capital, and other forms of capital, that CT strategies require. Many CT strategies included a focus on training and deploying community health care workers, disease intervention specialists, or community health nurses, or culturally embedded local volunteers. These may not be available at speed and in the required quantities in a pandemic context.

Outcomes

This review did not focus on contact yield as a key outcome, preferring instead measures of disease prevalence and incidence, transmission, and treatment, alongside a parallel focus on unintended consequences. Contact yield is an important outcome because it reflects engagement of contacts with public health services, but it does not capture meaningful impacts on health and well-being of contacts or on disease control. It is an important challenge to the field to consider why an outcome that is primarily about health system outputs, rather than population health impact, is so paramount. This is also important because anecdotally, even where interventions reported improved contact yield, this did not always translate to improved prevalence, incidence, transmission, or treatment outcomes.

About half of all studies in this review reported on a case detection–specific outcome, with treatment rates among contacts a close second. Treatment rates were particularly well described in evidence from CT for bacterial or parasitic STIs but will be less relevant for infectious diseases where treatments either do not exist or are not widely available. The ultimate test of CT effectiveness is whether it reduces, or even eliminates, longer-term incidence or prevalence of the condition, or whether it reduces morbidity or mortality where this is a risk. These outcomes were comparatively rare, and mortality analyses specifically were linked to TB and COVID-19 evidence. In addition, the link between treatment rates and improved mortality is driven by adherence to treatment regimens. In conditions such as TB that require extended courses of therapy, CT-driven improvements in treatment uptake may not translate into improvements in incidence, prevalence, and disease-specific mortality [110].

Unintended Consequences

An innovative feature of our review was the inclusion of unintended consequences of CT. It is notable that despite an exhaustive search for a wide range of unintended consequences, only 2 studies presented evidence in this respect. These studies evaluated CT in sexual health contexts and focused on intimate partner violence or relationship breakdown, with neither finding a difference in intimate partner violence but one finding an increase in relationship abandonment attributable to PN. While these are important, it is likely that the generalizability of these findings to other disease contexts is limited. A broader focus on the potential unintended consequences of health and social interventions is important to fully account for the health impacts of CT, especially for conditions with airborne transmission. For example, in situations where enhanced enrollment into TB treatment is rolled out because of CT, suboptimal concordance with treatment regimens may amplify the challenge of antimicrobial resistance in communities [111].

Unanswered Questions and Future Research

CT strategies in a future pandemic will need to be designed from an understanding of disease transmission pathways. In this light, then, what is the relevance of CT strategies for other conditions, particularly given that we were only able to identify relevant evidence for 2 routes of transmission? Learning from CT for STIs and HIV could support management of stigma and

XSL•FO
RenderX

linkage into treatment, and the use of contact-delivered testing where feasible. Learning from CT for TB could support appropriate services for household contacts, especially where those household contacts might be more susceptible. Finally, learning from CT for COVID-19 could focus on rapidly accelerating CT in the context of a pandemic, in addition to management of contacts where these are an ill-defined set. An important consideration for policymakers and commissioners will be to identify how innovations in CT could support new approaches for conditions, including how these relate to learning from the Global South.

This review will be used to inform the development of UK Health Security Agency advice and guidance as part of pandemic preparedness and to help inform research priorities and future evaluation and improvement approaches to CT. For future preparedness, this means developing flexible CT strategies tailored to specific diseases, investing in higher-quality evaluation methods, and ensuring fairness considerations are built into program design from the outset.

Strengths and Limitations of the Review

An extensive and comprehensive search, selection, and appraisal were undertaken to the highest standards, though as with any systematic review, eligible studies may have been missed. Due to the quantity of identified evidence, we prioritized evidence from evaluations with a concurrent control. Although this was a post hoc decision, we believe this strengthened analysis because it focused on the best available evidence, and the deprioritized studies provided no unique coverage. As a general point, while CT is often presented as a highly prescriptive, medicalized intervention, it was not our experience that this was reflected in included studies. This was both in situations where CT was bundled as part of multicomponent interventions and where specific comparisons undertaken in the context of CT were judged by our team as not relevant. An example of this is partner-delivered treatment, which is largely applicable to bacterial STIs alone. In addition, studies compared CT against heterogeneous comparators, such as usual care and other forms of CT. Usual care was frequently described in scant detail, precluding precision in defining the comparator. Some element of subjectivity, shaped by the challenging and complex nature of CT interventions and their descriptions, may mean that an alternative review group would have included different studies with possibly different overall conclusions. However, we have sought to mitigate this with an auditable and comprehensive account of how judgments were undertaken throughout the review process.

Conclusion

CT is a mainstay of public health response to infectious diseases. Our review found inconsistent evidence for the effectiveness of CT, focused primarily on TB and on contrasts between PICT and patient-led referral in STIs and HIV. While most outcomes did not demonstrate a difference between CT and the comparator intervention, there was some evidence of reductions in disease prevalence in TB and for PICT to be superior to patient-led approaches in STIs. These results do not necessarily reflect a lack of effectiveness of CT strategies. However, they highlight the need to build evaluation into the implementation of CT strategies to improve evidence generation. Heterogeneity of intervention, comparator, and outcome precluded any clear assertions as to optimal strategies for CT, including with respect to relevant subgroups. Future work should consider generalizability of learning across contexts, including by route of transmission and from the Global South; a clearer focus on health equity; and a more thorough account of unintended consequences of CT.

Acknowledgments

We would like to thank Sue Whiffin and Jenny Lowe for administrative support, and the members of PERSPEX for their valuable insights. The authors declare that no generative artificial intelligence was used in any portion of the manuscript generation.

Funding

This review was funded by the National Institute for Health and Care Research (NIHR) Evidence Synthesis Programme (grant number NIHR153420). The views expressed are those of the authors and not necessarily those of the NIHR, the Department of Health and Social Care, or the UK Health Security Agency. Several authors were supported by NIHR Applied Research Collaboration South West Peninsula (PenARC). GJMT is an NIHR senior investigator.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA checklist.
[PDF File (Adobe PDF File), 162 KB - publichealth_v12i1e84805_app1.pdf]

XSL•FO
RenderX

Multimedia Appendix 2
Search strategy.
[DOCX File , 27 KB - publichealth_v12i1e84805_app2.docx]

Multimedia Appendix 3
Inclusion criteria.
[DOCX File , 28 KB - publichealth_v12i1e84805_app3.docx]

Multimedia Appendix 4
Reasons for exclusion at full text.
[DOCX File , 52 KB - publichealth_v12i1e84805_app4.docx]

Multimedia Appendix 5
Search summary table.
[XLSX File (Microsoft Excel File), 27 KB - publichealth_v12i1e84805_app5.xlsx]

Multimedia Appendix 6
Study details.
[DOCX File , 276 KB - publichealth_v12i1e84805_app6.docx]

Multimedia Appendix 7
Tables relating to airborne transmission.
[DOCX File , 223 KB - publichealth_v12i1e84805_app7.docx]

Multimedia Appendix 8
Tables relating to sexual or bloodborne transmission.
[DOCX File , 379 KB - publichealth_v12i1e84805_app8.docx]

References

1. Contact tracing in the context of COVID-19. World Health Organization. 2020. URL: https://www.who.int/publications/m/item/contact-tracing-in-the-context-of-covid-19 [accessed 2024-10-27]
2. Hossain AD, Jarolimova J, Elnaiem A, Huang CX, Richterman A, Ivers LC. Effectiveness of contact tracing in the control of infectious diseases: a systematic review. Lancet Public Health 2022;7(3):e259-e273 [FREE Full text] [doi: 10.1016/S2468-2667(22)00001-9] [Medline: 35180434]
3. Guy D, Kodjamanova P, Woldmann L, Sahota J, Bannister-Tyrrell M, Elouard Y. Contact tracing strategies for infectious diseases: a systematic literature review. PLOS Glob Public Health 2025;5(5):e0004579. [doi: 10.1371/journal.pgph.0004579] [Medline: 40343962]
4. Amicosante AMV, Rosso A, Bernardini F, Guglielmi E, Eugeni E, Da Re F. COVID-19 contact tracing strategies during the first wave of the pandemic: systematic review of published studies. JMIR Public Health Surveill 2023;9:e42678 [FREE Full text] [doi: 10.2196/42678] [Medline: 37351939]
5. Juneau C, Briand A, Collazzo P, Siebert U, Pueyo T. Effective contact tracing for COVID-19: a systematic review. Glob Epidemiol 2023;5:100103 [FREE Full text] [doi: 10.1016/j.gloepi.2023.100103] [Medline: 36959868]
6. Fadlallah R, El-Jardali F, Karroum L, Kalach N, Hoteit R, Aoun A. The effects of public health and social measures (PHSM) implemented during the COVID-19 pandemic: an overview of systematic reviews. Cochrane Evid Synth Methods 2024;2(5):e12055. [doi: 10.1002/cesm.12055] [Medline: 40476261]
7. Ashcroft T, McSwiggan E, Agyei-Manu E, Nundy M, Atkins N, Kirkwood J. Effectiveness of non-pharmaceutical interventions as implemented in the UK during the COVID-19 pandemic: a rapid review. J Public Health (Oxf) 2025;47(2):268-302. [doi: 10.1093/pubmed/fdaf017] [Medline: 40037637]
8. Littlecott H, Herd C, O'Rourke J, Chaparro LT, Keeling M, James Rubin G. Effectiveness of testing, contact tracing and isolation interventions among the general population on reducing transmission of SARS-CoV-2: a systematic review. Philos Trans A Math Phys Eng Sci 2023;381(2257):20230131 [FREE Full text] [doi: 10.1098/rsta.2023.0131] [Medline: 37611628]
9. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. bmj 2020;372:n71. [doi: 10.31222/osf.io/jb4dx]
10. Effectiveness of contact tracing to reduce transmission of infectious disease as part of epidemic or pandemic response: a rapid systematic review. zenodo.org. URL: https://doi.org/10.5281/zenodo.17599575 [accessed 2026-03-11]
11. Glanville J, Lefebvre C, Manson P, Robinson S, Brbre I. ISSG Search Filter Resource.: InterTASC Information Specialists' Sub-Group; 2024. URL: https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home [accessed 2024-11-29]

XSL•FO
RenderX

12. Lefebvre C, Glanville J, Briscoe S, Featherstone R, Littlewood A, Metzendorf M. Cochrane RCT filters. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 65. London: Cochrane; 2024.
13. Glanville J, Dooley G, Wisniewski S, Foxlee R, Noel-Storr A. Development of a search filter to identify reports of controlled clinical trials within CINAHL Plus. Health Info Libr J 2019;36(1):73-90 [FREE Full text] [doi: 10.1111/hir.12251] [Medline: 30737884]
14. Search filters database. Canada's Drug Agency. 2024. URL: https://searchfilters.cda-amc.ca [accessed 2024-11-29]
15. Ovid MEDLINE search filters. University of Pittsburgh. 2024. URL: https://hsls.libguides.com/Ovid-Medline-search-filters/experimental-studies [accessed 2024-11-29]
16. World Health Organization (WHO). Glossary of health emergency and disaster risk management terminology. World Health Organization. 2020. URL: https://www.who.int/publications/i/item/9789240003699 [accessed 2026-03-18]
17. Armijo-Olivo S, Stiles CR, Hagen NA, Biondo PD, Cummings GG. Assessment of study quality for systematic reviews: a comparison of the Cochrane Collaboration Risk of Bias Tool and the Effective Public Health Practice Project Quality Assessment Tool: methodological research. J Eval Clin Pract 2012;18(1):12-18. [doi: 10.1111/j.1365-2753.2010.01516.x] [Medline: 20698919]
18. Operational guidance for investigating and managing STI outbreaks. UK Health Security Agency. 2024. URL: https://www.gov.uk/government/publications/sexually-transmitted-infections-stis-managing-outbreaks/operational-guidance-for-investigating-and-managing-sti-outbreaks [accessed 2025-04-29]
19. Campbell MAO, McKenzie JE, Sowden A, Katikireddi SV, Brennan SE, Ellis S. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. BMJ 2020;368:l6890 [FREE Full text] [doi: 10.1136/bmj.l6890] [Medline: 31948937]
20. Bethel AC, Rogers M, Abbott R. Use of a search summary table to improve systematic review search methods, results, and efficiency. J Med Libr Assoc 2021;109(1):97-106 [FREE Full text] [doi: 10.5195/jmla.2021.809] [Medline: 33424470]
21. Andersen B, Ostergaard L, Møller JK, Olesen F. Home sampling versus conventional contact tracing for detecting Chlamydia trachomatis infection in male partners of infected women: randomised study. BMJ 1998;316(7128):350-351 [FREE Full text] [doi: 10.1136/bmj.316.7128.350] [Medline: 9487169]
22. Apoola A, Beardsley J. Does the addition of a urine testing kit to use of contact slips increase the partner notification rates for genital chlamydial infection? Int J STD AIDS 2009;20(11):775-777. [doi: 10.1258/ijsa.2009.009196] [Medline: 19833688]
23. Bai LQ, Yang HL, Jian XW, He XG, Chen YF, Tang Y. Increasing tuberculosis case detection through intensive referral and tracing in Hunan, China. Int J Tuberc Lung Dis 2008;12(12):1431-1435. [Medline: 19017453]
24. Brown L, Miller W, Kamanga G, Nyirenda N, Mmodzi P, Pettifor A. HIV partner notification is effective and feasible in sub-Saharan Africa: opportunities for HIV treatment and prevention. J Acquir Immune Defic Syndr 2011;56(5):437-442 [FREE Full text] [doi: 10.1097/qai.0b013e318202bf7d] [Medline: 22046601]
25. Cavalcante SC, Durovni B, Barnes GL, Souza FBA, Silva RF, Barroso PF. Community-randomized trial of enhanced DOTS for tuberculosis control in Rio de Janeiro, Brazil. Int J Tuberc Lung Dis 2010;14(2):203-209 [FREE Full text] [Medline: 20074412]
26. Chen JS, Matoga M, Pence BW, Powers KA, Maierhofer CN, Jere E. A randomized controlled trial evaluating combination detection of HIV in Malawian sexually transmitted infections clinics. J Int AIDS Soc 2021;24(4):e25701 [FREE Full text] [doi: 10.1002/jia2.25701] [Medline: 33929094]
27. Cherutich P, Golden MR, Wamuti B, Richardson BA, Ásbjörnsdóttir KH, Otieno FA. Assisted partner services for HIV in Kenya: a cluster randomised controlled trial. Lancet HIV 2017;4(2):e74-e82 [FREE Full text] [doi: 10.1016/S2352-3018(16)30214-4] [Medline: 27913227]
28. Chiou PY, Lin LC, Chen YM, Wu SC, Lew-Ting CY, Yen HW. The effects of early multiple-time PN counseling on newly HIV-diagnosed men who have sex with men in Taiwan. AIDS Behav 2015;19(10):1773-1781. [doi: 10.1007/s10461-015-1007-0] [Medline: 25645329]
29. Choko AT, Fielding K, Johnson CC, Kumwenda MK, Chilongosi R, Baggaley RC. Partner-delivered HIV self-test kits with and without financial incentives in antenatal care and index patients with HIV in Malawi: a three-arm, cluster-randomised controlled trial. Lancet Glob Health 2021;9(7):e977-e988 [FREE Full text] [doi: 10.1016/S2214-109X(21)00175-3] [Medline: 34143996]
30. Clark JL, Segura ER, Oldenburg CE, Salvatierra HJ, Rios J, Perez-Brumer AG, et al. raditional and web-based technologies to improve partner notification following syphilis diagnosis among men who have sex with men in Lima, Peru: pilot randomized controlled trial. J Med Internet Res 2018;20(7):e232 [FREE Full text] [doi: 10.2196/jmir.9821] [Medline: 29970355]
31. Culbert GJ, Levy JA, Steffen AD, Waluyo A, Earnshaw VA, Rahadi A. Impart: findings from a prison-based model of HIV assisted partner notification in Indonesia. J Int AIDS Soc 2023;26(6):e26132 [FREE Full text] [doi: 10.1002/jia2.26132] [Medline: 37339342]
32. Dibia CC, Nwaokoro P, Akpan U, Toyo O, Cartier S, Sanwo O. Innovations in providing HIV index testing services: a retrospective evaluation of partner elicitation models in Southern Nigeria. Glob Health Sci Pract 2024;12(5):e2400013 [FREE Full text] [doi: 10.9745/GHSP-D-24-00013] [Medline: 39317450]

XSL•FO
RenderX

33. England DO, Currie MJ, Bowden FJ. An audit of contact tracing for cases of chlamydia in the Australian Capital Territory. Sex Health 2005;2(4):255-258. [doi: 10.1071/sh05021] [Medline: 16402675]
34. Estcourt C, Sutcliffe L, Cassell J, Mercer CH, Copas A, James L. Can we improve partner notification rates through expedited partner therapy in the UK? Findings from an exploratory trial of Accelerated Partner Therapy (APT). Sex Transm Infect 2012;88(1):21-26. [doi: 10.1136/sti.2010.047258] [Medline: 21795290]
35. Estcourt CS, Sutcliffe LJ, Copas A, Mercer CH, Roberts TE, Jackson LJ. Developing and testing accelerated partner therapy for partner notification for people with genital Chlamydia trachomatis diagnosed in primary care: a pilot randomised controlled trial. Sex Transm Infect 2015;91(8):548-554 [FREE Full text] [doi: 10.1136/sextrans-2014-051994] [Medline: 26019232]
36. Estcourt C, Sutcliffe L, Mercer CH, Copas A, Saunders J, Roberts TE. The Ballseye programme: a mixed-methods programme of research in traditional sexual health and alternative community settings to improve the sexual health of men in the UK. Programme Grants Appl Res 2016;4(20). [doi: 10.3310/pgfar04200]
37. Estcourt CS, Stirrup O, Copas A, Low N, Mapp F, Saunders J. Accelerated partner therapy contact tracing for people with chlamydia (LUSTRUM): a crossover cluster-randomised controlled trial. Lancet Public Health 2022;7(10):e853-e865 [FREE Full text] [doi: 10.1016/S2468-2667(22)00204-3] [Medline: 36182235]
38. Estcourt CS, Mapp F, Woode Owusu M, Low N, Flowers P, Copas A. Improving sexual health through partner notification: the LUSTRUM mixed-methods research programme including RCT of accelerated partner therapy. Programme Grants Appl Res 2024. [doi: 10.3310/trqw3886]
39. Faxelid E, Tembo G, Ndulo J, Krantz I. Individual counseling of patients with sexually transmitted diseases: a way to improve partner notification in a Zambian setting? Sex Transm Dis 1996;23(4):289-292. [doi: 10.1097/00007435-199607000-00008]
40. Fetzer T, Graeber T. Measuring the scientific effectiveness of contact tracing: evidence from a natural experiment. Proc Natl Acad Sci U S A 2021;118(33):e2100814118 [FREE Full text] [doi: 10.1073/pnas.2100814118] [Medline: 34385318]
41. Findlater L, Pierotti L, Turner C, Wensley A, Chen C, Seaman S. Evaluating the impact on health outcomes of an event that resulted in a delay in contact tracing of COVID-19 cases in England, September 2020: an observational study. BMJ Open 2023;13(10):e064982 [FREE Full text] [doi: 10.1136/bmjopen-2022-064982] [Medline: 37827740]
42. Hanrahan CF, Nonyane BAS, Mmolawa L, West NS, Siwelana T, Lebina L. Contact tracing versus facility-based screening for active TB case finding in rural South Africa: a pragmatic cluster-randomized trial (Kharitode TB). PLoS Med 2019;16(4):e1002796 [FREE Full text] [doi: 10.1371/journal.pmed.1002796] [Medline: 31039165]
43. Hanrahan CF, Nonyane BAS, Lebina L, Mmolawa L, Siwelana T, West NS. Household- versus incentive-based contact investigation for tuberculosis in rural South Africa: a cluster-randomized trial. Clin Infect Dis 2023;76(7):1164-1172 [FREE Full text] [doi: 10.1093/cid/ciac920] [Medline: 36458857]
44. Heumann CL, Katz DA, Dombrowski JC, Bennett AB, Manhart LE, Golden MR. Comparison of in-person versus telephone interviews for early syphilis and human immunodeficiency virus partner services in King County, Washington (2010-2014). Sex Transm Dis 2017;44(4):249-254 [FREE Full text] [doi: 10.1097/OLQ.0000000000000583] [Medline: 28282653]
45. Hu QH, Qian HZ, Li JM, Leuba SI, Chu ZX, Turner D. Assisted partner notification and uptake of HIV testing among men who have sex with men: a randomized controlled trial in China. Lancet Reg Health West Pac 2021;12:100171 [FREE Full text] [doi: 10.1016/j.lanwpc.2021.100171] [Medline: 34527967]
46. Huang YW, Tsao CY, Chen WW, Yen PS, Lee JJ, Huang SS. Enhanced directly-observed treatment short-course for tuberculosis control program in mountain areas of Taiwan. J Infect Dev Ctries 2019;13(2):123-129 [FREE Full text] [doi: 10.3855/jidc.10219] [Medline: 32036347]
47. Jerene D, Assefa D, Tesfaye K, Bayu S, Seid S, Aberra F. Effectiveness of women-led community interventions in improving tuberculosis preventive treatment in children: results from a comparative, before-after study in Ethiopia. BMJ Open 2022;12(7):e062298 [FREE Full text] [doi: 10.1136/bmjopen-2022-062298] [Medline: 35863840]
48. José B, Manhiça I, Jones J, Mutaquiha C, Zindoga P, Eduardo I, et al. Using community health workers for facility and community based TB case finding: an evaluation in central Mozambique. PLoS One 2020;15(7):e0236262 [FREE Full text] [doi: 10.1371/journal.pone.0236262] [Medline: 32702073]
49. Joshi B, Chinnakali P, Shrestha A, Das M, Kumar AMV, Pant R. Impact of intensified case-finding strategies on childhood TB case registration in Nepal. Public Health Action 2015;5(2):93-98 [FREE Full text] [doi: 10.5588/pha.15.0004] [Medline: 26400376]
50. Kagujje M, Nyangu S, Maimbolwa MM, Shuma B, Mutti L, Somwe P. Strategies to increase childhood tuberculosis case detection at the primary health care level: lessons from an active case finding study in Zambia. PLoS One 2023;18(7):e0288643 [FREE Full text] [doi: 10.1371/journal.pone.0288643] [Medline: 37467209]
51. Kaswaswa K, MacPherson P, Kumwenda M, Mpunga J, Thindwa D, Nliwasa M. Effect of patient-delivered household contact tracing and prevention for tuberculosis: a household cluster-randomised trial in Malawi. PLoS One 2022;17(9):e0269219 [FREE Full text] [doi: 10.1371/journal.pone.0269219] [Medline: 36074775]
52. Katz BP, Danos CS, Quinn TS, Caine V, Jones RB. Efficiency and cost-effectiveness of field follow-up for patients with Chlamydia trachomatis infection in a sexually transmitted diseases clinic. Sex Transm Dis 1988;15(1):11-16. [doi: 10.1097/00007435-198801000-00003] [Medline: 3358237]

XSL•FO
RenderX

53. Kerani RP, Fleming M, DeYoung B, Golden MR. A randomized, controlled trial of inSPOT and patient-delivered partner therapy for gonorrhea and chlamydial infection among men who have sex with men. Sex Transm Dis 2011;38(10):941-946. [doi: 10.1097/OLQ.0b013e318223fcbc] [Medline: 21934569]
54. Ketema L, Dememew ZG, Assefa D, Gudina T, Kassa A, Letta T. Evaluating the integration of tuberculosis screening and contact investigation in tuberculosis clinics in Ethiopia: a mixed method study. PLoS One 2020;15(11):e0241977 [FREE Full text] [doi: 10.1371/journal.pone.0241977] [Medline: 33211710]
55. Khatana GH, Haq I, Khan SS. Effectiveness, acceptance and feasibility of for tuberculosis contact tracing in Kashmir. J Clin Tuberc Other Mycobact Dis 2019;14:19-25 [FREE Full text] [doi: 10.1016/j.jctube.2019.01.001] [Medline: 31720414]
56. Kissinger P, Mohammed H, Richardson-Alston G, Leichliter JS, Taylor SN, Martin DH. Patient-delivered partner treatment for male urethritis: a randomized, controlled trial. Clin Infect Dis 2005;41(5):623-629. [doi: 10.1086/432476] [Medline: 16080084]
57. Kissinger P, Schmidt N, Mohammed H, Leichliter JS, Gift TL, Meadors B. Patient-delivered partner treatment for Trichomonas vaginalis infection: a randomized controlled trial. Sex Transm Dis 2006;33(7):445-450. [doi: 10.1097/01.olq.0000204511.84485.4c] [Medline: 16531939]
58. Landis SE, Schoenbach VJ, Weber DJ, Mittal M, Krishan B, Lewis K. Results of a randomized trial of partner notification in cases of HIV infection in North Carolina. N Engl J Med 1992;326(2):101-106. [doi: 10.1056/NEJM199201093260205] [Medline: 1445500]
59. Low N, McCarthy A, Roberts TE, Huengsberg M, Sanford E, Sterne JAC. Partner notification of chlamydia infection in primary care: randomised controlled trial and analysis of resource use. BMJ 2006;332(7532):14-19 [FREE Full text] [doi: 10.1136/bmj.38678.405370.7C] [Medline: 16356945]
60. Lugada E, Levin J, Abang B, Mermin J, Mugalanzi E, Namara G. Comparison of home and clinic-based HIV testing among household members of persons taking antiretroviral therapy in Uganda: results from a randomized trial. J Acquir Immune Defic Syndr 2010;55(2):245-252. [doi: 10.1097/QAI.0b013e3181e9e069] [Medline: 20714273]
61. Lukac CD, Consolacion T, Ryan V, Cumming E, Mercado J, Ford G. Population-level outcomes of partner notification among gay, bisexual, and other men who report sex with men diagnosed with infectious syphilis in British Columbia, Canada. Sex Transm Dis 2021;48(12):901-908. [doi: 10.1097/OLQ.0000000000001477] [Medline: 34030159]
62. Luo M, Hann K, Zhang G, Pan X, Ma Q, Jiang J. HIV testing uptake and yield among sexual partners of HIV-positive men who have sex with men in Zhejiang Province, China, 2014-2016: a cross-sectional pilot study of a choice-based partner tracing and testing package. PLoS One 2020;15(6):e0232268 [FREE Full text] [doi: 10.1371/journal.pone.0232268] [Medline: 32497114]
63. Malave MC, Shah D, Sackoff JE, Rubin S, Begier EM. Human immunodeficiency virus partner elicitation and notification in New York City: public health does it better. Sex Transm Dis 2008;35(10):869-876. [doi: 10.1097/OLQ.0b013e31817d2f82] [Medline: 18641535]
64. Malheiro R, Figueiredo A, Magalhães JP, Teixeira P, Moita I, Moutinho M. Effectiveness of contact tracing and quarantine on reducing COVID-19 transmission: a retrospective cohort study. Public Health 2020;189:54-59 [FREE Full text] [doi: 10.1016/j.puhe.2020.09.012] [Medline: 33160088]
65. Martinson NA, Lebina L, Webb EL, Ratsela A, Varavia E, Kinghorn A. Household contact tracing with intensified tuberculosis and human immunodeficiency virus screening in South Africa: a cluster-randomized trial. Clin Infect Dis 2022;75(5):849-856 [FREE Full text] [doi: 10.1093/cid/ciab1047] [Medline: 34950944]
66. Mathews C, Lombard C, Kalichman M, Dewing S, Banas E, Dumile S. Effects of enhanced STI partner notification counselling and provider-assisted partner services on partner referral and the incidence of STI diagnosis in Cape Town, South Africa: randomised controlled trial. Sex Transm Infect 2021;97(1):38-44 [FREE Full text] [doi: 10.1136/sextrans-2020-054499] [Medline: 32482641]
67. Morishita F, Eang MT, Nishikiori N, Yadav RP. Increased case notification through active case finding of tuberculosis among household and neighbourhood contacts in Cambodia. PLoS One 2016;11(3):e0150405 [FREE Full text] [doi: 10.1371/journal.pone.0150405] [Medline: 26930415]
68. Oh MK, Boker JR, Genuardi FJ, Cloud GA, Reynolds J, Hodgens JB. Sexual contact tracing outcome in adolescent chlamydial and gonococcal cervicitis cases. J Adolesc Health 1996;18(1):4-9. [doi: 10.1016/1054-139X(95)00109-6] [Medline: 8750422]
69. Østergaard L, Andersen B, Møller JK, Olesen F, Worm AM. Managing partners of people diagnosed with Chlamydia trachomatis: a comparison of two partner testing methods. Sex Transm Infect 2003;79(5):358-361 [FREE Full text] [doi: 10.1136/sti.79.5.358] [Medline: 14573827]
70. Parkes-Ratanshi R, Mbazira Kimeze J, Nakku-Joloba E, Hamill M, Namawejje M, Kiragga A. Low male partner attendance after syphilis screening in pregnant women leads to worse birth outcomes: the syphilis treatment of partners (STOP) randomised control trial. Sex Health 2020;17(3):214-222 [FREE Full text] [doi: 10.1071/SH19092] [Medline: 32527365]
71. Peterman TA, Toomey KE, Dicker LW, Zaidi AA, Wroten JE, Carolina J. Partner notification for syphilis: a randomized, controlled trial of three approaches. Sex Transm Dis 1997;24(9):511-518. [doi: 10.1097/00007435-199710000-00003] [Medline: 9339968]

XSL•FO
RenderX

72. Potterat JJ, Rothenberg R. The case-finding effectiveness of self-referral system for gonorrhea: a preliminary report. Am J Public Health 1977;67(2):174-176. [doi: 10.2105/ajph.67.2.174] [Medline: 835764]
73. Raymenants J, Geenen C, Thibaut J, Nelissen K, Gorissen S, Andre E. Empirical evidence on the efficiency of backward contact tracing in COVID-19. Nat Commun 2022;13(1):4750 [FREE Full text] [doi: 10.1038/s41467-022-32531-6] [Medline: 35963872]
74. Schwebke J, Desmond R. A randomized controlled trial of partner notification methods for prevention of trichomoniasis in women. Sex Transm Dis 2010;37(6):392-396. [doi: 10.1097/OLQ.0b013e3181dd1691] [Medline: 20453720]
75. Shah L, Rojas Peña M, Mori O, Zamudio C, Kaufman JS, Otero L. A pragmatic stepped-wedge cluster randomized trial to evaluate the effectiveness and cost-effectiveness of active case finding for household contacts within a routine tuberculosis program, San Juan de Lurigancho, Lima, Peru. Int J Infect Dis 2020;100:95-103 [FREE Full text] [doi: 10.1016/j.ijid.2020.09.034] [Medline: 32950737]
76. Tchakounte Youngui B, Mambo A, Machekano R, Kana R, Epée E, Tenkeu SZ. Improving COVID-19 contact tracing and testing of exposed individuals in Cameroon using digital health technology: a cluster randomised trial. EClinicalMedicine 2024;74:102730 [FREE Full text] [doi: 10.1016/j.eclinm.2024.102730] [Medline: 39109192]
77. Wilson TE, Hogben M, Malka ES, Liddon N, McCormack WM, Rubin SR. A randomized controlled trial for reducing risks for sexually transmitted infections through enhanced patient-based partner notification. Am J Public Health 2009;99 Suppl 1(Suppl 1):S104-S110. [doi: 10.2105/AJPH.2007.112128] [Medline: 18556619]
78. Edmiston N, Merritt T, Ooi C. Make contact: a comparative study of contact tracing strategies. Int J STD AIDS 2010;21(6):431-434. [doi: 10.1258/ijsa.2010.010118] [Medline: 20606225]
79. Engelgau MM, Woernle CH, Rolfs RT, Greenspan JR, O'Cain M, Gorsky RD. Control of epidemic early syphilis. Sex Transm Dis 1995;22(4):203-209. [doi: 10.1097/00007435-199507000-00001]
80. Fiumara N, Segal J, Jolly J. Venereal disease contact investigation; a combined military-civilian program. Public Health Rep 1953;68(3):289-294. [doi: 10.2307/4588400]
81. Jones AT, Craig-Kuhn MC, Schmidt N, Gomes G, Scott G, Watson S. Adapting index/partner services for the treatment of chlamydia among young African American men in a community screening program. Sex Transm Dis 2021;48(5):323-328 [FREE Full text] [doi: 10.1097/OLQ.0000000000001325] [Medline: 33137012]
82. Learner ER, Schlanger K, Mauk K, Pham CD, Mukai R, Mulleavey L. Outcomes of traditional and enhanced gonorrhea partner services in the strengthening the US response to resistant gonorrhea project, 2017 to 2019. Sex Transm Dis 2021;48(12S Suppl 2):S124-S130 [FREE Full text] [doi: 10.1097/OLQ.0000000000001527] [Medline: 34407012]
83. Osterlund A, Persson T, Persson I, Lysén M, Herrmann B. mproved contact tracing of Chlamydia trachomatis in a Swedish county-is genotyping worthwhile? Int J STD AIDS 2005;16(1):9-13. [doi: 10.1258/0956462052932809] [Medline: 15705265]
84. Rudy ET, Aynalem G, Cross J, Ramirez F, Bolan RK, Kerndt PR. Community-embedded disease intervention specialist program for syphilis partner notification in a clinic serving men who have sex with men. Sex Transm Dis 2012;39(9):701-705. [doi: 10.1097/OLQ.0b013e3182593b51] [Medline: 22902665]
85. Schleihauf E, Leonard E, Phillips C, Hatchette T, Haldane D, Arnason T. Increase in gonorrhea incidence associated with enhanced partner notification strategy. Sex Transm Dis 2019;46(11):706-712. [doi: 10.1097/OLQ.0000000000001060] [Medline: 31644498]
86. Taylor MM, Mickey T, Winscott M, James H, Kenney K, England B. Improving partner services by embedding disease intervention specialists in HIV-clinics. Sex Transm Dis 2010;37(12):767-770 [FREE Full text] [doi: 10.1097/OLQ.0b013e3181e65e8b] [Medline: 20693936]
87. Fu X, Qi J, Hu Y, Pan X, Li Y, Liu H. Partner notification in cooperation with community-based organizations among HIV-positive men who have sex with men in two Chinese cities. Int J STD AIDS 2016;27(10):821-831 [FREE Full text] [doi: 10.1177/0956462416648827] [Medline: 27140946]
88. Grande M, Mawandia S, Bakae O, Tau L, Mogomotsi G, Ngombo M. Intensified assisted partner notification implementation in Botswana increased partner identification but not HIV case-finding: findings highlight the need for improved data monitoring. J Acquir Immune Defic Syndr 2021;87(3):951-958. [doi: 10.1097/QAI.0000000000002673] [Medline: 34110311]
89. Kantor R, Steingrimsson J, Fulton J, Novitsky V, Howison M, Gillani F. Prospective evaluation of routine statewide integration of molecular epidemiology and contact tracing to disrupt human immunodeficiency virus transmission. Open Forum Infect Dis 2024;11(10):ofae599. [doi: 10.1093/ofid/ofae599] [Medline: 39474444]
90. Kingbo MHKA, Isaakidis P, Lasry A, Takarinda KC, Manzi M, Pringle J. Partner notification approaches for sex partners and children of human immunodeficiency virus index cases in Côte d'Ivoire. Sex Transm Dis 2020;47(7):450-457 [FREE Full text] [doi: 10.1097/OLQ.0000000000001180] [Medline: 32541304]
91. Mahachi N, Muchedzi A, Tafuma TA, Mawora P, Kariuki L, Semo B. Sustained high HIV case-finding through index testing and partner notification services: experiences from three provinces in Zimbabwe. J Int AIDS Soc 2019;22 Suppl 3(Suppl Suppl 3):e25321 [FREE Full text] [doi: 10.1002/jia2.25321] [Medline: 31321918]
92. Tembo TA, Kim MH, Simon KR, Ahmed S, Beyene T, Wetzel E. Enhancing an HIV index case testing passive referral model through a behavioural skills-building training for healthcare providers: a pre-/post-assessment in Mangochi District, Malawi. J Int AIDS Soc 2019;22 Suppl 3(Suppl Suppl 3):e25292 [FREE Full text] [doi: 10.1002/jia2.25292] [Medline: 31321917]

XSL•FO
RenderX

93. Tembo T, Simon K, Kim M, Chikoti C, Huffstetler H, Ahmed S. Pilot-testing a blended learning package for health care workers to improve index testing services in southern Malawi: an implementation science study. J Acquir Immune Defic Syndr 2021;88(5):470-476 [FREE Full text] [doi: 10.1097/QAI.0000000000002796] [Medline: 34483296]
94. Thorpe L, Udeagu CCN, Shah, D, Shepard CW, Bocour A, Guiterrez R. Impact of a New York City health department initiative to expand HIV partner services outside STD clinics. Public Health Rep 2012;127(1):107-114 [FREE Full text] [doi: 10.1177/003335491212700112] [Medline: 22298929]
95. Duarte R, Neto M, Carvalho A, Barros H. Improving tuberculosis contact tracing: the role of evaluations in the home and workplace. Int J Tuberc Lung Dis 2012;16(1):55-59. [doi: 10.5588/ijtld.10.0511] [Medline: 22236846]
96. Dzhangaziev B, Kulzhabaeva A, Truzyan N, Zhoroev A, Otorbaeva D, Temirbekov S. New approach for tuberculosis contact tracing implemented in the two regions of Kyrgyz republic during 2017-2018. J Infect Dev Ctries 2020;14(11.1):109S-115S [FREE Full text] [doi: 10.3855/jidc.11978] [Medline: 33226968]
97. Fatima R, Qadeer E, Yaqoob A, Haq MU, Majumdar SS, Shewade HD, et al. Extending 'contact tracing' into the community within a 50-metre radius of an index tuberculosis patient using Xpert MTB/RIF in urban Pakistan: did it increase case detection? PLoS One 2016;11(11):e0165813 [FREE Full text] [doi: 10.1371/journal.pone.0165813] [Medline: 27898665]
98. Martínez-Campreciós J, Gil E, Aixut S, Moreno M, Zacarias A, Nindia A. Tuberculosis contact tracing, Angola. Bull World Health Organ 2024;102(3):196-203 [FREE Full text] [doi: 10.2471/BLT.23.290068] [Medline: 38420572]
99. Ospina JE, Orcau A, Millet J, Sánchez F, Casals M, Caylà JA. Community health workers improve contact tracing among immigrants with tuberculosis in Barcelona. BMC Public Health 2012;12:158 [FREE Full text] [doi: 10.1186/1471-2458-12-158] [Medline: 22394990]
100. Qader G, Hamim A, Sayedi M, Rashidi M, Manzoor L, Seddiq MK. Addressing tuberculosis control in fragile states: urban DOTS experience in Kabul, Afghanistan, 2009-2015. PLoS One 2017;12(5):e0178053 [FREE Full text] [doi: 10.1371/journal.pone.0178053] [Medline: 28562675]
101. Sanaie A, Mergenthaler C, Nasrat A, Seddiq MK, Mahmoodi SD, Stevens RH. An evaluation of passive and active approaches to improve tuberculosis notifications in Afghanistan. PLoS One 2016;11(10):e0163813 [FREE Full text] [doi: 10.1371/journal.pone.0163813] [Medline: 27701446]
102. Sulis G, Combary A, Getahun H, Gnanou S, Giorgetti PF, Konseimbo A. Implementation of tuberculosis prevention for exposed children, Burkina Faso. Bull World Health Organ 2018;96(6):386-392 [FREE Full text] [doi: 10.2471/BLT.17.201343] [Medline: 29904221]
103. van de Berg S, Erkens C, Mulder C. Tuberculosis contact investigation following the stone-in-the-pond principle in the Netherlands — did adjusted guidelines improve efficiency? Euro Surveill 2021;26(45):2001828 [FREE Full text] [doi: 10.2807/1560-7917.ES.2021.26.45.2001828] [Medline: 34763753]
104. Volkmann T, Okelloh D, Agaya J, Cain K, Ooko B, Malika T. Pilot implementation of a contact tracing intervention for tuberculosis case detection in Kisumu County, Kenya. Public Health Action 2016;6(4):217-219 [FREE Full text] [doi: 10.5588/pha.16.0032] [Medline: 28123956]
105. Zachariah R, Spielmann MP, Harries AD, Gomani P, Graham SM, Bakali E. Passive versus active tuberculosis case finding and isoniazid preventive therapy among household contacts in a rural district of Malawi. Int J Tuberc Lung Dis 2003;7(11):1033-1039. [Medline: 14598961]
106. Adejugbagbe AM, Fatiregun AA, Dosumu MO, Itse O, Akanbiemu FA, Fagbemi S. Lessons learnt from decentralization of COVID-19 response in a southwest state of Nigeria. Global Biosecurity 2022;4. [doi: 10.31646/gbio.158]
107. Djuric O, Larosa E, Cassinadri M, Cilloni S, Bisaccia E, Pepe D. Effect of an enhanced public health contact tracing intervention on the secondary transmission of SARS-CoV-2 in educational settings: the four-way decomposition analysis. Elife 2024;13:e85802 [FREE Full text] [doi: 10.7554/eLife.85802] [Medline: 38416129]
108. Beebeejaun K, Amin-Chowdhury Z, Letley L, Kara E, Mahange B, Harrington K. Impact of a nurse-led enhanced monitoring, management and contact tracing intervention for chronic hepatitis B in England, 2015-2017. J Viral Hepat 2021;28(1):72-79. [doi: 10.1111/jvh.13403] [Medline: 32926589]
109. Corredor Jimenez J, Grimm HM, Ceesay LO, Wondirad M. Learned: what African countries could teach the global north about one health during the pandemics. J Comp Policy Anal Res Pract 2023;25(5):487-505. [doi: 10.1080/13876988.2023.2187698]
110. Sandoval M, Mtetwa G, Devezin T, Vambe D, Sibanda J, Dube GS. Community-based tuberculosis contact management: caregiver experience and factors promoting adherence to preventive therapy. PLOS Glob Public Health 2023;3(7):e0001920 [FREE Full text] [doi: 10.1371/journal.pgph.0001920] [Medline: 37450473]
111. Xi Y, Zhang W, Qiao RJ, Tang J. Risk factors for multidrug-resistant tuberculosis: a worldwide systematic review and meta-analysis. PLoS One 2022;17(6):e0270003 [FREE Full text] [doi: 10.1371/journal.pone.0270003] [Medline: 35709161]

Abbreviations

CT: contact tracing
DOTS: directly observed treatment, short-course
HI: household investigation

XSL•FO
RenderX

IRR: incidence rate ratio
ISSG: InterTASC Information Specialists' Sub-Group
MeSH: Medical Subject Headings
OR: odds ratio
PERSPEX: Public Engagement in Research for Health and Social Policy at Exeter
PHSM: Public Health and Social Measure
PICT: provider-initiated contact tracing
PN: partner notification
PR: partner referral
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO: International Prospective Register of Systematic Reviews
RCT: randomized controlled trial
RR: risk ratio
STI: sexually transmitted infection
TB: tuberculosis
TECH: technology-based contact tracing
WHO: World Health Organization

Edited by A Mavragani; submitted 25.Sep.2025; peer-reviewed by K Li, A Girma, CG Udensi; comments to author 07.Jan.2026; revised version received 12.Feb.2026; accepted 24.Feb.2026; published 31.Mar.2026.

Please cite as:
Nunns M, Febrey S, Becker K, Weiland M, Buckland J, Abbott R, Whear R, Bethel A, Shaw L, Boddy K, Carville S, Harris T, Thompson Coon J, Melendez-Torres GJ
The Effectiveness of Contact Tracing to Reduce Transmission of Infectious Diseases During Epidemic or Pandemic Response: Rapid Systematic Review
JMIR Public Health Surveill 2026;12:e84805
URL: https://publichealth.jmir.org/2026/1/e84805
doi:10.2196/84805
PMID:

©Michael Nunns, Samantha Febrey, Kieran Becker, Morgan Weiland, Jill Buckland, Rebecca Abbott, Rebecca Whear, Alison Bethel, Liz Shaw, Kate Boddy, Serena Carville, Tamsyn Harris, Jo Thompson Coon, G J Melendez-Torres. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 31.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Examining the Factor Structure of Objective Health Literacy and Numeracy Scales: Large-Scale Cross-Sectional Study

Chihiro Moriishi[1], PhD; Keisuke Takano[1], PhD; Takeyuki Oba[1], PhD; Naoki Konishi[1], PhD; Kentaro Katahira[1], PhD; Kenta Kimura[1,2], PhD

[1]Human Informatics and Interaction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Higashi, Tsukuba, Ibaraki, Japan

[2]Faculty of Letters, Hosei University, Chiyoda-ku, Japan

Corresponding Author:
Keisuke Takano, PhD
Human Informatics and Interaction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Higashi, Tsukuba, Ibaraki, Japan

Abstract

Background: Scales for measuring health literacy and numeracy have been broadly classified into performance-based (objective) and self-reported (subjective) scales. Both types of scales have been widely used in research and practice; however, they are not always consistent and may assess different latent constructs. Furthermore, an increasing number of objective measures have been developed, and it is unclear how many latent factors should be assumed.

Objective: This study aimed to examine the psychometric properties and factor structure of items assessing objective health literacy across multiple scales and to clarify which aspects of objective health literacy would be correlated with subjective measures, as well as health behaviors and lifestyles.

Methods: A total of 5 objective scales (72 items in total) were administered to Japanese-speaking adults (N=16,097; women: 7722/16,097, 48%; mean age 54.89, SD 16.46 years). The analyzed scales included items assessing the numeracy, comprehension, and application of health information, some of which were contextualized for specific diseases, such as diabetes and cancer. Participants' responses were submitted to exploratory factor analysis, and individual factor scores were calculated to test correlations with subjective health literacy, health behavior, and lifestyle.

Results: Exploratory factor analysis identified 3 factors, which were interpreted as conceptual knowledge, numeracy, and synthesis. The conceptual knowledge factor consisted of items about medical word comprehension. All numeracy items loaded onto the same factor, even when contextualized for different diseases. The synthesis factor was characterized by items assessing the ability to read and understand health-related information and make judgments on it using one's own knowledge. The identified factors showed high interfactor correlations (r values 0.53 - 0.64) and small-to-moderate correlations with subjective health literacy (r values 0.14 - 0.45). Additionally, each factor indicated small positive correlations with healthy diet and nutrition and lower substance use (r values 0.17 - 0.26).

Conclusions: Our findings suggest that scales of objective health literacy have at least three latent constructs (ie, conceptual knowledge, numeracy, and synthesis) and that disease specificity is not psychometrically prominent. Each factor has some overlap with subjective health literacy, but overall, subjective and objective health literacy should be interpreted as independent constructs, given the small-to-modest correlations.

(JMIR Public Health Surveill 2026;12:e71701) doi:10.2196/71701

KEYWORDS

objective health literacy; objective health numeracy; health-related behaviors; exploratory factor analysis; confirmatory factor analysis

Introduction

Background

Health literacy plays a pivotal role in acquiring and maintaining healthy lifestyles, which help individuals prevent diseases and maintain their well-being [1]. Although the definition of health literacy varies across studies, the core concept refers to the ability of an individual to obtain, process, understand, and use health information and services [2]. This conceptualization covers health numeracy, namely, applying numerical and quantitative reasoning skills to navigate a health care environment, access care, engage in treatment, and make informed health decisions [3]. Empirical studies have demonstrated that lower health literacy, including lower health

XSL•FO
RenderX

numeracy, is associated with lower autonomy and self-control in health behaviors as well as negative health outcomes, such as higher older adult mortality, increased emergency and inpatient facility use, lower medication compliance, and lower preventive service use [3,4].

Health literacy assessment has long been a research target, and hundreds of measures have been developed and published over the past 3 decades [5-8]. As Nguyen [6] noted, a typical assessment approach is to ask respondents to self-report about their experience on Likert scales (ie, subjective measurement), whereas it is also common to challenge individuals using standardized test stimuli to evaluate their underlying traits, knowledge, skills, and numeracy [9-12] (ie, objective measurement). For example, the Lipkus Numeracy Scale (henceforth, Lipkus) requires respondents to perform numeracy tests in general (eg, "Imagine that we rolled a fair, six-sided die 1,000 times. Out of 1,000 rolls, how many times do you think the die would come up even (2, 4, or 6)?") [12]. Another typical approach is to assess word comprehension of health-related and medical terms [13]. It is also common to present responders with hypothetical scenarios or visual materials, such as nutrition labels [14,15] or maps of hospitals [16], to assess their ability to read, interpret, and process relevant information. Objective measures have been suggested to be suitable for estimating individual skills guiding actual health behavior [6]—an experimental study showed that individuals with high levels of objective (but not subjective) health literacy were able to critically evaluate health information on websites, which further helped them to choose an appropriate treatment option [17]. In addition, a prospective cohort study on patients with cardiovascular-related diseases showed that the lack of objective health literacy predicted poor refill adherence [18].

In contrast, most subjective measures ask respondents to self-report their perceptions and experiences of handling health information, typically using a Likert scale [6]. The 47-item European Health Literacy Survey Questionnaire (HLS-EU-Q47) [19] is one of the most widely used measures to assess individuals' perceived abilities to access, understand, appraise, and apply health information (eg, "Finding information on symptoms of illnesses that concern you is..."; respondents indicate from very easy to very difficult) [19]. Another example is the Subjective Numeracy Scale (SNS), which assesses individuals' beliefs about their skill in performing various mathematical operations (eg, "How good are you at working with fractions?") and individuals' preferences regarding the presentation of numerical information (eg, "When reading the newspaper, how helpful do you find tables and graphs that are parts of a story?") [20]. Subjective measures typically assess individuals' self-perceived ability to find and understand health-related information as well as their confidence in doing so [17]. Also, some measures cover a wider range of psychological (eg, motivation and self-efficacy) aspects of health literacy [21]. A study suggested that individuals with lower levels of subjective numeracy are less motivated and less confident in numeric tasks [22]. Furthermore, the European Health Literacy Survey showed that subjective (but not objective) health literacy is predictive of self-perceived health [23], which might suggest that subjective measures may be more suited to studying perception and beliefs about health status and behavior.

The objective and subjective measures appeared to tap into the same latent construct, that is, the ability to process health information. However, Waters et al [24] suggested that these 2 types of measures assess conceptually related but psychometrically distinct constructs and that numeracy should be separated from general health literacy. Begoray and Kwan [25] found almost null correlations between objective (word recognition and reading comprehension) and subjective (self-reporting of skills to access and communicate health information) assessments. Marks et al [26] suggested that objective measures may reflect medication knowledge, whereas subjective measures may not. For the associations with health outcomes and behaviors, a systematic review [27] concluded that the evidence is mixed. Several studies observed no differences between performance-based and self-reported health literacy for the associations with relevant health outcomes (eg, diabetes, stroke, and hypertension), whereas others documented objective-subjective discrepancies (eg, for cancer screening use). Hirsh et al [28] noticed that the self-reported disease severity of rheumatoid arthritis was associated with subjective health literacy but not with objective health literacy, including the ability to read and pronounce medical terms.

The possibility that objective and subjective measures assess different constructs of health literacy may make it difficult for researchers and practitioners to determine which type (or both) to include in their assessment batteries. Another challenge when building an assessment battery for health literacy research is that an enormous number of measures have been developed; thus far, there is no clear guidance on which to use and when [8]. Recently, we conducted an exploratory factor analysis of 219 items across 11 subjective measures (encompassing 45 subscales), indicating that dimension reduction was effective, as the items were well explained by 7 latent factors [29].

Objectives

In this study, we aimed to expand these findings to objective health literacy measures; namely, we conducted an exploratory factor analysis on 5 performance-based measures of health literacy and numeracy (see the *Methods* section for the selection criteria of the analyzed scales), including general and disease-specific (ie, chronic pain, cancer, and diabetes) scales. Through the analyses, we explored how many and what factors would emerge. In addition to the number of factors identified, we were also interested in whether disease-specific items would be recognized as independent factors or factors that reflect common skills and performances regardless of target diseases. Simultaneously, the identified factors were tested for their correlations with lifestyle and health status, as well as subjective health literacy and numeracy, to explore the consistencies and inconsistencies (or validity) with perceived health literacy and behaviors.

XSL•FO
RenderX

Methods

Data

Data from a larger longitudinal survey on the health behaviors, psychological characteristics, and lifestyles of Japanese-speaking adults (aged >18 y living in Japan) were used. We used quota sampling to represent the population distribution for age and gender in Japan, and thus, we did not use a survey weight in the analysis. The overarching project (still ongoing) is a 3-year longitudinal study that includes multiple waves with different focuses: wave 1 (N=20,573; early 2023) for physical activity (PA) and psychological characteristics [30] and for mobile health technology use [31], wave 2 (conducted in 2023; 6 mo after wave 1) for changes in PA and digital health behaviors [32], and wave 3 (conducted in early 2024) for health literacy and lifestyle. Wave 3 included both subjective and objective health literacy scales; the psychometric properties of the subjective scales have been reported elsewhere [26]. This study used the wave 3 data (N=16,097; women: 7722/16,097, 48%; mean age 54.89, SD 16.46 years), of which 87% (14,064/16,097) participated in wave 1. As the dropout rate was high, an additional sample of 2033 participants was recruited at wave 3. This addition was for the overarching project but not for this study specifically. Although we could not use quotas in this extra sampling due to the time pressure that we had, we found that the age and gender distributions were similar to those of the general population, so we included this additional sample in the analysis. This study focused exclusively on objective scales. We used data from 5 objective health literacy (or numeracy) scales together with the validation measures of subjective health literacy, health behavior, and lifestyle (refer to the *Measures* section).

Ethical Considerations

Participants were paid for online panels recruited by a survey firm. Interested individuals followed a link to the survey site, and on the top page, they received study information (written) and provided informed consent to proceed to individual questionnaire pages. Each participant was assigned a study ID, which was used as the key in merging their responses across different waves. No personal information was obtained throughout the study. The study was approved by the Ethics Committee of the National Institute of Advanced Industrial Science and Technology (approval ID: 2022 - 1279).

Measures

Objective Health Literacy and Numeracy Scales

We selected the scales for inclusion in this study following published reviews (eg, [8,33,34], including Tavousi et al [8], the latest review on health literacy scales over the past 3 decades when the study was conceptualized, and Nakadai et al [34], a narrative review of the scales available in Japanese). Among the scales listed, we included those that met the following criteria: the scale was available in English or Japanese and could be implemented on a static online survey (ie, did not require audiovisual materials or in-person interactions), and specific instructions and items were available from published articles, supporting materials, or personal correspondence with the authors of the scales. This selection process resulted in four objective health literacy or numeracy scales: the Lipkus [12,35], Newest Vital Sign (NVS) scale [14,15], Functional Health Literacy Scale for Young Adults (funHLS) [13], and Cancer Health Literacy Test (CHLT) scale [16]. An additional database search (Google Scholar and PubMed) identified the Diabetes Health Numeracy (DHN) scale [36], which was eligible for this study. Table 1 summarizes the characteristics of each included scale; most objective health literacy scales are not Likert type. For example, the funHLS presented medical stem terms (eg, caries) and asked participants to indicate the most relevant words for each stem term among 3 response options (eg, virus, bacteria, and fungus). Across the scales, each response was binary coded to represent 1 (correct) and 0 (incorrect), and the total score was calculated for each scale, with higher values indicating higher levels of objective health literacy or numeracy. It should be noted that the current analyses included translated versions of the scales, and responses to some of the items were potentially affected by cultural differences. For example, the NVS and CHLT included items assessing comprehension of food nutrition and prescription medication labels. These stimuli were modified to be familiar to Japanese respondents—particularly for the NVS, the translation and adjustment were conducted rigorously in accordance with the established cross-cultural adaptation guidelines [37,38].

XSL•FO
RenderX

Table . Overview of the objective health literacy and numeracy scales.

Scale name (abbreviation)	Items, n (Cronbach α)	Test format	Description
Lipkus Numeracy Scale (Lipkus) [12,35]	10 (0.79)	Numeric response questions	Measures the ability to understand and use numeric information, particularly for probability: (eg, "Imagine that we rolled a fair, six-sided die 1,000 times. Of 1,000 rolls, how many times do you think the die would come up even (2, 4, or 6)")?
Newest Vital Sign (NVS) scale [14,15]	6 (0.63)	Numeric response questions and open-ended questions	Measures comprehension, numeracy, and application and evaluation skills. Responders are presented with a nutrition label of ice cream, from which they are required to extract necessary information for calculation (eg, "If you eat the entire container, how many calories will you eat?") and evaluation (eg, "Pretend that you are allergic to the following substances: penicillin, peanuts, latex gloves, and bee stings. Is it safe for you to eat this ice cream?")
Functional Health Literacy Scale for Young Adults (funHLS) [13]	19 (0.93)	Multiple choice questions	Measures knowledge and comprehension of health-related and medical terms. Responders are presented with stem words, for each of which they are asked to indicate the most relevant among 3 response options (eg, stem=caries: response options=virus, bacteria, and fungus).
Diabetes Health Numeracy (DHN) scale [36]	7 (0.85)	Multiple choice questions	Measures numeracy skills, contextualized for diabetes (eg, "If you walk for about 30 minutes you can burn 100 calories. If you want to burn 150 calories, how long do you have to walk?"). Several items tap into interpretation skills (eg, "read a table about diagnostic criteria for diabetes and indicate the stage of an example patient").
Cancer Health Literacy Test scale (CHLT) [16]	30 (0.85)	Multiple choice questions	Measures knowledge (eg, "Which is the highest in calories and protein? – French fries, cheeseburger, hard-boiled egg"), comprehension skills (eg, "In people who develop oral cancers, 25% of these cases occur in the tongue. Oral cancer occurs in the tongue..."), and their synthesis, contextualized for cancer.

Subjective Health Literacy Scale

The HLS-EU-Q47 [19,39] was used to assess subjective health literacy. The HLS-EU-Q47 and other self-report scales (see below) were used as validation measures to test for correlations with objective health literacy measures. The HLS-EU-Q47 measures 4 information-processing competencies (ie, how easy it is to access, understand, appraise, and apply health information) for 3 health-relevant domains (ie, health care, disease prevention, and health promotion). Participants indicated how applicable each item was to them using a 4-point scale (1=very easy and 4=very difficult). For ease of interpretation, each item was reverse scored, with higher values indicating higher health literacy levels, and the total score was normalized to a range between 0 and 50 using the following formula: (mean−1)×(50/3). This scale has shown good reliability in the current data (Cronbach α=0.97).

Subjective Health Numeracy Scale

The SNS was used to assess subjective health numeracy levels [20]. The SNS measures one's perceived ability to perform mathematical tasks (eg, How good are you at working with fractions?) and preferences for the use of numerical (vs prose) information (eg, When reading the newspaper, how helpful do

XSL•FO
RenderX

you find tables and graphs that are parts of a story?). Participants indicated how applicable each item was to them using a 4-point scale (1=not good at all, not helpful at all; 4=very good, very helpful). This scale has shown good reliability in the current data (Cronbach α=0.75).

Physical Activity

The International Physical Activity Questionnaire Short Form [40,41] was used to assess PA levels. Respondents were asked to indicate the number of days and minutes per day spent walking, engaging in moderate-intensity activities, and engaging in vigorous-intensity activities. We did not use sedentary time for the current analyses. The weighted sum of the reported durations was calculated across the 3 activity categories, representing the total PA in the form of metabolic equivalents (METs hours per week). According to the Ministry of Health, Labour and Welfare in Japan, the recommended amount is 23 METs hours per week or higher for adults aged <65 years and 10 METs hours per week for older people [42].

Quality of Life and Health State

Quality of life (QoL) and health status were assessed using the 5-level EuroQol 5-Dimension (EQ-5D) version [43]. Participants indicated their health status by selecting the most appropriate statement (ie, no problems to extreme problems) for the following five dimensions: mobility, self-care, usual activities, pain or discomfort, and anxiety or depression. Participants' responses were combined into a 5-digit code, which was then converted into a numerical QoL score. The QoL score ranges from −0.025 to 1, where a negative value signifies a condition worse than death, 0 represents a state equivalent to death, and 1 denotes the highest possible health utility. At the end of the EQ-5D questions, participants were asked to rate their health status using a visual analog scale, ranging from 0 to 100, with 0 representing the worst health condition they could imagine and 100 representing the best health condition they could imagine.

Health-Related Lifestyles

The Short Multidimensional Inventory Lifestyle Evaluation (SMILE; [44]) consists of 45 items covering seven domains of health-related lifestyles: diet and nutrition, substance use, PA, strategies to deal with stress, sleep pattern, social support, and environmental exposure. Items asking about the use of illegal drugs (ie, items 10 and 11) were excluded to adhere to the ethics standards of the administering survey firm, and the remaining 43 items were used in the survey. Participants rated each item on a 4-point scale (1=always and 4=not at all). Summed scores were calculated for each domain, whereas items were reverse scored (with higher values indicating healthier lifestyles). The global score (sum of the 7 domains) demonstrated good internal consistency in the current data (Cronbach α=0.88).

Statistical Analysis

An exploratory factor analysis was conducted on the 5 objective health literacy and numeracy scales. We excluded from the analysis (1) an item (funHLS12) exhibiting high correlations with other items (r values 0.72 - 0.80) and (2) 5 items to which >90% of participants responded correctly (ie, items 2, 4, 14, and 27 of the CHLT and item 5 of the Lipkus). The final dataset consisted of 66 items. As each item was binary scored (correct vs incorrect), polychoric correlations were calculated and used in factor analysis. The number of factors was determined based on the reduction in eigenvalues (ie, a scree plot), as well as on the interpretability of the identified factors. Exploratory factor analysis was conducted on randomly sampled 70% of the data (n=11,268), and the remaining 30% (n=4829) was used for confirmatory factor analysis as testing data. Before factor analyses, each dataset was tested with the Kaiser-Meyer-Olkin sampling adequacy measure (≥0.8; Kaiser 1970; [45]) and Bartlett sphericity test (P≤.05; [46]). Confirmatory factor analysis was conducted with maximum likelihood estimation to replicate the factor structure obtained in the exploratory factor analysis. Our focus was on the goodness of fit of the model to the data, evaluated by the following indices: chi-square [47], comparative fit index [48], root mean square error of approximation [47,49], and the standardized root mean square residual [47]. For each factor, items with factor loadings of 0.40 or greater (a commonly used threshold for identifying meaningful loadings; eg, see [50]) were interpreted and were used to calculate a factor score (as the mean of raw item scores). For each factor, items with factor loadings of 0.40 or greater were interpreted and were used to calculate a factor score (as the mean of raw item scores). These factor scores were tested for correlations with validation measures (ie, subjective health literacy and numeracy scales, PA, QoL, health status, and health-related lifestyles). All analyses were performed using R (version 4.3.3; R Foundation for Statistical Computing). The *factanal* function was used for the exploratory factor analysis, and the *cfa* function of the *lavaan* package [51] was used for the confirmatory factor analysis.

Results

Descriptive Information

Table 2 presents the descriptive statistics. For the objective measures, the mean scores were comparable to those reported in previous studies—for example, in the general Japanese population (Lipkus, mean 9.6) [35], an Italian population-based sample (NVS, mean 4.1) [52], and a sample from the United States (CHLT, mean 22.3) [16]. The total score on the HLS-EU-Q47 was slightly higher than that reported among Japanese people (mean 25.3) but lower than that reported among Europeans in the literature (mean 33.8) [39].

XSL•FO
RenderX

Table . Descriptive statistics (n=16,097).

Variable	Values
Age (y), mean (SD)	54.89 (16.46)
Gender (women), n (%)	7722 (48)
Objective health literacy and numeracy scales, mean (SD)	
Lipkus[a]	7.80 (2.34)
NVS[b]	3.50 (1.69)
funHLS[c]	14.14 (5.21)
DHN[d]	5.38 (2.09)
CHLT[e]	24.67 (4.89)
Subjective health literacy and numeracy scales, mean (SD)	
HLS-EU-Q47[f]	28.23 (8.07)
SNS[g]	3.24 (0.67)

[a]Lipkus: Lipkus Numeracy Scale.

[b]NVS: Newest Vital Sign scale.

[c]funHLS: Functional Health Literacy Scale for Young Adults.

[d]DHN: Diabetes Health Numeracy scale.

[e]CHLT: Cancer Health Literacy Test scale.

[f]HLS-EU-Q47: 47-item European Health Literacy Survey Questionnaire. A general health literacy index score comprising all items was standardized on a metric between 0 and 50, using the following formula: (mean − 1) × (50/3).

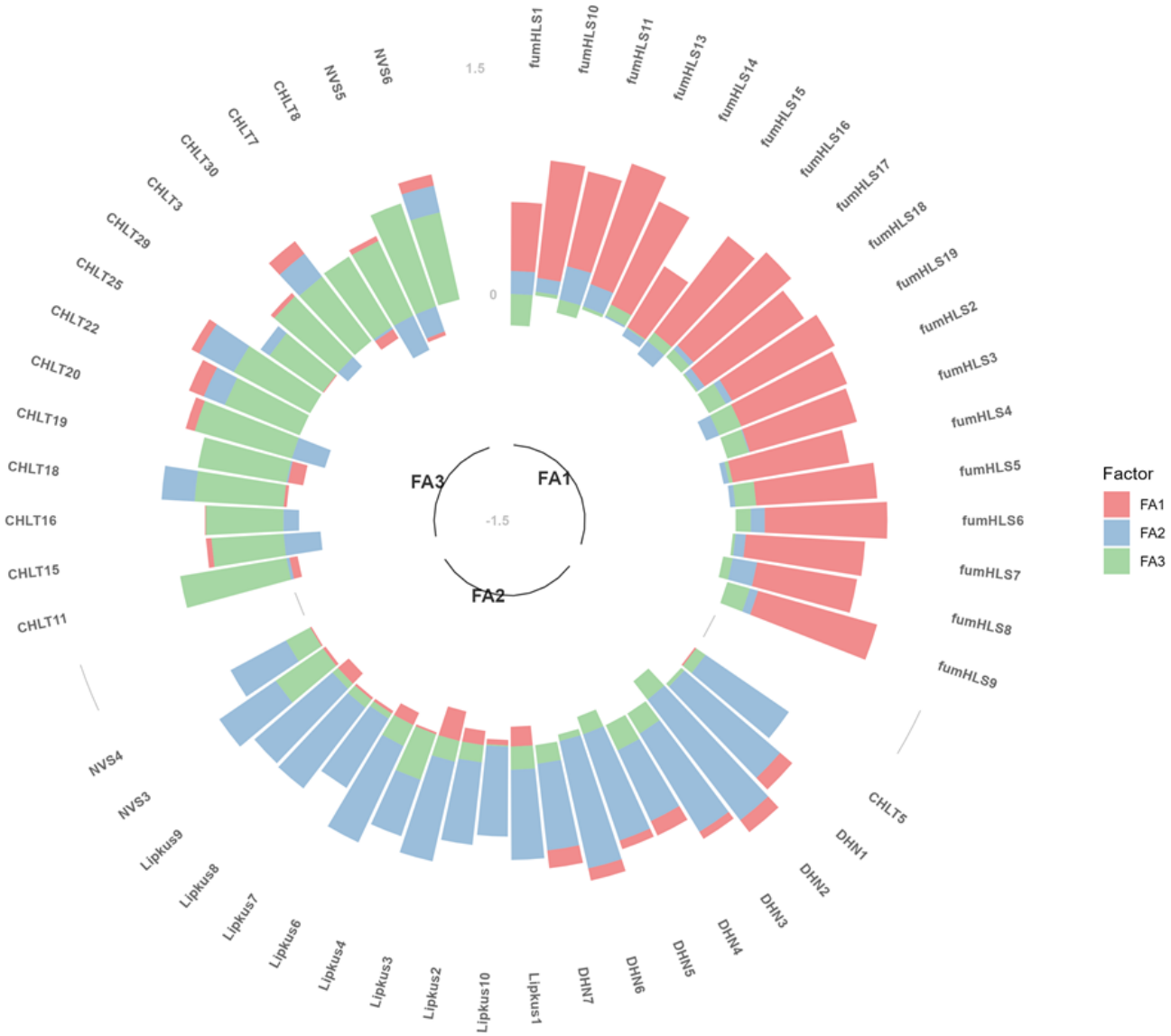
[g]SNS: Subjective Numeracy Scale.

Exploratory Factor Analysis

The factor analysis performed on 66 items across 5 scales revealed eigenvalues of 16.43, 3.09, 1.93, and 1.59 for the 1- to 4-factor solutions. The reduction in the eigenvalue supported the 3-factor solution, with explained variances of 0.16, 0.15, and 0.10 for the 3 factors (total explained variance: 0.41). Additionally, the 3-factor solution had good interpretability; the factor loadings are visualized in Figure 1, which confirms that no items had double or triple loadings. The exact factor loadings for each item are listed in Table S1 in Multimedia Appendix 1. Table 3 summarizes the characteristics of each factor.

XSL•FO
RenderX

Figure 1. Items' factor loadings on each factor. CHLT: Cancer Health Literacy Test scale; DHN: Diabetes Health Numeracy scale; FA: factor; funHLS: Functional Health Literacy Scale for Young Adults; Lipkus: Lipkus Numeracy Scale; NVS: Newest Vital Sign scale.



XSL•FO
RenderX

Table . Interpretations of identified factors.

Factor[a] and item with a factor loading of 0.40 or higher	Example item
FA1 (conceptual knowledge)	
funHLS[b] (items 1 - 11 and 13 - 19)	funHLS 6: "Indicate the most relevant word for *Vitamin C*. Response options*: Vegetables, Fat, Grain, and I Don't know.*"
FA2 (numeracy)	
Lipkus[c] (items 1 - 4, 6 - 10)	Lipkus 6: "If Person A's risk of getting a disease is 1% in ten years, and person B's risk is double that of A's, what is B's risk?"
NVS[d] (items 3 and 4)	NVS 3: "Your doctor advises you to reduce the amount of saturated fat in your diet. You usually have 42 g of saturated fat each day, which includes 1 serving of ice cream. If you stop eating ice cream, how many grams of saturated fat would you be consuming each day?"
DHN[e] (items 1 - 7)	DHN 2: "A male diabetic patient weighs 80 kilograms (kg). The doctor advised this patient to lose 10% of his weight. How much weight does this patient need to lose?"
CHLT[f] (item 5)	CHLT 5: "In people who develop oral cancers, 25% of these cases occur in the tongue. Oral cancer occurs in the tongue..."
FA3 (synthesis)	
NVS (items 5 and 6)	NVS 5: "Pretend that you are allergic to the following substances: Penicillin, peanuts, latex gloves, and bee stings. Is it safe for you to eat this ice cream?"
CHLT (items 3, 7 - 8, 11, 15 - 16, 18 - 20, 22, 25, and 29 - 30)	CHLT 18: "An appointment card says not to eat or drink anything 9 hours prior to the appointment. Sally has an appointment at 11:15 a.m. on Friday. What time should she stop eating or drinking?"

[a]The means and SDs of each item as well as their factor loadings are provided in the supplementary materials in Multimedia Appendix 1.

[b]funHLS: Functional Health Literacy Scale for Young Adults.

[c]Lipkus: Lipkus scale.

[d]NVS: Newest Vital Sign scale.

[e]DHN: Diabetes Health Numeracy scale.

[f]CHLT: Cancer Health Literacy Test scale.

Factor 1 (FA1) consisted exclusively of items from the funHLS, which asked participants to indicate the word most relevant to a stem (medical) word. Items from the funHLS assess word comprehension and knowledge about diseases and symptoms that young adults often experience, as well as nutrition, diet, and human biology. All items of the funHLS showed loadings of >0.40 on FA1. Although the funHLS items covered a range of topics (eg, caries, depression, and BMI), most items loaded on the same factor, and items from other scales were not included in FA1. This factor could be interpreted as a conceptual knowledge of health-related and medical terms in general (ie, not limited to a particular disease or health condition); however, it is still possible that the factor may reflect the unique test format, as the other scales require binary (true-false) responses or numeric responses, for example, to calculate a probability or health risk.

Factor 2 (FA2) included items from 4 of the 5 analyzed scales (ie, Lipkus, NVS, DHN, and CHLT), representing performance-based health numeracy in general (eg, "Imagine that we rolled a fair, six-sided die 1,000 times. Out of 1,000 rolls, how many times do you think the die would come up even (2, 4, or 6)?"). These 4 scales target different populations—Lipkus was designed for the general population, whereas the other 3 were contextualized for particular diseases and health conditions (DHN for diabetes and CHLT for cancer). The test format also differed across the 4 scales; the CHLT used multiple-choice questions, whereas the NVS and DHN included numeric response questions. These results suggest that the items assessing performance-based numeracy correlate well with each other, regardless of heterogeneity in the target diseases and test format.

Factor 3 (FA3) included items from 2 scales, the NVS and CHLT, which assess the ability to process and synthesize

XSL•FO
RenderX

health-related information. For example, item 5 of the NVS concerns abstract reasoning, integrating reading, comprehending, and interpreting skills as applied to material with health content [15]. Respondents were presented with a hypothetical nutrition label of ice cream and asked to judge whether the ice cream would be safe if the respondents were allergic to the indicated substances. Similarly, many of the items loaded onto FA3 required respondents to comprehend and synthesize the presented information (eg, the nutrition label) to make the correct response. Items from the CHLT are contextualized in a daily cancer patient routine at a clinic (eg, instructions for the use of medicines and reading a floor map of a hospital), assessing respondents' knowledge, numeracy, navigation, and synthesis [16]. Therefore, compared to FA1 (word comprehension and knowledge) and FA2 (numeracy), FA3 is distinguished in that it broadly measures higher order skills that require the synthesis of multiple skills (eg, reading, comprehension, and interpretation) to apply in a daily health context.

Confirmatory Factor Analysis

The testing dataset was found suitable for factor analysis: Kaiser-Meyer-Olkin=0.98 and the Bartlett Test of Sphericity, $P \leq .001$. We built a confirmatory factor analysis model with the 3 factors identified through exploratory factor analysis. This model showed an excellent fit to the testing data, $\chi^2_{1271}=7015.7$, comparative fit index=0.97, root mean square error of approximation=0.03, and standardized root mean square residual=0.05, which reassures that the analyzed scales can be reduced to the 3 factors.

Correlation Analysis

The 3 identified factors were tested for their correlations with subjective health literacy and numeracy, as well as with health status and lifestyle (Table 4). Each correlation was interpreted for magnitude but not for statistical significance, given the large sample size of the analyzed dataset. The Cohen guideline was used, with r=0.10, 0.30, and 0.50 being interpreted as small, moderate, and large effects, respectively [53,54]. FA1 to FA3 showed large interfactor correlations. However, these factors showed small-to-moderate correlations with the HLS-EU-Q47 (subjective health literacy), SNS (subjective numeracy), and SMILE (subscales of diet, nutrition, and substance use). Moreover, FA1 and FA2 showed small correlations with SMILE sleep and social support (r values 0.10 - 0.13). None of the factors showed interpretable size correlations with the International Physical Activity Questionnaire Short Form (total PA) or EQ-5D (QoL and subjective health) scores. The 2 subjective measures, the HLS-EU-Q47 and SNS, presented stronger correlations with the SMILE subscales, except for substance use, than FA1 to FA3.

XSL•FO
RenderX

Table . Correlations between each factor and comprehensive health status.

	Values, mean (SD)	FA1[a]	FA2[b]	FA3[c]	HLS-EU[d]	SNS[e]
FA1	0.74 (0.27)	—[f]	—	—	—	—
FA2	0.75 (0.24)	0.63	—	—	—	—
FA3	0.80 (0.19)	0.53	0.64	—	—	—
HLS-EU-Q47	28.23 (8.07)	0.24	0.19	0.14	—	—
SNS	3.24 (0.67)	0.33	0.45	0.32	—	—
Total physical activity (METs[g] hours per week)	34.20 (55.21)	−0.00	−0.02	−0.05	0.09	0.06
EQ-5D[h] quality of life	0.82 (0.14)	−0.01	0.03	0.02	0.10	0.07
EQ-5D health status	76.17 (17.59)	0.05	0.06	0.02	0.18	0.12
SMILE[i] diet	2.88 (0.49)	0.26	0.24	0.19	0.33	0.27
SMILE substance use	3.29 (0.82)	0.21	0.17	0.20	0.08	0.07
SMILE physical activity	2.29 (0.62)	0.05	0.05	−0.02	0.24	0.19
SMILE stress management	2.40 (0.48)	0.11	0.09	0.02	0.32	0.21
SMILE sleep	2.77 (0.58)	0.10	0.11	0.05	0.23	0.17
SMILE social support	2.59 (0.63)	0.13	0.11	0.06	0.31	0.23
SMILE environment	2.44 (0.53)	0.04	0.03	−0.02	0.16	0.13

[a]FA1: factor 1 (conceptual knowledge).

[b]FA2: factor 2 (numeracy).

[c]FA3: factor 3 (synthesis).

[d]HLS-EU-Q47: 47-item European Health Literacy Survey Questionnaire. A general health literacy index score comprising all items was standardized on a metric between 0 and 50, using the following formula: (mean − 1) × (50/3).

[e]SNS: Subjective Numeracy Scale.

[f]Not available.

[g]MET: metabolic equivalent.

[h]EQ-5L: EuroQol 5-dimension.

[i]SMILE: Short Multidimensional Inventory Lifestyle Evaluation.

Discussion

Principal Findings

This study examined the factor structure of the multiobjective health literacy and numeracy scales among Japanese-speaking adults. Specifically, we explored how many factors would emerge in the pool of 72 items extracted from 5 scales, with or without being contextualized for specific diseases. The exploratory factor analysis indicated that the items could be categorized into three factors: performance-based conceptual knowledge (FA1), numeracy (FA2), and synthesis (FA3).

Most funHLS items loaded on FA1, assessing the conceptual knowledge of health-related and medical terms. FA2 consisted of items from 4 scales targeting people with different health conditions and diseases that typically assess their ability to perform mathematical calculations. The NVS and CHLT items not included in FA1 were identified as FA3, which required the synthesis of multiple skills to handle health information, such as reading, knowledge, navigation, and interpretation skills, to provide a correct response. A correlation analysis indicated that all factors had weak correlations with subjective health literacy, moderate correlations with subjective health numeracy, and weak correlations with lifestyle (eg, diet, nutrition, and substance use). Lifestyles concerning sleep and social support demonstrated small correlations only with FA1 and FA2 but not with FA3.

In line with Altin et al [9] and Wu et al [55], we observed small-to-moderate correlations between the 3 factors and the

XSL•FO
RenderX

subjective scales (ie, HLS-EU-Q47 and SNS). Furthermore, the 3 identified factors were highly correlated with each other, yet were recognized as independent factors. These findings echo Waters et al's [24] argument—health literacy and numeracy are related but distinct constructs, each of which can be psychometrically divided into performance-based (objective) and self-reported (subjective) constructs. Another important point is that our analysis did not identify disease-specific factors, although we included cancer- and diabetes-specific items in the item pool. Therefore, it is plausible to assume that the 3 identified factors—conceptual knowledge, numeracy, and synthesis—form a common basis for processing health information in general. Health literacy covers a range of skills from basic to advanced levels. Basic skills include reading and writing (ie, literacy), which allow individuals to function effectively in everyday situations. These skills serve as a foundation for more advanced ones, for example, extracting information, deriving meaning from different sources of communication, and applying new information to changing circumstances [1]. We assume a similar hierarchical structure for the identified 3 factors, which may explain the interfactor correlations; that is, synthesis represents higher order skills that require more basic ones, such as numeracy and knowledge, along with other cognitive and literacy skills (eg, reading, comprehension, and interpretation).

Regarding the associations with health behaviors and lifestyles, each factor presented small correlations with diet and substance use but not with PA. Some overlaps were noticed at the item content level; for example, the NVS includes items about caloric calculation as well as reading and interpreting a nutrition label, whereas the SMILE asks how often respondents eat high-calorie sweet or fatty foods and how frequently they check the food ingredient labels. A similar association was found in patients with diabetes; performance-based numeracy is positively correlated with a healthy diet [56]. These findings suggest that skills and abilities assessed by objective measures underlie perceived health behaviors (eg, individuals are able to read and interpret ingredient labels and check them regularly when shopping for food). However, the size of the correlations was modest, and the results should be interpreted carefully, particularly for the practical significance.

Compared with objective measures, subjective measures demonstrated overall larger correlations with health behaviors and lifestyles. The conceptual knowledge and numeracy factors (FA1 and FA2) had small correlations with sleep and social support of the SMILE (r values 0.11 - 0.13) but subjective health literacy (HLS-EU-Q47) and numeracy (SNS) presented slightly larger correlations with sleep and nutrition (r values 0.17 - 0.33) as well as with other subscales (eg, PA, r=0.24; stress management, r=0.33). Higher levels of objective health literacy are thought to be associated with an inclination to behave in a manner that is beneficial to one's own and others' health (eg, choosing beneficial treatments for a disease) [17]. However, subjective health literacy may share even larger variance with the perception of health behaviors; that is, how people perceive their ability to process health information may overlap with how they believe to behave in a context where their health matters. It is too early to conclude that subjective measures are more suited for studying health behaviors based only on the correlations found in this study. Instead, it is fair to argue that objective and subjective measures reflect different psychological processes, and further research is warranted to clarify which type (or both) of health literacy measure is associated with actual health behaviors that can be assessed using sensors and devices, such as accelerometers for PA.

Limitations

This study has several methodological limitations. First, the item pool was neither exhaustive nor comprehensive. Importantly, we did not include Test of Functional Health Literacy in Adults (TOFHLA) [10] and Rapid Estimate of Adult Literacy in Medicine (REALM) [11], which are the most widely used objective measures, because of language and cultural differences (all materials had to be in the Japanese language) and technical limitations of the survey platform (audio-visual recording could not be implemented). Both tools are closely bound to the English language (or even to the culture and health care system of the country where the scales were developed). For example, the REALM evaluates whether respondents pronounce medical terms correctly, and the TOFHLA assesses the ability to read and understand health-related materials contextualized in the US health care system. Yet, our analyses covered the scales and items conceptually overlapping with the REALM and TOFHLA; the funHLS is a word recognition test for medical terms, the CHLT and NVS assess reading comprehension of texts and tables, and the Lipkus evaluates numerical ability. However, we acknowledge that the exact items of the REALM and TOFHLA were not included here, and this may affect the interpretation of the results, particularly for the generalizability of the study findings. Furthermore, it is highly likely that the results of the factor analysis and subsequent analyses might differ if the item pool were expanded. Second, the exploratory factor analysis showed that the 3-factor structure explained less than half of the total item variance. A possible explanation is that measurement invariance might not be assumed in subgroups of participants as the data covered a diverse range of people in terms of demographics and other psychosocial variables. Different factor structures could be found across participants with different backgrounds, which should be clarified in future research. Third, participants were recruited using quota sampling to match the known population distribution in Japan for age and gender. Quota sampling is useful to ensure broad coverage of different groups and to prevent overrepresentation of a particular group in data. However, this approach is known to be vulnerable to sampling bias within a subgroup, which could be addressed by the use of self-weighted sampling if the cost of random sampling does not matter. Fourth, diagnostic information on physical or mental disorders was not collected. Testing patients with a particular disease or disorder was out of our focus, as we set a community sample as our target population. Health literacy is essential in maintaining one's health and preventing future diseases. However, it is important to widen the focus to include patient care and disease management, for which health literacy and assessments are highly relevant. Fifth, convenient self-reporting tools were used to assess PA and lifestyle habits. Health behaviors can be assessed using wearable devices and e-diaries

XSL•FO
RenderX

(eg, food recordings), which may allow for a more reliable estimation of healthy lifestyles [57]. It was technically impossible for us to use device- or sensor-based assessments, given the sample size of this study, but objective assessment tools could be considered when a focused sample is the research target.

Conclusions

Despite these limitations, our findings contribute to the psychometric evidence base of objective health literacy and numeracy scales. The results of the exploratory factor analysis identified 3 factors—conceptual knowledge, numeracy, and synthesis—among 66 items from 5 scales, independent of disease specificity and different contextualizations of the items. These 3 factors showed marginal correlations with subjective measures of health literacy and numeracy, highlighting the distinction between performance-based and self-reported assessment approaches [58]. Researchers and practitioners should be aware that self-report measures do not always reflect the skills and abilities reflected in performance on tests assessing conceptual knowledge, numeracy, and more integrated information processing skills. In other words, both subjective and objective measures should be considered if one wishes to assess different aspects of health literacy. In general, subjective measures are easier to administer and less cognitively demanding [6,7]; also, these measures are more suitable for assessing meta-cognitive, emotional, or motivational aspects of health literacy rather than knowledge and numeracy [22,27]. However, self-reported measures are vulnerable to social desirability and other biases owing to health beliefs [20], which may reduce the accuracy of assessing health information skills [9]. In contrast, objective measures are less affected by response biases [6,17] but may feel like examinations and evoke a sense of shame and stigma. This aspect is particularly relevant for individuals feeling uncomfortable with examinations and not confident in their skills (eg, test anxiety). Also, objective measures often cover a limited, highly contextualized range of skills [6]. Given these advantages and disadvantages, it is not readily possible to uniformly determine the best measures to assess health literacy. It is important for individual researchers to be aware of what aspects of health literacy they want to assess, which then helps them select appropriate scales and items in line with their objectives.

Acknowledgments

The authors attest that no generative AI tools were used in the preparation of this manuscript.

Funding

This work was supported by Council for Science, Technology and Innovation, Cross-ministerial Strategic Innovation Promotion Program, “Innovation of Inclusive Community Platform” grant JPJ012248 (funding agency: National Institutes of Biomedical Innovation, Health and Nutrition).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: KT
Data curation: KT, TO
Formal analysis: CM
Methodology: CM, KT, NK
Project administration: K Kimura
Supervision: KT
Writing – original draft: CM
Writing – reviewing & editing: KT, TO, NK, K Katahira, K Kimura
All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The means and SDs of each item as well as their factor loadings.
[XLSX File, 19 KB - publichealth_v12i1e71701_app1.xlsx]

References

1. Nutbeam D. Health literacy as a public health goal: a challenge for contemporary health education and communication strategies into the 21st century. Health Promot Int 2000 Sep 1;15(3):259-267. [doi: 10.1093/heapro/15.3.259]
2. Nutbeam D. Health promotion glossary. Health Promot Int 1998 Dec 1;13(4):349-364. [doi: 10.1093/heapro/13.4.349]

XSL•FO
RenderX

3. Reyna VF, Nelson WL, Han PK, Dieckmann NF. How numeracy influences risk comprehension and medical decision making. Psychol Bull 2009 Nov;135(6):943-973. [doi: 10.1037/a0017327] [Medline: 19883143]
4. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. Ann Intern Med 2011 Jul 19;155(2):97-107. [doi: 10.7326/0003-4819-155-2-201107190-00005] [Medline: 21768583]
5. Haun JN, Valerio MA, McCormack LA, Sørensen K, Paasche-Orlow MK. Health literacy measurement: an inventory and descriptive summary of 51 instruments. J Health Commun 2014;19(Suppl 2):302-333. [doi: 10.1080/10810730.2014.936571] [Medline: 25315600]
6. Nguyen TH, Paasche-Orlow MK, McCormack LA. The state of the science of health literacy measurement. Stud Health Technol Inform 2017;240(2):17-33. [doi: 10.3233/ISU-170827] [Medline: 28972507]
7. O Neill B, Gonçalves D, Ricci-Cabello I, Ziebland S, Valderas J. An overview of self-administered health literacy instruments. PLoS One 2014;9(12):e109110. [doi: 10.1371/journal.pone.0109110] [Medline: 25478813]
8. Tavousi M, Mohammadi S, Sadighi J, et al. Measuring health literacy: a systematic review and bibliometric analysis of instruments from 1993 to 2021. PLoS One 2022;17(7):e0271524. [doi: 10.1371/journal.pone.0271524] [Medline: 35839272]
9. Altin SV, Finke I, Kautz-Freimuth S, Stock S. The evolution of health literacy assessment tools: a systematic review. BMC Public Health 2014 Nov 24;14(1):1207. [doi: 10.1186/1471-2458-14-1207] [Medline: 25418011]
10. Parker RM, Baker DW, Williams MV, Nurss JR. The test of functional health literacy in adults: a new instrument for measuring patients' literacy skills. J Gen Intern Med 1995 Oct;10(10):537-541. [doi: 10.1007/BF02640361] [Medline: 8576769]
11. Davis TC, Crouch MA, Long SW, et al. Rapid assessment of literacy levels of adult primary care patients. Fam Med 1991 Aug;23(6):433-435. [Medline: 1936717]
12. Lipkus IM, Samsa G, Rimer BK. General performance on a numeracy scale among highly educated samples. Med Decis Making 2001;21(1):37-44. [doi: 10.1177/0272989X0102100105] [Medline: 11206945]
13. Tsubakita T, Kawazoe N, Kasano E. A new functional health literacy scale for Japanese young adults based on item response theory. Asia Pac J Public Health 2017 Mar;29(2):149-158. [doi: 10.1177/1010539517690226] [Medline: 28201939]
14. Kogure T, Sumitani M, Suka M, et al. Validity and reliability of the Japanese version of the Newest Vital Sign: a preliminary study. PLoS One 2014;9(4):e94582. [doi: 10.1371/journal.pone.0094582] [Medline: 24762459]
15. Weiss BD, Mays MZ, Martz W, et al. Quick assessment of literacy in primary care: the Newest Vital Sign. Ann Fam Med 2005;3(6):514-522. [doi: 10.1370/afm.405] [Medline: 16338915]
16. Dumenci L, Matsuyama R, Riddle DL, et al. Measurement of cancer health literacy and identification of patients with limited cancer health literacy. J Health Commun 2014;19(2):205-224. [doi: 10.1080/10810730.2014.943377] [Medline: 25315594]
17. Schulz PJ, Pessina A, Hartung U, Petrocchi S. Effects of objective and subjective health literacy on patients' accurate judgment of health information and decision-making ability: survey study. J Med Internet Res 2021 Jan 21;23(1):e20457. [doi: 10.2196/20457] [Medline: 33475519]
18. Gazmararian JA, Kripalani S, Miller MJ, Echt KV, Ren J, Rask K. Factors associated with medication refill adherence in cardiovascular-related diseases: a focus on health literacy. J Gen Intern Med 2006 Dec;21(12):1215-1221. [doi: 10.1111/j.1525-1497.2006.00591.x] [Medline: 17105519]
19. Sørensen K, Pelikan JM, Röthlin F, et al. Health literacy in Europe: comparative results of the European Health Literacy Survey (HLS-EU). Eur J Public Health 2015 Dec;25(6):1053-1058. [doi: 10.1093/eurpub/ckv043] [Medline: 25843827]
20. Fagerlin A, Zikmund-Fisher BJ, Ubel PA, Jankovic A, Derry HA, Smith DM. Measuring numeracy without a math test: development of the subjective numeracy scale. Med Decis Making 2007;27(5):672-680. [doi: 10.1177/0272989X07304449] [Medline: 17641137]
21. Song I. Differential relationships between self-reported and performance-based health literacy measures and health-related quality of life: a cross-sectional study. Medicine 2025 Jul 18;104(29):e43458. [doi: 10.1097/MD.0000000000043458] [Medline: 40696663]
22. Peters E, Bjalkebring P. Multiple numeric competencies: when a number is not just a number. J Pers Soc Psychol 2015 May;108(5):802-822. [doi: 10.1037/pspp0000019] [Medline: 25285966]
23. HLS-EU Consortium. Comparative Report of Health Literacy in Eight EU Member States: The European Health Literacy Survey HLS-EU; 2012. URL: https://oepgk.at/website2023/wp-content/uploads/2012/10/neu-rev-hls-eu-report-2015-05-13-lit.pdf [accessed 2025-09-02]
24. Waters EA, Biddle C, Kaphingst KA, et al. Examining the interrelations among objective and subjective health literacy and numeracy and their associations with health knowledge. J Gen Intern Med 2018 Nov;33(11):1945-1953. [doi: 10.1007/s11606-018-4624-2] [Medline: 30120636]
25. Begoray DL, Kwan B. A Canadian exploratory study to define a measure of health literacy. Health Promot Int 2012 Mar;27(1):23-32. [doi: 10.1093/heapro/dar015] [Medline: 21436143]
26. Marks JR, Schectman JM, Groninger H, Plews-Ogan ML. The association of health literacy and socio-demographic factors with medication knowledge. Patient Educ Couns 2010 Mar;78(3):372-376. [doi: 10.1016/j.pec.2009.06.017] [Medline: 19773144]

XSL•FO
RenderX

27. Kiechle ES, Bailey SC, Hedlund LA, Viera AJ, Sheridan SL. Different measures, different outcomes? A systematic review of performance-based versus self-reported measures of health literacy and numeracy. J Gen Intern Med 2015 Oct;30(10):1538-1546. [doi: 10.1007/s11606-015-3288-4] [Medline: 25917656]
28. Hirsh JM, Boyle DJ, Collier DH, et al. Limited health literacy is a common finding in a public health hospital's rheumatology clinic and is predictive of disease severity. J Clin Rheumatol 2011 Aug;17(5):236-241. [doi: 10.1097/RHU.0b013e318226a01f] [Medline: 21778910]
29. Moriishi C, Takano K, Oba T, Konishi N, Katahira K, Kimura K. Factor structure of self-reported health literacy scales: a large-scale cross-sectional study among Japanese-speaking adults. Health Promot Int 2025 Jul 1;40(4):daaf135. [doi: 10.1093/heapro/daaf135] [Medline: 40795183]
30. Oba T, Takano K, Katahira K, Kimura K. Exploring individual, social and environmental factors related to physical activity: a network analysis. BMJ Open Sport Exerc Med 2024;10(3):e001983. [doi: 10.1136/bmjsem-2024-001983] [Medline: 39104375]
31. Oba T, Takano K, Katahira K, Kimura K. Use patterns of smartphone apps and wearable devices supporting physical activity and exercise: large-scale cross-sectional survey. JMIR Mhealth Uhealth 2023 Nov 22;11:e49148. [doi: 10.2196/49148] [Medline: 37997790]
32. Konishi N, Oba T, Takano K, Katahira K, Kimura K. Functions of smartphone apps and wearable devices promoting physical activity: six-month longitudinal study on Japanese-speaking adults. JMIR Mhealth Uhealth 2024 Dec 10;12:e59708. [doi: 10.2196/59708] [Medline: 39658011]
33. Lee J, Lee EH, Chae D. eHealth literacy instruments: systematic review of measurement properties. J Med Internet Res 2021 Nov 15;23(11):e30644. [doi: 10.2196/30644] [Medline: 34779781]
34. Nakadai K, Kasamaki J, Shimofure T. Levels of the preventive medicine and the competence measured by health literacy scale in Japan—the literature review. Health Behavior Sci 2018;16(2):73-93. [doi: 10.32269/hbs.16.2_73]
35. Okamoto M, Kyutoku Y, Sawada M, et al. Health numeracy in Japan: measures of basic numeracy account for framing bias in a highly numerate population. BMC Med Inform Decis Mak 2012 Sep 11;12(1):104. [doi: 10.1186/1472-6947-12-104] [Medline: 22967200]
36. Lee EH, Lee YW, Lee KW, Hong S, Kim SH. A new objective health numeracy test for patients with type 2 diabetes: development and evaluation of psychometric properties. Asian Nurs Res (Korean Soc Nurs Sci) 2020 May;14(2):66-72. [doi: 10.1016/j.anr.2020.01.006] [Medline: 32044378]
37. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine 2000 Dec 15;25(24):3186-3191. [doi: 10.1097/00007632-200012150-00014] [Medline: 11124735]
38. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol 1993 Dec;46(12):1417-1432. [doi: 10.1016/0895-4356(93)90142-n] [Medline: 8263569]
39. Nakayama K, Osaka W, Togari T, et al. Comprehensive health literacy in Japan is lower than in Europe: a validated Japanese-language assessment of health literacy. BMC Public Health 2015 May 23;15:505. [doi: 10.1186/s12889-015-1835-x] [Medline: 26001385]
40. Craig CL, Marshall AL, Sjöström M, et al. International Physical Activity Questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
41. Murase N, Katsumura T, Ueda C, Inoue S, Shimomitsu T. Validity and reliability of Japanese version of International Physical Activity Questionnaire. J Health Welfare Stat 2002;49(11):1-9 [FREE Full text]
42. Physical activity standards for health promotion 2013. Ministry of Health, Labor and Welfare. 2013. URL: https://kennet.mhlw.go.jp/information/information/exercise/s-00-002 [accessed 2024-04-22]
43. Shiroiwa T, Ikeda S, Noto S, et al. Comparison of value set based on DCE and/or TTO data: scoring for EQ-5D-5L health states in Japan. Value Health 2016;19(5):648-654. [doi: 10.1016/j.jval.2016.03.1834] [Medline: 27565282]
44. Balanzá-Martínez V, Kapczinski F, de Azevedo Cardoso T, et al. The assessment of lifestyle changes during the COVID-19 pandemic using a multidimensional scale. Rev Psiquiatr Salud Ment 2021;14(1):16-26. [doi: 10.1016/j.rpsm.2020.07.003] [Medline: 38620670]
45. Kaiser HF, Rice J. Little Jiffy, Mark Iv. Educ Psychol Meas 1974 Apr;34(1):111-117. [doi: 10.1177/001316447403400115]
46. Bartlett MS. The effect of standardization on a χ2 approximation in factor analysis. Biometrika 1951 Dec;38(3/4):337-344. [doi: 10.2307/2332580]
47. Hooper D, Coughlan J, Mullen MR. Structural equation modelling: guidelines for determining model fit. Electron J Bus Res Methods 2008;6(1):53-60. [doi: 10.21427/D7CF7R]
48. Bentler PM. Comparative fit indexes in structural models. Psychol Bull 1990 Mar;107(2):238-246. [doi: 10.1037/0033-2909.107.2.238] [Medline: 2320703]
49. Steiger JH. Structural model evaluation and modification: an interval estimation approach. Multivariate Behav Res 1990 Apr 1;25(2):173-180. [doi: 10.1207/s15327906mbr2502_4] [Medline: 26794479]
50. Stevens JP. Applied Multivariate Statistics for the Social Sciences, 2nd edition: Lawrence Erlbaum Associates; 1992.
51. Rosseel Y. lavaan: an R package for structural equation modeling. J Stat Softw 2012;48(2):1-36. [doi: 10.18637/jss.v048.i02]

XSL•FO
RenderX

52. Bonaccorsi G, Lastrucci V, Vettori V, Lorini C, Florence Health Literacy Research Group. Functional health literacy in a population-based sample in Florence: a cross-sectional study using the newest vital sign. BMJ Open 2019 Jun 19;9(6):e026356. [doi: 10.1136/bmjopen-2018-026356] [Medline: 31221877]
53. Cohen J. A power primer. Psychol Bull 1992 Jul;112(1):155-159. [doi: 10.1037//0033-2909.112.1.155] [Medline: 19565683]
54. Cohen J. Statistical Power Analysis for the Behavioral Sciences: Lawrence Erlbaum Associates; 1988.
55. Wu AD, Begoray DL, Macdonald M, et al. Developing and evaluating a relevant and feasible instrument for measuring health literacy of Canadian high school students. Health Promot Int 2010 Dec;25(4):444-452. [doi: 10.1093/heapro/daq032] [Medline: 20466776]
56. Bouclaous C, Azar LJ, Barmo N, et al. Levels and correlates of numeracy skills in Lebanese adults with diabetes: a cross-sectional study. Int J Environ Res Public Health 2022 Aug 24;19(17):10557. [doi: 10.3390/ijerph191710557] [Medline: 36078271]
57. Rudolf K, Biallas B, Dejonghe LAL, et al. Influence of health literacy on the physical activity of working adults: a cross-sectional analysis of the TRISEARCH trial. Int J Environ Res Public Health 2019 Dec 6;16(24):4948. [doi: 10.3390/ijerph16244948] [Medline: 31817592]
58. Marciano L, Camerini AL, Schulz PJ. The role of health literacy in diabetes knowledge, self-care, and glycemic control: a meta-analysis. J Gen Intern Med 2019 Jun;34(6):1007-1017. [doi: 10.1007/s11606-019-04832-y] [Medline: 30877457]

Abbreviations

CHLT: Cancer Health Literacy Test
DHN: Diabetes Health Numeracy
EQ-5D: EuroQol 5-Dimension
funHLS: Functional Health Literacy Scale for Young Adults
HLS-EU-Q47: 47-item European Health Literacy Survey Questionnaire
Lipkus: Lipkus Numeracy Scale
MET: metabolic equivalent
NVS: Newest Vital Sign
PA: physical activity
QoL: quality of life
REALM: Rapid Estimate of Adult Literacy in Medicine
SMILE: Short Multidimensional Inventory Lifestyle Evaluation
SNS: Subjective Numeracy Scale
TOFHLA: Test of Functional Health Literacy in Adults

Edited by A Mavragani, T Sanchez; submitted 27.Jan.2025; peer-reviewed by F Carrouel, W Wu; revised version received 25.Sep.2025; accepted 07.Nov.2025; published 06.Jan.2026.

Please cite as:
Moriishi C, Takano K, Oba T, Konishi N, Katahira K, Kimura K
Examining the Factor Structure of Objective Health Literacy and Numeracy Scales: Large-Scale Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e71701
URL: https://publichealth.jmir.org/2026/1/e71701
doi:10.2196/71701

© Chihiro Moriishi, Keisuke Takano, Takeyuki Oba, Naoki Konishi, Kentaro Katahira, Kenta Kimura. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 6.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Data Completeness and Acceptability of Annual Feedback Reports and Tuberculosis Surveillance in Spain, 2018-2024: Mixed Methods Analysis

Allegra Chatterjee[1,2], MSc; Alvaro Roy[2,3], PhD; Rocio Amillategui[2,3], BS; María Sastre[2,3], BS; Christine Schwarz[4], MPH; Rosa Cano[2,3], MPH; Laura Herrera-León[3,5], PhD; Zaida Herrador[2,3], PhD; Spanish Working Group on Tuberculosis[6]

[1]European Centre for Disease Prevention and Control, Fellowship Programme/EPIET, Stockholm, Sweden
[2]Centro Nacional de Epidemiología, Instituto de Salud Carlos III, calle Sinesio Delgado nº6, Pabellon 12, Madrid, Madrid, Spain
[3]Network Biomedical Research Centre in Epidemiology and Public Health (CIBERESP), Madrid, Spain
[4]Health Division for HIV, STI, Viral Hepatitis and Tuberculosis Control, Ministry of Health, Madrid, Madrid, Spain
[5]Centro Nacional de Microbiologia, Instituto de Salud Carlos III, Majadahonda, Madrid, Spain
[6]See Acknowledgments

Corresponding Author:
Zaida Herrador, PhD
Centro Nacional de Epidemiología, Instituto de Salud Carlos III, calle Sinesio Delgado nº6, Pabellon 12, Madrid, Madrid, Spain

Abstract

Background: Surveillance of tuberculosis (TB) is challenging given the typically long course of the disease and treatment as well as the often complex socioeconomic factors that are associated with it. In Spain, surveillance is decentralized across autonomous communities (CCAAs), leading to additional challenges in heterogeneity and data quality.

Objective: We aimed to quantify changes in internal data completeness following introduction of annual regional surveillance quality reports and assess the acceptability and perceived usefulness among CCAAs.

Methods: This mixed methods assessment consisted of 2 stages. First, we quantitatively assessed the internal completeness of 40 key TB variables for all reported cases from 2018 to 2023, comparing 2 periods: before (2018‑2020) and after (2021‑2023) the implementation of feedback reports. Mean completeness for each variable was calculated for each period, and the differences were compared using either the paired *t* test or Wilcoxon signed rank test depending on whether differences were normally distributed. Analyses were conducted nationally and by CCAA as well as for different variable groups (patient-, illness-, and laboratory-related groups). For the second stage, we circulated the results to regional TB surveillance focal points, along with a survey to assess acceptability, perceived usefulness of reports, and barriers to data completeness.

Results: There were 25,229 reported cases across the study period: 13,328 in period 1 (2018‑2020) and 11,901 in period 2 (2021‑2023). Nationally, mean completeness increased from 66% in period 1 to 76.2% in period 2 (+10.2%, $P<.001$). Improvements were greatest for laboratory-related (+13.9%, $P=.001$) and illness-related variables (+12.8%, $P<.001$), while patient-related variables showed minimal change that was not statistically significant (+2.4%, $P=.48$). Most CCAAs (16/19, 84%) demonstrated improved completeness in period 2, though substantial regional variation persisted. Of the 19 CCAAs, 18 responded to the survey. Respondents reported high acceptability of the surveillance system and considered feedback reports useful. Challenges included resource constraints, system interoperability between laboratory and surveillance, and patient follow-up.

Conclusions: Following the introduction of annual feedback reports, which were well accepted by all regional stakeholders, Spain saw significant improvements in TB data completeness. Sustained feedback mechanisms, streamlined reporting requirements, and improved integration between epidemiological and laboratory systems are key to strengthening TB surveillance.

(JMIR Public Health Surveill 2026;12:e93719) doi:10.2196/93719

KEYWORDS

tuberculosis; surveillance; Spain; public health; infectious diseases

Introduction

Tuberculosis (TB) remains the world's deadliest infectious disease, despite being both preventable and usually curable [1]. In 2024, an estimated 10.7 million people worldwide developed TB, corresponding to an incidence rate of 131 cases per 100,000 population [2]. The same year, 11.5% of those diagnosed with TB died from the disease, a total of 1.2 million people.

XSL•FO
RenderX

To address the global TB burden, in 2014, the World Health Organization (WHO) launched the End TB Strategy, aiming to reduce incidence by 90% and deaths by 95% by 2035, compared with 2015 levels [3]. Although there is large regional variation, the number of TB cases and deaths globally decreased in 2024, following 3 consecutive years of increases due to disruptions caused by the COVID-19 pandemic [2]. The WHO European Region has achieved substantial progress since WHO targets were introduced, with a 39% reduction in incidence rate observed by 2024 [2]. However, significant challenges remain in the region; according to a 2026 report, multidrug-resistant TB rates are 7 times higher in Europe than the global average, 21% of TB cases are undiagnosed or unreported, and 22% of people who start treatment are not followed up after 1 year [4].

As countries continue to strive to meet these goals, robust TB surveillance systems are essential, enabling monitoring of trends in incidence and mortality, detecting outbreaks, identifying affected populations to create targeted interventions, and evaluating the impact of TB control measures to ensure accountability for national and international commitments. However, achieving comprehensive and accurate TB surveillance remains challenging. Common barriers identified in European countries include fragmented integration between clinical and laboratory information systems and heterogeneity in reporting across jurisdictions [5], as well as loss to follow-up during the long treatment courses required, particularly for vulnerable populations such as migrants or people experiencing homelessness [6]. Regular assessments of TB surveillance systems are therefore critical to ensure their effectiveness, identify gaps, and prioritize improvements.

In Spain, reported TB cases and hospitalizations showed a gradually declining trend between 2012 to 2020 [7]; however, recent surveillance reports show a slight rebound, with the notification rate rising from 7.7 to 8.8 per 100,000 population from 2020 to 2024 [8]. This trend was observed among children as well as adults in Spain; although a lower incidence is generally observed among children, following a decrease after 2012, increases have been observed in both notification and hospitalization rates among this group since 2021 [9]. Nevertheless, compared with global figures, the incidence and deaths rates are low, with WHO recategorizing the country from incidence group “low moderate” in 2015 to “low” by 2025 [2].

However, the decentralized nature of surveillance presents a challenge, with each of the 17 autonomous communities (CCAAs) and 2 autonomous cities (Ceuta and Melilla) having their own surveillance systems, with annual integration of data at the national level. In 2021, a rapid internal review by the National Center for Epidemiology (CNE) recognized that there were deficiencies in the completeness of some variables and a large degree of variation among CCAAs. To support improvements in TB surveillance data quality, the CNE began sharing annual personalized feedback reports to each CCAA applying to submissions of consolidated annual data from the year 2021 onward. These reports summarize completeness across TB surveillance variables, highlighting any inconsistencies identified and identifying areas for improvement.

We performed a mixed methods study aiming to quantify changes in internal data completeness following the introduction of annual regional feedback reports on the quality of surveillance data and to assess acceptability and perceived usefulness among CCAAs. The purpose of this was to evaluate whether this quality improvement initiative of annual feedback reports should be recommended to continue and, if so, whether they should be refined. The study comprised 4 objectives: (1) quantitatively evaluate the internal completeness of TB surveillance data before and after implementation of feedback reports, (2) qualitatively assess the acceptability and perceived usefulness among CCAAs of both the current surveillance system and the feedback reports, (3) update annual feedback reports based on findings from objectives 1 and 3, and (4) assess the impact of the newly updated reports on completeness of TB surveillance data for 2024.

Methods

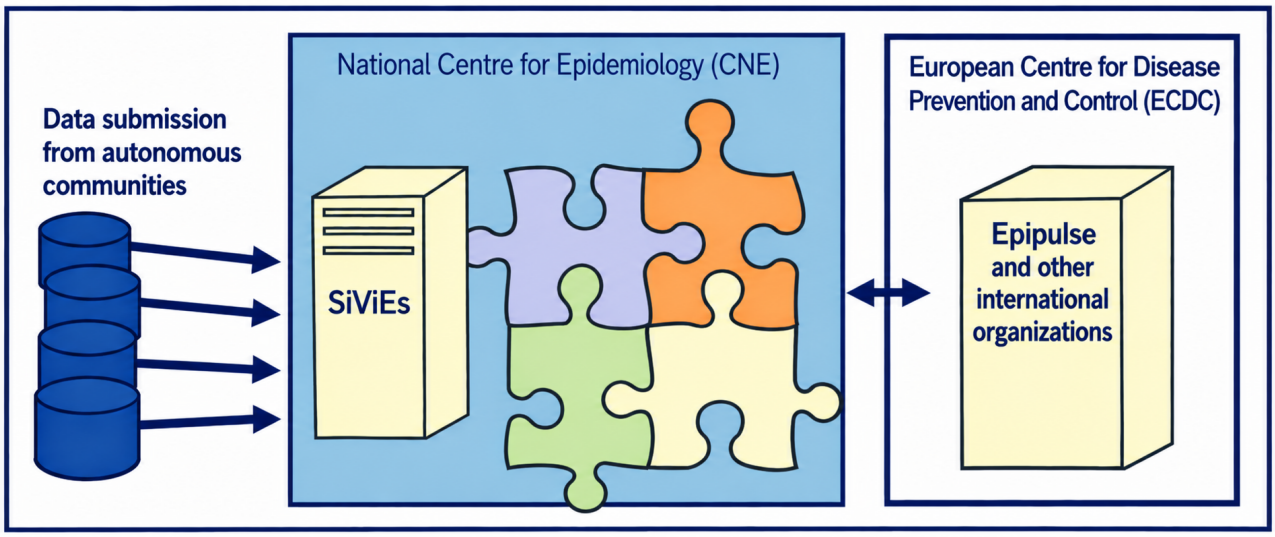
Data Sources

The Spanish TB national surveillance system is passive, case-based, and at the population level; each CCAA is responsible for collecting and annually reporting patient-level data on predefined variables into the Spanish Surveillance System (SiViEs Plus) according to the TB protocol published by The National Epidemiological Surveillance Network (La Red Española de Vigilancia Epidemiológica [RENAVE]; Figure 1) [10].

We extracted data on 40 key variables from all reported cases in Spain from the SiViEs Plus database between January 1, 2018, and December 31, 2023. We divided variables into 3 categories: (1) patient-related information (11 variables), (2) illness-related information (18 variables), and (3) laboratory-related information (11 variables).

XSL•FO
RenderX

Figure 1. Flow of data from autonomous communities into the national Spanish Surveillance System (SiViEs Plus) and the ECDC.



For the second aim, we developed a questionnaire based on a literature review and consulting with TB and surveillance experts at the CNE. The 17-question online questionnaire included both closed and open-ended questions to gather information on the perceived usefulness of the annual feedback reports, challenges around TB surveillance, and suggestions on how to improve TB surveillance in Spain. We piloted the questionnaire with the TB surveillance lead from 1 CCAA. In April 2025 and May 2025, we circulated the final version via Google Forms among the Spanish TB Surveillance Working Group. The survey was completed by the lead individual responsible for TB surveillance in each CCAA; these individuals were considered best placed to respond on behalf of their CCAA as they have an overview of all areas of TB surveillance in their region, they are responsible for submission of data at the national level, and they are the person who receives the annual feedback reports. The survey invitation emphasized that results would be reported anonymously and not attributable to any CCAA or individual.

Ethical Considerations

This study used anonymized data from RENAVE in Spain. RENAVE collects data on cases notified through the national reporting electronic platform SiViEs Plus, and it is hosted by the CNE. These data are generated under a legal public health mandate and are routinely used for epidemiological monitoring. As the study involved secondary analysis of anonymized administrative surveillance data accessed under legally mandated confidentiality safeguards, formal ethics committee approval was not required. This exemption is consistent with Spanish and European data protection regulations governing the processing of anonymized public health surveillance data, including Ley Orgánica 3/2018 de Protección de Datos Personales y Garantía de los Derechos Digitales (LOPDGDD) and Reglamento (UE) 2016/679 (GDPR).

Survey data were submitted by a regional focal point from CCAAs in Spain. For survey participation, informed consent was obtained from participants prior to beginning the survey. Participants were made aware that taking part was voluntary, that there would be no consequences for not taking part, and that they could withdraw from participating at any time without consequence. One response was obtained from each CCAA; however, responses were reported anonymously in the analysis and write-up to maintain confidentiality. All data were stored and analyzed within secure, password-protected ISCIII systems in compliance with institutional data protection policies.

Definitions

Internal completeness and acceptability surveillance attributes were chosen as they align with those set out by the European Center for Disease Control and Prevention (ECDC) technical handbook “Data quality monitoring and surveillance system evaluation” [11] as key for high-functioning surveillance systems.

We defined internal completeness according to ECDC [11] as “the number of completed data fields out of the total number of data fields (unknown and missing items should be included in the denominator).” The denominator was not always equal to the total number of cases; for example, the variable “number of years living in Spain” was only applicable to cases born outside of Spain.

Surveillance Cycle and Feedback Reports

TB data are submitted continuously by CCAAs to SiViEs Plus; however, each annual dataset is only considered complete once regions confirm that all records for the previous year have been uploaded, typically in Q2 of the following year (eg, annual data for 2021 would be consolidated in Q2 of 2022).

Since 2022, following the annual submissions from the CCAAs, the national level (CNE) performs data validation and generates feedback reports highlighting missing or inconsistent information (see Multimedia Appendix 1 for an example report in the original Spanish and in an English translation that was carried out using the generative artificial intelligence [AI] tool, Copilot). These reports are typically shared with CCAAs between May and July, with variation in timing depending on when each region finalizes its initial submission. CCAAs subsequently update their annual data before final national consolidation, which usually occurs in the third quarter (Q3).

XSL•FO
RenderX

The feedback reports disseminated in a given year are therefore designed to inform and improve data reporting practices for the year prior. As feedback reports began being disseminated in 2022, the year 2021 was considered the first year in which reporting could be influenced by the feedback process.

Data Analysis

Part 1: Internal Completeness Analysis

We first calculated the percentage internal completeness for each of the 40 variables across each year from 2018 to 2023. For each variable, we then calculated the mean percentage completeness (see the definition in the "Definitions" section) for two 3-year periods: (1) before reports were implemented (period 1): 2018 - 2020 and (2) after reports were implemented (period 2): 2021 - 2023.

For each period, we then calculated a mean across all variables by summing the mean percentage completeness for each variable and dividing by 40; we used the same approach to calculate means for period 1 and period 2 for each of the 3 subcategories (patient-, illness- and laboratory-related variables).

For each variable, we calculated the difference in completeness before and after implementation of feedback reports by subtracting the mean percentage completeness of that variable in period 1 from that of period 2.

For all variables combined and for each of the 3 subcategories, we used the Shapiro-Wilk test to examine whether the differences in completeness between the 2 periods were normally distributed or not; if found to be normally distributed, the paired *t* test was used to test for statistical significance (set at 5%). Otherwise, the Wilcoxon signed-rank test was used. This analysis was carried out at the national level and separately for each CCAA. All analyses were performed using R version 4.4.1.

We also carried out a sensitivity analysis to examine whether the COVID-19 pandemic may have impacted the results. We repeated the internal completeness analysis after removing the 2 years of 2020 and 2021; therefore, the comparison periods each contained 2 years (2018 - 2019 and 2022 - 2023), and we compared the results with those obtained for the original comparison periods.

Part 2: Survey Analysis

Together with the questionnaire, we circulated to each CCAA a preliminary analysis of national-level data compared with the regional data. This included a written summary and figures showing changes in completeness of surveillance data with respect to the 3 categories of variables across the 2 periods, comparing the values for their CCAA with the national averages (see Multimedia Appendix 1).

We calculated the proportion of respondents in each category of closed-ended questions. We reviewed free-text responses and grouped them into categories based on recurring content and meaning.

Part 3: Annual Report Updates and Continuous Quality Improvement

In June 2025, to support the CCAAs to make continual improvements in TB surveillance, we shared with the regions a summary of the qualitative analysis, along with a description of changes made to the upcoming feedback report (for 2024 data submission) based on the findings. Following this, we later performed an analysis of the 2024 TB surveillance data to determine whether the updated feedback report supported further improvements, as part of a continuous improvement cycle.

Results

Part 1: Internal Completeness Analysis

TB surveillance data for 25,229 reported cases were extracted across the study period. In period 1 (2018 - 2020), there was a total of 13,328 cases, and in period 2 (2021 - 2023), there was a total of 11,901 cases. Comparing internal completeness of 40 variables across period 1 and period 2, there was a mean improvement of 10.2% (Wilcoxon signed rank test $P<.001$), from 66% in period 1 to 76.2% in period 2, with 78% (31/40) variables showing an increase (Table 1). No change was observed for 4 (10%) of the 40 variables (age, sex, TB classification, date of case report), which all had a completeness rates ≥99.9% in period 1, while 5 (13%) of the 40 variables had a decrease in completeness (country, CCAA, province and municipality of residence, and mother number of years in Spain; Table 1).

Patient-related variables had the highest overall completeness for both periods (80.3% for period 1 and 82.7% for period 2). However, the mean increases of 2.4% across the 11 variables was lower than that of other variable groups and was not statistically significant (Wilcoxon signed rank test $P=.48$). All 5 variables with a decrease in completeness were within this category.

Illness-related variables significantly increased by 12.8% (paired *t* test: $t_{17}=-5.185$, $P<.001$), from 64.6% in period 1 to 77.4% in period 2. Of the 18 variables within this group, 2 did not increase (TB classification and date of case report) as they were already at 100% completion during period 1.

The greatest improvement was seen with laboratory-related variables, with an increase of 13.9% (paired *t* test: $t_{10}=-4.392$, $P=.001$) from 53.9% in period 1 to 67.7% in period 2. All 11 variables increased, the largest of which was for "date of laboratory diagnosis," which increased by 37.5%.

XSL•FO
RenderX

Table . Completeness of tuberculosis (TB) surveillance variables at the national level comparing period 1 (2018 - 2020) with period 2 (2021 - 2023) in Spain.

Variable	Mean completeness period 1, %	Mean completeness period 2, %	Difference (period 2 – period 1), %	*P* value for the difference
Patient variables				
Date of birth	85.8	98.6	+12.8	N/A[a]
Age	99.9	99.9	0	N/A
Sex	100	100	0	N/A
Country of residence	99.6	98.2	−1.4	N/A
Autonomous community of residence	99.5	98	−1.5	N/A
Province of residence	99.4	97.2	−2.2	N/A
Municipality of residence	99.1	97.6	−1.5	N/A
Country of birth	87.4	90.8	+3.4	N/A
Number of years in Spain	47.3	62.3	+15.0	N/A
Mother country of birth	7.6	12.3	+4.6	N/A
Mother number of years in Spain	57.3	55.3	−2.0	N/A
Mean of all patient variables	80.3	82.7	+2.4	.48
Illness variables				
TB classification	100	100	0	N/A
Date of case report	100	100	0	N/A
Date of symptom onset	79.4	84.9	+5.5	N/A
Country of case report	88.9	89.5	+0.6	N/A
Autonomous community of case report	70.1	71.9	+1.8	N/A
Province of case report	35.2	66.1	+30.9	N/A
Municipality of case report	32.6	64.5	+31.9	N/A
Importation	85	92.4	+7.4	N/A
TB location	97.4	98	+0.7	N/A
Previous treatment	64.9	82.1	+17.2	N/A
Directly observed treatment	64.9	82.1	+17.2	N/A
Associated with outbreak	27.7	54	+26.3	N/A
Date of treatment initiation	70.8	85.1	+14.4	N/A
Date of treatment end	44.8	57.1	+12.3	N/A
Treatment result	63	77.8	+14.8	N/A
Follow-up result	15.4	39.3	+23.8	N/A
Hospitalization	76.9	90.8	+13.9	N/A
Death	63.6	79	+15.4	N/A
Mean of all illness variables	64.6	77.4	+12.8	<.001
Laboratory variables				
Date of lab diagnosis	41.7	79.2	+37.5	N/A

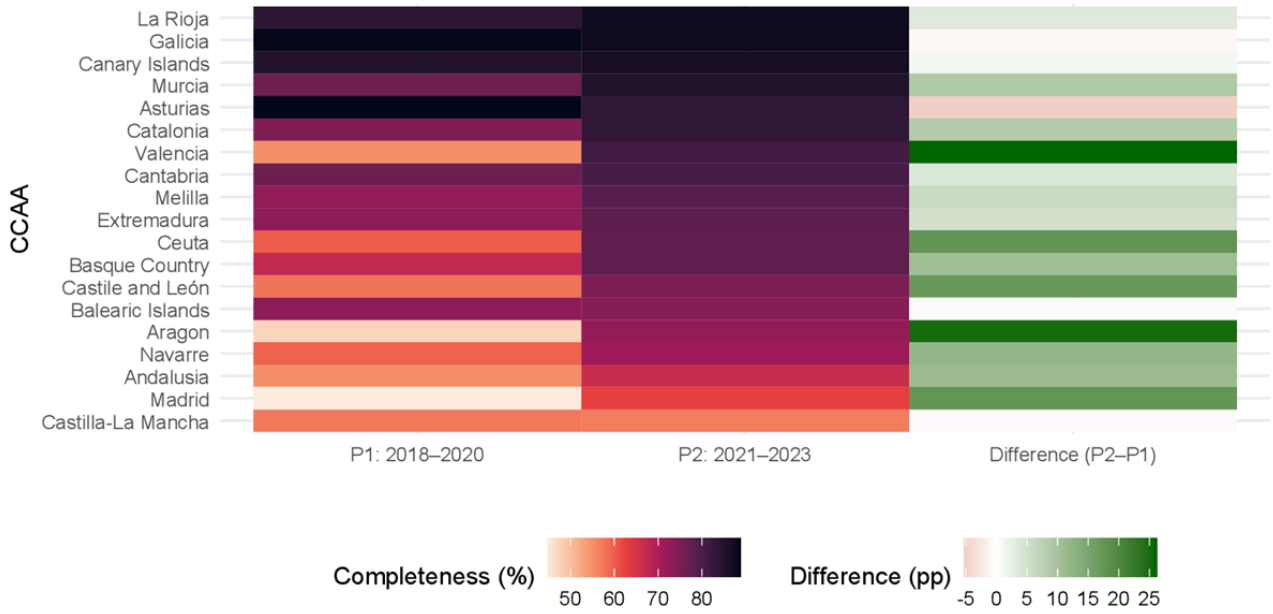
XSL•FO
RenderX

Variable	Mean completeness period 1, %	Mean completeness period 2, %	Difference (period 2 – period 1), %	*P* value for the difference
Causal agent	73.3	76.6	+3.4	N/A
Lab test carried out	35.2	50.6	+15.4	N/A
Lab result: isolation	93.3	93.6	+0.3	N/A
Lab result: visualization	92.9	93.0	+0.1	N/A
Lab result: antimicrobial susceptibility	47.8	66.2	+18.4	N/A
Etambutol resistance test	31.9	50.2	+18.3	N/A
Isoniazide resistance test	35.4	49.9	+14.5	N/A
Pirazinamide resistance test	33.6	46.8	+13.2	N/A
Rifampicine resistance test	35.8	50.4	+14.6	N/A
HIV test	71.7	88.7	+17.1	N/A
Mean of all laboratory variables	53.9	67.7	+13.9	.001
All variables	66	76.2	+10.2	<.001

[a]N/A: not applicable.

There was a large degree of variation across the 17 CCAAs and the 2 autonomous cities (Figure 2). For period 1, mean annual percentage completeness across all variables ranged from 44.9% to 88.7%, and for period 2, it ranged between 56.9% and 87.5%. Comparing period 1 with period 2, differences in completeness ranged from a decrease of 5.2% to an increase of 26.2% in period 2. For all regions, 84% (16/19) had an improvement in period 2, 69% (11/16) of which had statistically significant improvements (*P*<.05). Of the 3 (3/19, 16%) CCAAs where a decrease in completeness was observed, only 1 was statistically significant.

Figure 2. Heat map showing mean percentage completeness of 40 tuberculosis (TB) variables reported by autonomous communities (CCAAs) in period 1 (P1: 2018 - 2020) and period 2 (P2: 2021 - 2023) in Spain.



Results obtained were similar across all analyses when the pandemic years of 2020 and 2021 were excluded (Table S2 in Multimedia Appendix 2). Across all 40 variables, there was a statistically significant increase of 12.5% (*P*<.001), compared with 10.2% (*P*<.001) in the original analysis. For patient variables, results were identical to the original analysis, with a nonstatistically significant increase of 2.4% observed. Differences between the 2 periods were slightly greater for illness-related variables and laboratory-related variables when excluding the pandemic years: 14.8% (*P*<.001) compared with 12.8% (*P*<.001) and 18.9% (*P*<.001) compared with 13.9% (*P*=.001), respectively.

XSL•FO
RenderX

Part 2: Survey Analysis

Satisfaction With Surveillance System

Overall, 95% (18/19) of CCAAs completed the online questionnaire. All 18 respondents reported that having high-quality TB surveillance data at the national level is either "very important" (17/18, 94%) or "important" (1/18, 6%). However, satisfaction with the current system varied; although most reported feeling either "very satisfied" (3/18, 17%) or "satisfied" (8/18, 44%), 33% (6/18) were neutral, and 1 (1/18, 6%) CCAA reported feeling unsatisfied. Regarding the completeness of data reported from the CCAAs, results were similar, with most reporting being either very satisfied (2/18, 11%) or satisfied (9/18, 50%), while 33% (6/18) were neutral and 1 respondent was unsatisfied. Satisfaction with the current system did not appear to vary depending on the completeness of data submitted by CCAAs.

Reported issues included the challenging context of TB surveillance given the complex interactions between the social, cultural, and health care–related aspects of the disease. Some respondents felt that the amount of information requested for each case was too high and that some variables may not be necessary, particularly those that are hard to define or calculate such as time to follow-up.

Respondents suggested that the current system could be streamlined to focus on high-priority indicators for TB control. Several respondents called for clearer guidance on variable definitions and their categories, as well as integration of case follow-up information given the long disease course and, in particular, how to manage surveillance and declaration of cases that transfer between CCAAs.

In addition, the need for a strong system for identifying unreported cases and the extent of under-ascertainment in Spain was highlighted; suggestions included examining differences in proportions (eg, confirmed or extrapulmonary TB) across regions to detect inconsistencies that could indicate under-ascertainment.

Usefulness of Annual Feedback Reports

The annual feedback reports were considered to be useful by all respondents; 17% (3/18) felt they are "extremely useful," 61% (11/18) "very useful," and 22% (4/18) "moderately useful." . As with the attribute of usefulness, this did not seem to be associated with how well the region performed with respect to data completeness. Free-text responses indicated that CCAAs felt that reports help improve data quality by identifying missing or inconsistent entries. Some respondents noted that reports contribute to an environment of continuous quality improvement, stimulating internal reviews and priority areas to improve regional surveillance, as well as helping to evaluate interventions. Several also highlighted their value in raising awareness among TB units, nurses, IT staff, and public health leadership.

Of respondents, 78% (14/18) and 83% (15/18) felt that the format and content of the feedback reports, respectively, were appropriate, while 22% (4/18) and 17% (3/18) felt that the format and content of the feedback reports, respectively, were partially appropriate. None of the participants reported that they were not appropriate. Suggestions for improvements included highlighting the poorest performing variables within each section, considering simplifying data, and including fewer variables so that those of higher importance can be prioritized. In addition, one respondent suggested that coherence and internal validity between variables could be highlighted as well as completeness. Of all respondents, 72% (13/18) thought that the timing of report distribution allowed for improvements before the next data submission cycle; however, 22% (4/18) felt that they were sent too late, and 1 respondent thought they arrived too early. One person commented that they would prefer to receive reports more frequently, every 4 months. In terms of dissemination, 50% (9/18) of respondents reported sharing the reports within their regional teams, including with epidemiologists, TB units, IT teams, primary care, hospital services, and regional public health directors and leadership. Better-performing CCAAs appear to be more likely to circulate reports to colleagues; among the 5 regions with highest period 2 completeness, 4 reported sharing reports with colleagues; however, none of the bottom 5 performing regions reported sharing the reports.

Barriers and Enablers

Barriers to complete TB data reporting varied across regions but commonly included limited human resources and time (15/18, 83%); technical restrictions (9/18, 50%); insufficient training (5/18, 28%), which is exacerbated by high staff turnover; and lack of clear surveillance guidelines (4/18, 22%). Many CCAAs had already implemented measures to improve surveillance, including integration of laboratory and microbiological data into their surveillance systems, local data validation prior to submission, feedback loops between regional TB units and central teams, automated quality checks to flag inconsistencies within information systems, and continuous improvements to electronic reporting platforms.

Suggestions for broader system improvements included online training resources for surveillance staff and continuous quality improvement initiatives, including maintaining the annual feedback reports and possibly increasing their frequency, improved variable definitions in the national protocol to ensure consistency, and national guidelines for follow up of cases who move between regions.

Part 3: Annual Report Updates and Continuous Quality Improvement

In June 2025, to support CCAAs with making continuous improvements in TB surveillance, we shared with regions a summary of Part 1 and Part 2 of the assessment, along with a description of changes made to the upcoming feedback reports (for 2024 data submission) based on the findings:

- Number of variables included reduced to the 40 key variables included in the analysis
- Variables separated into 3 groups to coincide with the analysis: patient, illness, laboratory
- Addition of various cross-checks, including infective agent classification and diagnostic test used, date of case registration and result of treatment, date of treatment

XSL•FO
RenderX

completion, and result of treatment; also, the addition of several new cross-checks related to antimicrobial resistance testing
- Time to diagnostic and time to treatment added to reports as a new variable, calculated using the difference between the date of diagnosis and date of symptom onset and between the date of initiation of treatment and date of symptom onset, respectively

Following this, we analyzed the preliminary 2024 TB surveillance data to determine whether the updated feedback report supported further improvements, as part of a continuous improvement cycle. We found that, at the national level, mean completeness across all variables increased to 78% in 2024 (compared with 66% in 2018 - 2020 and 76% in 2021 - 2023).

Discussion

Summary of Principal Findings

This study aimed to quantitively assess the internal completeness of TB surveillance data before and after implementation of feedback reports in Spain and to qualitatively assess the perceived usefulness among CCAAs of both the current surveillance system and the feedback reports. We identified significant improvements in completeness across most variables and regions following the introduction of annual feedback reports, although there was some variation across both type of variable and CCAAs. At the national level, mean completeness increased from 66% in 2018 - 2020 to 76.2% in 2021 - 2023 (with nearly four-fifths of variables indicating improvements) and increased further to 77.8% in 2024. Both the surveillance system and feedback reports were generally viewed as acceptable and useful among regional stakeholders included in the survey. The additional study aiming to update the feedback reports for the 2024 surveillance cycle based on our findings and assess the potential impact of the updated reports on TB surveillance data completeness found that mean completeness continued to improve in 2024 relative to previous years.

Principal Findings in Context

Although the highest completeness levels were seen among patient-related variables, no statistically significant improvement was observed in this group. This is partly due to many variables being at or close to 100% completeness in period 1, leaving minimal opportunity for further improvement. This is in line with findings from other European Union (EU) and European Economic Area (EEA) countries, where, in 2023, most countries had complete or near complete data (>95%) for variables requiring mandatory reporting to the ECDC, such as age and sex [12]. Patient-related variables that continued to have low completeness rates in period 2 were those that applied to individuals born outside of Spain: number of years in Spain, mother country of birth, and mother number of years in Spain (62.3%, 12.3%, and 55.3% period 2 completeness, respectively). This is pertinent, as individuals from countries of origin with higher TB prevalences may have an increased risk of importation; therefore, efforts should be made to improve data collection for these variables.

Although illness-related variables had similar percentage completeness (77.4%) to that of patient-related variables in period 2, a larger improvement of 12.8% in illness-related variables was observed relative to period 1. However, some variables remained poorly completed, including follow-up result, which, despite a 23.8% increase, had 39.3% completeness in period 2. Supporting this, survey respondents called for improved protocols and better integration of case follow-up and laboratory information, particularly for patients who move between CCAAs. The respondents' call is in the context of the WHO recommendations for a 6-month regimen of 4 first-line drugs (and longer treatment and additional drugs for resistant TB) [13], which complicates patient follow-up. Continuity of TB treatment is essential for effective control of the disease, particularly with increasing movement of individuals both within and between countries across Europe [14]. In particular, interruption of treatment among mobile populations (including migrants, refugees, and asylum seekers) is a recognized challenge in Europe, often leading to poorer health outcomes, higher risk of transmission, and the emergence of resistant strains [15]. In line with this study, evidence from other countries has found that variability regarding migrant health policies and unclear or absent referral mechanisms can hinder treatment continuity [6,14,16].

Improvements were most pronounced for laboratory-related variables, with a mean increase of 13.9%; however, increases were particularly pronounced for those with lower baseline completeness, such as for date of laboratory diagnosis and variables related to antimicrobial susceptibility testing. Although this highlights substantial improvements in capturing microbiological information, challenges persist due to the lack of interoperability between clinical and laboratory systems. Despite improvements, some laboratory variables continue to have low completeness in period 2, including lab tests carried out (50.6%) and results for the 4 included antimicrobial resistance testing variables, which ranged from 31.9% to 50.4%. This finding reinforces the need to consolidate mechanisms for the automatic integration of microbiological data into epidemiological surveillance, as demonstrated effectively in some CCAAs. In Catalonia, for example, which ranks higher than most CCAAs with respect to completeness in period 1 (7th) and period 2 (6th), laboratory results are linked to the regional surveillance system [17]. At the European level, models such as the MiBa (Danish Microbiology Database) in Denmark [18], Osiris system in the Netherlands [19], and MycobNet network in the United Kingdom [20] enable automated laboratory result reporting and the genomic integration of *Mycobacterium tuberculosis* strains, improving outbreak detection and resistance surveillance.

This finding that variables requiring integration between clinical and laboratory systems or ongoing follow-up were the most incomplete is consistent with the ECDC [11] and United States Centers for Disease Control and Prevention (CDC) guidance [21], emphasizing that data involving multiple sources or complex definitions are generally more vulnerable to missingness. In line with this, across EU/EEA countries in 2023, the availability of data related to clinical and laboratory information was lower than for patient-related variables, for

XSL•FO
RenderX

example for previous treatment history (86%), treatment outcome at 12 months (61.5%), and HIV status (39.3%) [12].

At the regional level, 16 of Spain's 19 CCAAs (84%) saw an improvement in period 2, although the magnitude varied widely, ranging from modest gains to more than 25% increases. The one region where a statistically significant decrease was observed already had a high level of completeness in period 1 (the 5th highest among all CCAAs). These disparities reflect the decentralized nature of the Spanish surveillance network RENAVE. This is in line with other countries with decentralized surveillance systems, where regional disparities have been found in evaluations of the surveillance systems, for example in Italy where northern regions were found to report acute viral hepatitis cases in a more timely manner than central and southern regions [22].

Survey findings provided important explanatory context. Respondents were generally satisfied with TB surveillance in Spain and considered it to be important, indicating high level of acceptability and perceived usefulness of the system, providing a strong foundation for future quality improvement initiatives. There appeared to be a possible association between higher-performing CCAAs and the extent to which the feedback reports are disseminated among colleagues within the region; further investigation would be required to understand whether this is because the reports are more effective at driving improvements if disseminated more widely or simply that higher-performing CCAAs have both better communication channels and better surveillance systems.

Overall, our findings show an improvement in internal completeness of TB surveillance data following implementation of feedback reports and positive perceptions of these reports among regional stakeholders. Although a causal link is difficult to establish, our findings suggest that systematic data quality monitoring and feedback may help to strength surveillance, consistent with international guidance that emphasizes regular evaluation as a cornerstone of high-performing systems [11,21]. The observed improvements, together with respondents' endorsement of feedback reports, suggest that feedback cycles may contribute to improved completeness and system acceptability and continuous quality improvement.

Limitations

This analysis is subject to several limitations. First, due to the descriptive nature of the analysis, it was difficult to establish a causal link between the observed improvements in data quality and the implementation of the feedback reports. For this reason, we included the qualitative element of the survey, but it is not possible to rule out other explanations for the observed results, such as improvements in surveillance across all areas or increased focus on infectious diseases after the COVID-19 pandemic. Furthermore, although improved completeness likely reflects better data capture, it does not guarantee data accuracy; although the feedback reports were designed to examine this by including validity cross-checks, there are variables for which this was not possible. Case ascertainment and the sensitivity of the system—the number of cases captured by the system divided by the true number of cases in the community [11]—were flagged as a potential area of risk by survey respondents and remains an important attribute of surveillance that has yet to be explored.

There were additional limitations related to the survey component. The survey was circulated to the lead individual responsible for TB surveillance in each CCAA, meaning that there was only one response per region. This individual is the person receiving the feedback reports so they were considered best placed to respond to our survey; however, it is possible that this could introduce bias and may not fully reflect the breadth of opinions of individuals in the CCAA. In addition, although it was emphasized that survey responses were anonymous, there may remain some degree of social desirability bias or underrepresentation of more critical perspectives.

Comparisons with other European countries with similar decentralized systems remains a challenge given the limited number of similar studies published internationally. It is hoped that this study will serve as a useful tool for other countries wishing to carry out future assessments examining the attributes of internal completeness, acceptability, and usefulness.

Recommendations and Conclusions

Surveillance of TB is inherently challenging given the typically long-term course of the disease and treatment within the context of complex socioeconomic risk factors. This study indicated that completeness of TB surveillance data in Spain has improved following the introduction of annual feedback reports in 2022, although variation persists by variable and region. Importantly, the combination of quantitative and qualitative analyses highlights not only the measurable progress in surveillance quality but also the perceptions, barriers, and enablers that shape the system's performance. Although the reasons underlying observed improvements can be uncertain and are likely multifaceted, feedback reports were valued by regional stakeholders and appear to support a culture of continuous quality improvement.

Our findings suggest that the CNE should continue to send annual feedback reports to regions; however, they could be refined to be of even more benefit, for example, by focusing on a smaller number of key variables and indicators that align with those set out by the WHO guidance on TB surveillance published in 2024 [23]. This guidance defines a set of core indicators for countries with a case-based digital surveillance system as well as recommended reporting intervals, visualization, and interpretation for these indicators. This would support regions to target efforts where they are most effective and facilitate comparison with other countries. More broadly, strengthening alignment with international standards may help Spain position itself within global TB monitoring frameworks and contribute more effectively to regional benchmarking efforts.

In addition, the CNE should provide improved guidance on follow up of individuals who move between CCAAs or internationally during the long treatment course and how information from these individuals should be reported in the surveillance system. This could be supported by encouraging use of the ECDC "International tuberculosis care transfer form" [24]. This tool aims to support cross-border collaboration

XSL•FO
RenderX

between health care providers to ensure the care needs for people with TB are transferred securely and efficiently, minimizing delays and reducing the risk of treatment interruption. In addition, the availability of a national personal identifier in Spain would facilitate interoperability of laboratory and surveillance systems within regions, as well as support easier transfer of care of patients between regions. These improvements would not only enhance data quality but also strengthen continuity of care, an essential component of TB control in decentralized health systems.

To fully realize the potential of TB surveillance, future assessments could consider expanding beyond completeness and perceptions of feedback reports, for example using standardized methods to carry out a formal evaluation and examining a more comprehensive list of attributes, such as validity, timeliness, representativeness, and sensitivity, in line with established frameworks [11,21,25]. Future evaluations could be also carried out using the WHO TB surveillance guidelines [23] as a framework, with completeness of data fields measured against those required to calculate the recommended indicators. In addition, to examine the possible determinants of changes observed beyond descriptive analysis, approaches such as multilevel logistic regression analysis with adjustment for covariates including CCAA could be considered to complement future evaluations. Finally, embedding assessments into a continuous improvement cycle, as already initiated by revising the most recent feedback reports, will help ensure that the system evolves and continues to provide reliable and actionable information that meet both national and international TB control objectives.

Acknowledgments

We would like to thank all the professionals involved in data collecting, from clinicians, coders, administrative staff, computer scientists, epidemiologists, microbiologists, and the community participating in the National Surveillance Network. Finally, we would also like to acknowledge the important work of the members of the Spanish Working Group on Tuberculosis:

Consejería de Salud de Andalucía: Verónica Álamo; Departamento de Sanidad de Aragón: Cristina Navarro; Consejería de Salud de Asturias: Ana Fernandez Ibañez; Consejería de Salud de Baleares: Jaume Giménez Duran; Consejería de Salud de Canarias: Laura García Hernández; Consejería de Salud de Cantabria: Marta Pacheco Gorostiaga; Consejería de Salud de Castilla y León: María del Henar Marcos; Consejería de Salud de Castilla y la Mancha: Pilar Peces; Consejería de Salud de Cataluña: Laura Gavalda Mestre; Conselleria de Sanidad de la Comunitat Valenciana: Ana Sofia Lameiras Azevedo; Consejería de Salud de Extremadura: María del Mar López-Tercero Torvisco; Consejería de Salud de Galicia: Elena Cruz Ferro; Consejería de Salud de La Rioja: Carmen Quiñones; Consejería de Salud de Madrid: Mercedes Rumayor; Consejería de Salud de Murcia: Blanca Andreu Ivorra; Instituto de Salud Pública de Navarra – IdiSNA -CIBERESP: Jesús Castilla; Consejería de Salud del País Vasco: Patricia Sancho; Consejería de Salud del Ceuta: Violeta Ramos; Consejería de Salud del Melilla: Daniel Castrillejo.

The authors declare the use of generative artificial intelligence (GAI) in the research and writing process.

According to the GAIDeT taxonomy (2025), the following tasks were delegated to GAI tools under full human supervision: code optimization, proofreading and editing, and translation. The GAI tool used was: Microsoft Copilot. Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Funding

This research was supported by CIBER (Consorcio Centro de Investigación Biomédica en Red), Instituto de Salud Carlos III, Ministerio de Ciencia e Innovación and Unión Europea, European Regional Development Fund. AR is the recipient of a postdoctoral Sara Borrell contract funded by the Instituto de Salud Carlos III, Ministerio de Ciencia, Innovación y Universidades (CD23CIII/00008). The authors declare no other additional financial support was received for this work.

Data Availability

The datasets analyzed in this study are not publicly available because they contain sensitive patient information reported to the National Epidemiological Surveillance Network (RENAVE) in Spain. RENAVE data are submitted through the national electronic reporting platform SiViEs Plus and are hosted by the National Centre for Epidemiology (CNE). Access to RENAVE data can be requested formally through the CNE website (subject to institutional approval and data protection regulations) [26].

Authors' Contributions

Conceptualization: AC, AR, RA, MS, CS, RC, LHL, ZH
Data curation: AC, ZH
Formal analysis: AC, ZH
Investigation: AC, ZH
Methdology: AC, ZH
Project administration: ZH
Visualization: AC

XSL•FO
RenderX

Writing – original draft: AC, ZH
Writing – review and editing: AC, AR, RA, MS, CS, RC, LHL, ZH

Conflicts of Interest

None declared.

Multimedia Appendix 1
Example feedback report, example letter, and survey.
[DOCX File, 39 KB - publichealth_v12i1e93719_app1.docx]

Multimedia Appendix 2
Mean percentage completeness and percentage of internal completeness.
[DOCX File, 31 KB - publichealth_v12i1e93719_app2.docx]

References

1. The Lancet Microbe. Despite available tools tuberculosis is the top killer again. Lancet Microbe 2024 Dec;5(12):101044. [doi: 10.1016/j.lanmic.2024.101044] [Medline: 39608383]
2. Global tuberculosis report 2025. : World Health Organization; 2025 URL: https://www.who.int/teams/global-programme-on-tuberculosis-and-lung-health/tb-reports/global-tuberculosis-report-2025 [accessed 2026-08-21]
3. The end TB strategy. : The World Health Organization (WHO); 2014 URL: https://www.who.int/publications/i/item/WHO-HTM-TB-2015.19 [accessed 2026-08-21]
4. Tuberculosis surveillance and monitoring in Europe 2026: 2024 data. : World Health Organization; 2026 URL: https://www.who.int/europe/publications/i/item/9789289062879 [accessed 2026-08-21]
5. van Hest R, Ködmön C, Verver S, et al. Tuberculosis treatment outcome monitoring in European Union countries: systematic review. Eur Respir J 2013 Mar;41(3):635-643. [doi: 10.1183/09031936.00030612] [Medline: 22790913]
6. Scandurra G, Degeling C, Douglas P, Dobler CC, Marais B. Tuberculosis in migrants - screening, surveillance and ethics. Pneumonia (Nathan) 2020;12(1):9. [doi: 10.1186/s41479-020-00072-5] [Medline: 32923311]
7. Pedraz T, Herrera L, Vazquez MC, et al. The epidemiological situation of tuberculosis in Spain according to surveillance and hospitalization data, 2012-2020. PLOS ONE 2024;19(1):e0295918. [doi: 10.1371/journal.pone.0295918] [Medline: 38165979]
8. Vigilancia de la tuberculosis año 2024: resultados de la Red Nacional de Vigilancia Epidemiológica. : Ministerio de Ciencia, Innovacion y Universidades, Centro Nacional de Epidemiología; 2024 URL: https://cne.isciii.es/documents/d/cne/renave_informe_vigilancia-tb_-2024 [accessed 2026-08-21]
9. Fontana Campos M, Guerras JM, Chatterjee A, et al. Epidemiological characteristics and trends of paediatric tuberculosis cases in Spain, 2012-2023. Ann Med 2026 Dec;58(1):2656557. [doi: 10.1080/07853890.2026.2656557] [Medline: 41969056]
10. Protocolo de vigilancia de tuberculosis. : Centro Nacional de Epidemiologia URL: https://cne.isciii.es/documents/d/cne/protocolo-20de-20vigilancia-20de-20tuberculosis-pdf [accessed 2026-08-21]
11. Data quality monitoring and surveillance system evaluation: a handbook of methods and applications. : European Centre for Disease Prevention and Control; 2014 URL: https://data.europa.eu/doi/10.2900/35329 [accessed 2026-08-21]
12. European Centre for Disease Prevention and Control, World Health Organization. Tuberculosis surveillance and monitoring in europe 2025: 2023 data. : European Union; 2025 URL: https://data.europa.eu/doi/10.2900/2650058 [accessed 2026-08-21]
13. WHO consolidated guidelines on tuberculosis: module 4: treatment: drug-susceptible tuberculosis treatment: web annexes. : World Health Organization; 2022 URL: https://iris.who.int/items/b9dde581-515d-4283-94a1-f3f99125700a [accessed 2026-08-21]
14. Lönnroth K, Migliori GB, Abubakar I, et al. Towards tuberculosis elimination: an action framework for low-incidence countries. Eur Respir J 2015 Apr;45(4):928-952. [doi: 10.1183/09031936.00214014] [Medline: 25792630]
15. Lönnroth K, Mor Z, Erkens C, et al. Tuberculosis in migrants in low-incidence countries: epidemiology and intervention entry points. Int J Tuberc Lung Dis 2017 Jun 1;21(6):624-637. [doi: 10.5588/ijtld.16.0845] [Medline: 28482956]
16. Hargreaves S, Rustage K, Nellums LB, et al. What Constitutes an Effective and Efficient Package of Services for the Prevention, Diagnosis, Treatment and Care of Tuberculosis among Refugees and Migrants in the WHO European Region? Themed Issues on Migration and Health, VIII: WHO Regional Office for Europe; 2018. URL: http://www.ncbi.nlm.nih.gov/books/NBK519087
17. Gavaldà Mestre L, López Espinilla M, Martínez Alguacil H, Medina Maestro S, Sicart Torres E. Situació epidemiològica i tendència de l'endèmia tuberculosa a Catalunya - 2023.: Subdirecció General de Vigilància i Resposta a Emergències de Salut Pública; 2025 Feb. URL: https://scientiasalut.gencat.cat/handle/11351/12767 [accessed 2026-01-09]
18. Voldstedlund M, Haarh M, Mølbak K, MiBa Board of Representatives. The Danish Microbiology Database (MiBa) 2010 to 2013. Euro Surveill 2014 Jan 9;19(1):20667. [doi: 10.2807/1560-7917.es2014.19.1.20667] [Medline: 24434175]

XSL•FO
RenderX

19. Ward M, Brandsema P, van Straten E, Bosman A. Electronic reporting improves timeliness and completeness of infectious disease notification, The Netherlands, 2003. Euro Surveill 2005 Jan;10(1):7-8. [doi: 10.2807/esm.10.01.00513-en] [Medline: 29183539]
20. Djuretic T, Herbert J, Drobniewski F, et al. Antibiotic resistant tuberculosis in the United Kingdom: 1993-1999. Thorax 2002 Jun;57(6):477-482. [doi: 10.1136/thorax.57.6.477] [Medline: 12037221]
21. Klaucke DN, Buehler JW, Thacker SB, Parrish RG, et al. Guidelines for evaluating surveillance systems. MMWR 1998;37(S-5):1-18 [FREE Full text]
22. Tosti ME, Longhi S, de Waure C, et al. Assessment of timeliness, representativeness and quality of data reported to Italy's national integrated surveillance system for acute viral hepatitis (SEIEVA). Public Health 2015 May;129(5):561-568. [doi: 10.1016/j.puhe.2015.02.015] [Medline: 25795017]
23. Consolidated guidance on tuberculosis data generation and use module 1 tuberculosis surveillance. : World Health Organization; 2024 URL: https://www.who.int/publications/i/item/9789240075290 [accessed 2026-08-21]
24. International tuberculosis care transfer form. The European Centre for Disease Prevention and Control (ECDC). 2025 Sep 29. URL: https://www.ecdc.europa.eu/en/publications-data/international-tuberculosis-care-transfer-form [accessed 2025-10-27]
25. Calba C, Goutard FL, Hoinville L, et al. Surveillance systems evaluation: a systematic review of the existing approaches. BMC Public Health 2015 May 1;15(1):448. [doi: 10.1186/s12889-015-1791-5] [Medline: 25928645]
26. Contact. Centro Nacional de Epidemiología, ISCIII. URL: https://cne.isciii.es/en/contact [accessed 2026-08-21]

Abbreviations

AI: artificial intelligence
CCAA: autonomous community of Spain
CDC: Centers for Disease Control and Prevention
CNE: National Epidemiology Center (Centro Nacional de Epidemiología in Spanish)
ECDC: European Center for Disease Prevention and Control
EEA: European Economic Area
LOPDGDD: Ley Orgánica 3/2018 de Protección de Datos Personales y Garantía de los Derechos Digitales
RENAVE: The National Epidemiological Surveillance Network (La Red Española de Vigilancia Epidemiológica in Spanish)
TB: tuberculosis
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 18.Feb.2026; peer-reviewed by B Mastrovito, H Kato, P Stefanoff, S Ryu; revised version received 14.Jul.2026; accepted 18.Jul.2026; published 16.Sep.2026.

Please cite as:
Chatterjee A, Roy A, Amillategui R, Sastre M, Schwarz C, Cano R, Herrera-León L, Herrador Z, Spanish Working Group on Tuberculosis Data Completeness and Acceptability of Annual Feedback Reports and Tuberculosis Surveillance in Spain, 2018-2024: Mixed Methods Analysis
JMIR Public Health Surveill 2026;12:e93719
URL: https://publichealth.jmir.org/2026/1/e93719
doi:10.2196/93719

© Allegra Chatterjee, Alvaro Roy, Rocio Amillategui, María Sastre, Christine Schwarz, Rosa Cano, Laura Herrera-León, Zaida Herrador, Spanish Working Group on Tuberculosis. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 16.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluating Tuberculosis Surveillance Using Global Standards and Benchmarks in the Philippines: Mixed Methods Study

Minal Ahson[1,2*], MD, MSPH; Donna Mae Gaviola[3], BA; Stephanie O'Connor[1], MPH; Andro Gutierrez[4], BSIT; Marvin Apas[3], BA; Maria Alezandra Eguia[3], BS; Devon Ray Pacial[3], MPH; Arianne Zamora[5], MPH, MSc; Romel Lacson[5], MPH, PhD; Alethea De Guzman[3], MD, MCHM, PHSAE; Patrick K Moonan[1*], DPH, MMS

[1]Division of Global HIV and Tuberculosis, Centers for Disease Control and Prevention, 1600 Clifton Road NE, Atlanta, GA, United States

[2]Epidemic Intelligence Service, Office of Workplace Development, Centers for Disease Control and Prevention, Atlanta, GA, United States

[3]Department of Health, Epidemiology Bureau, Manila, Philippines

[4]Department of Health, Knowledge Management and Information Technology Service, Manila, Philippines

[5]Division of Global HIV and Tuberculosis, Centers for Disease Control and Prevention, Manila, Philippines

[*]these authors contributed equally

Corresponding Author:
Minal Ahson, MD, MSPH
Division of Global HIV and Tuberculosis, Centers for Disease Control and Prevention, 1600 Clifton Road NE, Atlanta, GA, United States

Abstract

Background: The Philippines accounts for 7% of the global tuberculosis (TB) burden. In 2022, an estimated 741,000 Filipinos developed TB, and 40,000 died as a result. Approximately 350,000 new TB diagnoses remain unreported annually.

Objective: In October 2023, we undertook an assessment of the accuracy, completeness, and timeliness of the national TB surveillance system using globally accepted TB surveillance standards and benchmarks.

Methods: A team of Filipino and international evaluators conducted field visits to 9 purposively selected health facilities across 4 regions. Health records were reviewed, and key informants were interviewed to assess essential activities and best practices for TB surveillance.

Results: Of 16 applicable performance standards, 4 (25%) were fully met, 9 (56.3%) were partially met, and 3 (18.8%) were not met. This was an improvement from a similar assessment conducted in 2019. Although electronic case-based recording and reporting had expanded greatly since 2019, large local variations in the adoption of quality assurance practices were evident. More than 50% of persons with reported TB had no bacteriological confirmation despite the presence of a functional national laboratory network. Underreporting due to underdiagnosis was also noted. Delayed reporting was common, especially in locations with limited internet connectivity or limited access to anti-TB medications.

Conclusions: Issues related to data quality assurance, gaps in case verification, and timeliness emerged as potential threats to data integrity and surveillance fidelity. Enhanced monitoring and evaluation, along with tailored studies, such as a national TB inventory study, could assist in quantifying potential underreporting and clinical overdiagnosis, guiding future funding, and assessing progress toward elimination targets.

(JMIR Public Health Surveill 2026;12:e77058) doi:10.2196/77058

KEYWORDS

surveillance; tuberculosis; tuberculosis surveillance; disease surveillance; public health surveillance; data quality; case notification; electronic health records; health information systems; Philippines; surveillance evaluation

Introduction

Surveillance is a key component of global efforts to eliminate tuberculosis (TB) [1,2]. To support countries in strengthening their surveillance systems and achieving TB management goals, the World Health Organization (WHO) developed a standardized assessment tool that countries can use to evaluate the performance of their TB surveillance and vital registration systems [3].

The Philippines currently ranks fourth in the list of countries with the highest TB incidence and accounts for 7% of the global TB burden [4]. WHO incidence estimates for 2022 indicated approximately 741,000 incident TB cases in the Philippines compared with 435,890 notified new and retreatment cases, yielding a treatment coverage of 59% [5]. The estimated number

XSL•FO
RenderX

of unreported TB cases reflects the residual between WHO-modeled incidence and observed notifications (derived from the WHO's standardized incidence estimation framework rather than from calculations performed by the authors) [5]. The Philippines Department of Health introduced an internet-based TB surveillance system, the Integrated TB Information System (ITIS), in 2016 [6]. This facilitates the systematic recording and reporting of case-based information in accordance with National Tuberculosis Control Program (NTP) guidance [7,8]. Currently, TB case data are captured from all primary public facilities and several tertiary facilities. A growing number of private healthcare practitioners also participate in mandatory TB notification [9]. All participating network facilities collect TB information in logs and then use ITIS to enter data electronically. Quality assurance procedures guide systematic data entry. Data are reviewed and validated by provincial and regional NTP offices each quarter. The national NTP office reviews data annually and produces an annual TB report to document programmatic performance over time [10].

We aimed to assess the ability of the NTP to accurately enumerate TB cases and deaths using the WHO's TB surveillance assessment tool, which comprises a checklist of standards and benchmarks covering metrics such as TB data quality, surveillance system coverage, and treatment coverage [11]. The previous surveillance evaluation conducted in 2019 using the same standards and benchmarksfound that a total of 5 benchmarks were met, 1 was partially met, 6 were not met, and 5 were not applicable (Multimedia Appendix 1).

Methods

Overview

In October 2023, a team of Filipino and international evaluators from the Epidemiology Bureau of the Philippines Department of Health and the US Centers for Disease Control and Prevention (CDC) conducted field visits to 9 health facilities in 4 of the country's 18 regions: Calabarzon, Cordillera Administrative Region, Ilocos Region, and National Capital Region. The facilities were selected to gain perspectives from a variety of settings and included urban and rural, primary and tertiary care, and private and government settings. These included regional and provincial NTP offices. We interviewed stakeholders (including clinicians, program managers, and frontline workers) to learn about essential activities and best practices regarding TB surveillance. Each visit included a presentation from the facility, discussion of surveillance processes, and a walkthrough of TB-specific areas. Selected paper-based registries, clinical records, and electronic-based data (ITIS) were reviewed for accuracy, completeness, and timeliness. Additional publicly available data sources were also consulted [12-14]. After site visits and discussion with local and national partners, the WHO's assessment tool checklist was completed and compared with the previous assessment conducted in 2019 [11,12]. The previous TB surveillance assessment used the same WHO standards and benchmarks [12]. That assessment included field visits to a comparable number and mix of health facilities across selected regions of the Philippines, representing varied geographic, administrative, and health system contexts. While the specific facilities and regions assessed in 2019 were not identical to those included in the 2023 assessment, both evaluations were designed to capture perspectives from urban and rural settings and from multiple levels of the health system, allowing for broad comparability of findings over time. This evaluation focused on 5 standards: TB surveillance system data quality, system coverage and vital registration, specific populations, TB treatment outcomes, and programmatic management of TB preventive treatment. We used the same set of 16 standards and benchmarks as in the previous assessment; changes between successive versions of the checklist were relatively minor. An indicator was considered "partially met" if at least one, but not all, of the outcome measures for each benchmark was met.

Ethical Considerations

This activity was reviewed by the CDC and deemed not research and was conducted consistent with applicable federal law and CDC policy (see, eg, Title 45, part 46.102(l)(2), and Title 21, part 56, of the Code of Federal Regulations and Title 42, §241(d); Title 5, §552a; and Title 44, §3501 et seq of the United States Code [15]). As this activity was classified as a public health surveillance evaluation rather than human subjects research, formal written informed consent was not required. Key informants were verbally informed of the voluntary nature of their participation prior to being interviewed. No personally identifiable patient information was collected or retained. Participants received no compensation.

Results

On the basis of 2022 data, of the 17 total checklist indicators, of which 1 (5.9%) was not applicable, 4 (23.5%) were fully met, 9 (52.9%) were partially met, and 3 (17.6%) were not met (Multimedia Appendix 1).

TB Surveillance System Data Quality

Benchmarks B1.1 to 1.3 (B1.4 was not applicable to electronic systems), which relate to case-level data capture, were fully met (Multimedia Appendix 1). Case definitions were consistent with WHO guidelines and the TB case-based reporting framework (B1.1) [2,3]. Persons with TB were identified using unique, anonymized registration numbers. The surveillance system captured a minimum set of variables for all cases (B1.2). These included sociodemographic characteristics (ie, age, sex, and residential address), date of registration, clinical and diagnostic information (ie, anatomical site of disease, radiographic interpretations, and treatment history), laboratory test results, and current treatment regimens. Test results and treatment status were recorded as patients moved through the care pathway. Information was captured in near real time; once data was entered at local levels, they were immediately available for view by the higher-level offices. We observed that ITIS was widely used and that sufficient training and infrastructure were in place. This included training on data collection and entry into the electronic system. Infrastructure included computer availability, electricity, and internet connection at various facilities. However, several users reported that, on occasion, ITIS was slow to load and unsaved data could be lost due to unstable internet connectivity. In facilities with insufficient internet

XSL•FO
RenderX

bandwidth, paper-based registers were transferred for entry at hubs with adequate connectivity. In addition, ITIS was not fully integrated with laboratory information systems. Laboratory staff manually entered results into ITIS; automated reporting of molecular-based testing (ie, Xpert MTB/RIF; Cepheid) was not in place.

The B1.5 benchmark, which measures the accuracy, completeness, and internal consistency of data in the national database, was partially met. ITIS has integrated mechanisms to ensure data quality, such as checking for duplicate names with the same spelling. However, although less than 1% of records were invalid or incomplete (judged against a minimal set of variables, including sociodemographic characteristics, date of registration, and clinical and diagnostic variables), the internal mechanisms did not include checking for duplicate entries with different identifiers or spelling, and no evidence of addressing these potential duplications was observed. Furthermore, data validation checks were inconsistent and widely varying. ITIS features a simple, one-click validation step, which users were able to demonstrate during site visits. While ease of use can be an advantage, the current validation process may be insufficient to verify the accuracy of the entered data. Because HIV-associated information is restricted at some levels (ie, patient records are confidential and not integrated into ITIS), validation of records for people with HIV-TB appeared limited. Electronic surveillance systems provide opportunities to automate a portion of the data quality assurance process, but only minimal automated checks for internal validity were in place.

The B1.6 and B1.7 benchmarks related to the internal and external consistency of TB surveillance data, respectively, were also only partially met. Among newly notified persons with TB in 2022, a total of 8% were children aged 0 to 14 years. This was within the recommended acceptable range standard (5% - 15% for low- and middle-income countries). Across all ages, only 50% of reported persons with TB had bacteriological confirmation, well below the 70% to 90% recommended standard [4]. Surveillance data appeared to be internally consistent over time even after considering year-to-year differences experienced during the COVID-19 pandemic. The ratio of pulmonary to extrapulmonary TB was 37:8, and for male to female patients, it was 26:14, which was consistent with pre–COVID-19 levels. The proportion of newly notified TB cases had steadily increased to 49.6% (Table 1; [13,14,16]).

Table . Selected tuberculosis (TB) surveillance system indicators (the Philippines, 2018 - 2022)[a].

	2018	2019	2020	2021	2022
TB programmatic indicator					
Ratio of notified pulmonary to extrapulmonary TB cases	32:7	36:7	22:5	28:6	37:8
Ratio of male TB cases to female TB cases	21:11	24:13	2:1	2:1	26:14
Childhood TB cases out of all TB cases (%)	13.9	11.6	7.7	7.0	8.1
Proportion of newly notified TB cases with bacteriological confirmation (%)	36.1	33.9	43.9	45.0	49.6
Year-to-year change in new case notification rate (%)	8.4	9.2	–59.5	20.2	26.2
Vital registration indicator					
Proportion of documented primary cause of death among all registered deaths (%)	96.3	96.3	95.7	93.5	94.2
Deaths attributed to TB—all forms, n (%)	23,409 (4.0)	23,877 (3.8)	19,430 (3.2)	19,952 (2.3)	20,588 (3.0)

[a]As a population-level health system indicator, the universal health coverage service index was 48 in 2010, 57 in 2015, 60 in 2017, 60 in 2019, and 58 in 2021 [13,14,16].

System Coverage and Vital Registration

B1.8 assesses how well the surveillance system captures TB cases. By default, ITIS only captures persons who received treatment at NTP facilities. Despite mandating notifications from all private health care providers who diagnose persons with TB in 2018, only selected health care providers complied, leading to occasional nonreporting among private sector facilities. Without a complete census or representative survey of private facilities, it is unknown how many facilities are not reporting in ITIS.

XSL•FO
RenderX

In a few facilities, staff were uncertain of who was responsible for reporting TB cases and validating records; in others, staff asserted that they were only responsible for entering data into ITIS for TB treatment initiations, not TB diagnoses. Both scenarios may result in underreporting. A few facilities reported waiting to encode diagnosed patients until treatment was initiated. This practice had major implications for reporting timeliness in settings where supply chain issues have caused months-long treatment initiation delays. Facilities without the human resource capacity to enter data into ITIS on-site or with limited internet connectivity sent TB registers to regional or provincial offices for encoding, resulting in reporting and patient referral delays. In light of such evidence, the B1.8 benchmark was deemed as only partially met.

Although access to health care in the Philippines has steadily improved over the past decade (Table 1), the universal health coverage index was 58 in 2021, well below the recommended benchmark of 80. Thus, the B1.9 benchmark was not met (Multimedia Appendix 1). However, benchmark B1.10, which relates to the coverage and quality of vital registration data provided by the civil registration and vital statistics system, was partially met. According to the Philippine Statistics Authority, the primary cause of death was documented for 640,175 (94.2%) of 679,766 registered deaths in 2022, including 20,588 (3%) attributed to all forms of TB (Table 1). However, the extent of mortality reporting and recording was variable between regions.

Specific Populations

Of the 3 supplementary indicators pertaining to specific TB populations (people with drug-resistant TB, people with HIV coinfection, and children), only the B2.1 benchmark was met. In 2022, 52% of new cases were diagnosed without laboratory confirmation despite increased Xpert MTB/RIF testing. However, among patients with bacteriologically confirmed pulmonary TB, the vast majority (97%) had drug susceptibility test results, and a total of 9791 rifampin-resistant cases were detected. This meant that benchmark B2.1 was met.

Benchmark B2.2 was not met because, in 2022, the proportion of patients with newly reported TB tested for HIV (70%) remained below the recommended target of 80%. The data in Table 2 show that not only has the annual proportion of patients with newly notified TB tested for HIV increased in recent years but the proportion of patients with TB who tested positive for HIV has also remained low (range 0.8% - 1.5%). Moreover, the proportion of persons living with HIV who initiated antiretroviral treatment after testing positive in the last 3 years was exceptional; only 1 person with newly discovered TB and HIV coinfection did not start antiretroviral treatment.

Table . HIV testing and percentage of positivity among persons with newly reported tuberculosis (TB; the Philippines, 2018 - 2022)[a].

Indicator	2018	2019	2020	2021	2022
Persons with TB with a documented HIV test result, n (% of total)	101,681 (31)	136,749 (38)	85,672 (39)	109,952 (39)	263,846 (70)
Persons with TB who tested positive, n (% of persons tested)	1477 (1.5)	1615 (1.2)	898 (1.0)	1350 (1.2)	2201 (0.8)
Persons with TB who tested positive and initiated ART[b], n (% of positive cases)	1350 (91)	1509 (93)	898 (100)	1349 (99)	2201 (100)

[a]Source: Global TB Data Repository [14] .

[b]ART: antiretroviral therapy.

In 2022, children and adolescents aged 0 to 14 years accounted for 8% of all reported persons with TB (Table 1). Benchmark B2.3 requires a rate ratio of cases notified among children aged 0 to 4 years to those aged 5 to 14 years in the range of 1.5 to 3.0; the observed ratio of 0.82 fell outside the expected range, and therefore benchmark B2.3 was not met. A ratio lower than expected suggests that TB cases are proportionally higher among older children, which may reflect either underdiagnosis in children below the age of 5 years, overdiagnosis in children and adolescents aged 5 to 14 years, or both. Formal investigations of childhood TB epidemiology to determine the extent of under- or overreporting have not been conducted.

TB Treatment Outcomes

Both treatment outcome standards (B3.1 and B3.2) were partially met. We observed variations in the accuracy, completeness, and consistency of TB treatment recording and reporting by both setting and region. However, individual treatment outcomes were disaggregated by treatment history, HIV status, and drug resistance test results. Overall, there was limited evidence of local-level surveillance data use. Facilities generally focused on variables associated with targets and on measuring success based primarily on target achievement. We found no evidence that checks had been made to verify that the number of patients with an assigned treatment outcome in any given year matched the number of patients notified the preceding year.

Programmatic Management of TB Preventive Treatment

While monitoring and evaluation indicators for TB preventive therapy (TPT) were consistent with global guidance [17] and the data collected included the minimum set of variables (people living with HIV, children aged <5 years, individuals aged ≥5 years, and TPT completion), monitoring of TPT was noted in a few settings, and there was significant variation. Thus, B4.2

XSL•FO
RenderX

(TPT programmatic management data were accurate, complete, and consistent) was only partially met (Multimedia Appendix 1).

Discussion

Principal Findings

The global standards and benchmarks for TB surveillance and vital registration systems were developed to be applicable to TB control programs in various contexts [3]. By implementing this tool, we demonstrated that it is both applicable and feasible in the Filipino context. Overall, this assessment was an improvement from the previous assessment in 2019, in which 5 benchmarks were met, 1 was partially met, 6 were not met, and 5 were not applicable [12]. In this assessment, fewer indicators were not met, and more were fully or partially assessed.

This assessment revealed progress in several key areas of TB surveillance since the last review. Most notable was the impressive expansion and implementation of electronic case-based recording and reporting. However, we also identified several opportunities for further improvement.

Improving Data Quality

Data quality issues predominantly stemmed from staff time constraints and a lack of coordination between facilities. Procedures were in place to encourage data quality assurance prior to reporting. However, greater use of automated checks for internal validity, data completeness, and deduplication could reduce the burden on data entry staff and the risk of human error in the reporting process. Losses to follow-up and the number of duplicate records could be reduced by implementing monitoring at higher administrative levels of patients who are transferred but unclaimed in ITIS.

It should be noted that, while the proportion of bacteriologically confirmed pulmonary TB with drug susceptibility test results was considerably higher than the 80% recommended standard (Multimedia Appendix 1), the proportion of all TB cases with bacteriological confirmation still ranked among the lowest in the world [5]. Proposed explanations for the low rates of bacteriological confirmation include health care provider mistrust of molecular testing results, underuse of molecular diagnostics, and the high number of specimens rejected by laboratories due to low-quality sputum collection.

Hastening the Transition From Aggregated Paper-Based to Electronic Case-Based Recording and Reporting

ITIS represents a triumph in the transition to electronic, person-level reporting. However, it is hampered by slow loading speeds, which prevents users from accessing the system. Taking steps to reduce loading times, such as increasing the number of servers, would significantly improve both user satisfaction and the timeliness of data entry and is highly recommended. A mobile-based platform, which does not require a continuous internet connection or expensive hardware, would also help overcome connectivity barriers. Initiating automated transmission of molecular test results could also improve efficiency and data quality. Shifting from manual to automated validation measures offers the dual benefit of reducing the burden on staff and improving data quality. Increasing the number of dedicated personnel at each health facility would expedite implementation of these changes and ensure continuity.

Reducing Underreporting and Underdiagnosis

Ensuring that all persons with TB have timely access to treatment will help reduce household contact transmission risk and case reporting delays; resolving supply chain issues will further support this. Increasing the proportion of people with bacteriologically confirmed TB would also help tackle the problem of underreporting and underdiagnosis; while clinical diagnoses are important in the absence of laboratory-based alternatives, they can result in under- or misdiagnosis, especially for drug-resistant TB. Pooling samples and providing standardized sample collection instructions to facilities could improve bacteriological diagnosis and reduce the number of rejected sputum samples. Thus, studies to determine the causes of overreliance on clinical diagnosis and the extent of underreporting in the private sector are strongly recommended.

Promoting the Use of Surveillance Data to Drive Public Health Decision-Making

Informing public health action is a central objective of conducting disease surveillance activities. Training public health workers to analyze and use their own data could be undertaken in collaboration with epidemiology and surveillance units as well as the NTP.

Limitations

This evaluation took place over a short period and included a small number of field observations at 9 representative facilities. This limitation is particularly important given the geographic and socioeconomic heterogeneity of the Philippines. In addition, we focused on direct observation in reporting facilities and interviews with key informants rather than the more in-depth data evaluation that traditionally takes place during a larger and more formal epidemiologic or TB program review.

Conclusions

Data quality assurance, case verification, and timeliness were identified as potential risks to data integrity and TB surveillance fidelity in the Philippines. Enhanced monitoring and evaluation, along with tailored studies, such as a national TB inventory study, are proposed as possible measures to quantify underreporting and clinical overdiagnosis, guide future funding, and assess progress toward TB elimination targets.

Funding

This report was funded by the Department of Health Epidemiology Bureau of the Government of the Philippines, with technical assistance, personnel, and travel funds from the US Centers for Disease Control and Prevention.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: MA, SOC, PKM
Data curation: MA, SOC, PKM
Formal analysis: MA, SOC, PKM
Supervision: PKM
Writing—original draft: MA, SOC, PKM
Writing—review and editing: all authors

Conflicts of Interest

None declared.

Multimedia Appendix 1
Benchmarks for tuberculosis surveillance in the Philippines (2019 and 2023).
[DOCX File, 39 KB - publichealth_v12i1e77058_app1.docx]

References

1. Castro KG. Tuberculosis surveillance: data for decision-making. Clin Infect Dis 2007 May 15;44(10):1268-1270. [doi: 10.1086/514351] [Medline: 17443461]
2. Uplekar M, Weil D, Lonnroth K, et al. WHO's new end TB strategy. Lancet 2015 May 2;385(9979):1799-1801. [doi: 10.1016/S0140-6736(15)60570-0] [Medline: 25814376]
3. Standards and benchmarks for tuberculosis surveillance and vital registration systems. In: Checklist and User Guide, 1st edition: World Health Organization; 2014. URL: https://www.who.int/publications/i/item/9789241506724 [accessed 2026-03-30]
4. Global tuberculosis report 2023. World Health Organization. 2023. URL: https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2023 [accessed 2024-05-01]
5. Global tuberculosis report 2022. World Health Organization. 2022. URL: https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2022 [accessed 2024-05-01]
6. Implementing guidelines on Integrated Tuberculosis Information System (ITIS). National Tuberculosis Program, Department of Health. 2015. URL: https://ntp.doh.gov.ph/download/ao2015-0024 [accessed 2024-05-01]
7. Use of Integrated TB Information System (ITIS). National Tuberculosis Program, Department of Health. 2021. URL: https://ntp.doh.gov.ph/download/dc2021-0573 [accessed 2024-05-01]
8. Guidelines for the National Tuberculosis Control Program Manual of Procedures, 6th edition: National Tuberculosis Program, Department of Health; 2021. URL: https://ntp.doh.gov.ph/download/guidelines-for-the-national-tuberculosis-control-program-manual-of-procedures-6-ed-ntp-mop-online-trainings [accessed 2026-03-30]
9. Implementing rules and regulations of Republic Act No. 10767 otherwise known as "Comprehensive Tuberculosis Elimination Plan Act of 2016". National Tuberculosis Program, Department of Health. 2017. URL: https://ntp.doh.gov.ph/download/mc2017-0014 [accessed 2024-05-01]
10. Philippines National Tuberculosis Report 2022. : National Tuberculosis Program, Department of Health; 2022 URL: https://ntp.doh.gov.ph/download/philippines-national-tuberculosis-report-2022 [accessed 2024-05-01]
11. Consolidated Guidance on Tuberculosis Data Generation and Use: Module 1: Tuberculosis Surveillance: Web Annex B: Standards and Benchmarks for Tuberculosis Surveillance and Vital Registration Systems: Checklist, 2nd edition: World Health Organization; 2024. URL: https://iris.who.int/handle/10665/376483 [accessed 2024-05-01]
12. 2022 Joint Program Review for the Philippines National TB and HIV Programs. Republic of the Philippines National Tuberculosis Control Program. 2023. URL: https://ntp.doh.gov.ph/download/2022-jpr-report/ [accessed 2026-01-02]
13. Philippine national health account. Philippine Statistics Authority. 2024. URL: https://psa.gov.ph/statistics/pnha [accessed 2024-05-01]
14. Tuberculosis profile: Philippines. World Health Organization. 2024. URL: https://worldhealthorg.shinyapps.io/tb_profiles/?_inputs_&tab=%22charts%22&lan=%22EN%22&iso3=%22PHL%22&entity_type=%22country%22 [accessed 2024-05-01]
15. Distinguishing public health research and public health nonresearch. Centers for Disease Control and Prevention. 2010. URL: https://stacks.cdc.gov/view/cdc/24235 [accessed 2026-03-30]
16. Universal health coverage (UHC) service coverage index. World Health Organization. URL: https://data.who.int/indicators/i/9A706FD [accessed 2026-03-20]
17. WHO operational handbook on tuberculosis: module 1: prevention: tuberculosis preventive treatment. : World Health Organization; 2020. URL: https://www.who.int/publications/i/item/9789240002906 [accessed 2024-05-01]

XSL•FO
RenderX

Abbreviations

CDC: Centers for Disease Control and Prevention
ITIS: Integrated Tuberculosis Information System
NTP: National Tuberculosis Control Program
TB: tuberculosis
TPT: tuberculosis preventive therapy
WHO: World Health Organization

Edited by A Mavragani; submitted 07.May.2025; peer-reviewed by C Obianyo, M Macaraig, M Donadel; revised version received 30.Jan.2026; accepted 24.Feb.2026; published 06.Apr.2026.

Please cite as:
Ahson M, Gaviola DM, O'Connor S, Gutierrez A, Apas M, Eguia MA, Pacial DR, Zamora A, Lacson R, De Guzman A, Moonan PK
Evaluating Tuberculosis Surveillance Using Global Standards and Benchmarks in the Philippines: Mixed Methods Study
JMIR Public Health Surveill 2026;12:e77058
URL: https://publichealth.jmir.org/2026/1/e77058
doi:10.2196/77058

© Minal Ahson, Donna Mae Gaviola, Stephanie O'Connor, Andro Gutierrez, Marvin Apas, Maria Alezandra Eguia, Devon Ray Pacial, Arianne Zamora, Romel Lacson, Alethea De Guzman, Patrick K Moonan. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 6.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Global Heated Tobacco Product User Estimates, 2014-2024: Descriptive Surveillance Study Using Manufacturer Disclosures

Giorgi Mzhavanadze[1,2], MA; Gerry V Stimson[1], PhD; Tomasz Jerzyński[1,3], PhD

[1]Knowledge-Action-Change, 5 Heatherfold Way, London, United Kingdom

[2]Healthy Initiatives (GO Zdorovi Iniciativi), Kyiv, Ukraine

[3]Center for Social Research Methodology, Robert Zajonc Institute for Social Studies (ISS), University of Warsaw, Warsaw, Poland

Corresponding Author:
Giorgi Mzhavanadze, MA
Knowledge-Action-Change, 5 Heatherfold Way, London, United Kingdom

Abstract

Background: Heated tobacco products (HTPs) expanded rapidly after 2014; yet globally, comparable estimates of the number of HTP users remain limited. National surveys rarely include standardized HTP measures, and international surveillance systems do not provide harmonized global user counts.

Objective: The aim of the study is to construct a reproducible annual global time series of HTP user estimates for 2014‑2024 using publicly available manufacturer disclosures, quantify uncertainty, and compare these estimates with user counts derived from nationally representative survey prevalence data where available.

Methods: We compiled annual HTP user counts and heated-tobacco stick shipment volumes from publicly available reports of major transnational tobacco companies. In the primary series, we used company-reported user counts when available. For company-year observations without reported user counts, we estimated user numbers by converting reported shipment volumes into implied user counts, applying brand-specific assumptions on average daily consumption per user. As an alternative sensitivity analysis, all company user counts were excluded, and all shipments were converted using a common literature-based consumption parameter. As a complementary check, nationally representative current-use prevalence data from 35 countries were converted to user counts.

Results: Global HTP users reached 48.9 million in 2024 in the primary series (range 45.6‑52.1 million). The shipment-only series showed higher totals (67.9 million; range 59.7‑78.7 million). Survey-based estimates yielded 21.8 million users across 35 countries, reflecting partial geographic coverage.

Conclusions: A transparent, updateable global time series documents rapid HTP growth from 2014 to 2024, with 2024 totals ranging from 45.6 to 78.7 million users depending on the estimation approach. As standardized HTP measures are incorporated into national surveys, future estimates can be refined, and uncertainty reduced.

(JMIR Public Health Surveill 2026;12:e88761) doi:10.2196/88761

KEYWORDS

heated tobacco products; global estimates; tobacco and nicotine market; tobacco industry data; user prevalence; surveillance and monitoring

Introduction

Heated tobacco products (HTPs) are electronic devices designed to heat processed tobacco and generate an aerosol without burning the tobacco. Unlike conventional cigarettes, which combust tobacco at high temperatures (>600 °C), HTPs typically operate at lower temperatures (250‑350 °C) using an electrically powered heating element. HTP systems generally consist of a specially manufactured tobacco consumable (commonly in the form of sticks), a heating mechanism, and a battery-powered control unit that regulates device operation [1-3].

Introduced commercially in 2014, HTPs have become a notable segment of the global nicotine market, expanding rapidly from initial launches in Japan and South Korea [4-9]. Despite this expansion, globally comparable estimates of the number of HTP users remain scarce. National tobacco surveys rarely include HTP questions, and where they do, definitions and measurement approaches vary, limiting cross-country comparability and time-series analysis. The most comprehensive systematic reviews to date highlight persistent surveillance gaps and heterogeneity in measures [4,5]. These reviews reflect prevalence data from surveys conducted up to 2020 and identify current adult HTP prevalence estimates from 34 countries; only a small number of countries (Japan, Italy, the United States, and the United Kingdom) contribute 3 or more prevalence

XSL•FO
RenderX

estimates across different years [5]. By contrast, HTPs are legally available in at least 72 countries as of 2025 [10], indicating a substantial mismatch between product availability and the coverage of nationally representative prevalence evidence. To date, global surveillance has not produced an annual time series of the number of HTP users that spans the full period of commercial uptake.

Similar measurement challenges previously existed for other noncombustible nicotine products, e-cigarettes, where global user estimates were initially constructed by compiling heterogeneous sources and then iteratively updated as evidence improved [11-13]. The World Health Organization has recently reported adult e-cigarette user estimates based on nationally representative surveys. However, a comparable evidence base does not yet exist for HTP, which are not consistently separated from broader smoked tobacco categories [14]. A systematic assessment of Eurobarometer tobacco surveys documented that HTP-specific questions were not introduced until 2020 (6 years after HTPs entered European markets) and identified significant inconsistencies in question wording and product definitions across survey waves, further limiting comparisons [15].

This study addresses that empirical gap by generating a global time series of HTP user estimates for 2014 - 2024. In contrast with existing estimates, which are primarily survey-based, we systematically compile manufacturer-reported user counts and sales volumes and convert volumes to user numbers when necessary using explicit assumptions about daily consumption. The rationale for using company data is pragmatic: such disclosures are often the only publicly available and regularly updated source of information on global HTP uptake. Furthermore, the HTP market is highly concentrated among a small number of multinational firms [16,17], which allows the construction of consistent and comprehensive annual estimates.

The aim of this study is to construct a transparent and reproducible global time series estimate of the number of HTP users from 2014 to 2024 using publicly available manufacturer disclosures. Specifically, we address the following research questions: (1) What is the estimated global number of HTP users annually from 2014 to 2024 based on manufacturer-reported data? (2) How do estimates differ across alternative estimation approaches: company-reported user counts, shipment-based conversions, and survey-based prevalence? (3) What are the principal sources of uncertainty in these estimates, and how sensitive are results to key methodological assumptions?

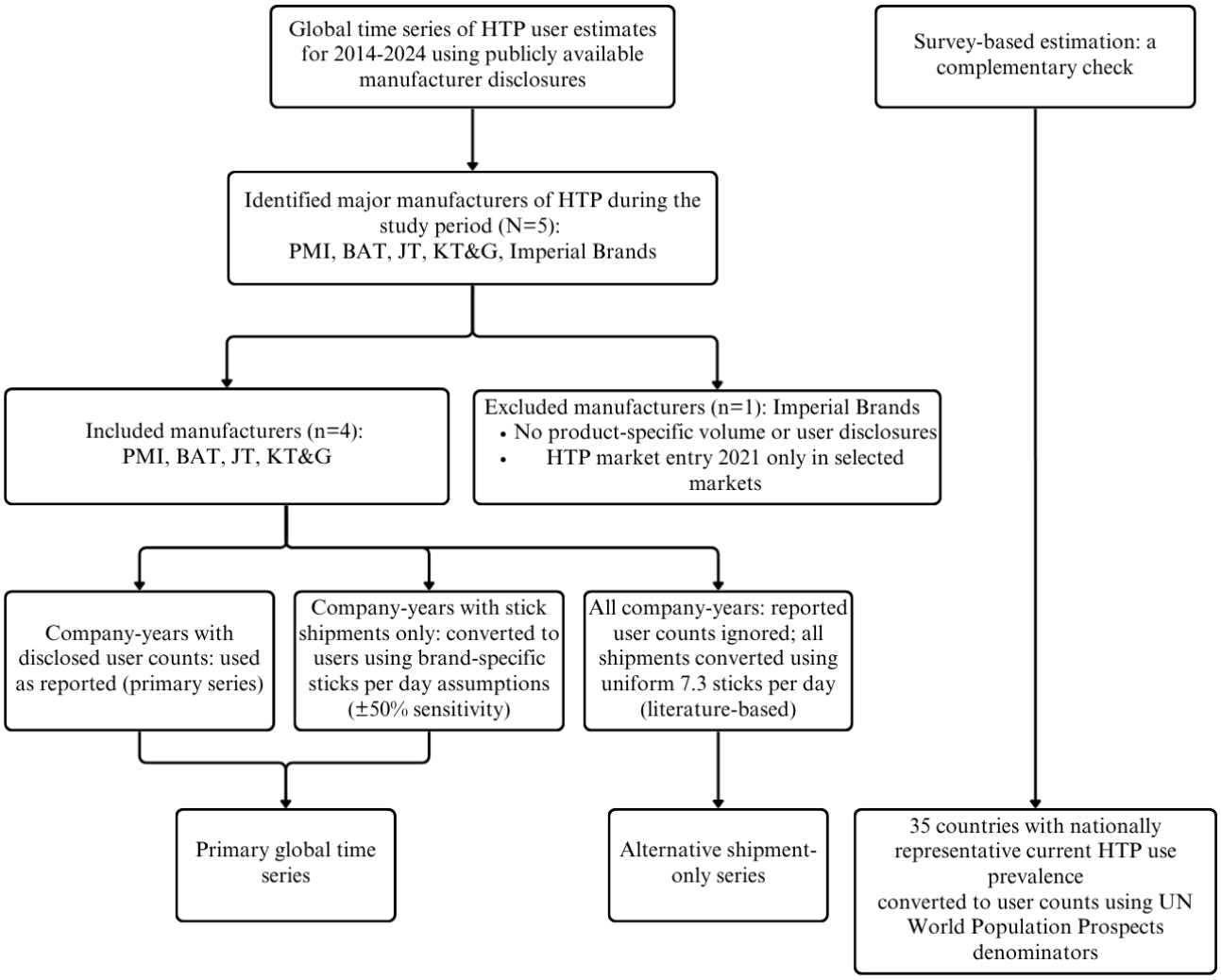
Methods

Overview

This section describes the study design, data sources, analytical procedures, and ethical considerations. The study is a descriptive global surveillance analysis based on secondary data from publicly available manufacturer disclosures. No primary data were collected. Figure 1 summarizes the data identification and estimation workflow.

XSL•FO
RenderX

Figure 1. Data identification and estimation workflow for global HTP user estimates, 2014 - 2024. BAT: British American Tobacco; HTP: heated tobacco product; JT: Japan Tobacco; KT&G: Korea Tobacco & Ginseng Corporation; PMI: Philip Morris International.



Study Design

This study estimates the global number of HTP users from 2014 to 2024 by synthesizing secondary data from publicly available manufacturer disclosures to construct a replicable annual time series. Two parallel disclosure-based global series were developed to address definitional heterogeneity and reporting differences across manufacturers: a primary series combining company-reported user counts with shipment-based conversions, and an alternative series relying solely on shipment-based conversions using a common literature-based consumption parameter. As a complementary check, nationally representative survey prevalence data were converted to user counts independently.

Data

Manufacturer-reported data were extracted from publicly available corporate disclosures (annual reports, investor presentations, regulatory filings, and press releases). These sources provided information on (1) HTP user or consumer counts and (2) annual shipment volumes of HTP consumables (sticks), typically reported at the global level [18-22].

Companies were included if they (1) manufactured and sold HTPs at a material global scale during 2014 - 2024 and (2) publicly disclosed annual user counts or stick shipment volumes sufficient to construct year-by-year estimates. Five major manufacturers of HTP were identified during the study period: Philip Morris International (PMI; IQOS), British American Tobacco (BAT; glo), Japan Tobacco (JT; Ploom), Korea Tobacco & Ginseng Corporation (KT&G; lil), and Imperial Brands. Four companies met the inclusion criteria: PMI, BAT, JT, and KT&G. Imperial Brands was excluded because the company entered the HTP market only in 2021 in selected European markets and does not disclose product-specific volumes or user counts [23]. No other manufacturers met the inclusion threshold of material global scale and public disclosure. For each included company, all available annual observations from 2014 to 2024 were extracted. Where multiple disclosures existed for a single company-year, the most recent published figure was used.

Reporting conventions differed across companies. PMI and BAT publish headline user or consumer counts alongside annual HTP stick shipments. PMI reports “total IQOS users,” defined as adults who used PMI HTPs for at least part of their daily tobacco consumption in the past 7 days, and in some reports, also distinguishes “switched” users. BAT reports user counts only in selected years and does not provide a formal definition. JT does not disclose global user counts; instead, it reports stick

XSL•FO
RenderX

volumes in selected years and, more recently, market-share indicators based on stick volumes. KT&G generally reports stick shipments and does not publish global user counts.

Notably, under the PMI-KT&G distribution or supply agreement, initiated in January 2020 (extended in January 2023), PMI commercializes KT&G's HTPs outside South Korea [24]. PMI's disclosures state that, from December 2020, its reported figures include licensed KT&G products in non-Korean markets [25]. To prevent double-counting, we therefore treat PMI's 2021 onward published user totals as already inclusive of KT&G users outside South Korea. KT&G's domestic (South Korea) stick shipments are converted to user estimates and counted separately, while KT&G volumes outside South Korea are not added on top of PMI's user totals.

Table 1 summarizes the availability of company-reported user counts and stick shipment data by company and year. It documents the availability and structure of input data rather than estimated results and functions as a data inventory.

Table . Manufacturer-reported heated tobacco product (HTP) user counts and heated-tobacco stick shipment volumes used to construct global HTP user estimates, by company and year, 2014 - 2024[a].

Year	PMI[b] (IQOS)		KT&G[c] (lil)		BAT[d] (glo)		JT[e] (Ploom)	
	Users (million)	HTP sticks (billion)	Users (million)	HTP sticks (billion)	Users (million)	HTP sticks (billion)	Users (million)	HTP sticks (billion)
2014	N/A[f]	N/A	0.0	0.0	0.0	0.0	0.0	0.0
2015	N/A	N/A	0.0	0.0	0.0	0.0	0.0	0.0
2016	2.1	7.4	0.0	0.0	N/A	N/A	0.0	0.0
2017	6.9	36.2	N/A	0.0	N/A	2.2	0.0	0.0
2018	9.6	41.4	N/A	1.2	N/A	7.0	0.0	0.0
2019	13.6	59.7	N/A	2.5	3.0	9.0	N/A	N/A
2020	17.6	76.1	N/A	1.5	4.0	11.0	N/A	N/A
2021	21.2	95.0	N/A	3.7	6.8	19.0	N/A	3.0
2022	24.9	109.2	N/A	5.8	7.2	24.0	N/A	6.7
2023	28.6	125.3	N/A	8.2	8.6	24.0	N/A	11.1
2024	32.2	139.7	N/A	8.3	10.2	21.0	N/A	14.1

[a]IQOS, lil, glo, and Ploom are heated tobacco brand names.

[b]PMI: Philip Morris International.

[c]KT&G: Korea Tobacco & Ginseng Corporation.

[d]BAT: British American Tobacco.

[e]JT: Japan Tobacco

[f]N/A: not available.

Analysis

To address definitional heterogeneity and variable disclosure completeness across manufacturers, we constructed 2 parallel global time series.

Primary Time Series: Company-Reported Users and Shipment-Based Conversions

For company-years with disclosed HTP user counts, we used those values as reported. For company-years with only HTP stick shipment volumes, we converted shipments to implied user counts using:

$$\text{Usersy} = \frac{\text{HTP stick shipmentsy}}{\text{average number of sticks per user per day} \times 365}$$

Brand-specific assumptions were applied as follows: (1) KT&G—12 sticks per day (derived from years in which PMI disclosed both sticks and users) and (2) BAT and JT—7.6 sticks per day (based on BAT years reporting both sticks and users). For each year, we then aggregated company-level user counts (reported or converted) across firms to produce global totals from 2014 to 2024.

Additional assumptions were required where data were unavailable:

- JT's global heated tobacco stick share in 2017 - 2018 is assumed to be equal to its 2019 level (2.5%).
- BAT launch year: glo launched in December 2016; we assume 0 consumers for 2016 [26].
- BAT revisions: When BAT revises historical "consumer acquisition" figures in later reports, we use the most recent published figures for that year.
- PMI launch year: IQOS launched in 2014; given that the national prevalence of HTP use in Japan was 0.2% in 2015, we set 0 users in 2014 [27].
- KT&G 2020 market split: The domestic versus international split for KT&G in 2020 is not reported. We assume the proportion matched the 2021 level.

Because daily stick consumption might vary by brand, device generation, market maturity, and user profile, we applied a

XSL•FO
RenderX

±50% sensitivity around each brand's central sticks-per-day assumption when converting stick volumes to users. Lower and upper bounds were reported alongside central values for any company-year derived from shipments. No sensitivity was applied to directly reported user counts.

Alternative Time Series: Shipment-Only Conversions Using Literature-Based Consumption

To address uncertainty arising from inconsistent or noncomparable manufacturer user definitions, we constructed an alternative series that ignores all manufacturer-reported user counts and converts all companies' stick shipments to users using a common consumption parameter derived from the literature.

External survey evidence from Germany (2018 - 2023) indicates that current HTP users reported a mean consumption of 7.3 sticks per day, with IQOS accounting for 84% of current use [28]. Comparable estimates were observed in Italy (2018 - 2020), where IQOS users reported a mean consumption of 7.9 sticks per day [29].

Based on these findings, we applied a central estimate of 7.3 sticks per day and used the corresponding 95% CI (6.33 - 8.27 sticks per day) to generate lower and upper bound user estimates, respectively.

Estimation Using National Survey Prevalence

To complement the industry disclosure-based series, we developed an independent survey-based estimate of HTP users derived from nationally representative prevalence data where available. We extracted country-level estimates of current HTP use from 35 countries compiled in the Global State of Tobacco Harm Reduction database, which combines data from Eurobarometer 2023, Global Adult Tobacco Surveys conducted between 2019 and 2023, World Health Organization's STEPwise approach to noncommunicable diseases risk factor surveillance surveys conducted between 2019 and 2022, and recent nationally representative surveys [10].

For each country, the reported prevalence of current HTP use was multiplied by the corresponding adult population for the survey year, aligned with the survey's defined age range (eg, 15+, 18+, 16 - 74, or 19+ years), using population denominators from the United Nations World Population Prospects [30]. Country-specific user counts were then aggregated to produce a combined total across the 35 countries. All data sources used for this estimation are presented in Table 2.

XSL•FO
RenderX

Table . Nationally representative survey inputs used to estimate heated tobacco product (HTP) users in 35 countries.

Country	Current HTP use prevalence	Year	Survey	Cohort	Population (thousands), n	Estimated HTP users, n
Austria	3.00	2023	Eurobarometer	15+	7764.2	232,927
Belarus	3.00	2020	STEPS[a]	18+	7571.0	227,131
Belgium	1.00	2023	Eurobarometer	15+	9717.5	97,175
Bulgaria	2.00	2023	Eurobarometer	15+	5833.4	116,669
Croatia	2.00	2023	Eurobarometer	15+	3348.3	66,966
Cyprus	5.00	2023	Eurobarometer	15+	1117.9	55,895
Czechia	4.00	2023	Eurobarometer	15+	8971.1	358,843
Estonia	2.00	2023	Eurobarometer	15+	1130.9	22,619
Georgia	1.60	2021	Local	18+	2873.7	45,980
Germany	1.00	2023	Eurobarometer	15+	72,387.7	723,877
Greece	3.00	2023	Eurobarometer	15+	8970.5	269,116
Hungary	3.00	2023	Eurobarometer	15+	8275.2	248,255
Ireland	1.00	2023	Eurobarometer	15+	4115.8	41,158
Italy	4.00	2023	Eurobarometer	15+	52,232.4	2,089,297
Japan	11.80	2022	Local	16 - 74	90,240.2	10,648,339
Jordan	4.10	2025	Local	15+	7670.8	314,503
Latvia	3.00	2023	Eurobarometer	15+	1581.5	47,445
Lithuania	5.00	2023	Eurobarometer	15+	2393.0	119,650
Luxembourg	1.00	2023	Eurobarometer	15+	549.6	5496
Netherlands	1.00	2023	Eurobarometer	15+	15,155.9	151,559
Philippines	0.10	2021	GATS[b]	15+	77,692.2	77,692
Poland	1.00	2023	Eurobarometer	15+	32,503.6	325,036
Portugal	5.00	2023	Eurobarometer	15+	9062.6	453,131
Romania	2.00	2023	Eurobarometer	15+	16,097.0	321,941
Slovakia	4.00	2023	Eurobarometer	15+	4599.2	183,969
Slovenia	2.00	2023	Eurobarometer	15+	1795.9	35,918
South Korea	5.90	2022	Local	19+	44,374.2	2,618,081
Spain	1.00	2023	Eurobarometer	15+	41,325.7	413,257
Mexico	0.20	2023	GATS	15+	96,059.4	192,119
Costa Rica	0.04	2022	GATS	15+	4033.0	1613
Ukraine	3.00	2023	Local	18+	33,723.9	1,011,717
Moldova	2.60	2021	STEPS	18+	2354.7	61,223
Kazakhstan	1.00	2019	GATS	15+	13,478.2	134,782
Armenia	0.80	2022	Local	15+	2294.9	18,359
Uzbekistan	0.24	2019	STEPS	18+	21,492.9	50,508

[a]STEPS: World Health Organization STEPwise approach to noncommunicable diseases risk factor surveillance.

[b]GATS: Global Adult Tobacco Survey.

Ethical Considerations

This study used only publicly available, aggregated secondary data and did not involve identifiable personal information. Therefore, it does not meet the definition of human subjects research under applicable regulatory frameworks, and formal ethics committee approval was not required. This determination is consistent with the US Common Rule (45 CFR 46.102) [31], which defines human subjects research as involving identifiable private information or interaction with individuals, and with

XSL•FO
RenderX

the European Union General Data Protection Regulation (Regulation [EU] 2016/679) [32], which applies only to personal data.

Results

Overview

This section presents the findings from the 3 estimation approaches. We first report the primary disclosure-based time series, then the alternative shipment-only series, and finally the survey-based estimation. Results are organized to address the 3 research questions outlined in the Introduction section.

The estimation drew on 44 possible company-year observations (4 companies×11 years, 2014‐2024). User counts were directly reported by manufacturers for 25 (56.8%) of these observations, while stick shipment volumes were available for an additional 14 (31.8%) observations, resulting in 39 (88.6%) company-year observations with at least 1 usable data point. For the remaining 5 (11.4%) observations, primarily early years before product launch or before public disclosure began, values were set to 0 or imputed using the documented assumptions described in the Methods section. No company-year observations were excluded from the analysis due to missing data; all gaps were addressed through the documented imputation and conversion procedures.

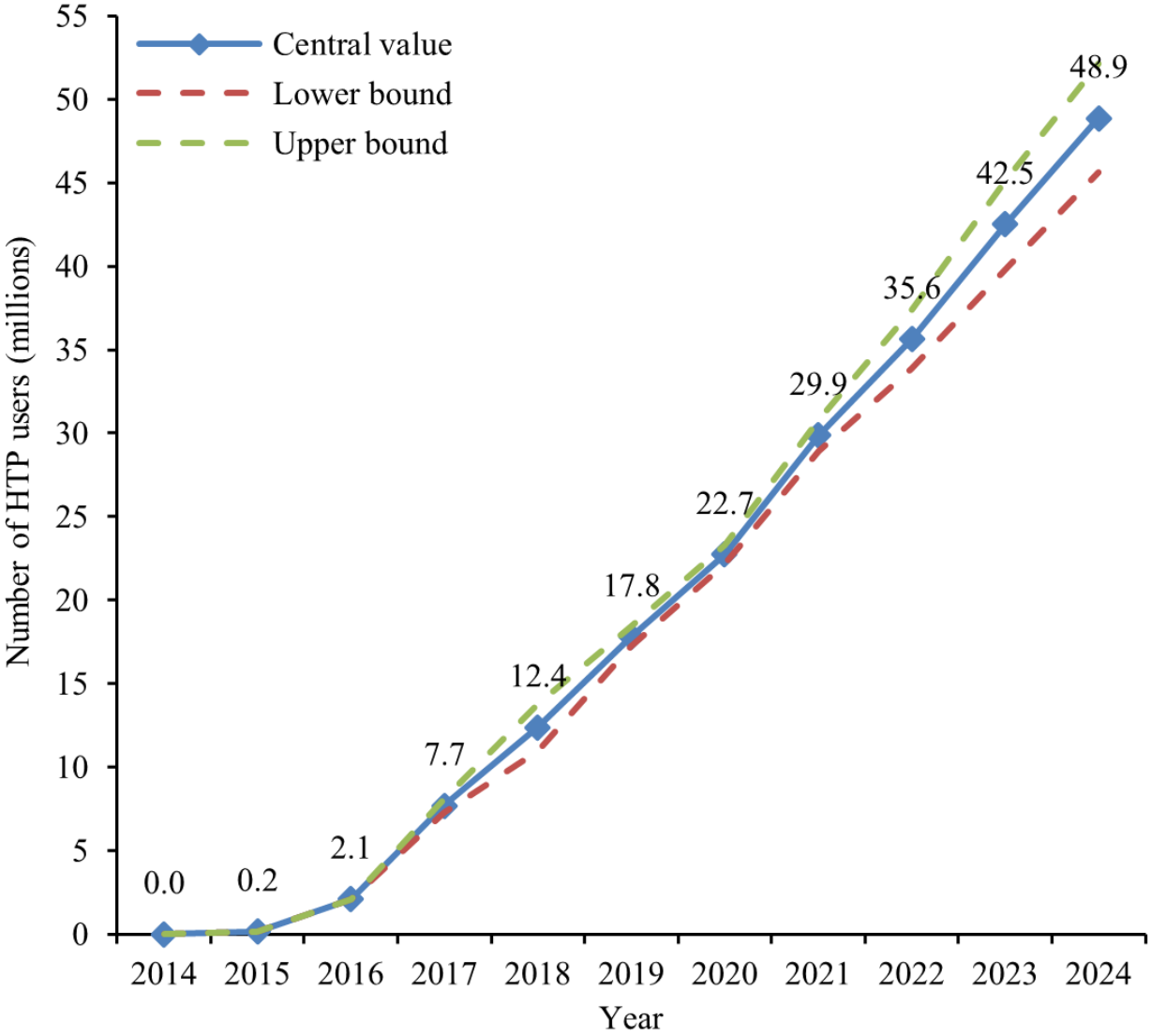
Primary Time Series: Company-Reported Users and Shipment-Based Conversions

The number of HTP users grew from negligible levels in 2014 to an estimated 48.9 million in 2024 (Figure 2). The estimates indicate a consistent upward trend across the decade, with the lower and upper bounds for 2024 ranging from 45.6 to 52.1 million, respectively, reflecting the ±50% sensitivity applied to shipment-based conversion assumptions. Because more than half of the 2024 global total is anchored in directly reported user counts, the overall uncertainty range remains relatively narrow.

Growth was driven primarily by PMI's IQOS, which accounted for the majority of users throughout the period (Figure 3). BAT's glo gained traction from 2017 onward, while JT's Ploom contributed smaller but steadily increasing shares. KT&G's lil maintained a modest share.

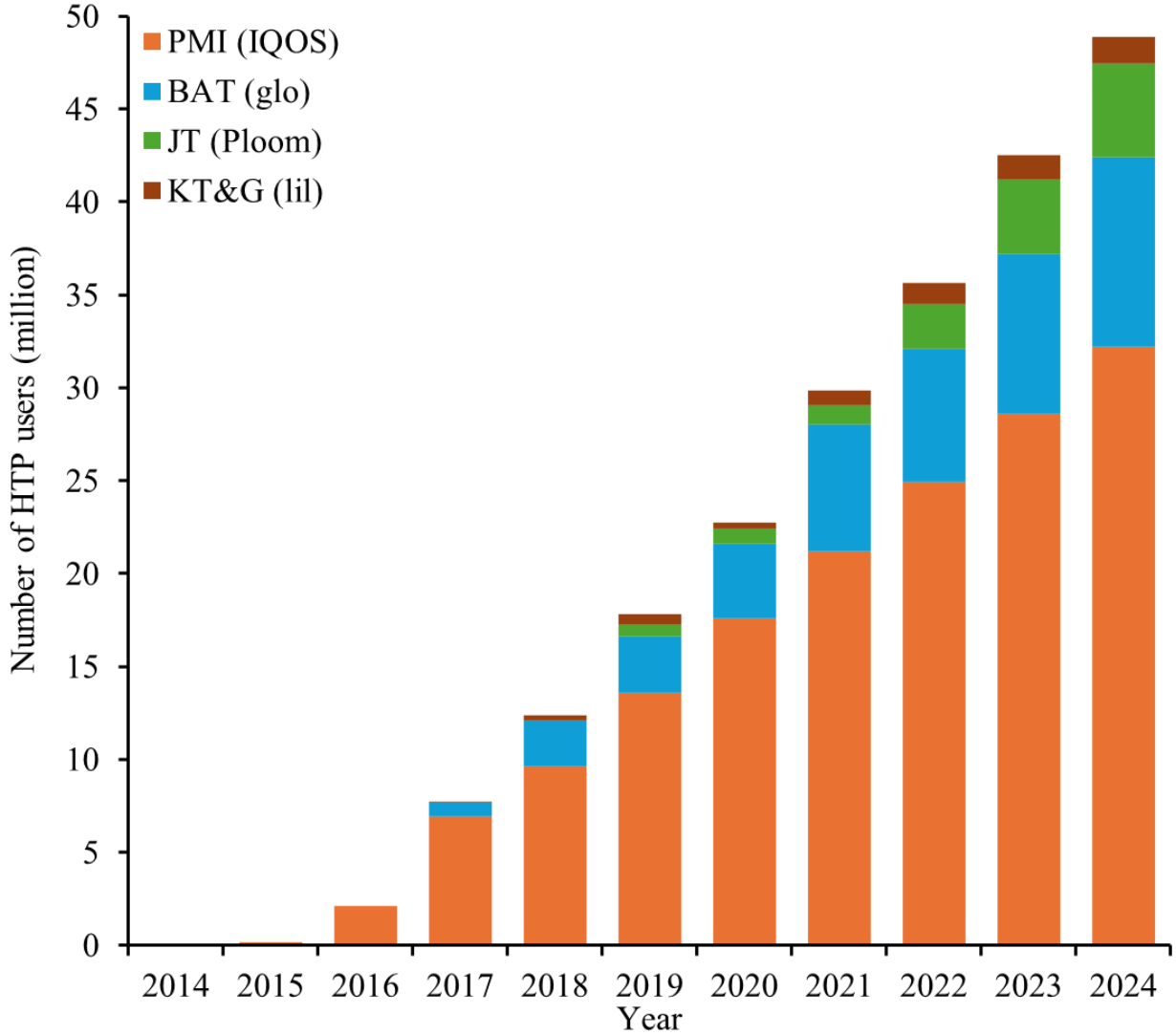
XSL•FO
RenderX

Figure 2. Global number of HTP users, 2014 - 2024—primary disclosure-based time series. Dashed lines indicate lower and upper bounds based on ±50% sensitivity on daily consumption assumptions. HTP: heated tobacco product.



XSL•FO
RenderX

Figure 3. Estimated global HTP users by manufacturer, 2014 - 2024—primary disclosure-based series. IQOS, lil, glo, and Ploom are heated tobacco brand names. BAT: British American Tobacco; HTP: heated tobacco product; JT: Japan Tobacco; KT&G: Korea Tobacco & Ginseng Corporation; PMI: Philip Morris International.

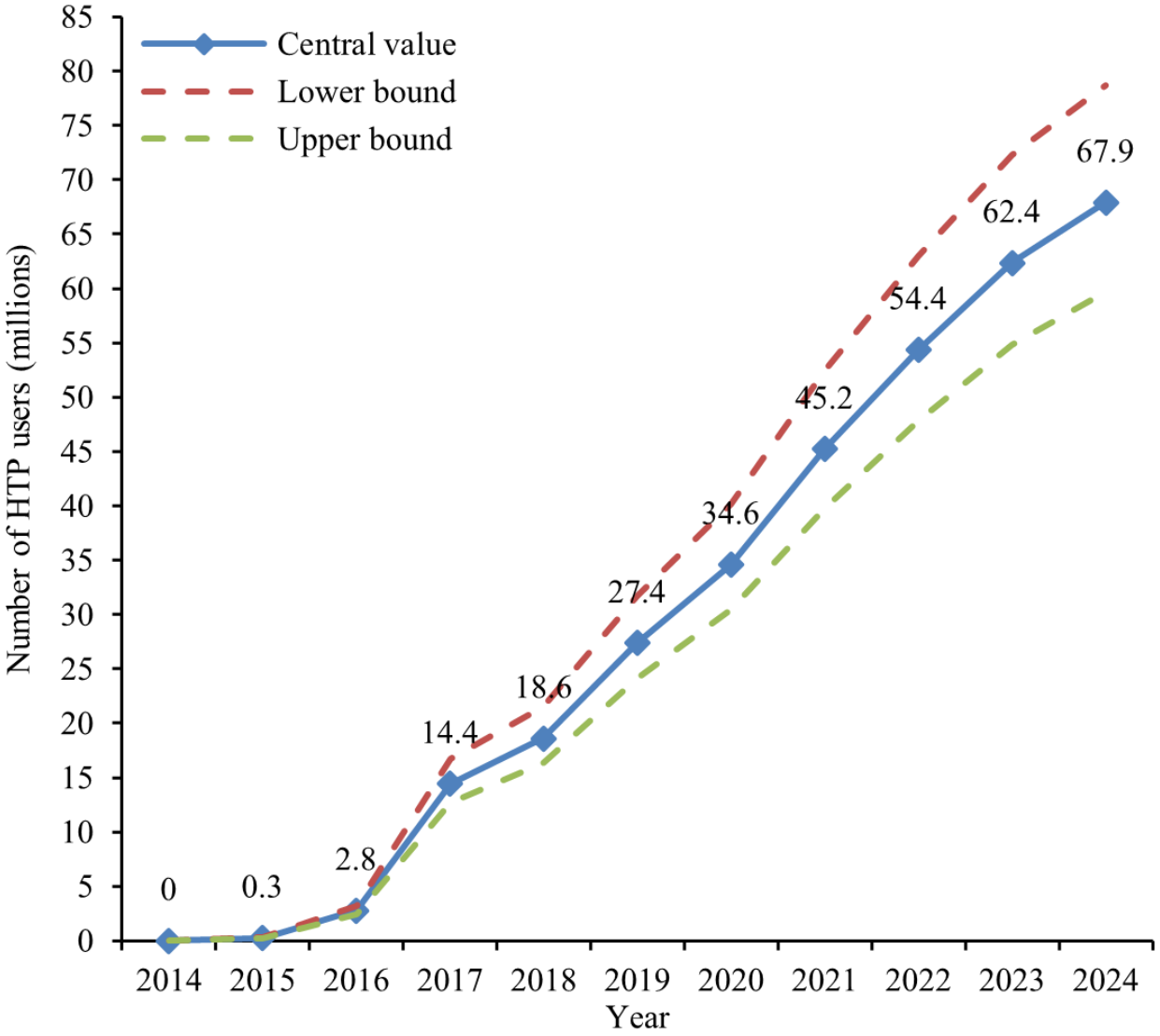


Alternative Time Series: Shipment-Only Conversions Using Literature-Based Consumption

Applying a uniform 7.3 (95% CI 6.33 - 8.27) sticks per day, consumption assumption to all firms' shipment volumes yields a consistently higher series, reflecting the more conservative daily consumption assumption and the removal of unclear company user definitions (Figure 4). In 2024, this approach produces an estimated 67.9 million users (range 59.7 - 78.7 million). The wider uncertainty interval compared with the primary series is expected, because all company-years are derived from shipment conversions, and none are anchored by directly reported user counts. Despite higher absolute levels, the trajectory closely mirrors the primary series.

XSL•FO
RenderX

Figure 4. Global number of HTP users, 2014 - 2024—alternative shipment-only time series using literature-based consumption assumptions. Alternative series converting all manufacturer stick shipments to user estimates using a uniform 7.3 sticks per day assumption (95% CI 6.33 - 8.27). Dashed lines indicate lower and upper bounds. HTP: heated tobacco product.



Estimation Using National Survey Prevalence

Nationally representative prevalence data yielded an estimated 21.8 million HTP users across 35 countries (Table 2). The largest numbers were observed in Japan (10.6 million in 2023), South Korea (2.6 million in 2022), Italy (2.1 million in 2023), Ukraine (1.0 million in 2023), and Germany (0.7 million in 2023). These 5 countries accounted for 78.5% (17.1 million) of the survey-based total. Most available surveys were conducted in 2022 or 2023, with some dating to 2019 - 2021 and only 1 in 2025, meaning that they may not capture more recent market expansion. The survey-based total of 21.8 million users across 35 countries is substantially lower than the disclosure-based estimates for the same period. Given partial geographic coverage, this should be interpreted as a conservative lower bound, rather than an alternative estimate of the global total.

Discussion

Principal Findings

This study addressed 3 research questions. First, using publicly available manufacturer disclosures combined with explicit shipment-to-user conversions where required, we constructed an annual global time series of HTP user estimates for 2014 - 2024. The primary series indicates that global HTP users increased from negligible levels to 48.9 million in 2024 (range 45.6 - 52.1 million).

Second, when all shipment volumes were converted using a standardized literature-based consumption parameter, the shipment-only sensitivity series yielded higher 2024 totals (67.9 million; range 59.7 - 78.7 million) but preserved a similar growth trajectory. This upward shift is expected because the shipment-only approach imposes a uniform consumption conversion factor and removes manufacturer-specific user definitions, some of which may reflect partial use, different reference periods, or varying inclusion rules.

XSL•FO
RenderX

Across both disclosure-based series, the direction and magnitude of growth since 2014 were consistent: near 0 user numbers in early launch years, accelerated expansion after 2016, and sustained increases through 2024. While absolute totals varied depending on conversion assumptions and reporting definitions, the upward trajectory was robust across estimation strategies.

Third, survey-derived user counts based on nationally representative prevalence data in 35 countries yielded a combined total of 21.8 million users. This estimate should be interpreted as a conservative lower bound rather than a competing global total. This estimate covers fewer than half of the countries where HTPs are legally available and relies on survey years that are mostly 2023 or earlier, before full expansion in some markets. Moreover, prevalence surveys vary in age frames, question wording, and whether they capture "current use" in a manner comparable across countries. These limitations restrict the survey series' ability to represent the full global market, but the estimates nonetheless provide an important external check for the disclosure-based totals and illustrate how incomplete surveillance coverage remains despite the rapid market expansion documented here.

These findings extend the existing literature on HTP adoption and surveillance. Prior work has primarily focused on country-specific prevalence, particularly in early-adopting markets such as Japan and South Korea, documenting rapid substitution dynamics and substantial category growth within a relatively short period [4-9,28,29,33]. Sun et al [4] synthesized 45 studies across 42 countries but identified adult current-use prevalence data for only 12 countries. More recently, Scala et al [5] conducted the most comprehensive systematic review to date, covering 76 studies published between 2015 and February 2022 and reporting current adult HTP prevalence from 34 countries based on surveys conducted prior to 2021.

Our survey-based analysis covers 35 countries and incorporates prevalence data from surveys conducted in years including 2025. The results align with prior reviews, which have documented that HTP use remains concentrated in certain markets. However, this approach demonstrates the structural limitation of prevalence-based aggregation when nationally representative data are available for only a subset of countries where HTPs are legally marketed.

Unlike prior syntheses based exclusively on survey data, the disclosure-based series presented here generates global totals that are not constrained by survey coverage, thereby addressing a key measurement gap in global HTP surveillance. Methodological precedent for combining survey and commercial data sources exists in the broader tobacco surveillance literature. A comparison of self-reported cigarette consumption from a national survey in England with retail sales data for 2011 - 2018 showed closely aligned trends, supporting the validity of combining survey-based and sales-based approaches to monitor population-level tobacco product dynamics [34].

As of 2025, HTPs are legally available in at least 72 countries, with uptake concentrated in early-adopting and predominantly higher-income markets [10]. Consequently, global estimates must be interpreted in this context. Although tens of millions of HTP users remain small relative to the 1 billion global population of people who smoke cigarettes [35], HTPs can capture substantial shares of the tobacco market within countries where they are widely available. For example, HTPs represented 48.5% of PMI's combined cigarette and HTP market in Japan, 31.8% in Hungary, 28.7% in Lithuania, and 23.9% in Greece, according to company disclosures from the third quarter of 2025 [36].

From a public health perspective, the implications of HTP growth depend on behavioral transitions rather than uptake alone. A decision-theoretic framework indicates that population-level benefits are most plausible when HTP uptake substitutes for cigarette smoking among adults who would otherwise continue to smoke, whereas dual use or uptake among never-smokers may reduce or negate potential benefits [37]. The evidence base summarized in recent reviews remains mixed and is shaped by heterogeneity in study designs and regulatory contexts. The Cochrane review concluded that evidence for HTP as cessation tools remains limited relative to other interventions and highlighted substantial uncertainty [38]. Scala et al [5] emphasized variability in dual use and switching patterns, indicating that aggregate uptake alone does not resolve whether an increase in HTP use represents displacement of cigarette smoking or supplementation. Evidence from Japan suggests partial displacement of cigarette sales after HTP introduction, while also showing heterogeneity in exclusive use versus dual use across survey designs and population subgroups [7,8,27,33]. These findings reinforce that the key policy question is not simply how many people use HTP, but how HTP uptake changes cigarette smoking trajectories over time.

Strengths

This study has several strengths. First, it relies exclusively on publicly verifiable data sources and documents extraction, harmonization, and revision procedures, enhancing transparency and reproducibility. Second, the dual-series approach (combining a primary disclosure-based series with an alternative shipment-only sensitivity) allows readers to assess how estimation choices affect results and provides bounded rather than single-point estimates. Third, all data sources, assumptions, and conversion formulas are fully documented, enabling independent replication and updating as new disclosures become available. Finally, the analysis covers the full commercially relevant period for HTP (2014 - 2024) and all manufacturers operating at a material global scale, providing comprehensive market coverage.

Limitations

This study has several limitations that should be considered when interpreting the findings. First, the estimates presented here are derived from secondary data sources rather than from a single harmonized global survey. Manufacturer-reported user counts and shipment volumes were used because nationally representative HTP prevalence data remain unavailable for most countries. Although these disclosures are among the few consistently updated global sources, they may differ in definitions, reporting granularity, and geographic aggregation across firms and over time. Listed firms face strong incentives for accuracy: disclosures are governed by securities laws, audited or reviewed through established internal controls, and

XSL•FO
RenderX

subject to penalties for material misstatement, while reputational and investor-relations risks further discourage misreporting. However, these safeguards do not eliminate error or definitional ambiguity, and they reduce the likelihood of systematic overstatement.

Second, where shipment volumes were converted into user estimates, the analysis relied on explicit assumptions regarding average daily consumption of HTP sticks per user. While these assumptions were transparently documented and subjected to sensitivity analysis, true consumption patterns may vary across countries, over time, and by user characteristics. As a result, the estimates should be interpreted as approximations rather than precise counts.

Third, the presence of illicit or gray-market distribution may affect the allocation of HTP use at the country level, particularly in jurisdictions where HTP sales are restricted or prohibited. However, HTPs are manufactured by a small number of multinational firms, and HTP consumables are technologically standardized and brand-locked, creating substantial barriers to informal or counterfeit production [16]. Consequently, most HTP sticks consumed globally are produced by these firms and are captured in shipment-based disclosures at the point of manufacture and are therefore included in the estimates presented here.

Fourth, the study does not distinguish between exclusive HTP use and dual use with cigarettes. Manufacturer disclosures and most national surveys do not provide information to allow consistent differentiation of these patterns. As a result, the estimates quantify the number of people using HTP but do not directly inform whether HTP uptake represents substitution for cigarette smoking or supplementation alongside continued smoking.

Scope

The scope of this study is the global population of HTP users as estimated from manufacturer disclosures and a subset of national surveys. The estimates apply to the period 2014 - 2024 and to the 4 major HTP manufacturers. They are most informative for understanding the scale and growth of the HTP category in a global market characterized by concentration among a small number of multinational manufacturers. They are less informative for geographic allocation, intensity of use distributions, or product-specific switching dynamics. Generalizability to country-level prevalence trends is therefore limited and depends on the extent to which manufacturer-reported volumes reflect geographically diverse consumption patterns.

The survey-based series is limited to 35 countries with available nationally representative prevalence data, representing fewer than half of the countries where HTPs are legally sold.

Recommendations for Future Research

Several directions for future research emerge from this analysis. First, as national surveys increasingly incorporate standardized HTP measures, researchers should validate disclosure-based estimates against survey-derived counts at the country level to calibrate conversion assumptions and assess the reliability of manufacturer-reported user definitions. Second, future studies should link HTP uptake estimates with longitudinal data on behavioral transitions (switching, dual use, relapse, and initiation), as aggregate uptake alone does not resolve the key policy question of whether HTP growth displaces or supplements cigarette smoking. Third, the estimation framework introduced here could be extended to incorporate subnational data and to distinguish exclusive HTP use from concurrent cigarette smoking where data permit. Fourth, as additional consumption data become available from diverse markets, the brand-specific and literature-based sticks-per-day assumptions used in this study should be updated and validated against observed use patterns in a wider range of regulatory and cultural contexts.

Significance

This study contributes to the tobacco and nicotine surveillance literature by providing a reproducible global time series of HTP users, offering population estimates that are currently missing from the evidence base and that are required for epidemiologic modeling, burden-of-disease estimation, and regulatory and fiscal evaluation. The transparent methodology and publicly available data sources enable regular updating as new disclosures become available, supporting ongoing monitoring of the evolving global nicotine product landscape. By generating bounded estimates across multiple approaches, this work provides a more complete empirical foundation for interpreting shifts in nicotine consumption patterns and for informing evidence-based tobacco policy at the global level. Overall, this work complements, rather than replaces, survey-based epidemiology and helps bridge the measurement gap until national surveillance systems expand and standardize HTP measures sufficiently to produce harmonized global estimates.

Conclusions

This study provides a transparent and replicable global time series estimate of HTP use from 2014 to 2024, addressing a notable gap in international tobacco and nicotine surveillance. Depending on the estimation approach, global HTP user totals in 2024 range from 45.6 to 78.7 million users in the disclosure-based series, with survey-based estimates across 35 countries providing a conservative lower bound of 21.8 million. Across all approaches, the direction and timing of growth are consistent, indicating rapid and sustained expansion of HTP use over the past decade. The analysis documents the scale and trajectory of this growth and provides essential population denominators for interpreting market transitions, supporting epidemiologic modeling, and informing regulatory and fiscal policy.

XSL•FO
RenderX

Acknowledgments

The authors disclose that the artificial intelligence chatbot ChatGPT was used solely for language editing purposes in the preparation of this manuscript. The model was not used for data analysis, data interpretation, or content generation. All intellectual content, analysis, and conclusions are of the authors, and they take full responsibility for the integrity and accuracy of the work.

Funding

This study was funded within a grant from Global Action to End Smoking (formerly known as the Foundation for a Smoke-Free World), an independent US nonprofit 501(c)(3) grantmaking organization. The funder had no role in the conception, design, implementation, data analysis, interpretation of the study results, or preparation of this manuscript.

Data Availability

All data generated or analyzed during this study are included in this published paper and its supplementary materials. Manufacturer-reported user counts and heated tobacco stick shipment volumes were extracted from publicly available corporate disclosures (annual reports, investor presentations, and regulatory filings) accessible through the investor relations sections of the respective companies' websites [18-22]. The survey-based prevalence inputs and corresponding derived user counts are provided in Table 2.

Authors' Contributions

GM contributed to the conceptualization, data collection, data analysis, and drafting of the manuscript. GVS and TJ contributed to the conceptualization of the study. All authors read and approved the final manuscript.

Conflicts of Interest

GM works as an independent consultant on various research projects undertaken by multiple research institutions and nongovernmental organizations (NGOs). At the time this research was conducted, the author had a service agreement as a data scientist with Knowledge-Action-Change (KAC), a UK-based private organization that focuses on harm reduction as a key public health strategy. In parallel, the author has a service agreement with the Ukraine-based NGO Healthy Initiatives as an economist. Healthy Initiatives is a nonprofit organization aimed at promoting and strengthening public health and well-being, tackling the most urgent questions in need of an answer. The organization supports and implements projects in Eastern Europe and Central Asia to address the growing risks of noncommunicable diseases. Both organizations receive grants from Global Action to End Smoking. Global Action to End Smoking, formerly known as Foundation for a Smoke-Free World, is an independent, US nonprofit 501(c)(3) grantmaking organization accelerating science-based efforts worldwide to end the smoking epidemic. Its mission is to end combustible tobacco use, which remains the leading preventable cause of death globally. The organization collaborates with academic and research centers and others to accelerate life-saving research and educational projects. Global Action does not seek or accept funding from companies that produce tobacco or nonmedicinal nicotine products. The charitable gift agreement (the "Pledge Agreement") between the organization and its prior funder, PMI Global Services Inc, was terminated in September 2023. To complement the termination of the Pledge Agreement, the organization's board of directors established a new policy to not accept or seek any tobacco or nonmedicinal nicotine industry funding. Global Action to End Smoking's Certificate of Incorporation states that Global Action was formed to, among other things, "support independent scientific research free from the influence of any commercial entity that may be affected by the research outcome." Global Action's organizational documents, including the Certificate of Incorporation and bylaws, are publicly available on its website and fully disclosed in public materials. TJ works as a researcher, sociologist, and data scientist at The Robert Zajonc Institute for Social Studies, University of Warsaw, Poland. At the time this research was conducted, the author had a service agreement as a data scientist with KAC. GVS has been a consultant to KAC since 2023. He was a founder of KAC and director from 2011 until 2022. He was a principal grant holder for grants from the Foundation for a Smoke-Free World (subsequently named Global Action to End Smoking) from 2018 until 2023. The statements above apply regarding the Foundation for a Smoke-Free World and Global Action to End Smoking. He is a shareholder in the Global Forum on Nicotine Ltd, which holds events on nicotine. Authors declare no other financial or nonfinancial competing interests in the past 36 months. This study was conducted as part of the authors' collaboration with KAC.

References

1. O'Connor R, Schneller LM, Felicione NJ, Talhout R, Goniewicz ML, Ashley DL. Evolution of tobacco products: recent history and future directions. Tob Control 2022 Mar;31(2):175-182. [doi: 10.1136/tobaccocontrol-2021-056544] [Medline: 35241585]
2. Bitzer ZT, Goel R, Trushin N, Muscat J, Richie JP Jr. Free radical production and characterization of heat-not-burn cigarettes in comparison to conventional and electronic cigarettes. Chem Res Toxicol 2020 Jul 20;33(7):1882-1887. [doi: 10.1021/acs.chemrestox.0c00088] [Medline: 32432464]
3. Heated tobacco products: a brief. World Health Organization. 2020. URL: https://www.who.int/europe/publications/i/item/WHO-EURO-2020-4571-44334-62636 [accessed 2026-03-31]

XSL•FO
RenderX

4. Sun T, Anandan A, Lim CCW, et al. Global prevalence of heated tobacco product use, 2015-22: a systematic review and meta-analysis. Addiction 2023 Aug;118(8):1430-1444. [doi: 10.1111/add.16199] [Medline: 37005862]
5. Scala M, Dallera G, Gorini G, et al. Patterns of use of heated tobacco products: a comprehensive systematic review. J Epidemiol 2025 May 5;35(5):213-221. [doi: 10.2188/jea.JE20240189] [Medline: 39805598]
6. Cummings KM, Roberson A, Levy DT, et al. Transformation of the tobacco product market in Japan, 2011-2023. Tob Control 2026 Jan 8;35(1):79-82. [doi: 10.1136/tc-2024-058734] [Medline: 39472049]
7. Issabakhsh M, Gravely S, Katanoda K, et al. Recent trends in cigarette and HTP use in Japan: a scoping review. Nicotine Tob Res 2026 Mar 24;28(4):511-525. [doi: 10.1093/ntr/ntaf216] [Medline: 41134207]
8. Levy DT, Issabakhsh M, Warner KE, Liber AC, Meza R, Cummings KM. Evaluating trends in cigarette and HTP use in Japan and measurement issues in the National Health and Nutrition Survey. Tob Control 2025 Oct 3;34(5):659-663. [doi: 10.1136/tc-2023-058526] [Medline: 38906697]
9. Lee S, Xu SS, Yan M, et al. Patterns of cigarette, heated tobacco product, and nicotine vaping product use among Korean adults: findings from the 2020 ITC Korea Survey. Tob Induc Dis 2024;22(April):63. [doi: 10.18332/tid/186273] [Medline: 38638419]
10. Global smoking and Tobacco Harm Reduction database. Global State of Tobacco Harm Reduction. 2025. URL: https://gsthr.org/countries/ [accessed 2026-04-09]
11. Jerzyński T, Stimson GV, Shapiro H, Król G. Estimation of the global number of e-cigarette users in 2020. Harm Reduct J 2021 Oct 23;18(1):109. [doi: 10.1186/s12954-021-00556-7] [Medline: 34688284]
12. Jerzyński T, Stimson GV. Estimation of the global number of vapers: 82 million worldwide in 2021. Drugs Habits Soc Policy 2023 May 25;24(2):91-103. [doi: 10.1108/DHS-07-2022-0028]
13. The global state of tobacco harm reduction 2024: a situation report. : Global State of Tobacco Harm Reduction; 2024 URL: https://gsthr.org/resources/thr-reports/the-global-state-of-tobacco-harm-reduction-2024-a-situation-report/ [accessed 2026-03-31]
14. WHO report on the global tobacco epidemic, 2025: warning about the dangers of tobacco. World Health Organization. 2025. URL: https://www.who.int/publications/i/item/9789240112063 [accessed 2026-03-31]
15. Teshima A, Martínez C, Filippidis FT, et al. Mapping indicators of tobacco and related product use: unveiling challenges and variations in the Eurobarometer surveys over three decades. Tob Induc Dis 2025;23:49. [doi: 10.18332/tid/202651] [Medline: 40241742]
16. Simonavicius E, McNeill A, Shahab L, Brose LS. Heat-not-burn tobacco products: a systematic literature review. Tob Control 2019 Sep;28(5):582-594. [doi: 10.1136/tobaccocontrol-2018-054419] [Medline: 30181382]
17. Heated tobacco products. Tobacco Tactics. 2025. URL: https://www.tobaccotactics.org/article/heated-tobacco-products/ [accessed 2025-10-22]
18. Investor relations. Philip Morris International. URL: https://www.pmi.com/investor-relations/overview/ [accessed 2026-03-31]
19. Investor relations overview. KT&G Corporation. URL: https://en.ktng.com/ir/overview [accessed 2026-03-31]
20. Previous annual reporting. British American Tobacco. URL: https://www.bat.com/investors-and-reporting/reporting/previous-annual-reporting [accessed 2026-03-31]
21. Combined annual and sustainability report. British American Tobacco. URL: https://www.bat.com/annualreport [accessed 2026-03-31]
22. Investor briefing materials. Japan Tobacco Inc. URL: https://www.jti.co.jp/investors/library/presentation/index.html [accessed 2026-03-31]
23. Results, reports and presentations. Imperial Brands. URL: https://www.imperialbrandsplc.com/investor-hub/results-reports-and-presentations [accessed 2026-04-09]
24. Philip Morris International 2024 annual report. : Philip Morris International; 2024 URL: https://www.pmi.com/resources/docs/default-source/investor_relation/pmi_2024_annualreport.pdf [accessed 2026-03-31]
25. Philip Morris International 2020 annual report. : Philip Morris International; 2020 URL: https://www.annualreports.com/HostedData/AnnualReportArchive/p/NYSE_PM_2020.pdf [accessed 2026-03-31]
26. Annual report and Form 20-F 2017. : British American Tobacco p.l.c; 2018 URL: https://www.bat.com/content/dam/batcom/global/main-nav/investors-and-reporting/reporting/combined-annual-and-esg-report/previous-annual-reporting/Annual_Report_and_Form_20-F_2017.pdf [accessed 2026-03-31]
27. Craig LV, Yoshimi I, Fong GT, et al. Awareness of marketing of heated tobacco products and cigarettes and support for tobacco marketing restrictions in Japan: findings from the 2018 International Tobacco Control (ITC) Japan Survey. Int J Environ Res Public Health 2020 Nov 13;17(22):8418. [doi: 10.3390/ijerph17228418] [Medline: 33202995]
28. Keller AM, Klosterhalfen S, Kastaun S, Kotz D. Heated tobacco product use prevalence and temporal trends among the German population: a series of representative cross-sectional household surveys, 2018-2023. BMJ Open 2025 Jun 17;15(6):e094610. [doi: 10.1136/bmjopen-2024-094610] [Medline: 40527558]
29. Roulet S, Pellegatti T, Fischer K, Magnani P, di Luzio Paparatti U. Tobacco- and nicotine-containing product use in Italy: results from two cross-sectional studies. Qeios. Preprint posted online on Aug 22, 2024. [doi: 10.32388/F3LAEC]
30. World Population Prospects 2024. United Nations, Department of Economic and Social Affairs, Population Division. 2024. URL: https://population.un.org/wpp/ [accessed 2026-03-31]

XSL•FO
RenderX

31. Protection of human subjects, 45 CFR part 46. US Department of Health and Human Services. 2018. URL: https://www.ecfr.gov/current/title-45/subtitle-A/subchapter-A/part-46 [accessed 2026-04-09]
32. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). European Parliament and Council of the European Union. 2016. URL: https://eur-lex.europa.eu/eli/reg/2016/679/oj [accessed 2026-04-09]
33. Cigarette sales halved: heated tobacco products and the Japanese experience. Knowledge•Action•Change | Global State of Tobacco Harm Reduction. 2024 May. URL: https://gsthr.org/resources/briefing-papers/cigarette-sales-halved-heated-tobacco-products-and-the-japanese-experience/ [accessed 2026-03-31]
34. Jackson SE, Beard E, Kujawski B, et al. Comparison of trends in self-reported cigarette consumption and sales in England, 2011 to 2018. JAMA Netw Open 2019 Aug 2;2(8):e1910161. [doi: 10.1001/jamanetworkopen.2019.10161] [Medline: 31461148]
35. WHO global report on trends in prevalence of tobacco use 2000–2024 and projections 2025–2030 (6th ed). World Health Organization. 2025. URL: https://www.who.int/publications/i/item/9789240116276 [accessed 2026-03-31]
36. 2025 third-quarter results: october 21, 2025. Philip Morris International Inc. 2025. URL: https://philipmorrisinternational.gcs-web.com/static-files/cbb6105c-9cc6-4b3b-8850-307cc2aef996 [accessed 2026-04-09]
37. Levy DT, Cadham CJ, Li Y, et al. A decision-theoretic public health framework for heated tobacco and nicotine vaping products. Int J Environ Res Public Health 2022 Oct 18;19(20):13431. [doi: 10.3390/ijerph192013431] [Medline: 36294011]
38. Tattan-Birch H, Hartmann-Boyce J, Kock L, et al. Heated tobacco products for smoking cessation and reducing smoking prevalence. Cochrane Database Syst Rev 2022 Jan 6;1(1):CD013790. [doi: 10.1002/14651858.CD013790.pub2] [Medline: 34988969]

Abbreviations

BAT: British American Tobacco
HTP: heated tobacco product
JT: Japan Tobacco
KT&G: Korea Tobacco & Ginseng Corporation
PMI: Philip Morris International

Edited by A Mavragani, T Sanchez; submitted 01.Dec.2025; peer-reviewed by NR Geyer, Y Jain; revised version received 26.Feb.2026; accepted 12.Mar.2026; published 20.Apr.2026.

Please cite as:
Mzhavanadze G, Stimson GV, Jerzyński T
Global Heated Tobacco Product User Estimates, 2014-2024: Descriptive Surveillance Study Using Manufacturer Disclosures
JMIR Public Health Surveill 2026;12:e88761
URL: https://publichealth.jmir.org/2026/1/e88761
doi:10.2196/88761

© Giorgi Mzhavanadze, Gerry V Stimson, Tomasz Jerzyński. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 20.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Systematic Assessment of Flavor Cues and Additives in Cigarettes and Heated Tobacco Products in Korea: Cross-Sectional Surveillance Study

Geon Heo[1], BSPH; Doyoon Kim[1], BSN; Yehyun Kim[1], BSPH; Ara Kim[2], MPH; Nagyeong Cho[1], MPH; Naeun Kang[1], MPH; Jungmi Park[1], MPH; Eunsil Cheon[1], MPH; Seulgi Kim[3], PhD; Susan Park[4], PhD; Sungmin Park[1], PhD; Sung-il Cho[1,3], MD, SCD; Heewon Kang[5], PhD

[1]Department of Public Health Sciences, Graduate School of Public Health, Seoul National University, Seoul, Republic of Korea

[2]Armed Forces Capital Hospital, Seongnam, Republic of Korea

[3]Institute of Health and Environment, Seoul National University, Seoul, Republic of Korea

[4]School of Medicine, Inha University, Incheon, Republic of Korea

[5]Department of Health Administration, Daejin University, 1007 Hoguk-ro, Pocheon-si, Gyeonggi-do, Gyeonggi, Republic of Korea

Corresponding Author:
Heewon Kang, PhD
Department of Health Administration, Daejin University, 1007 Hoguk-ro, Pocheon-si, Gyeonggi-do, Gyeonggi, Republic of Korea

Abstract

Background: In South Korea, where plain packaging has not been adopted, tobacco packaging continues to function as a key marketing tool for the tobacco industry, using texts, colors, and imagery to attract consumers. Among these, flavor cues are especially important as they enhance product appeal. Cigarette sticks also serve marketing functions through design features such as colors and capsule indicators.

Objective: This study aimed to examine flavor-related cues on cigarettes and heated tobacco products (HTPs) packaging and stick design and to assess the presence of flavor additives in these products.

Methods: This surveillance study was conducted in November 2024. Tobacco products were purchased from convenience stores located in Seoul, supplemented by cross-referencing with national market monitoring data. Of 353 identified products, 214 products (150 cigarettes and 64 HTPs) were collected. Flavor cues were categorized by pack and stick design features, and additives were identified through sensory analysis of product components.

Results: Among the collected products, 63.1% (54% for cigarettes and 84.4% for HTPs) had both flavor cues and flavor additives, while 20.6% (27.3% for cigarettes and 4.7% for HTPs) had neither. Flavor cues were found in 67.3% of cigarettes and 95.3% of HTPs ($P<.001$), and flavor additives in 59.3% of cigarettes and 84.4% of HTPs ($P<.001$). Pack color was the most common cue, and additives were most often delivered through crushable capsules. HTPs used a wider range of flavoring methods, including flavoring in tobacco leaves and inner wrappers.

Conclusions: Tobacco packaging and stick design in South Korea remain important marketing tools for the tobacco industry. Flavor cues and additives are widely used in tobacco products, particularly in HTPs. These findings highlight the need for plain packaging policies and bans on flavor additives in tobacco products.

(JMIR Public Health Surveill 2026;12:e87537) doi:10.2196/87537

KEYWORDS

cigarettes; heated tobacco products; packaging; flavor cues; flavor additives; regulation

Introduction

Tobacco packaging is the first element of the product that people encounter, and it functions as a key marketing tool for the tobacco industry [1], particularly in settings where plain packaging has not been implemented [2]. Packaging elements such as color, imagery, and text can evoke positive perceptions and stimulate tobacco use, particularly among youth and individuals who have quit smoking [3-6]. Accordingly, Article 11 of the Framework Convention on Tobacco Control (FCTC) emphasizes strict regulation of promotional features on tobacco packs and labels [7], and several countries have adopted plain packaging and enlarged health warnings to reduce product appeal [8]. These measures have also been extended to cigarette sticks, restricting the use of stick appearance as a marketing tool [9]. Experimental evidence shows that unattractive stick

XSL•FO
RenderX

colors or warning messages on sticks significantly reduce appeal and purchase intentions, suggesting that plain packaging should extend beyond packs to include sticks [9,10].

Flavor-related experiences are frequently communicated through tobacco packaging using contextual and visual cues [11-15]. In settings without plain packaging, manufacturers use concept descriptors (eg, “Ice” and “Blue”) in combination with color schemes to signal taste and sensation without explicitly naming a characterizing flavor [12,16]. These cues shape consumer perceptions, and products with flavor cues are often perceived as more appealing, better tasting, and smoother than nonflavored or standardized variants [13]. This sensation transfer reduces perceived harshness and creates a health halo that leads consumers, especially youth and young adults, to perceive these products as less harmful [11,14].

To effectively implement and evaluate packaging regulations, systematic surveillance of tobacco packs is crucial. Much of the evidence base for assessing tobacco packaging has been shaped by the Tobacco Pack Surveillance System (TPackSS), developed by the Institute for Global Tobacco Control [11,12,17]. These studies demonstrate how systematic observation can identify flavor cues and other tactics used to enhance product appeal, providing evidence to support stronger regulations [11,12,17]. Recent pack surveillance in the Philippines showed that nearly 60% of tobacco products had explicit flavor cues, often reinforced by imagery, highlighting the persistent appeal of flavored products in the absence of regulation [11].

Building on this evidence, systematic surveillance is particularly important for addressing flavored tobacco products. This aligns with the recommendations of Articles 9 and 10 of the FCTC, which emphasize the need to monitor and regulate product contents, including flavoring [18]. Flavored tobacco products mask the harshness of nicotine and attract youth, women, and racial and ethnic minority groups [19-22]. They make tobacco use more appealing, increase nicotine dependence, and undermine quitting [21,23]. In response to this evidence, an increasing number of countries have banned flavored cigarettes [24]. However, most surveillance studies have primarily focused on cigarettes, with limited assessment of flavor cues in heated tobacco products (HTPs).

In South Korea (hereafter referred to as Korea), this issue has become particularly pressing. Following 3 decades of decline, smoking prevalence rebounded to 19.6% among adults in 2023. Although it declined to 16.7% in 2024, it remains high (28.5% among men and 4.2% among women) [25]. This trend has coincided with increased sales of flavored tobacco products, driven predominantly by flavor capsule cigarettes and HTPs [26]. In 2023, 1.68 billion packs of flavored cigarettes and HTPs were sold, accounting for 46.5% of total tobacco sales (3.61 billion packs). This marks an increase from 270 million packs (6.1%) in 2011 and 860 million packs (24.4%) in 2017 [26].

Despite the rapidly increasing sales of flavored cigarettes and HTPs, Korea still lacks comprehensive regulatory measures addressing flavor-related marketing. Although flavored tobacco products remain legal to manufacture and sell, Article 9 - 3 of the National Health Promotion Act (NHPA) prohibits the use of explicit text or imagery indicating flavors on tobacco packaging [27]. However, enforcement has been limited. Indirect flavor cues are also not regulated, allowing manufacturers to imply flavors through concept descriptors, colors, and other design elements [27].

In addition, while the NHPA mandates health warnings covering at least 50% of the main surfaces, with pictorial warnings accounting for at least 30% [27], Korea has not adopted plain packaging, despite its inclusion in the 2019 Tobacco Control Plan [28]. Other measures outlined in the Plan, such as a ban on flavored additives, also remain unimplemented. Moreover, there is no display ban at retail points of sale, and tobacco packs are predominantly showcased on power walls [27]. Together, these gaps leave visitors to approximately 140,000 tobacco retailers (primarily convenience stores and supermarkets) [29] continuously exposed to tobacco marketing, thereby increasing product appeal and the likelihood of tobacco use, particularly among youth [1,4,5].

In this study, we conducted a systematic assessment of cigarettes and HTPs sold in Korea, examining pack- and stick-level flavor cues as well as flavor additives at the component level. By combining pack- and stick-level surveillance with component-based sensory screening, we address a key gap in the literature on how flavor marketing cues align with flavoring and how these patterns differ between cigarettes and HTPs. We aimed to generate evidence to inform regulation and strengthen surveillance in Korea and internationally.

Methods

Study Design

To assess the packaging of cigarettes and HTPs sold in Korea, we used a market-based purposive sampling strategy designed to capture the diversity of unique brands and pack variants available on the market, rather than to estimate sales volume or market share [17]. Our pack documentation and photography procedures were informed by the TPackSS protocol [17], and we coded pack-level flavor cues such as descriptors, imagery, and capsules based on prior pack surveillance research [11,12]. Furthermore, drawing on prior studies [30-32], we developed an exploratory protocol to cut open individual cigarettes and HTP sticks (hereafter referred to as sticks) to examine internal flavor-related features. We also documented external flavor cues, including capsule indicators on sticks, before disassembly.

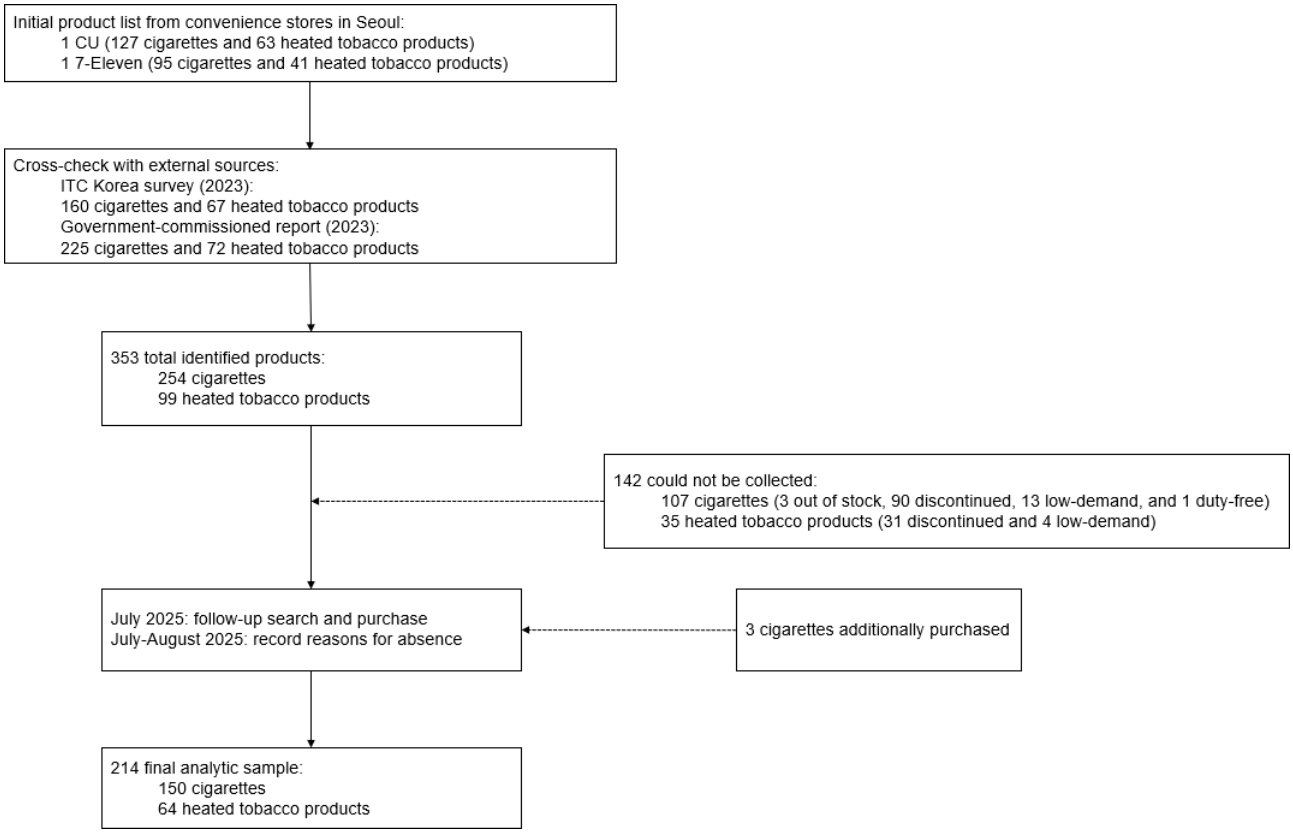
Sample

The product identification process involved four steps: (1) compiling an initial product list, (2) purchasing available products, (3) cross-referencing with external data sources, and (4) investigating reasons for unavailable products (Figure 1). To construct an initial product list, we visited 1 retail store from each of the 2 major convenience store chains in Seoul (CU and 7-Eleven) in November 2024 to collect all available tobacco products. We focused exclusively on convenience stores because they represent the predominant retail channel for tobacco sales in South Korea [29]. Among these, CU and 7-Eleven were selected due to their market dominance; together, they account for 56.3% of all convenience stores nationwide [33]. Within

XSL•FO
RenderX

each chain, 1 store in Seoul was selected based on geographic proximity to the research team. Observational checks at additional outlets suggested that product availability was comparable across stores.

Figure 1. Product identification and inclusion flow.



All tobacco products available in the visited stores were purchased, including all unique retail pack variants available at the point of sale. The initial list of purchased products was cross-checked against 2 external sources: the Wave 3 (2023) International Tobacco Control Korea Survey [34], which listed 160 cigarettes and 67 HTP brands, and a government-commissioned tobacco product market report [35], which identified 225 cigarettes and 72 HTP products purchased from 46 convenience stores nationwide in October 2023.

Products missing from the initial list (142 products: 107 cigarettes and 35 HTPs) were assessed using external sources to determine their market status. Products were classified as discontinued or low-demand based on Euromonitor reports (market share <0.3%) [26] and corroborating evidence from news articles and websites describing products as “discontinued,” “low-demand,” or “unpopular.” Duty-free exclusive products were verified through media reports. Among missing cigarettes, 90 (84.1%) were discontinued, 13 (12.1%) were not stocked due to low-demand, 3 (2.8%) were out of stock at the time of purchase, and 1 (0.9%) was duty-free exclusive and therefore excluded. Among missing HTPs, 31 (88.6%) were discontinued, and 4 (11.4%) were not stocked due to low demand.

Products not classified as discontinued, low-demand, or duty-free exclusives were subsequently sought through additional visits in July 2025, resulting in the collection of 3 additional packs. In total, 254 cigarettes and 99 HTPs were identified across all sources, of which 104 (40.9%) cigarettes and 35 (35.3%) HTPs were not collected. Reasons for noncollection are documented in Tables S1 and S2 in Multimedia Appendix 1.

Coding Procedure and Measures

Overview

After collection, we extracted and coded relevant information based on the assessment guidelines. HK developed a draft codebook, informed by previous pack- and stick-level assessment frameworks, to capture flavoring components in cigarettes and HTPs [11,12,17,32]. The codebook was reviewed and refined by SK and Susan Park. The final codebook included predefined variables with fixed response options and illustrative examples (where applicable).

Three coding teams (team 1: GH, YK, and NC; team 2: DK, AK, and JP; and team 3: EC and NK) assessed the products, with each product coded by 1 team. Each team consisted of 2 or 3 coders who jointly reviewed each product and recorded a single consensus code. Unresolved cases were cross-checked by another team. For data quality control, the final dataset was checked for internal consistency (logical validity across variables) and missingness. As coding was consensus-based and only final codes were retained, intercoder reliability statistics were not calculated. All coding was conducted collaboratively using a shared online spreadsheet.

XSL•FO
RenderX

Visual Measures

Visual measures captured visible pack [11,12,36] and stick features. Packs were measured for size and photographed before and after opening using the TPackSS photography protocol [17], and sticks were measured and photographed using a study-developed protocol (Figures S1 and S2 in Multimedia Appendix 1). Coding was conducted using photographs alongside direct inspection of the physical packs [17] and sticks.

We recorded whether tar and nicotine yields were displayed on the exterior of the cigarette package. Compliance with mandated health warnings was assessed by checking their presence and placement in accordance with national regulatory requirements [37]. Pack shape was coded based on the overall external form of the package. Marketing text was coded as text promoting product features or technological attributes, excluding legally required warnings.

Pack-level flavor cues were identified across 5 elements (Figure S3 in Multimedia Appendix 1): product name, pack color, capsule imagery, pack imagery, and other features [11]. Stick-level flavor cues were additionally coded based on the presence of capsule imagery. Coders were not blinded to brand names or pack descriptors and marketing language, because the study aimed to document consumer-facing packaging cues.

Flavor Cue Coding

Flavor cues were coded using a 2-step approach. First, coders recorded the presence or absence of a cue across the 5 pack elements. Second, for cued elements, coders assigned a predefined descriptive category (eg, product name: “cool or minty,” “color,” “flavor,” “functional,” and “others”) to document how the cue was communicated. Thus, the primary quantitative measure was the binary presence or absence of each cue element, while descriptive labels were used to characterize how cues were expressed across products.

Product names were coded as flavor cues when they contained semantic or lexical descriptors implying a specific flavor(“eg, “Raison Ice Presso”), color-associated sensation (eg, “Mevius Mix Green”), or sensory experience (eg, coolness or sweetness), rather than functioning solely as brand or variant names (eg, “Esse Prime”). This distinction clarified that “cool or minty” terms represented perceived freshness rather than literal colors (eg, aqua, cool shot, and ice). Within the product name category, “cool or minty” cues were treated as a distinct subgroup given the particularly prominent impact on perceptions of reduced sensory harshness and greater appeal [38].

Pack color and capsule imagery were coded by dominant color tones on the principal display panel and grouped into 4 sensory color groups: cool or minty, fruity, nutty, and others [11,12,16,39]. Previous experimental research has shown that the appeal and interpretation of pack color and capsule imagery may vary across flavor contexts, including fruity, nutty, and cool or minty categories, supporting the use of descriptive sensory groupings in packaging research [4,40]. Dominant color was defined as the most visually prominent color on the principal display panel, and only colors corresponding to flavor-related conventions reported in prior tobacco packaging research [4,16,41] were coded as flavor cues. These groupings descriptively captured packaging conventions in which color signals expected sensations [4]. More broadly, evidence on crossmodal correspondences shows that package color can shape flavor and sensory perceptions even without explicit flavor descriptors [42].

Additionally, we coded animal imagery as implicit flavor cues. Animal imagery (eg, penguins and polar bears) was included because such visuals function as established metaphors for coolness or freshness in tobacco marketing [43]. The “Others” category included textual statements promoting a cool or minty sensation (eg, Cuban coolness inside) or presence of flavor (eg, max flavor), which were evaluated as presented on the packs without interpretive modification. The complete coding framework with detailed criteria is presented in Table 1 and Figure S4 in Multimedia Appendix 1.

XSL•FO
RenderX

Table . Tobacco product and packaging characteristics codebook.

Main category and subcategory	Class	Evaluation criteria and examples	Categorization[a]
Basic characteristics			
Product type	• Cigarette • Heated tobacco product (HTP)	• ESSE, MARLBORO, DUNHILL, MEVIUS, etc. • TEREA, FiiT, NEO, etc.	—[b]
Manufacturers	• BAT • JTI • PMIK • KT&G	• DUNHILL, ROTHMAN, glo, etc. • MEVIUS, CAMEL, etc. • MARLBORO, PARLIAMENT, TEREA, etc. • ESSE, RAISON, FiiT, MIIX, etc.	—
Price	Continuous data	Fill in the price written on the receipt.	• <4500 KRW[c] • =4500 KRW • >4500 KRW
Number of sticks per pack	• 14 • 20	Fill in the form after opening the pack.	—
Packaging characteristic			
Tar content (mg)[d]	Continuous data	Fill in the form referring to the left or right side of the tobacco pack; not applicable for HTPs.	• <2.7 mg • ≥2.7 mg
Nicotine content (mg)[d]	Continuous data	Fill in the form referring to the left or right side of the tobacco pack; not applicable for HTPs.	• <0.2 mg • =0.2 mg • >0.2 mg
Cigarette stick length (mm)	Continuous data	Fill in the form after measuring the length of the cigarette stick.	• <79 mm • ≥79 mm
Health warning compliance	• Yes • No		—
Marketing text	• Yes • No	• “Smell Care,” “SENSIBLE SLIM,” “Innovation SUPER CARBON FILTER,” survey results, etc.	—
Dimension characteristic			
Pack depth (mm)	Continuous data	Fill in the form after measuring the length of the applicable part.	• <60 mm • ≥60 mm
Pack width (mm)	Continuous data	Fill in the form after measuring the length of the applicable part.	• <17 mm • ≥17 mm
Pack height (mm)	Continuous data	Fill in the form after measuring the length of the applicable part.	• <82 mm • ≥82 mm
Pack shape	• Rectangle • Round-edged rectangle • Half-round-edged rectangle		—
Flavor cue[e] [11,12,16,38]			

XSL•FO
RenderX

Main category and subcategory	Class	Evaluation criteria and examples	Categorization[a]
Product name	• Flavor cued • Flavor not cued	If flavor cued, classified into the following subtypes: • Cool or minty (eg, frozen and ice) • Color (eg, green and blue) • Flavor (eg, tropical and pearl), • Functional (eg, boost and hybrid) • Other (eg, beach and blossom)	—
Pack color	• Flavor cued • Flavor not cued	If flavor cued, classified into the following subtypes: • Cool or minty (eg, blue and green) • Fruity (eg, orange and pink) • Nutty (eg, brown and yellow) • Other (eg, purple)	—
Capsule imagery	• Flavor cued • Flavor not cued	If flavor cued, classified into the following subtypes: • Cool or minty (eg, blue and green) • Fruity (eg, orange and pink) • Nutty (eg, brown and yellow) • Other (eg, purple)	—
Pack imagery	• Flavor cued • Flavor not cued	If flavor cued, classified into the following subtypes: • Animal (eg, penguin and polar bear) • Nature (eg, glacier and palm tree) • Other (eg, switch button)	—
Others	• Flavor cued • Flavor not cued	If flavor cued, classified into the following subtypes: • Cool or minty (eg, cuban coolness and cooling taste) • Flavor (eg, max flavor and two tastes)	—
Number of flavor-cued elements	• 0 • 1 • 2 • 3 • 4 • 5	Sum the number of flavor-cued elements (responses)	—
Capsule imagery (stick)	• Flavor cued • Flavor not cued	If flavor cued, classified into the following subtypes: • Cool or minty (eg, blue and green) • Fruity (eg, orange and pink) • Nutty (eg, brown and yellow) • Other (eg, purple)	—
Flavoring method			

XSL•FO
RenderX

Main category and subcategory	Class	Evaluation criteria and examples	Categorization[a]
Inner wrapper	• Flavored • Not flavored	For each product, one team assessed the presence of flavoring using the orthonasal smelling method.	—
Cigarette wrapper	• Flavored • Not flavored	For each product, one team assessed the presence of flavoring using the orthonasal smelling method.	—
Tobacco leaves	• Flavored • Not flavored	For each product, one team assessed the presence of flavoring using the orthonasal smelling method.	—
Crushable capsule	• Flavored • Not flavored	For each product, one team assessed the presence of flavoring using the orthonasal smelling method.	—
Cellulose filter	• Flavored • Not flavored	For each product, one team assessed the presence of flavoring using the orthonasal smelling method.	—
Charcoal filter	• Flavored • Not flavored	For each product, one team assessed the presence of flavoring using the orthonasal smelling method.	—
Polylactic acid filter[f]	• Flavored • Not flavored	For each product, one team assessed the presence of flavoring using the orthonasal smelling method.	—
Hollow tube[f]	• Flavored • Not flavored	For each product, one team assessed the presence of flavoring using the orthonasal smelling method.	—
Others	• Flavored • Not flavored	For each product, one team assessed the presence of flavoring using the orthonasal smelling method.	—

[a]Only applicable for continuous data.

[b]Not available.

[c]All prices and currency values are reported in Korean won (KRW). The exchange rate as of November 2024 was approximately 1390 KRW to US $1.

[d]Not applicable for heated tobacco products.

[e]Flavor descriptors on tobacco packaging were coded based on prior tobacco packaging studies that typically distinguished menthol or mint, fruit or citrus, beverage, concept or functional, and other descriptors. Consistent with these studies, the further subdivision of flavor cue elements (eg, classifying blue and green as indicators of cool or minty) was informed by and closely aligned with prior studies.

[f]Only applicable for heated tobacco products.

For the stick-level flavor cue, we assessed the presence of capsule imagery on the stick surface. As illustrated in Figure S5 in Multimedia Appendix 1, capsule imagery was defined as a circular symbol resembling a crushable capsule printed on the stick. When present, it was classified using the same 4 color groups applied to pack color cues (ie, cool or minty, fruity, nutty, and others), based on the dominant color tone of the capsule symbol.

Sensory Measures

Sensory measures were collected to detect flavor additives and identify their location within the product. Orthonasal smelling [30], a validated method for detecting flavors through direct olfactory perception without inhalation [30,31], was used as a pragmatic screening method to detect the presence or absence of a perceptible nontobacco odor. This aligns with the definition of a “characterizing flavor” as a clearly noticeable smell or taste before or during consumption [44].

Using orthonasal smelling, each cigarette and HTP was assessed across 7 common components: inner wrapper, cigarette wrapper, tobacco leaves, crushable capsule, cellulose filter, charcoal filter, and others (eg, flavor card) [32]. For HTPs, 2 additional components (polylactic acid filter and hollow tube) were included [45]. Sticks were disassembled, and each component was evaluated separately. When the same odor was detectable in multiple components, the component with the strongest odor was recorded as the primary source. Because orthonasal smelling of tobacco products, even when unlit, may still involve minor sensory irritation or discomfort, all assessments were conducted in a ventilated room to minimize potential irritation and maintain olfactory sensitivity [46]. Coders also took breaks as needed to minimize olfactory fatigue and carryover effects before finalizing their judgments.

Analysis

Descriptive analyses and data tabulation were conducted using Microsoft Excel. All characteristics were assessed according to tobacco product type (cigarettes vs HTPs). Categorical variables were presented as frequencies and percentages, while continuous variables were summarized using the minimum, median, mean, and maximum values. Continuous variables were additionally

XSL•FO
RenderX

categorized, where applicable, to facilitate comparisons. Differences between cigarettes and HTPs were examined using chi-square tests in SAS (version 9.4; SAS Institute Inc), with a 2-sided significance level of $P<.05$. For variables with sparse expected cell counts (n<5), Fisher exact tests were used instead. For flavoring components, which allowed multiple responses, each component was treated as a binary variable (present vs absent). Proportional differences between cigarettes and HTPs were evaluated using separate chi-square tests.

Ethical Considerations

This study did not involve human participants or animals. The data were collected from publicly available commercial tobacco products. Therefore, ethics approval was not required and was not sought.

Results

Sample General Characteristics

A total of 214 tobacco products were included in the final analytic sample, comprising 150 cigarettes and 64 HTPs. The descriptive characteristics of the collected products are provided in Table 2 and Table S3 in Multimedia Appendix 1. More than half of cigarettes (50%) and HTPs (51.6%) were manufactured by KT&G. The average tar and nicotine contents displayed on cigarette packs were 2.70 mg (range 0.1 - 8.0 mg) and 0.24 mg (range: 0.01 - 0.7 mg), respectively. The display of these contents is not mandatory for HTPs, and none of the HTP packs included them. All products complied with mandated health warnings.

XSL•FO
RenderX

Table . General characteristics of samples (N=214).

Variable	Cigarette (n=150), n (%)	Heated tobacco product (n=64), n (%)	*P* value[a]
Manufacturer			<.001
British American Tobacco (BAT)	21 (14)	10 (15.6)	
Japan Tobacco International (JTI)	26 (17.3)	0 (0)	
Philip Morris International Korea (PMIK)	28 (18.7)	21 (32.8)	
KT&G	75 (50)	33 (51.6)	
Price (KRW[b])			<.001
<4500 KRW	19 (12.7)	0 (0)	
=4500 KRW	116 (77.3)	33 (51.6)	
>4500 KRW	15 (10)	31 (48.4)	
Number of sticks per pack			.70[c]
14	1 (0.7)	0 (0)	
20	149 (99.3)	64 (100)	
Tar content (mg)[d]			<.001
<2.7 mg	82 (54.7)	0 (0)	
≥2.7 mg	68 (45.3)	0 (0)	
N/A[e]	0 (0)	64 (100)	
Nicotine content (mg)[d]			<.001
<0.2 mg	79 (52.7)	0 (0)	
=0.2 mg	16 (10.7)	0 (0)	
>0.2 mg	55 (36.7)	0 (0)	
N/A[e]	0 (0)	64 (100)	
Pack width (mm)			<.001
<60 mm	150 (100)	10 (15.6)	
≥60 mm	0 (0)	54 (84.4)	
Pack depth (mm)			<.001
<17 mm	58 (38.7)	63 (98.4)	
≥17 mm	92 (61.3)	1 (1.6)	
Pack height (mm)			<.001
<82 mm	0 (0)	55 (85.9)	
≥82 mm	150 (100)	9 (14.1)	
Pack shape			<.001
Rectangle	82 (54.7)	54 (84.4)	
Round-edged rectangle	45 (30)	3 (4.7)	
Half round-edged rectangle	23 (15.3)	7 (10.9)	
Stick length (mm)			<.001
<79 mm	150 (100)	55 (85.9)	
≥79 mm	0 (0)	9 (14.1)	
Health warning compliance			N/A[e]
Yes	150 (100)	64 (100)	

XSL•FO
RenderX

Variable	Cigarette (n=150), n (%)	Heated tobacco product (n=64), n (%)	*P* value[a]
No	0 (0)	0 (0)	
Marketing text (excluding brand name)[f]			.03
Yes	103 (68.7)	34 (53.1)	
No	47 (31.3)	30 (46.9)	

[a]*P* values were obtained from chi-square tests comparing the distribution of characteristics between cigarettes and heated tobacco products.

[b]All prices and currency values are reported in Korean won (KRW). The exchange rate as of November 2024 was approximately 1390 KRW to US $1.

[c]The *P* value was calculated using the Fisher exact test.

[d]None of the heated tobacco product packs displayed nicotine or tar content.

[e]NA: not applicable.

[f]Marketing text includes “activated carbon filter,” “nano capsule technology,” and “smartcore induction system,” among others.

The main differences between cigarettes and HTPs were pack size (*P*<.001) and shape (*P*<.001) (Figure S6 in Multimedia Appendix 1). Cigarette sticks were longer, resulting in taller, vertically oriented rectangular packs with variation in depth. HTP packs were shorter, reflecting the shorter stick length, and had a consistent, landscape-oriented rectangular shape.

Although no distinct qualitative differences in promotional elements were observed between cigarettes and HTPs, a higher proportion of cigarette packs (68.7%) included marketing texts (*P*=.030) compared with HTPs (53.1%). Most of these statements highlighted hand and mouth odor reduction technology. For example, texts such as “activated carbon filter,” “nano capsule technology,” and “smartcore induction system” were frequently displayed, serving as marketing claims to emphasize the companies’ technological capabilities.

Flavor Cue

The distribution of flavor cues on packaging is summarized in Table 3. Among the 5 components of pack-level flavor cues, pack color was the most common, appearing on 97 (64.7%) cigarettes and 61 (95.3%) HTPs (*P*<.001). Within pack color cues, cool or minty shades (blue, green, and sky blue) were most frequent, followed by fruity (red, pink, and orange), nutty (brown, dark brown, and yellow), and others (dark green and purple). The second most frequent cues varied by product type, but product names and capsule images were frequently used in both. Product names often implied coolness, such as “aqua,” “cool shot,” “frozen,” and “ice.” Capsule images, typically placed on the front of the pack, signaled flavor through colors associated with specific tastes. For example, blue tones for menthol or ice and red tones for cherry or berry. Pack imagery was less frequently used but featured cool- or minty-related images such as animals (eg, penguins) or natural scenes (eg, glaciers), reinforcing coolness, and other cues included promotional terms such as “double capsule” and “refreshing scent.” In addition, capsule imagery on the stick was observed in 71 (47.3%) cigarettes and 37 (57.8%) HTPs (*P*=.16). Among products with capsule imagery, cigarettes were predominantly cool or minty (59.2%), whereas HTPs were more evenly distributed across cool or minty (27%), fruity (29.7%), and other categories (27%).

XSL•FO
RenderX

Table . Flavor cues and flavoring methods according to tobacco product type.

Variable	Cigarettes (n=150), n (%)	Heated tobacco products (n=64), n (%)	*P* value
Flavor cues (pack)	101 (67.3)	61 (95.3)	<.001
Product name	61 (40.7)	44 (68.8)	<.001
Cool or minty	24 (39.3)	15 (34.1)	
Color	20 (32.8)	14 (31.8)	
Flavor	5 (8.2)	3 (6.8)	
Technology	6 (9.8)	7 (15.9)	
Other	6 (9.8)	5 (11.4)	
Pack color	97 (64.7)	61 (95.3)	<.001
Cool or minty	63 (64.9)	32 (52.5)	
Fruity	21 (21.6)	12 (19.7)	
Nutty	9 (9.3)	5 (8.2)	
Other	4 (4.1)	12 (19.7)	
Capsule imagery	70 (46.7)	39 (60.9)	.06
Cool or minty	41 (58.6)	18 (46.2)	
Fruity	20 (28.6)	14 (35.9)	
Nutty	4 (5.7)	1 (2.6)	
Other	5 (7.1)	6 (15.5)	
Pack imagery	22 (14.7)	9 (14.1)	.91
Animal	4 (18.2)	0 (0)	
Nature	16 (72.7)	6 (66.7)	
Other	2 (9.1)	3 (33.3)	
Other cues	21 (14)	13 (20.3)	.81
Cool or minty	11 (52.4)	9 (69.2)	
Flavor	10 (47.6)	4 (30.8)	
Number of flavor-cued elements[a]			<.001[b]
0	49 (32.7)	3 (4.7)	
1	11 (7.3)	3 (4.7)	
2	35 (23.3)	25 (39.1)	
3	33 (22)	22 (34.4)	
4	19 (12.7)	8 (12.5)	
5	3 (2)	3 (4.7)	
Capsule imagery (stick)	71 (47.3)	37 (57.8)	.16
Cool or minty	42 (59.2)	10 (27)	
Fruity	11 (15.5)	11 (29.7)	
Nutty	11 (15.5)	6 (16.2)	
Others	7 (9.8)	10 (27)	
Pack and stick flavor cue combination	150 (100)	64 (100)	<.001
Pack and stick cues	69 (46)	37 (57.8)	
Pack-only cues	32 (21.3)	24 (37.5)	
Stick-only cues	2 (1.3)	0 (0)	

XSL•FO
RenderX

Variable	Cigarettes (n=150), n (%)	Heated tobacco products (n=64), n (%)	*P* value
Neither	47 (31.3)	3 (4.7)	
Flavored products[c]	89 (59.3)	54 (84.4)	<.001
Inner wrapper	16 (10.7)	8 (12.5)	.67
Cigarette wrapper	10 (6.7)	15 (23.4)	<.001
Tobacco leaves	18 (12)	29 (45.3)	<.001
Crushable capsule	73 (48.7)	37 (57.8)	.22
Cellulose filter	22 (14.7)	7 (10.9)	.47
Charcoal filter	1 (0.7)	0 (0)	≥.99
Polylactic acid filter[d]	—[e]	1 (1.6)	—
Hollow tube[d]	—	15 (23.4)	—
Others	11 (7.3)	2 (3.1)	.35
Number of flavoring methods[f]			<.001
0	61 (40.7)	10 (15.6)	
1	51 (34)	19 (29.7)	
2	19 (12.7)	17 (26.6)	
3	15 (10)	11 (17.2)	
4	4 (2.7)	7 (10.9)	
Products with crushable capsules[g]	73 (100)	37 (100)	<.001
Only a crushable capsule	40 (54.8)	17 (45.9)	
Crushable capsule + 1 method	16 (21.9)	13 (35.1)	
Crushable capsule + 2 methods	14 (19.2)	3 (8.2)	
Crushable capsule + 3 methods	3 (4.2)	4 (10.8)	
Flavoring methods used jointly with crushable capsules[h]			<.001
Inner wrapper	12 (36.4)	5 (25)	
Cigarette wrapper	3 (9.1)	3 (15)	
Tobacco leaves	13 (39.4)	14 (70)	
Cellulose filter	17 (51.5)	4 (20)	
Charcoal filter	0 (0)	0 (0)	
Polylactic acid filter	—	1 (5)	
Hollow tube	—	3 (15)	
Other	8 (24.2)	1 (5)	

[a]Total count of all flavor cues identified on the pack.

[b]*P* values were calculated using the Fisher exact test where applicable.

[c]Percentages may sum to more than 100% because categories are not mutually exclusive; a single product may contain flavor additives in multiple components.

[d]Applicable only to heated tobacco products.

[e]Not applicable.

[f]Total count of components with detectable flavor additives among products.

[g]Percentages are calculated among products with crushable capsules.

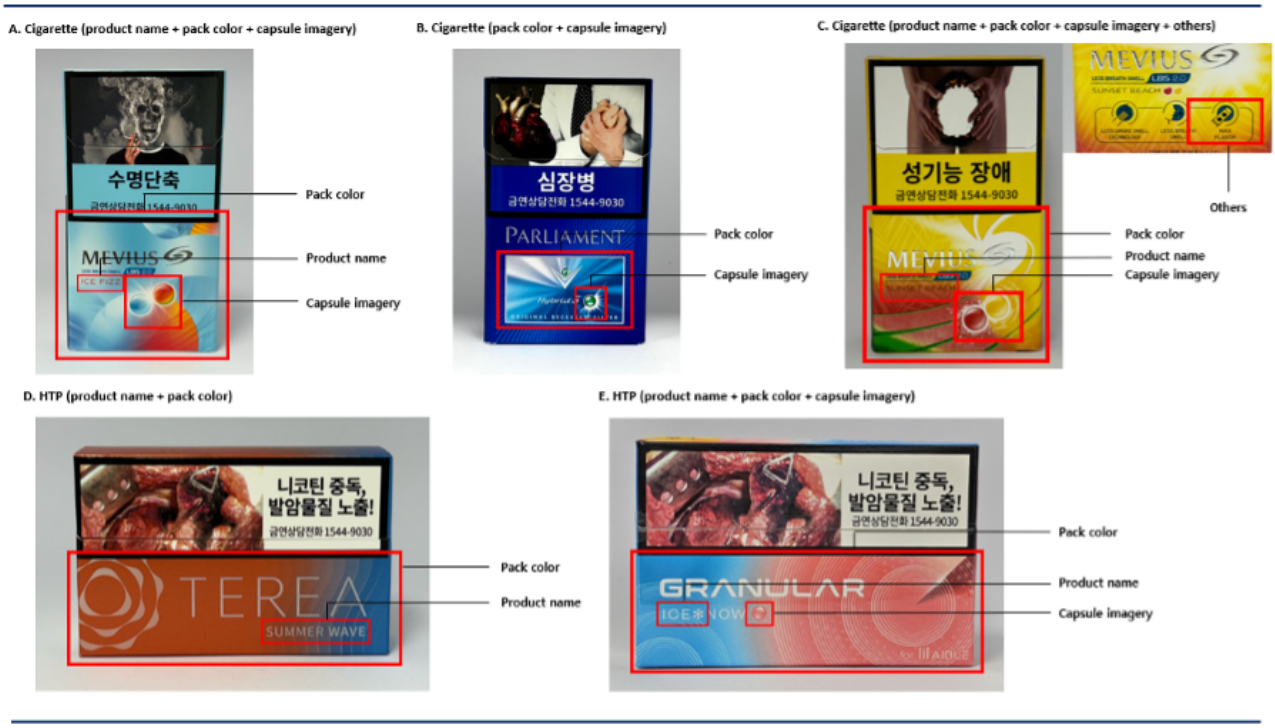
[h]Total count of components with detectable flavor additives among products containing a crushable capsule combined with at least 1 additional flavoring method.

Figure 2 presents examples of flavor cue combinations by product. Overall, 101 (67.3%) cigarettes and 61 (95.3%) HTPs included at least 1 flavor cue (*P*<.001). For cigarettes, color and capsule images were commonly used together, with the most

XSL•FO
RenderX

frequent combinations being "product name + color + capsule image" (11.3%; Figure 2A) and "color + capsule image" (10.7%; Figure 2B). All HTPs with flavor cues included pack color as one of the cue elements, often reinforced by additional cues such as product name and capsule imagery (Figures 2D and 2E).

Figure 2. Common combinations of flavor cues in cigarettes and heated tobacco products.



Flavoring Method

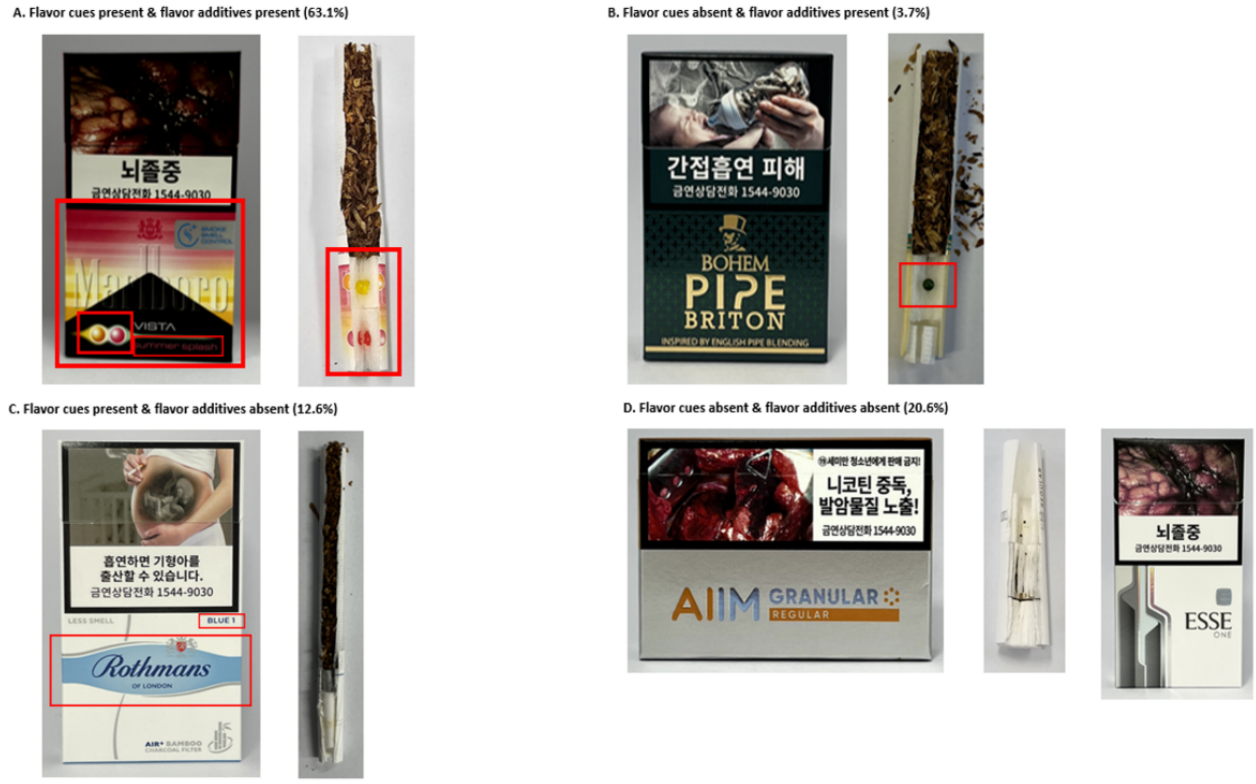
Of the 214 products analyzed, 143 (66.8%) contained at least 1 flavoring component (Table 3). Flavoring was identified in 89 (59.3%) cigarettes and 54 (84.4%) HTPs (P<.001). Crushable capsules were the most common flavoring method in both cigarettes (48.7%) and HTPs (57.8%) (P=.22). In cigarettes, flavor additives were also found in cellulose filters (14.7%) and tobacco leaves (12%). In contrast, HTPs used a wider range of flavoring methods, including tobacco leaves (45.3%), cigarette wrappers (23.4%), and hollow tubes (23.4%).

Although up to 9 flavoring components (7 for cigarettes) were possible, both cigarettes and HTPs showed a maximum of 4 components per product. Overall, 25.4% of cigarettes and 54.7% of HTPs used 2 or more flavoring methods. Among products with crushable capsules, 54.8% of cigarettes and 45.9% of HTPs used capsules as the sole flavoring method. In both product types, up to 3 additional components were combined with capsules. Among capsule-containing products that used at least 1 additional flavoring component, the most common paired components were cellulose filters (51.5%) and tobacco leaves (39.4%) in cigarettes. For HTPs, tobacco leaves (70%) and inner wrappers (25%) were the most frequently paired. The distribution of paired components differed significantly between cigarettes and HTPs (P<.001), with a higher proportion of HTPs containing flavoring in tobacco leaves (70%) compared with cigarettes (39.4%).

Across all products analyzed, 63.1% (54% of cigarettes and 84.4% of HTPs) both had flavor cues on the pack and contained flavor additives (Figure 3A), while 3.7% (5.3% of cigarettes and 0% of HTPs) contained flavor additives without any packaging cues (Figure 3B). Conversely, 12.6% (13.3% of cigarettes and 10.9% of HTPs) displayed cues but lacked detectable flavor additives (Figure 3C), and 20.6% (27.3% of cigarettes and 4.7% of HTPs) showed neither cues nor additives (Figure 3D).

XSL•FO
RenderX

Figure 3. Comparison of flavor cues and actual flavoring in tobacco products.



Discussion

Principal Findings and Implications

We conducted systematic surveillance of cigarettes and HTPs sold in Korea, integrating pack- and stick-level coding of flavor cues with component-based assessment of flavor additives. Our analysis showed that pack sizes and shapes were not standardized and varied across both product types. Pack color was the most common flavor cue, followed by capsule imagery and product names, often used in combination. Most products used 2-3 cues simultaneously, reinforcing flavor signals. Flavor cues and additives were more prevalent among HTPs than cigarettes. Although capsules were the dominant flavoring method, additives were also incorporated into other components.

The flavor cues identified align with previous tobacco pack surveillance studies documenting widespread use of packaging elements to signal flavor. However, regulatory environments shape how these cues manifest. In markets without explicit restrictions, such as the Philippines, 58.5% of tobacco products contained flavor cues through direct descriptors beneath the product name [11]. In contrast, Korea's Article 9‐3 of the NHPA prohibits explicit flavor descriptors such as "menthol" or "cherry," including text or imagery, on tobacco packaging [27]. Nevertheless, flavoring continued to be implied through indirect marketing strategies, using terms such as "Ice" or "Tropical" and capsule imagery on both packs and sticks. Consequently, the prevalence of flavor cues among the products collected in Korea (75.7%) was far higher than the average among countries without flavor bans (22.7%) [12]. These findings suggest that regulatory approaches limited to explicit flavor indicators may be insufficient to eliminate flavor-related marketing.

A key contribution of this study is the combined assessment of flavor cues across both packs and sticks, alongside component-based screening of flavoring additives. Over 60% of products combined flavor cues with detectable flavor additives, but cues were also present in products without detectable flavor additives. Pack-level cues were often reinforced at the point of use, as 47% of cigarettes and 58% of HTPs featured capsule imagery on the stick. This pattern highlights that flavor-related appeal can be generated independently or synergistically through perceptual mechanisms (pack and stick design) [1,47] and material product features (additives) [48]. Linking consumer-facing signals with product-level evidence, therefore, provides a more comprehensive characterization of flavored tobacco products.

These findings have practical implications for enforcing flavor ban policies. Comprehensive flavor bans alone may be insufficient, as they leave room for circumvention. Even without flavor additives, packaging elements that imply flavor can shape perceptions of taste, quality, and harm [41] and may even modulate craving-related neural responses, reinforcing product appeal [49]. Moreover, in jurisdictions where flavor bans have been implemented without plain packaging, the industry has introduced pack- and stick-level design features that insinuate flavor characteristics despite the absence of explicit additives [50]. Flavor bans and plain packaging covering both packs and sticks should therefore be implemented in tandem. However, enforcing flavor bans can be technically complex due to ambiguity in defining "characterizing flavor" [51] and limited

XSL•FO
RenderX

standardized chemical testing capacity [52]. In this context, plain packaging may serve as a more immediately implementable strategy while regulatory systems strengthen enforcement capacity for flavor bans.

Several areas for future research follow from our findings. The high prevalence of flavoring additives among sampled products is broadly consistent with reported sales patterns of flavored tobacco products in Korea [26], highlighting a discrepancy between market patterns and the relatively lower self-reported use of flavored tobacco noted in Korea [53]. Only 29.8% of Korean smokers reported using flavored tobacco products (excluding capsule products), whereas only 26.7% reported that their usual brand contained a flavor capsule [26], which is lower than both the market share of flavored products and the prevalence of additives detected in our assessment. Our findings show that while crushable capsules were common, flavor additives were also frequently incorporated into less visible components such as tobacco leaves, wrappers, and internal filters. Because flavoring in these components is less noticeable and not widely recognized as a marker of flavored tobacco, such products may go unrecognized by consumers as flavored. Future research could examine whether this gap reflects underrecognition of these less salient forms of flavor delivery, or other factors such as higher consumption intensity among flavored-product users [54]. Additionally, the high proportion of discontinued products identified from our sampling procedures suggests rapid market turnover. Continued surveillance research will therefore be important to track evolving product design and flavoring strategies.

Given that evidence on HTP flavoring remains limited globally [55], our findings provide early empirical evidence in this area. Compared with cigarettes, HTPs use more pervasive cues and diverse flavoring methods, often using multiple flavoring methods across various components rather than being limited to capsules alone. As HTP markets continue to expand [56], surveillance frameworks developed for cigarettes, including pack- and stick-level assessment and evaluation of flavor additives, should be extended to HTPs. Our methods and findings offer a preliminary framework that may be applied in other settings to systematically assess the characteristics of HTPs.

Another global implication of our study is that the patterns documented in Korea may extend beyond the domestic market. KT&G, a major Korean tobacco manufacturer and global exporter [57-60], has reported increasing exports of flavored products and capsules [59], with packaging designs in some export markets that are identical or highly similar to those sold in Korea [60,61]. Surveillance of products sold in Korea may therefore provide an early signal of evolving industry tactics that could emerge in other jurisdictions.

Limitations

Our study has several limitations. First, products were purchased in one metropolitan area and primarily from major convenience store chains, so availability may differ in other regions or retail channels (eg, supermarkets and duty-free outlets). Second, some identified products could not be collected due to discontinuation or low demand, which may have introduced selection bias and affected estimates of flavor cues and additives. For example, niche or experimental flavored variants are more likely to be dropped from retail shelves, potentially leading to an underestimation of the diversity or prevalence of flavor-related features. Third, coding was conducted through team-based consensus rather than independent double coding, which may have introduced bias and precluded calculation of intercoder reliability. Fourth, our sensory screening relied on a pragmatic, nonblinded orthonasal assessment conducted by trained research staff rather than professional sensory analysts. Subtle additives may have been missed, and component attribution should be interpreted as approximate due to potential odor diffusion. Future work should incorporate blinded sensory testing conducted by expert sensory panelists and chemical analyses to improve the detection and localization of flavoring additives.

Conclusions

To our knowledge, this study is among the first to simultaneously examine flavor cues on cigarette and HTP packaging and sticks, alongside the presence of flavor additives. By integrating pack- and stick-level assessments, we provide evidence on how flavor cues and additives work together to enhance product appeal. Our findings show that flavoring in tobacco products extends beyond visible pack design or crushable capsules to multiple structural components, highlighting the need for regulations that address both perceptual mechanisms and sources of flavoring. Incorporating stick-level monitoring and HTPs into existing surveillance systems will be crucial for identifying concealed flavoring cues and methods. These efforts can strengthen the implementation of World Health Organization (WHO) FCTC Articles 9 and 10 (regulation and disclosure of tobacco product contents and emissions) and Article 11 (packaging and labeling).

Acknowledgments

We would like to thank the International Tobacco Control (ITC) Project team for providing the list of response options for tobacco product brands, which supported the development of the product list used in this study. During the preparation of this work, the authors used ChatGPT (OpenAI) to improve readability and language. The authors reviewed and edited the content and take full responsibility for the final version of the paper.

XSL•FO
RenderX

Funding

This work was supported by the Ministry of Health and Welfare (MOHW) of Korea and the National Research Foundation of Korea (BK21 Center for Integrative Response to Health Disasters, Graduate School of Public Health, Seoul National University; 4199990514025).

Data Availability

The datasets generated and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

HK and GH conceptualized the study and were the principal investigators. GH drafted the first version of the paper. GH, DK, YK, AK, NC, NK, JP, and EC gathered and analyzed the data. HK and GH visualized the data. HK, SK, and SP administered the study project. SC acquired the funding source. Based on critical feedback, HK and GH revised the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
: Additional tables and figures.
[DOCX File, 7785 KB - publichealth_v12i1e87537_app1.docx]

References

1. Wakefield M, Morley C, Horan JK, Cummings KM. The cigarette pack as image: new evidence from tobacco industry documents. Tob Control 2002 Mar;11 Suppl 1(Suppl 1):I73-I80. [doi: 10.1136/tc.11.suppl_1.i73] [Medline: 11893817]
2. Moodie C, Hastings G. Tobacco packaging as promotion. Tob Control 2010 Apr;19(2):168-170. [doi: 10.1136/tc.2009.033449]
3. Felcione N, Norton K, Bansal-Travers M, Rees V, Cummings K, O'Connor R. Smokers' perceptions of different classes of cigarette brand descriptors. Tob Prev Cessation 2021 Feb 9;7(February):1-11. [doi: 10.18332/tpc/131243]
4. Brown JL, Grilo G, Cohen JE, et al. Colours, capsules and concept flavour names on cigarette packs appeal to youth in Mexico. Tob Control 2023 Apr;32(e1):e16-e22. [doi: 10.1136/tobaccocontrol-2021-056905]
5. Robertson L, McGee R, Marsh L, Hoek J. A systematic review on the impact of point-of-sale tobacco promotion on smoking. Nicotine Tob Res 2015 Jan;17(1):2-17. [doi: 10.1093/ntr/ntu168] [Medline: 25173775]
6. Kong AY, Henriksen L. Retail endgame strategies: reduce tobacco availability and visibility and promote health equity. Tob Control 2022 Mar;31(2):243-249. [doi: 10.1136/tobaccocontrol-2021-056555] [Medline: 35241596]
7. Guidelines for implementation of Article 11. : WHO Framework Convention on Tobacco Control; 2013 URL: https://fctc.who.int/resources/publications/m/item/packaging-and-labelling-of-tobacco-products [accessed 2025-07-01]
8. Moodie C, Hoek J, Hammond D, et al. Plain tobacco packaging: progress, challenges, learning and opportunities. Tob Control 2022 Mar;31(2):263-271. [doi: 10.1136/tobaccocontrol-2021-056559] [Medline: 35241599]
9. Hoek J, Gendall P, Eckert C, Louviere J. Dissuasive cigarette sticks: the next step in standardised ('plain') packaging? Tob Control 2016 Nov;25(6):699-705. [doi: 10.1136/tobaccocontrol-2015-052533]
10. Drovandi A, Teague PA, Glass B, Malau-Aduli B. A systematic review of smoker and non-smoker perceptions of visually unappealing cigarette sticks. Tob Induc Dis 2018;16(January):02. [doi: 10.18332/tid/82191] [Medline: 31516403]
11. Ackary SJ, Cabrera PJD, Santiago AJA, Amul GGH. An analysis of flavor descriptors on tobacco products in the Philippines: regulatory implications and lessons for low- and middle-income countries. Globalization and Health 2024 Sep 9;20(1):67. [doi: 10.1186/s12992-024-01072-6] [Medline: 39252089]
12. Cohen JE, Welding K, Erinoso O, Saraf S, Iacobelli M, Smith KC. The flavor train: the nature and extent of flavored cigarettes in low- and middle-income countries. Nicotine Tob Res 2021 Oct 7;23(11):1936-1941. [doi: 10.1093/ntr/ntab092] [Medline: 33983445]
13. Islam F, Thrasher JF, Szklo A, et al. Cigarette flavors, package shape, and cigarette brand perceptions: an experiment among young Brazilian women. Rev Panam Salud Pública 2018;42:e5. [doi: 10.26633/RPSP.2018.5] [Medline: 31093036]
14. Huang LL, Baker HM, Meernik C, Ranney LM, Richardson A, Goldstein AO. Impact of non-menthol flavours in tobacco products on perceptions and use among youth, young adults and adults: a systematic review. Tob Control 2017 Nov;26(6):709-719. [doi: 10.1136/tobaccocontrol-2016-053196] [Medline: 27872344]
15. Moodie C, Thrasher JF, Cho YJ, Barnoya J, Chaloupka FJ. Flavour capsule cigarettes continue to experience strong global growth. Tob Control 2019 Sep;28(5):595-596. [doi: 10.1136/tobaccocontrol-2018-054711]
16. Brown J, Zhu M, Moran M, Hoe C, Frejas F, Cohen JE. "It has candy. You need to press on it": young adults' perceptions of flavoured cigarettes in the Philippines. Tob Control 2021 May;30(3):293-298. [doi: 10.1136/tobaccocontrol-2019-055524] [Medline: 32447317]

XSL•FO
RenderX

17. Smith K, Washington C, Brown J, et al. The tobacco pack surveillance system: a protocol for assessing health warning compliance, design features, and appeals of tobacco packs sold in low- and middle-income countries. JMIR Public Health Surveill 2015;1(2):e8. [doi: 10.2196/publichealth.4616] [Medline: 27227142]
18. Partial guidelines for implementation of Articles 9 and 10. : WHO Framework Convention on Tobacco Control; 2017 URL: https://fctc.who.int/resources/publications/m/item/regulation-of-the-contents-of-tobacco-products-and-regulation-of-tobacco-product-disclosures [accessed 2025-07-01]
19. Krishnan-Sarin S, O'Malley SS, Green BG, Jordt SE. The science of flavour in tobacco products. World Health Organ Tech Rep Ser 2019 Oct 24;1015(125):125-142. [Medline: 36743396]
20. Villanti AC, Johnson AL, Glasser AM, et al. Association of flavored tobacco use with tobacco initiation and subsequent use among US youth and adults, 2013-2015. JAMA Netw Open 2019 Oct 2;2(10):e1913804. [doi: 10.1001/jamanetworkopen.2019.13804] [Medline: 31642927]
21. Kyriakos CN, Zatoński MZ, Filippidis FT. Flavour capsule cigarette use and perceptions: a systematic review. Tob Control 2023 Apr;32(e1):e83-e94. [doi: 10.1136/tobaccocontrol-2021-056837] [Medline: 34607888]
22. Watkins SL, Pieper F, Chaffee BW, Yerger VB, Ling PM, Max W. Flavored tobacco product use among young adults by race and ethnicity: evidence from the population assessment of tobacco and health study. J Adolesc Health 2022 Aug;71(2):226-232. [doi: 10.1016/j.jadohealth.2022.02.013] [Medline: 35550331]
23. Leas EC, Benmarhnia T, Strong DR, Pierce JP. Use of menthol cigarettes, smoking frequency, and nicotine dependence among US youth. JAMA Netw Open 2022 Jun 1;5(6):e2217144. [doi: 10.1001/jamanetworkopen.2022.17144] [Medline: 35666498]
24. Erinoso O, Clegg Smith K, Iacobelli M, Saraf S, Welding K, Cohen JE. Global review of tobacco product flavour policies. Tob Control 2020 May 15;30(4):373-379. [doi: 10.1136/tobaccocontrol-2019-055454] [Medline: 32414867]
25. Homepage. Korea National Health and Nutrition Examination Survey (KNHANES). 2025. URL: https://knhanes.kdca.go.kr/knhanes/archive/wsiStatsClct.do [accessed 2025-12-18]
26. Comprehensive analysis of flavored tobacco and research on establishing regulatory measures. : Ministry of Health and Welfare; 2025 URL: https://www.prism.go.kr/homepage/asmt/popup/1352000-202500173 [accessed 2025-02-28]
27. National Health Promotion Act. Korea Legislation Research Institute. 2021. URL: https://elaw.klri.re.kr/kor_service/lawView.do?hseq=57497&lang=ENG [accessed 2025-07-01]
28. Kang H, Cheon E, Kim HK, et al. Vision for tobacco endgame in Korea: suggestions for countries with endgame aspirations. Tob Control 2024 Sep 25;33(5):669-675. [doi: 10.1136/tc-2022-057691] [Medline: 37147128]
29. Kang H, Cheon E, Ha J, Cho SI. Use of geographically weighted regression models to inform retail endgame strategies in South Korea: application to cigarette and ENDS prevalence. Tob Control 2025 Apr 1;34(2):205-211. [doi: 10.1136/tc-2023-058117] [Medline: 37816592]
30. Bernat JK, Jackson KJ, Krüsemann EJZ, Boesveldt S, Rudy SF, Talhout R. Sensory methods to evaluate perception of flavours in tobacco and other nicotine-containing products: a review. Tob Control 2023 Apr;32(e1):e95-e102. [doi: 10.1136/tobaccocontrol-2021-056681] [Medline: 34615737]
31. Krüsemann EJZ, Wenng FM, Pennings JLA, de Graaf K, Talhout R, Boesveldt S. Sensory evaluation of e-liquid flavors by smelling and vaping yields similar results. Nicotine Tob Res 2020 Apr 21;22(5):798-805. [doi: 10.1093/ntr/ntz155] [Medline: 31437266]
32. Havermans A, Pauwels C, Bakker-'t Hart IME, et al. Across the world availability of flavour accessories for tobacco products. Tob Control 2025 Jul 31;34(4):532-538. [doi: 10.1136/tc-2023-058255] [Medline: 38580443]
33. Leading convenience store brands in South Korea in 2023, by number of stores. Statista. URL: https://www.statista.com/statistics/973276/south-korea-leading-convenience-stores-by-store-numbers/ [accessed 2024-05-30]
34. ITC Korea Wave 3 (2023) Technical Report. : International Tobacco Control Policy Evaluation Project (ITC Project); 2025 URL: https://itcproject.s3.amazonaws.com/uploads/documents/ITC_KRA3_Technical_Report_April_1_2025_SX.pdf [accessed 2025-03-31]
35. Research on tobacco product market analysis and usage behavior survey methods. : Ministry of Health and Welfare; 2023 URL: https://www.prism.go.kr/homepage/asmt/popup/1351000-202400076 [accessed 2025-07-15]
36. Cunningham R. Tobacco package health warnings: a global success story. Tob Control 2022 Mar;31(2):272-283. [doi: 10.1136/tobaccocontrol-2021-056560]
37. The 4th manual for marking the warning picture and phrases on packaging of tobacco. : Ministry of Health and Welfare, Republic of Korea; 2022 URL: https://assets.tobaccocontrollaws.org/uploads/legislation/Republic%20of%20Korea/Republic-of-Korea-HW-Manual-4th-round.pdf [accessed 2025-07-23]
38. Cohn AM, Alexander AC, Ehlke SJ. Affirming the abuse liability and addiction potential of menthol: differences in subjective appeal to smoking menthol versus non-menthol cigarettes across African American and white young adult smokers. Nicotine Tob Res 2022 Jan 1;24(1):20-27. [doi: 10.1093/ntr/ntab137] [Medline: 34405884]
39. Dewhirst T, Lee WB. Kent cigarette brand marketing in the Republic of Korea: the role of a pioneering image, flavour capsules and leader price promotions. Tob Control 2020;29(6):tobaccocontrol-tobaccocon2019. [doi: 10.1136/tobaccocontrol-2019-055346]

XSL•FO
RenderX

40. Grilo G, Brown JL, Cohen JE, Smith KC. Shared perceptions of flavored cigarette pack design among young adults who smoke in Mexico and the Philippines. Tob Induc Dis 2023;21(July):98. [doi: 10.18332/tid/168376] [Medline: 37502348]
41. Lempert LK, Glantz S. Packaging colour research by tobacco companies: the pack as a product characteristic. Tob Control 2017 May;26(3):307-315. [doi: 10.1136/tobaccocontrol-2015-052656] [Medline: 27255118]
42. Spence C, Velasco C. On the multiple effects of packaging colour on consumer behaviour and product experience in the 'food and beverage' and 'home and personal care' categories. Food Qual Prefer 2018 Sep;68:226-237. [doi: 10.1016/j.foodqual.2018.03.008]
43. Jackler RK, Eamamurthi D, Willett J, et al. Advertising created & continues to drive the menthol tobacco market: methods used by the industry to target youth, women, & Black Americans. : Stanford Research into the Impact of Tobacco Advertising; 2022 URL: https://tobacco-img.stanford.edu/wp-content/uploads/2022/10/02234723/SRITA-AHA-MentholAdvertising.pdf [accessed 2026-03-22]
44. Talhout R, van de Nobelen S, Kienhuis AS. An inventory of methods suitable to assess additive-induced characterising flavours of tobacco products. Drug Alcohol Depend 2016 Apr 1;161:9-14. [doi: 10.1016/j.drugalcdep.2015.12.019] [Medline: 26774948]
45. Pieper E, Mallock N, Henkler-Stephani F, Luch A. Tabakerhitzer als neues Produkt der Tabakindustrie: Gesundheitliche Risiken. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2018 Nov;61(11):1422-1428. [doi: 10.1007/s00103-018-2823-y] [Medline: 30284624]
46. Krüsemann EJZ, Cremers J, Visser WF, Punter PH, Talhout R. The sensory difference threshold of menthol odor in flavored tobacco determined by combining sensory and chemical analysis. Chem Senses 2017 Mar 1;42(3):233-238. [doi: 10.1093/chemse/bjw123] [Medline: 28034900]
47. Borland R, Savvas S. Effects of stick design features on perceptions of characteristics of cigarettes. Tob Control 2013 Sep;22(5):331-337. [doi: 10.1136/tobaccocontrol-2011-050199] [Medline: 22396209]
48. Krüsemann EJZ, Lasschuijt MP, de Graaf C, et al. Sensory analysis of characterising flavours: evaluating tobacco product odours using an expert panel. Tob Control 2019 Mar;28(2):152-160. [doi: 10.1136/tobaccocontrol-2017-054152] [Medline: 29792305]
49. Shi Z, Wang AL, Fairchild VP, et al. Addicted to green: priming effect of menthol cigarette packaging on brain response to smoking cues. Tob Control 2023 Apr;32(e1):e45-e52. [doi: 10.1136/tobaccocontrol-2021-056639] [Medline: 34599084]
50. Brink AL, Glahn AS, Kjaer NT. Tobacco companies' exploitation of loopholes in the EU ban on menthol cigarettes: a case study from Denmark. Tob Control 2023 Nov;32(6):809-812. [doi: 10.1136/tobaccocontrol-2021-057213] [Medline: 35314507]
51. Robson D, Kyriakos CN, McNeill A, et al. Menthol characterizing flavors in cigarettes on sale in England after a characterizing flavor ban: findings from sensory and chemical assessments. Nicotine Tob Res 2025 Sep 23;27(10):1805-1812. [doi: 10.1093/ntr/ntaf064] [Medline: 40227605]
52. Lim HH, Choi KY, Shin HS. Qualitative and quantitative comparison of flavor chemicals in tobacco heating products, traditional tobacco products and flavoring capsules. J Pharm Biomed Anal 2022 Jan 5;207:114397. [doi: 10.1016/j.jpba.2021.114397] [Medline: 34626940]
53. Kim H, Nam HK, Kang H. Tobacco consumption, sales, and output as monitoring indicators in the era of the tobacco endgame: a Korean example. Epidemiol Health 2023;45:e2023030. [doi: 10.4178/epih.e2023030] [Medline: 36915272]
54. Oliveira da Silva AL, Lempert LK, Glantz SA. More than a "characterizing flavor": menthol at subliminal levels in tobacco products. Drug Alcohol Depend 2024 Aug 1;261:111346. [doi: 10.1016/j.drugalcdep.2024.111346] [Medline: 38870568]
55. Ackary SJ, Cabrera P, Santiago AJA, Amul GGH. Examining the flavor descriptors of e-cigarettes, heated tobacco products, and nicotine pouches in the Philippines: regulatory challenges and opportunities. PLOS Glob Public Health 2025;5(2):e0004248. [doi: 10.1371/journal.pgph.0004248] [Medline: 39946399]
56. Sun T, Anandan A, Lim CCW, et al. Global prevalence of heated tobacco product use, 2015-22: a systematic review and meta-analysis. Addiction 2023 Aug;118(8):1430-1444. [doi: 10.1111/add.16199] [Medline: 37005862]
57. Lee K, Gong L, Eckhardt J, Holden C, Lee S. KT&G: from Korean monopoly to "a global name in the tobacco industry". Glob Public Health 2017 Mar;12(3):300-314. [doi: 10.1080/17441692.2016.1273367] [Medline: 28139963]
58. Lee K, Eckhardt J. Looming threat of Asian tobacco companies to global health. The Lancet 2017 May;389(10083):1958-1960. [doi: 10.1016/S0140-6736(17)31223-0]
59. 2025 2nd Quarter Business Report. KT&G. 2025. URL: https://www.ktng.com/report?cmsCd=CM0015 [accessed 2025-08-14]
60. No KM. K-cigarettes go global: KT&G lights up overseas markets. The Korea Herald. URL: https://www.koreaherald.com/article/10529614 [accessed 2025-07-12]
61. Global search. Tobacco Pack Surveillance System (TPackSS). URL: https://tpackss.globaltobaccocontrol.org/global-search/?search=esse [accessed 2024-09-30]

Abbreviations

FCTC: Framework Convention on Tobacco Control
HTP: heated tobacco product

XSL•FO
RenderX

NHPA: National Health Promotion Act
TPackSS: Tobacco Pack Surveillance System
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 10.Nov.2025; peer-reviewed by JL Brown, W Qin; revised version received 17.Mar.2026; accepted 20.Mar.2026; published 25.May.2026.

Please cite as:
Heo G, Kim D, Kim Y, Kim A, Cho N, Kang N, Park J, Cheon E, Kim S, Park S, Park S, Cho SI, Kang H
Systematic Assessment of Flavor Cues and Additives in Cigarettes and Heated Tobacco Products in Korea: Cross-Sectional Surveillance Study
JMIR Public Health Surveill 2026;12:e87537
URL: https://publichealth.jmir.org/2026/1/e87537
doi:10.2196/87537

© Geon Heo, Doyoon Kim, Yehyun Kim, Ara Kim, Nagyeong Cho, Naeun Kang, Jungmi Park, Eunsil Cheon, Seulgi Kim, Susan Park, Sungmin Park, Sung-il Cho, Heewon Kang. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 25.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Epidemiological Shifts and Trends From 2009 to 2023 in Hospital Admissions for Ruptured Aortic Aneurysms and Aortic Dissections in Austria: Retrospective, Population-Wide Study

Fadi Taher*, MD; Maria Elisabeth Leinweber*, MD; Afshin Assadian, MD; Amun Hofmann, MD

Department of Vascular and Endovascular Surgery, Klinik Ottakring, Montleartstraße 37, Vienna, Austria

*these authors contributed equally

Corresponding Author:
Amun Hofmann, MD
Department of Vascular and Endovascular Surgery, Klinik Ottakring, Montleartstraße 37, Vienna, Austria

Abstract

Background: Aortic pathologies in general and acute aortic syndromes in particular are persistent public health concerns worldwide. Continuous efforts to monitor and update incidence rates are necessary for guided public health interventions and health care policy adaptation.

Objective: This study aims to evaluate temporal patterns in the incidence and outcomes of hospital admissions for ruptured aortic aneurysms (rAAs) and aortic dissections (ADs) in Austria over a 15-year period, with a focus on population-wide trends.

Methods: This study provides a retrospective, population-wide analysis of aortic syndromes using national health care data from 2009 to 2023. The data were provided by a federal data warehouse that stores all recorded in-hospital services in Austria. We analyzed annual and population-adjusted hospital admissions, in-hospital mortality, comorbidities, complications, and demographic characteristics for rAA and AD cases. Temporal patterns were assessed for annual variations. Exploratory forecasts were made until 2030 using exponential smoothing time series models. Subgroup analyses by sex and age strata were incorporated to contextualize epidemiological shifts and assess differential trends across population segments.

Results: Corresponding to worldwide trends, the proportion of individuals aged 65 years and older slightly increased in Austria, from 17.4% in 2009 to 19.6% in 2023. Hospital admissions for rAA decreased from 357 in 2009 to 292 in 2023 (−18.2%), with admission rates falling from 4.3 to 3.2 per 100,000 (−25.2%) over the study period. In-hospital mortality among patients with rAA declined from 38.7% (138/357) to 34.6% (101/292), accompanied by a decrease in major comorbidities and recorded in-hospital complications. The incidence of hospital admissions for AD, however, showed a significant increase from 430 to 872 (+102.8%), with population-adjusted incidence rates rising from 5.2 to 9.6 per 100,000 (+85.6%). In-hospital mortality for aggregated AD cases decreased slightly, from 12.1% to 11.0%. Forecasting analyses suggest a continued rise in AD incidence through 2030, whereas rAA rates are expected to remain comparatively stable.

Conclusions: While rAA incidence and mortality declined, AD cases showed a substantial increase, potentially due to improved detection, diagnostic advances, or a true epidemiological shift that warrants further investigation. These diverging trends highlight the need for updated surveillance strategies, improved risk stratification, and adaptation of health care resources to a potentially increasing burden of AD.

(JMIR Public Health Surveill 2026;12:e81702) doi:10.2196/81702

KEYWORDS

aortic aneurysm; rupture; dissection; incidence; mortality

Introduction

The worldwide prevalence of both thoracic and abdominal aortic aneurysms (AAAs) has risen over the past two decades, attributable largely to increased detection through screening and the aging of populations in low- and middle-income countries [1-4]. However, high-income countries have experienced a decline in prevalence, rupture rates, and aneurysm-related mortality from historic peaks, reflecting the success of screening programs, reduced smoking prevalence, and improved management of hypertension [3,5,6]. A recent meta-analysis of 54 population-based studies suggested an AAA prevalence of 1.75% in men and 0.57% in women in Europe, which is in line with contemporary screening studies from this region [3,7,8]. Alongside the reduction in non-ruptured AAAs, driven by changes in cardiovascular risk factors, broader use of preventive elective surgeries through screening programs, and the widespread adoption of endovascular aneurysm repair, the prevalence and mortality of ruptured aortic aneurysms (rAAs) have markedly declined [9,10]. Nonetheless, rAAs remain a

XSL•FO
RenderX

relevant cause of vascular death, with pre-hospital mortality often exceeding 50% and reported in-hospital mortality rates of up to 25% even after successful surgical treatment [11-13]. The advent of endovascular aneurysm repair (EVAR) has markedly reduced mortality in patients with rAAs. Compared to open surgical repair, EVAR is associated with lower perioperative morbidity, reduced hospital stays, and improved short-term survival [9].

The epidemiology of aortic dissections (ADs) is well researched, with significant geographical and temporal variations in incidence. A population-based study in Olmsted County, Minnesota, from 1995 to 2015 reported an incidence of AD of 4.4 per 100,000 person-years, with men being more affected than women [14]. A systematic review and meta-analysis estimated a worldwide incidence of acute ADs of 4.8 per 100,000 individuals per year, with type A dissections occurring at 3.0 per 100,000 and type B dissections at 1.6 per 100,000 [2]. In contrast, a study in Miyazaki prefecture in Japan reported a much higher AD incidence of 17.6 per 100,000, likely due to an older population and comprehensive postmortem computed tomography analysis, which better captured pre-hospital deaths [15,16]. European studies have demonstrated variable incidence, with rates ranging from 2.9 per 100,000 in Hungary to 6 per 100,000 in the United Kingdom, while Sweden reported incidences between 3.3 and 7.2 per 100,000, depending on study methodology. Ontario in Canada reported an incidence of 4.6 per 100,000, whereas Medicare data in the United States indicated a higher hospitalization rate of 10 per 100,000. In Asia, estimates range from 2.8 per 100,000 in China to 10 per 100,000 in Tokyo, Japan [17]. Despite geographical differences, the high mortality associated with AD remains a significant concern, with a 30-day mortality rate as high as 57% in Japan [16].

This study provides a detailed temporal analysis of rAA and AD in the Austrian population over a 15-year observational period, leveraging *International Classification of Diseases, 10th Revision* (ICD-10)–coded health claims data. It thereby aims to investigate potentially contrasting trends in hospital admissions for rAA (ICD-10 codes I71.1, I71.3, I71.5, and I71.8) and AD (I71.0) and incidence, explore possible contributing factors, such as diagnostic advancements or a shifting risk profile of patients, and assess the impact of health care accessibility, including potential effects of the COVID-19 pandemic.

Methods

Data

Data included in this analysis are derived from a federal database that covers in-hospital medical services in Austria. The database in its current structure was established in 2009; therefore, data between 2009 and 2023 were extracted for analysis. ICD-10 codes were used to classify cases. rAAs were defined as I71.1, I71.3, I71.5, and I71.8; AD as I71.0; acute myocardial infarction as I21, stroke as I63, I64, and G45, respiratory complications as J00-J22, coronary heart disease as I25, diabetes mellitus as E10-E14, and pulmonary comorbidities (including chronic obstructive pulmonary disease, asthma, and chronic bronchitis) as J41-45. The current ICD-10 code I71.0 represents a composite code for all types of AD, encompassing types A and B as well as acute, subacute, and chronic presentations, which is a limitation inherent to ICD-10. Furthermore, the proportion of male and female patients was obtained, as well as age strata (60 years and younger, 60-69 years, 70 - 79 years, and 80 years and older) and length of hospital stay strata (less than 9 days, 10-14 days, 15-19 days, and more than 19 days). These data were obtained as aggregated cases per year. Mortality refers to all-cause in-hospital mortality for cases with rAA or AD recorded as the main diagnosis. Additionally, daily case numbers and associated surgeries were provided for rAA and AD. All data were obtained by the Austrian Ministry of Health through the DIAG (Dokumentations- und Informationssysteme für Analysen im Gesundheitswesen) data warehouse. Diagnoses and services were recorded by health care professionals and administrative staff in hospitals and underwent quality control checks at multiple stages. Data covering population size (at the start of the calendar year) and smoking habits were extracted from the publicly accessible databases of Statistics Austria, the federal Austrian statistical office.

Ethical Considerations

Requirement for institutional review board approval was waived by the ethics committee of the City of Vienna due to the exclusive use of anonymized, aggregated data.

Descriptive Analysis

This study is primarily an ecological study investigating national aggregated data. Descriptive analysis included inspection of measures of central tendency and dispersion. A qualitative and quantitative analysis of temporal patterns was conducted, including yearly changes and changes across the observational period. Incidence rates of hospital admissions reflect the number of cases per 100,000 person years. Since all data are aggregated per year, this rate is equal to per 100,000 inhabitants.

Forecasting Future Event Rates

To estimate future trends in the incidence of rAAs and ADs, we applied time series forecasting using exponential smoothing (Error, Trend, Seasonal [ETS]) models. Analyses were conducted separately for men and women, and rates were calculated as mentioned above. We generated univariate time series for each combination of event type (dissection or rupture) and sex. ETS models were fitted to each time series using the *forecast* package in R (version 4.4.2), which automatically selects the best-fitting model form (additive or multiplicative errors, trend, and seasonality) based on information criteria. Forecasts were projected from 2024 to 2030, including 95% CIs. This approach allowed us to estimate future sex-specific incidence trends under the assumption that past dynamics will continue.

Results

Population Characteristics

From 2009 to 2023 the Austrian population increased from 8.3 to 9.1 million inhabitants (Figure 1). The demographic structure remained rather stable over the observational period. The

XSL•FO
RenderX

population consisted of 50% to 51% female inhabitants. Corresponding to worldwide trends, the proportion of individuals aged 65 years or older slightly increased, from 17.4% in 2009 to 19.6% in 2023. As smoking and arterial hypertension constitute major risk factors for both pathologies, we investigated the datasets of Statistics Austria. While yearly data on smoking were not available, Statistics Austria reported that for men and women, daily smoking decreased from 27.5% and 19.4% in 2006 to 23.7% and 17.9% in 2019, respectively. The prevalence of arterial hypertension is also documented by Statistics Austria and did not change substantially between 2006 and 2019, affecting around a fifth of the adult population. The proportion of men with hypertension slightly increased, from 20% in 2006 to 22% in 2019, whereas in women it decreased from 23% to 21.6% [18,19]. Table 1 illustrates the main metrics of interest.

Figure 1. Demographics of the Austrian population from 2009 to 2023. (A) Population size (in millions). (B) Proportion of the population aged 65 years or older.

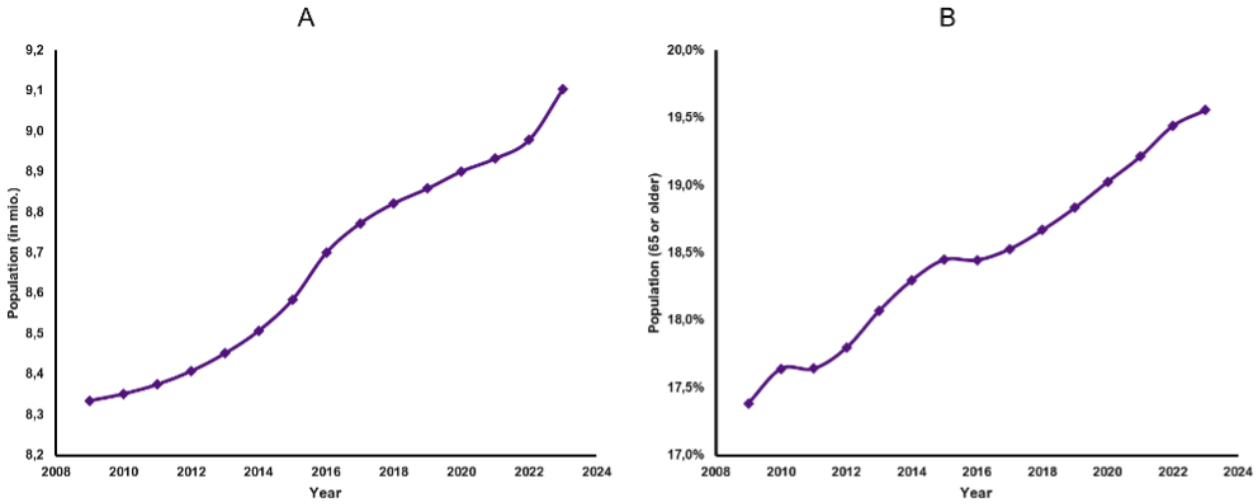


Table . Main metrics of interest regarding trends in population size, ruptured aortic aneurysms (rAAs), and aortic dissections (ADs) in Austria between 2009 and 2023.

Metrics	Median (IQR) between 2009 and 2023	2009	2023	Relative change
Population size (millions)	8.7 (8.4-8.9)	8.3	9.1	9.2%
Admissions for rAA (n)	303 (291-313)	357	292	−18.2%
Admissions for AD (n)	687 (507-720)	430	872	102.8%
Admission rate for rAA (per 100,000)	3.5 (3.2-3.7)	4.3	3.2	−25.2%
Admission rate for AD (per 100,000)	7.8 (6.0-8.1)	5.2	9.6	85.6%
In-hospital mortality rate for rAA (%)	36.7 (34.2-40.0)	38.7	34.6	−10.6%
In-hospital mortality rate for AD (%)	12.1 (11.7-13.1)	12.1	11.0	−9.0%

Trends in rAAs

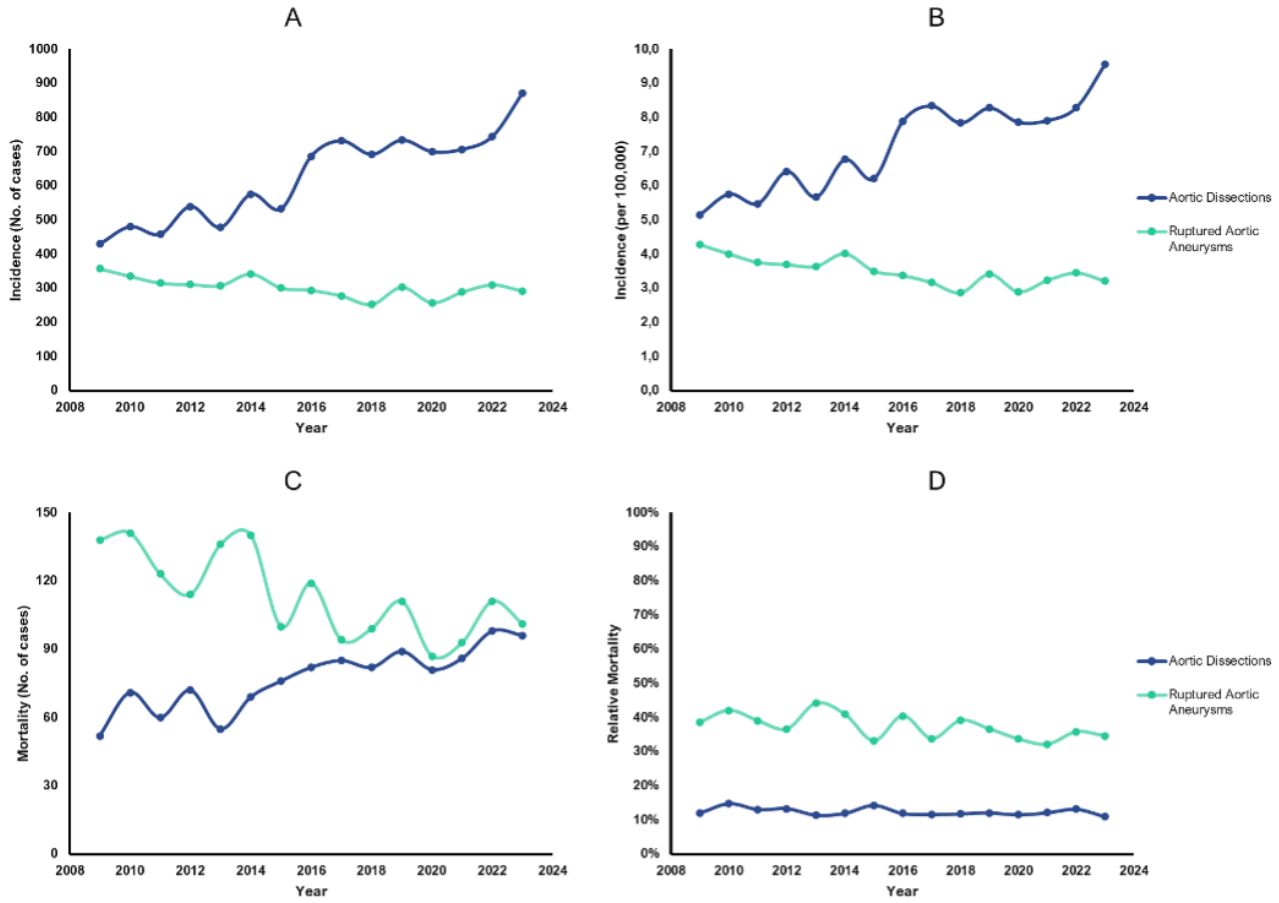
Hospital admissions for rAAs declined from 357 cases in 2009 to 292 cases in 2023 (–18.2%) during the observational period (Figure 2A). Adjusting for the increase in population size, admission rates fell from 4.3 to 3.2 per 100,000 (–25.2%; Figure 2B). Cases with a recorded in-hospital death decreased over the study period (Figure 2C). This corresponds to a relative in-hospital mortality decrease from 38.7% (138/357) of all cases in 2009 to 34.6% (101/292) in 2023, even though this was not a monotonic development, with a peak of 44.2% (136/308) in 2014 (Figure 2D). Of note, this was also accompanied by a decrease in major comorbidities (Figure 3A) and recorded in-hospital complications (Figure 3C).

There were no relevant changes in length of hospital stay during the observational period, with a median of 61.6% (range 56.1%-65.4%) of patients having a recorded length of stay of 9 days or less. Age structure also remained rather stable, with a median of 15.9% (range 12.6%-21.2%) of patients being aged 59 years or younger, a median of 21.3% (range 16.9%-26.0%) aged 60 to 69 years, a median of 31.3% (range 22.1%-36.8%) aged 70 to 79 years, and 30.3% (range 26.5%-39.8%) aged 80 years or older. Sex distribution of cases also remained stable, with a median of 29.5% (range 26.0%-38.0%) of rAAs occurring

XSL•FO
RenderX

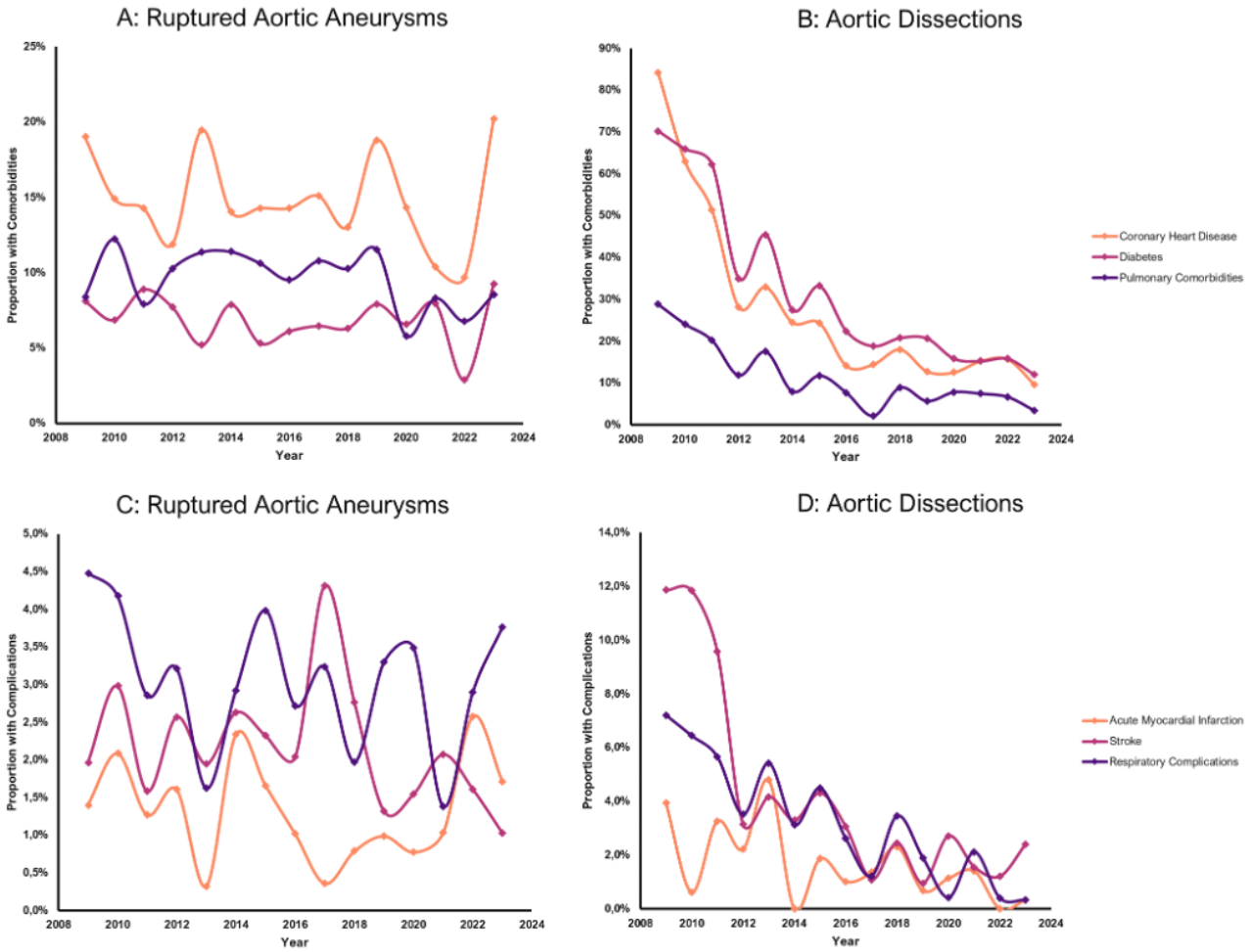
in female patients and correspondingly a median of 70.5% (range 62.0%-74.0%) in male patients.

Figure 2. Trends in ruptured aortic aneurysms and aortic dissections in Austria from 2009 to 2023. (A) Hospital admissions in absolute case numbers. (B) Admission rate per 100,000 for ruptured aortic aneurysms. (C) Mortality (absolute). (D) Mortality as a proportion of all cases (as a percentage).



XSL•FO
RenderX

Figure 3. Comorbidities and complications in cases of ruptured aortic aneurysms and aortic dissections in Austria between 2009 and 2023. (A) Proportion of ruptured aortic aneurysm cases with coronary heart disease, diabetes, or pulmonary comorbidities (B) Proportion of aortic dissection cases with coronary heart disease, diabetes, or pulmonary comorbidities (C) Proportion of ruptured aortic aneurysm cases with recorded in-hospital complications (D) Proportion of aortic dissection cases with recorded in-hospital complications.



Trends in ADs

Hospital admissions for ADs increased from 430 cases in 2009 to 872 in 2023 (+102.8%). The incidence rate rose from 5.2 to 9.6 per 100,000 (+85.6%; Figure 2A and B). Recorded in-hospital mortality decreased slightly from 12.1% to 11.0% over the study period (Figure 2C and D). In the early years of the observational period (2009 - 2013), AD cases were more frequently associated with cardiovascular and pulmonary comorbidities than in later years (Figure 3B and D).

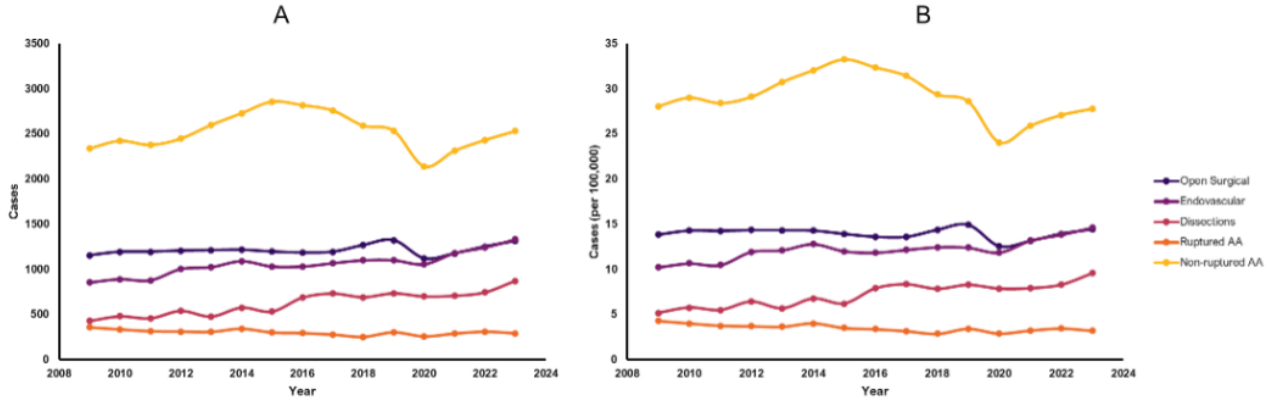
A median of 39.1% (range 34.5% - 44.5%) of patients were aged 59 years or younger, a median of 26.6% (range 19.3% - 31.2%) were aged 60 to 69 years, a median of 23.5% (range 19.1% - 29.0%) were aged 70 to 79 years, and a median of 10.2% (range 7.2% - 13.0%) were aged 80 years or older. A median of 35.6% (range 28.1% - 37.2%) of affected patients over the observational period were female and correspondingly a median of 64.4% (range 26.0% - 62.8%) were male.

Surgical Treatment of Aorta Pathologies

Figure 4A and B illustrate the temporal pattern of aortic surgeries in total and per 100,000 in Austria from 2009 to 2023. To explore a potential relationship between elective invasive treatments for nonruptured aortic aneurysms and the incidence of rAA, we investigated treatment patterns in elective settings. Elective treatments peaked in 2015 and steadily declined after that. With the start of the COVID-19 pandemic in 2020, both nonruptured aortic aneurysm admissions as well as surgical aortic aneurysm treatments showed a local minimum and a recovery of the caseload in the following years, returning to previous patterns. Endovascular aortic surgeries did not exhibit a corresponding trend, as they showed the steepest increase following the pandemic and a gradual increase over the study period. AD cases also showed a rather continual increase throughout the observational period. While no absolute increase in rAA cases was noted after the pandemic, the steepest increase in rAA cases did occur in the years after 2020.

XSL•FO
RenderX

Figure 4. Temporal patterns regarding surgical treatments in Austria between 2009 and 2023. (A) Aortic surgeries (absolute). (B) Aortic surgeries per 100,000. AA: aortic aneurysm.



Midterm Forecasts

In an exploratory analysis, forecasts for AD and rAA cases were made from 2024 through 2030, leveraging ETS models (see Methods). Incidence rate forecasts for AD cases among both male and female patients continuously increased, with a peak of 6.9 (95% CI 5.8 - 8.0) in male patients and 4.1 (95% CI 3.5 - 4.6) in female patients. This would result in a doubling of population-adjusted AD cases in about 20 years. Regarding rAA cases, estimated forecasts remained constant to the observational period from 2009 to 2023 in both sexes. These forecasts are illustrated in Figure 5.

Figure 5. Forecasted incidence rates of aortic dissection and aortic rupture by sex from 2009 to 2030. Solid lines represent historical rates per 100,000 population from 2009 to 2023, calculated separately for men and women. Shaded ribbons indicate 95% CIs of the forecasted rates for 2024 to 2030, derived using exponential smoothing time series models.

Discussion

This population-wide analysis highlights relevant changes in hospital admissions and outcomes of rAA and AD in a growing Austrian population over a 15-year period. The observed decline in rAA-related admissions and mortality contrasts a marked increase in AD cases, warranting further investigation into potential contributing factors. The admission rate for rAA decreased by 25.2% between 2009 and 2023, while the rate of AD increased by 85.6% over the same period. In-hospital mortality decreased slightly for both conditions, with relative reductions of around 10% over the observational period. The epidemiological profiles of affected patients correspond well with previous literature, with rAA occurring more often in older

XSL•FO
RenderX

patient groups, whereas AD was frequently observed in patients younger than 60 years. Both pathologies affected male patients almost twice as often as female patients [20,21].

Epidemiological shifts in aortic pathologies are accompanied by corresponding trends for associated risk factors. Smoking, a main contributor to aortic aneurysm occurrence and rupture, has declined in many high-income countries in recent decades. The World Health Organization (WHO) estimates that tobacco use prevalence has more than halved between 2000 and 2025 in Austria, a trend that can be observed in other European countries, such as Germany or the United Kingdom, to a comparable extent. During the same period, the WHO has reported a decrease in the prevalence of hypertension among many European countries, even though estimates commonly show reductions of 20% to 25% [22]. Nevertheless, it should be considered that the point prevalence of a pathology is merely one aspect, with disease management being at least equally important to understand the epidemiological burden. As for arterial hypertension, advancements in treatment strategies and patient care have resulted in vastly improved blood pressure control in high-income countries [23,24]. The same considerations should be applied when interpreting the present data.

Reports of increased occurrence of AD from different geographical regions are not uncommon in the current literature [14]. The rise may be attributed to a variety of different variables, which may include improved diagnostics and therapeutic advances [25] that allow invasive treatment in historically inoperable cases. Developments in imaging as well as availability and use of high-resolution imaging in Austria may also have led to more frequent detection of AD, especially in the earlier stages, akin to how screenings can temporarily increase incidence rates. Because of this, the critical question remains whether the rise reflects a true increase in disease occurrence or an artifact of improved case ascertainment. A possible explanation could therefore be an increase in elective cases that undergo invasive treatment. Nevertheless, the absence of information regarding urgency and morphology in ICD-10 codes for AD (I71.0) remains a challenge when analyzing health claims data and routine electronic health records. Literature estimates of AD mortality vary widely, being heavily influenced by dissection type, patient characteristics, treatment indications, and modalities. Type A dissections have a reported 30-day all-cause mortality of 16.9% to 22% with surgical intervention and up to 50% to 69% without it [26-28]. In contrast, type B dissections, managed medically or with endovascular repair, have lower mortality, typically 4.5% to 13.8% [28,29]. Patient factors such as age, comorbidities, and genetic predispositions further modulate outcomes, as do treatment settings (acute vs elective) and modalities (open surgery vs endovascular). The current study's aggregation of all AD cases, encompassing types A and B, acute and elective presentations, averaged mortality estimates, as the lower mortality of elective or type B cases may offset the higher lethality of acute type A dissections. Moreover, the exclusion of pre-hospital deaths introduced a significant survival bias, as patients who die before reaching the hospital were not captured. Nevertheless, considering this amalgamation and the reporting of in-hospital mortality rather than 30-day mortality, the present data correspond well to comparable investigations [30].

Regarding rAAs, there might be less scope for interpretation of the present data in detail. A decrease in hospital admissions for AAAs in England and a shift toward endovascular treatment has recently been reported [31]. While reductions in the incidence of rAA were already observed a decade ago [10], the underlying cause, although debated, was considered to likely be multifactorial and at least partially attributable to the implementation of screening programs and subsequent increases in elective repair [32]. The present data correspond to this hypothesis, with increasing elective cases until 2015. Although Austria does not have an established national AAA screening program, we hypothesize that the increased use of EVAR and greater awareness among primary care physicians may have contributed to a concomitant and subsequent decrease in the incidence of rAA. The absence of a national screening program is primarily a policy decision by the federal government and the association of national health insurance institutions and stands in contrast to current clinical practice guidelines [9]. Small decreases and an implied negative trend in mortality over the observational period are also likely multifactorial and certainly require further investigation. However, the fluctuating temporal pattern and overall high rates indicate that ruptures—when they occur—are still a major hazardous cardiovascular event.

Invasive treatment for nonruptured aortic aneurysms in Austria increased continuously from 2009 to 2015, when they peaked, with a subsequent steady decline. However, during the height of the global COVID-19 pandemic, a noticeable decrease was observed, with a recovery in 2023 to prepandemic caseloads. The fact that the COVID-19 pandemic led to delayed access to health care services was a phenomenon observed worldwide [33,34] that affected many pathologies, including cancer [35], stroke [36], and acute myocardial infarction [37]. Despite the challenges the COVID-19 pandemic presented to health care systems worldwide, its effects on aortic aneurysm treatment and associated outcomes were reported to be limited from a public health perspective [38-40]. Nevertheless, it was previously hypothesized that decreases in elective aortic aneurysm repair during the pandemic could have caused an accumulation of untreated aneurysms and an increased risk for rAA [41]. The slight decrease in elective aortic aneurysm treatments in 2020 was accompanied by small increases in ruptures in 2021 and 2022. However, this trend appeared to reverse in 2023, and nonruptured aortic aneurysm treatments were already in decline from 2015 onward. During the same time, both hospital admissions with rAA as well as associated mortality declined. While effects of the COVID-19 pandemic on the treatment of aortic aneurysms in Austria cannot be completely excluded, they appear rather limited for the time being.

The forecasts presented in this study should be interpreted as exploratory and illustrative rather than definitive predictions. Nevertheless, the estimated continuous increase in AD cases could require adaptation of infrastructure and policies, as forecasts indicate a potential doubling of caseloads in 2030 compared to 2009. However, while ETS models provide a

XSL•FO
RenderX

data-driven approach to estimate potential future trends in the incidence of ADs and ruptures, these projections are inherently uncertain and depend on the assumption that past patterns will continue unchanged. Consequently, these estimates should not be used as the sole basis for health care policy or resource allocation decisions. Instead, they should be seen as serving to provide an informed understanding of possible developments that may warrant attention. It is also essential to reassess and update these forecasts as new data become available, particularly because the most recent data included extend only through 2023, and additional information from subsequent years may significantly influence future trends and model accuracy.

The observational period, spanning 15 years, may not have fully captured longer-term trends, and more extensive time series analyses will be required in the future. Additionally, as the analysis was based on data from Austria, generalizing the findings to other countries or regions with differing health care systems, demographics, and lifestyle factors may be challenging. Furthermore, difficulties arise with the recording of ADs via ICD-10, since the code I71.0 condenses both acute and elective clinical presentations as well as variations in morphology, that is, type A, B, non-A, or non-B. Similarly, most rAA cases were coded as I71.8 (rupture of unspecified location), which prevented further analysis of affected anatomic segments. Potential unmeasured confounders, such as changes in diagnostic practices, population health behaviors, or health care access, could have influenced the observed trends. Moreover, limited data on smoking prevalence and other individual-level risk factors precluded a more granular analysis of their impact on AD and rAA. The available information on smoking trends restricts interpretations to broader temporal patterns rather than direct year-to-year associations with incidence changes, as annual fluctuations in smoking cannot be precisely correlated with yearly variations in rAA or AD rates. These limitations highlight the need for multicountry analyses and screening studies with more detailed individual-level data to corroborate and expand upon these findings. Furthermore, the included data are based on recorded in-hospital diagnoses and services, and misclassifications in claims data are a well-described limitation.

The database used here included and represented more than 99% of all in-hospital services in Austria, thereby truly enabling a nationwide evaluation of hospital admissions, practice patterns, and outcomes of aortic syndromes between 2009 and 2023. However, limitations of the available data within the database highlight the value of nationwide registries such as the Vascular Registry in Sweden. Nevertheless, the present analysis offers insights into the evolving patterns of aortic pathologies, which may deserve consideration in health care planning and patient management, while adding recent data to the cumulative body of evidence on aortic pathologies.

In light of shifting demographics, changing comorbidity profiles, and evolving diagnostic and treatment options, aortic syndromes continue to be relevant from a public health perspective.

Acknowledgments

Generative artificial intelligence (GPT-4.1) was used to improve structure and phrasing of the manuscript.

Funding

Article processing charges were funded by Gefäßforum Österreich.

Data Availability

Data are available only from the Austrian Ministry of Health.

Conflicts of Interest

None declared.

References

1. Ngetich E, Ward J, Cassimjee I, Lee R, Handa A, Oxford Abdominal Aortic Aneurysm OXAAA Group. Prevalence and epidemiological pattern of abdominal aortic aneurysms in Africa: a systematic review. J West Afr Coll Surg 2020;10(1):3-14. [doi: 10.4103/jwas.jwas_15_21] [Medline: 35531585]
2. Gouveia E Melo R, Silva Duarte G, Lopes A, et al. Incidence and prevalence of thoracic aortic aneurysms: a systematic review and meta-analysis of population-based studies. Semin Thorac Cardiovasc Surg 2022;34(1):1-16. [doi: 10.1053/j.semtcvs.2021.02.029] [Medline: 33705940]
3. Tyrovolas S, Tyrovola D, Giné-Vázquez I, et al. Global, regional, and national burden of aortic aneurysm, 1990-2017: a systematic analysis of the Global Burden of Disease Study 2017. Eur J Prev Cardiol 2022 May 27;29(8):1220-1232. [doi: 10.1093/eurjpc/zwab015] [Medline: 33783496]
4. Song P, He Y, Adeloye D, et al. The global and regional prevalence of abdominal aortic aneurysms: a systematic review and modeling analysis. Ann Surg 2023 Jun 1;277(6):912-919. [doi: 10.1097/SLA.0000000000005716] [Medline: 36177847]
5. Schermerhorn ML, Bensley RP, Giles KA, et al. Changes in abdominal aortic aneurysm rupture and short-term mortality, 1995-2008: a retrospective observational study. Ann Surg 2012 Oct;256(4):651-658. [doi: 10.1097/SLA.0b013e31826b4f91] [Medline: 22964737]

XSL•FO
RenderX

6. Png CYM, Wu J, Tang TY, Png IPL, Sheng TJ, Choke E. Editor's choice - decrease in mortality from abdominal aortic aneurysms (2001 to 2015): is it decreasing even faster? Eur J Vasc Endovasc Surg 2021 Jun;61(6):900-907. [doi: 10.1016/j.ejvs.2021.02.013] [Medline: 33773903]
7. Svensjö S, Björck M, Gürtelschmid M, Djavani Gidlund K, Hellberg A, Wanhainen A. Low prevalence of abdominal aortic aneurysm among 65-year-old Swedish men indicates a change in the epidemiology of the disease. Circulation 2011 Sep 6;124(10):1118-1123. [doi: 10.1161/CIRCULATIONAHA.111.030379] [Medline: 21844079]
8. Benson RA, Poole R, Murray S, Moxey P, Loftus IM. Screening results from a large United Kingdom abdominal aortic aneurysm screening center in the context of optimizing United Kingdom National Abdominal Aortic Aneurysm Screening Programme protocols. J Vasc Surg 2016 Feb;63(2):301-304. [doi: 10.1016/j.jvs.2015.08.091] [Medline: 26482996]
9. Wanhainen A, Van Herzeele I, Bastos Goncalves F, et al. Editor's choice -- European Society for Vascular Surgery (ESVS) 2024 Clinical Practice Guidelines on the Management of Abdominal Aorto-Iliac Artery Aneurysms. Eur J Vasc Endovasc Surg 2024 Feb;67(2):192-331. [doi: 10.1016/j.ejvs.2023.11.002] [Medline: 38307694]
10. Abdulameer H, Al Taii H, Al-Kindi SG, Milner R. Epidemiology of fatal ruptured aortic aneurysms in the United States (1999-2016). J Vasc Surg 2019 Feb;69(2):378-384. [doi: 10.1016/j.jvs.2018.03.435] [Medline: 29960790]
11. Reimerink JJ, van der Laan MJ, Koelemay MJ, Balm R, Legemate DA. Systematic review and meta-analysis of population-based mortality from ruptured abdominal aortic aneurysm. Br J Surg 2013 Oct;100(11):1405-1413. [doi: 10.1002/bjs.9235] [Medline: 24037558]
12. Karthikesalingam A, Holt PJ, Vidal-Diez A, et al. Mortality from ruptured abdominal aortic aneurysms: clinical lessons from a comparison of outcomes in England and the USA. Lancet 2014 Mar 15;383(9921):963-969. [doi: 10.1016/S0140-6736(14)60109-4] [Medline: 24629298]
13. Pál D, Szilágyi B, Berczeli M, et al. Ruptured aortic aneurysm and dissection related death: an autopsy database analysis. Pathol Oncol Res 2020 Oct;26(4):2391-2399. [doi: 10.1007/s12253-020-00835-x] [Medline: 32548697]
14. DeMartino RR, Sen I, Huang Y, et al. Population-based assessment of the incidence of aortic dissection, intramural hematoma, and penetrating ulcer, and its associated mortality from 1995 to 2015. Circ Cardiovasc Qual Outcomes 2018 Aug;11(8):e004689. [doi: 10.1161/CIRCOUTCOMES.118.004689] [Medline: 30354376]
15. Yamaguchi T, Nakai M, Sumita Y, et al. Current status of the management and outcomes of acute aortic dissection in Japan: analyses of nationwide japanese registry of all cardiac and vascular diseases-diagnostic procedure combination data. Eur Heart J Acute Cardiovasc Care 2020 Oct;9(3_suppl):S21-S31. [doi: 10.1177/2048872619872847] [Medline: 31460772]
16. Booth K, UK-AS, the UK Aortic Society. Acute aortic dissection (AAD) - a lethal disease: the epidemiology, pathophysiology and natural history. Br J Cardiol 2023;30(1):9. [doi: 10.5837/bjc.2023.009] [Medline: 37705834]
17. Yin J, Liu F, Wang J, Yuan P, Wang S, Guo W. Aortic dissection: global epidemiology. Cardiology Plus 2022;7(4):151-161. [doi: 10.1097/CP9.0000000000000028]
18. Klimont J, Kytir J, Leitner B. Österreichische Gesundheitsbefragung 2006/2007: Hauptergebnisse Und Methodische Dokumentation: BMGFJ; 2007.
19. Gesundheitszustand selbstberichtet. Statistik Austria. URL: https://www.statistik.at/statistiken/bevoelkerung-und-soziales/gesundheit/gesundheitszustand/gesundheitszustand-selbstberichtet [accessed 2025-04-13]
20. Schmitz-Rixen T, Keese M, Hakimi M, et al. Ruptured abdominal aortic aneurysm-epidemiology, predisposing factors, and biology. Langenbecks Arch Surg 2016 May;401(3):275-288. [doi: 10.1007/s00423-016-1401-8] [Medline: 27001684]
21. Sen I, Erben YM, Franco-Mesa C, DeMartino RR. Epidemiology of aortic dissection. Semin Vasc Surg 2021 Mar;34(1):10-17. [doi: 10.1053/j.semvascsurg.2021.02.003] [Medline: 33757630]
22. Tobacco control: monitor: current tobacco use, tobacco smoking and cigarette smoking, age-standardized. World Health Organization. URL: https://www.who.int/data/gho/data/indicators/indicator-details/GHO/gho-tobacco-control-monitor-current-tobaccouse-tobaccosmoking-cigarrettesmoking-agestd-tobagestdcurr [accessed 2025-04-19]
23. Parati G, Lombardi C, Pengo M, Bilo G, Ochoa JE. Current challenges for hypertension management: from better hypertension diagnosis to improved patients' adherence and blood pressure control. Int J Cardiol 2021 May 15;331:262-269. [doi: 10.1016/j.ijcard.2021.01.070] [Medline: 33548384]
24. Verdecchia P, Cavallini C, Angeli F. Advances in the treatment strategies in hypertension: present and future. J Cardiovasc Dev Dis 2022 Mar 3;9(3):72. [doi: 10.3390/jcdd9030072] [Medline: 35323620]
25. Juraszek A, Czerny M, Rylski B. Update in aortic dissection. Trends Cardiovasc Med 2022 Oct;32(7):456-461. [doi: 10.1016/j.tcm.2021.08.008] [Medline: 34411744]
26. Czerny M, Schoenhoff F, Etz C, et al. The impact of pre-operative malperfusion on outcome in acute type A aortic dissection: results from the GERAADA Registry. J Am Coll Cardiol 2015 Jun 23;65(24):2628-2635. [doi: 10.1016/j.jacc.2015.04.030] [Medline: 26088302]
27. Geirsson A, Shioda K, Olsson C, et al. Differential outcomes of open and clamp-on distal anastomosis techniques in acute type A aortic dissection. J Thorac Cardiovasc Surg 2019 May;157(5):1750-1758. [doi: 10.1016/j.jtcvs.2018.09.020] [Medline: 30401530]

XSL•FO
RenderX

28. Evangelista A, Isselbacher EM, Bossone E, et al. Insights from the International Registry of Acute Aortic Dissection: a 20-year experience of collaborative clinical research. Circulation 2018 Apr 24;137(17):1846-1860. [doi: 10.1161/CIRCULATIONAHA.117.031264] [Medline: 29685932]
29. Evangelista A, Rabasa JM, Mosquera VX, et al. Diagnosis, management and mortality in acute aortic syndrome: results of the Spanish Registry of Acute Aortic Syndrome (RESA-II). Eur Heart J Acute Cardiovasc Care 2018 Oct;7(7):602-608. [doi: 10.1177/2048872616682343] [Medline: 28029052]
30. Obel LM, Lindholt JS, Lasota AN, et al. Clinical characteristics, incidences, and mortality rates for type a and b aortic dissections: a nationwide Danish population-based cohort study from 1996 to 2016. Circulation 2022 Dec 20;146(25):1903-1917. [doi: 10.1161/CIRCULATIONAHA.122.061065] [Medline: 36321467]
31. Hanna L, Borsky K, Abdullah AA, et al. Trends in hospital admissions, operative approaches, and mortality related to abdominal aortic aneurysms in England between 1998 and 2020. Eur J Vasc Endovasc Surg 2023 Jul;66(1):68-76. [doi: 10.1016/j.ejvs.2023.03.015] [Medline: 36934837]
32. Leone N, Broda MA, Eiberg JP, Resch TA. Systematic review and meta-analysis of the incidence of rupture, repair, and death of small and large abdominal aortic aneurysms under surveillance. J Clin Med 2023 Oct 29;12(21):6837. [doi: 10.3390/jcm12216837] [Medline: 37959301]
33. de Oliveira MM, Fuller TL, Gabaglia CR, et al. Repercussions of the COVID-19 pandemic on preventive health services in Brazil. Prev Med 2022 Feb;155:106914. [doi: 10.1016/j.ypmed.2021.106914] [Medline: 34953811]
34. Mani SS, Schut RA. The impact of the COVID-19 pandemic on inequalities in preventive health screenings: trends and implications for U.S. population health. Soc Sci Med 2023 Jul;328:116003. [doi: 10.1016/j.socscimed.2023.116003] [Medline: 37301108]
35. Fox L, Monroy-Iglesias MJ, Aggarwal A, et al. Association between COVID-19 burden and delays to diagnosis and treatment of cancer patients in England. J Cancer Policy 2022 Mar;31:100316. [doi: 10.1016/j.jcpo.2021.100316] [Medline: 35559868]
36. Gu S, Li J, Shen H, et al. The impact of COVID-19 pandemic on treatment delay and short-term neurological functional prognosis for acute ischemic stroke during the lockdown period. Front Neurol 2022;13:998758. [doi: 10.3389/fneur.2022.998758] [Medline: 36341110]
37. Erol MK, Kayıkçıoğlu M, Kılıçkap M, et al. Treatment delays and in-hospital outcomes in acute myocardial infarction during the COVID-19 pandemic: a nationwide study. Anatol J Cardiol 2020 Nov;24(5):334-342. [doi: 10.14744/AnatolJCardiol.2020.98607] [Medline: 33122486]
38. Huang A, Henke P, Kasper G, Aplin B, Lurie F. The impact of the COVID-19 public health crisis on the surgical management of abdominal aortic aneurysm rupture. Ann Vasc Surg 2025 May;114:302-312. [doi: 10.1016/j.avsg.2024.12.049] [Medline: 39732327]
39. Grimsley EA, Janjua HM, Asirwatham M, et al. Endovascular aortic repair for ruptured abdominal aortic aneurysm during the COVID-19 era. JVS-Vascular Insights 2024;2:100073. [doi: 10.1016/j.jvsvi.2024.100073]
40. McGuinness B, Troncone M, James LP, Bisch SP, Iyer V. Reassessing the operative threshold for abdominal aortic aneurysm repair in the context of COVID-19. J Vasc Surg 2021 Mar;73(3):780-788. [doi: 10.1016/j.jvs.2020.08.115] [Medline: 32882347]
41. Bette S, Decker JA, Zerwes S, et al. German nation-wide in-patient treatment of abdominal aortic aneurysm-trends between 2005 and 2019 and impact of the SARS-CoV-2 pandemic. CVIR Endovasc 2023 Aug 29;6(1):44. [doi: 10.1186/s42155-023-00389-4] [Medline: 37642825]

Abbreviations

AAA: abdominal aortic aneurysm
AD: aortic dissection
EVAR: endovascular aneurysm repair
ICD-10: *International Classification of Diseases, 10th Revision*
rAA: ruptured aortic aneurysm
WHO: World Health Organization

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 01.Aug.2025; peer-reviewed by EO Mensah, M Sedighi; revised version received 25.Mar.2026; accepted 30.Mar.2026; published 12.Jun.2026.

Please cite as:
Taher F, Leinweber ME, Assadian A, Hofmann A
Epidemiological Shifts and Trends From 2009 to 2023 in Hospital Admissions for Ruptured Aortic Aneurysms and Aortic Dissections in Austria: Retrospective, Population-Wide Study
JMIR Public Health Surveill 2026;12:e81702
URL: https://publichealth.jmir.org/2026/1/e81702
doi:10.2196/81702

© Fadi Taher, Maria Elisabeth Leinweber, Afshin Assadian, Amun Hofmann. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 12.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Twenty-Five Years of Firearm Homicides and Suicides Among US Children and Young Adults, 1999 to 2024: Intersectional Surveillance Analysis

Chuka Emezue[1], MPA, MPH, PhD; Tipparat Udmuangpia[2], RN, MSN, PhD

[1]Women, Children, and Family Nursing, College of Nursing, Rush University, 600 S Paulina St, Chicago, IL, United States

[2]Boromarajonani College of Nursing, Khon Kaen, Faculty of Nursing, Praboromarajchanok Institute, Nonthaburi, Thailand

Corresponding Author:
Chuka Emezue, MPA, MPH, PhD
Women, Children, and Family Nursing, College of Nursing, Rush University, 600 S Paulina St, Chicago, IL, United States

Abstract

Background: Firearm mortality surveillance relies on aggregate demographic categories or low-dimensional stratifications, obscuring how race, ethnicity, sex, age, and intent intersect to shape risk among youth and young adults. The Centers for Disease Control and Prevention (CDC)'s transition to single-race population estimates created a structural break between historical and contemporary data series. Consequently, the 25-year trajectory of intersectional firearm mortality across changing federal classification frameworks remains uncharacterized.

Objective: We aimed to characterize 25-year national trajectories in firearm homicide and suicide across intersecting domains of race or ethnicity, sex, age, and intent among US youth and young adults (aged 0 - 24 y), identifying specific sociodemographic groups experiencing the highest and most rapidly changing burdens.

Methods: This study used the CDC WONDER Underlying Cause of Death database (25-year window from 1999 - 2024). To address the 2018 transition, a 2-period design was deployed: period 1 (1999 to 2020) used bridged 4-category race data, and period 2 (2018 to 2024) used 6-category single-race data. Crude death rates and disparity ratios (referenced to non-Hispanic White youth) were calculated using the Byar approximation. Average annual percentage changes (AAPCs) were modeled via log-linear regression. Analyses incorporated finalized 2024 CDC WONDER data in January 2026.

Results: Among 200,704 firearm deaths among youth aged 0 - 24 years from 1999 to 2024, overall mortality declined through 2013, reversed in 2014, and rose 28.9% from 2019 to 2020. In 2024, the overall rate declined 12% from 2023 but remained above the 2019 prepandemic baseline. During 2018 - 2024, Black non-Hispanic male youth experienced the highest pooled firearm homicide rate (44.03 per 100,000; 95% CI 43.46 - 44.61), 23.8 times the rate among White non-Hispanic males. The annual rates within this cohort increased from 33.28 in 2018 to 53.75 in 2021 before declining to 37.06 in 2024 (AAPC, +2.4%; P=.56), while the Black-to-White disparity ratio widened from 19.3 to 24.4. Black non-Hispanic female youth had the highest female firearm homicide rate (6.03 per 100,000), exceeding male rates in 7 of 12 other demographic groups. American Indian or Alaska Native non-Hispanic male youth experienced the highest pooled firearm suicide rate (11.74 per 100,000). Firearm suicide increased significantly among Black non-Hispanic male youth (AAPC, +9.2%; 95% CI +3.2% to +15.4%; P=.01), with a substantial but nonsignificant upward trend among Black non-Hispanic female youth (AAPC, +11.7%; P=.08), while White non-Hispanic male rates remained stable (AAPC, −0.1%; P=.91). Consequently, the Black-to-White non-Hispanic male suicide rate ratio shifted from 0.67 in 2018 to 1.07 in 2024, marking the first time Black youth rates surpassed White youth rates in modern public health surveillance history.

Conclusions: Firearm homicide and suicide among US youth exhibit distinct demographic concentration patterns that are systematically obscured by aggregated public health surveillance. Disaggregated surveillance remains imperative to design targeted prevention strategies.

(JMIR Public Health Surveill 2026;12:e96931) doi:10.2196/96931

KEYWORDS

firearm mortality; youth; racial disparities; health equity; intersectionality; homicide; suicide; surveillance; epidemiology; CDC WONDER; Centers for Disease Control and Prevention

XSL•FO
RenderX

Introduction

Firearm injury has emerged as a leading cause of death among children, adolescents, and young adults aged 1 to 24 years in the United States [1], surpassing motor vehicle crashes, cancer, and all other causes of mortality in recent years [2]. Within this composite burden, firearm homicides and suicides function as 2 epidemiologically distinct phenomena [1]. While both public health crises share upstream structural and social determinants, they diverge in their sociodemographic concentrations, proximal mechanisms, case-fatality rates, geographic distribution, and intervention entry points [3-6]. Two decades of national surveillance have structured prevention science and federal investment around a seemingly stable race-by-intent dichotomy. This paradigm is characterized by disproportionately high firearm homicide rates documented among non-Hispanic Black youth and elevated firearm suicide mortality among non-Hispanic White and American Indian or Alaska Native (AI/AN) youth [1,7-12]. While this framing has clarified the population-level magnitude of firearm injury among these groups, it has also concealed the intersectional gradients within and across other aggregated populations.

National surveillance reports rarely stratify age, sex, race or ethnicity, and injury intent jointly. Where such multiway stratifications are produced in subnational data, strict federal data suppression policies systematically suppress cells containing fewer than 10 deaths to protect privacy [13]. At the national level, the fundamental constraint is not the technical capacity of the underlying surveillance infrastructure, which permits 4-way disaggregation via the Centers for Disease Control and Prevention (CDC) WONDER database, but the reporting conventions that govern what enters published epidemiological records. Consequently, contemporary shifts in firearm violence and injury risk levels among AI/AN, Native Hawaiian or Other Pacific Islander (NH/OPI), multiracial, and non-Hispanic Black adolescent female populations remain underrepresented in standard surveillance literature.

Recent intersectional analyses have started to address these gaps. Mariño-Ramírez et al [9] stratified pediatric firearm mortality by race and ethnicity across the years 1999 - 2020 and called for more granular reporting. Likewise, Lee et al [10] simultaneously stratified by age, sex, race, and ethnicity through 2023 and documented the intersections of firearm and motor vehicle mortality; while Klein-Cloud et al [13] examined racial and ethnic disparities through 2023, raising the data suppression and classification concerns examined further in the present analysis.

However, 3 critical limitations of the current literature remain. First, the 2018 CDC transition from a bridged 4-category race classification scheme to a 6-category single-race scheme has effectively bifurcated the literature into isolated pre-2020 and post-2018 analyses, foreclosing an examination of the longitudinal arc of youth firearm mortality across the classification transition. Additionally, emerging intersectional signals remain underexamined relative to their epidemiological significance; notably, the recent convergence of firearm suicide mortality between non-Hispanic Black and White male youth, and the elevated firearm homicide burden among non-Hispanic Black adolescent and young adult female youth. Finally, this analysis incorporates finalized 2024 mortality data, released by the National Center for Health Statistics (NCHS) in January 2026. This additional year allows examination of whether pandemic-era increases in youth firearm mortality continued, stabilized, or declined.

Intersectionality frameworks, as articulated by Crenshaw [14] and adapted for population health by Bowleg [15] and Bauer [16], emphasize that social positions related to race, gender, class, and other axes are occupied simultaneously rather than additively, producing patterns of risk that are not visible when each axis is examined independently or only through two-dimensional comparisons. Applied to injury epidemiology, an intersectional surveillance lens directs analytic attention to specific demographic groups whose risks may be obscured when surveillance data are stratified by only 1 or 2 characteristics at a time. For example, the firearm mortality burden among Black non-Hispanic adolescent girls or AI/AN non-Hispanic young adult males may be less visible in race-only or sex-only analyses. Although vital statistics cannot directly measure the structural processes that produce these inequities, intersectionally disaggregated surveillance can identify where mortality is concentrated and help guide subsequent etiologic, policy, and prevention efforts [16,17].

This study examines national youth firearm mortality in the United States from 1999 through 2024 using the CDC WONDER Underlying Cause of Death database. To accommodate the federal classification transition, we adopt a 2-period analytic design. Period 1 (1999 to 2020) draws on the bridged-race data to characterize the 25-year temporal arc, while period 2 (2018 to 2024) leverages the single-race dataset to estimate contemporary intersectional patterns at a higher demographic resolution. The overlapping years (2018 to 2020) function as an internal consistency check to verify rate stability across both data structures rather than as an analytic break point. This design contributes to the current literature as it situates youth firearm mortality within the first complete year of postpandemic vital statistics; it treats the bridged-to-single-race transition as an analytic feature rather than truncating estimation at a single classification era, and it isolates several critical, underexamined intersectional signals across race or ethnicity, sex, age, and injury intent.

Methods

Study Design and Data Source

We conducted a serial cross-sectional, population-based descriptive analysis of firearm deaths among US children, adolescents, and young adults aged 0 to 24 years between 1999 and 2024, using the CDC WONDER Underlying Cause of Death database. WONDER compiles death certificate records through the National Vital Statistics System and provides annual counts, population denominators, and crude rates by demographic and cause-of-death categories [18]. Population denominators were drawn from CDC WONDER bridged-race population estimates for period 1 (1999 to 2020) and single-race population estimates for period 2 (2018 to 2024). The analytic population was

XSL•FO
RenderX

restricted to US residents of the 50 states and the District of Columbia.

The current study population was defined as individuals aged 0 to 24 years, comprising early childhood (0 to 12 y), adolescence (13 to 17 y), and young adulthood (18 to 24 y), consistent with established developmental and public health frameworks [19,20]. This upper boundary extends beyond the 0-to-17-year range commonly used in pediatric firearm surveillance reports [2,7,21], which exclude young adults aged 18 to 24 years, a developmental demographic that bears a disproportionate share of the national firearm homicide burden, particularly among Black and Hispanic male youth [3,21].

The 25-year analytic window was bounded by historical and administrative considerations of data continuity. The lower bound of 1999 represents the first full calendar year in which the United States implemented the *ICD-10* (*International Statistical Classification of Diseases, Tenth Revision*) for underlying cause-of-death coding, and the first year for which CDC WONDER bridged-race population denominators are available [18,22]. The upper bound of 2024 incorporates the most recently finalized national mortality file released by the NCHS in January 2026 [18]. The window spans 25 years, from 1999 through 2024, and comprises 26 annual observations. This window enables the contextualization of contemporary, postpandemic shifts within a quarter-century historical trajectory.

Ethical Considerations

This analysis used deidentified, publicly available, aggregated vital statistics data and was therefore determined not to constitute human participants research as defined under 45 CFR (Code of Federal Regulations) §46.102. Institutional review board review was not required under institutional policy governing the secondary analysis of public-use, fully anonymized federal data, and no application was submitted. The CDC WONDER public-use data files contain no direct or indirect personal identifiers. NCHS suppresses cells with fewer than 10 deaths, including finely stratified national cells, to protect confidentiality. The present analysis used visible cells and explicitly identified suppressed and statistically unreliable estimates that satisfied federal analytic reliability thresholds, with unstable rates explicitly flagged. As the underlying records are fully anonymous and collected via routine statutory vital statistics registration, informed consent was not applicable.

Firearm Fatality Classification

Firearm fatalities were identified using *ICD-10* underlying cause-of-death codes. In this analysis, the term “intent” refers to the manner of death recorded on the death certificate, representing the standard NCHS classification of the circumstances surrounding the fatal injury. Five mutually exclusive intent categories were retained: firearm homicide (X93-X95); firearm suicide (X72-X74); unintentional firearm injury, defined as a fatal accidental discharge of a firearm rather than a category of homicide or suicide (W32-W34); firearm injury of undetermined intent (Y22-Y24); and legal intervention involving a firearm (Y35.0).

Firearm homicide and suicide served as the primary intent categories for all intersectional analyses. Unintentional, undetermined, and legal intervention deaths were retained in cumulative *all-cause* totals but were excluded from intent-specific trend estimation, given high suppression rates across multiway demographic and intersectional strata. Cumulative 25-year death totals were calculated by aggregating all period 1 deaths (1999‑2020) and adding unique deaths from the nonoverlapping years of period 2 (2021‑2024) to avoid duplication. All 2024 data reflect finalized counts released by CDC WONDER in January 2026 [18].

Two-Period Analytic Design

In data year 2021, the CDC fully transitioned its multiple-cause-of-death files from a bridged 4-category race classification scheme (White, Black, AI/AN, and Asian or Pacific Islander, with multiple-race decedents algorithmically assigned to a single category) to a 6-category single-race scheme that separately identifies Asian, NH/OPI, and multi-racial decedents [22]. The legacy bridged-race data preserves long-term trend comparability back to 1999 but obscures contemporary diversity, whereas the single-race file allowed for greater contemporary sociodemographic comparability (back to 2018) but cannot be extended to the prebridging era. As these 2 population denominator systems introduce a structural trend break, a single continuous analytic framework spanning 1999 to 2024 was computationally untenable.

Accordingly, race and ethnicity were combined into a single mutually exclusive variable for all analyses (eg, non-Hispanic Black and Hispanic). The overlapping years (2018 to 2020) were used as an internal validation window to verify rate consistency and mathematical stability across both classification methods. The descriptive simultaneous stratification by race or ethnicity, sex, age, and injury intent used in this study constitutes a surveillance application of intersectional theory to isolate overlooked risk concentrations, rather than a formal quantitative intersectional decomposition [16].

Statistical Analysis

Crude firearm death rates per 100,000 population were calculated for each demographic subgroup. Crude rather than age-adjusted rates were prioritized because simultaneous multiway stratification renders within-stratum age adjustment statistically volatile and analytically untenable for low-denominator populations, such as AI/AN, NH/OPI, and multiracial youth [9,21,23]. Disparity ratios were computed using non-Hispanic White males as the reference group for male comparisons and non-Hispanic White females for female comparisons; absolute crude rates were also compared directly across sex lines where female burdens eclipsed male baselines. Further, 95% CIs for rates and ratios were derived using the Byar approximation, which provides robust coverage for the small expected cell counts characteristic of fine-grained intersectional strata [24].

Cells based on fewer than 20 deaths were flagged as analytically unreliable per NCHS standards. Within the primary pooled-age contemporary intersectional matrix for period 2 (6 racial groups x 2 Hispanic-origin categories x 2 sexes x 3 primary intent

XSL•FO
RenderX

categories [homicide, suicide, and unintentional injury]), there were 72 distinct cells. Of these, 19 of 72 (26.4%) cells had counts fewer than 10 and were omitted from detailed stratification, and 11 of 72 (15.3%) cells were flagged as unreliable, yielding fully stable estimates for 42 of 72 (58.3%) baseline cells. Rate instability was heavily concentrated among NH/OPI youth (10/12, 83.3% cells) and within specific unintentional injury strata.

To characterize directional trends within period 2 (2018-2024), average annual percentage changes (AAPCs) were estimated by log-linear regression of the natural logarithm of the annual crude rate on calendar year. These were calculated separately for firearm homicide and suicide across non-Hispanic race-sex strata that demonstrated 3 or more reliable annual observations. The AAPC was calculated as $(\exp[\beta] - 1)\times 100$, where β is the estimated slope parameter. Further, 95% CIs were derived from the t-distribution of β with $n - 2$ degrees of freedom. Statistical significance was assessed using 2-sided *t* tests with an a priori alpha =.05. As this study's period encompasses a nonlinear post-2020 mortality surge and subsequent decline, the log-linear AAPC is explicitly interpreted as the net annualized trend velocity across the interval rather than a monotonic trajectory. All analyses were executed in Python (version 3.11; Python Software Foundation) using the Pandas, NumPy, SciPy, Matplotlib, and Seaborn libraries. This study adhered strictly to the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) reporting guideline; see Checklist 1.

Results

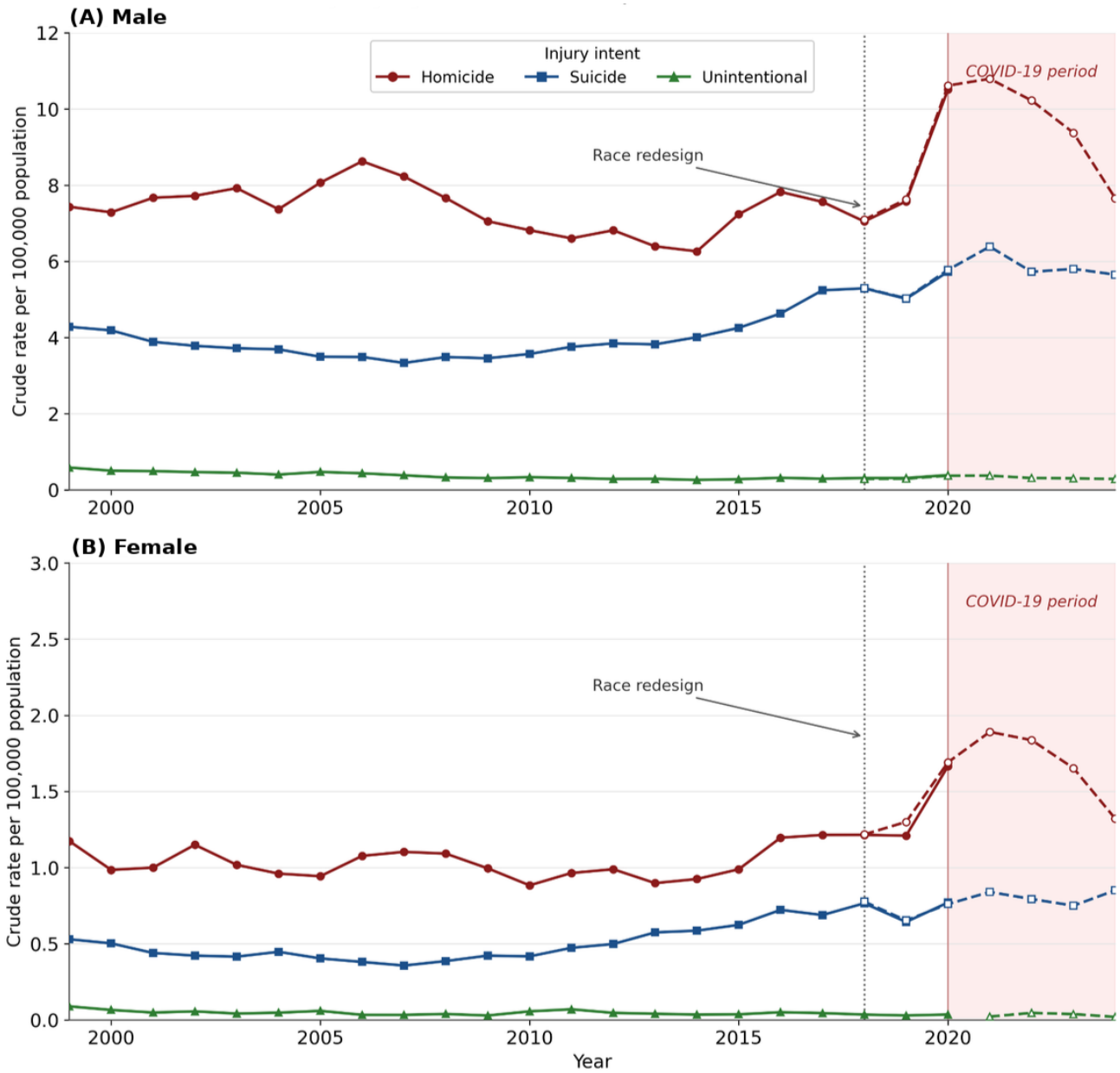
Cumulative Burden and Temporal Trends, 1999-2024

From 1999 through 2024, a total of 200,704 firearm deaths occurred among US youth and young adults aged 0 to 24 years (period 1: 161,313; period 2, nonoverlapping years 2021 - 2024: 39,391). Firearm mortality remained concentrated among male youth throughout the 25-year period, though the female share of the burden rose modestly, from 11.4% (18,459/161,313) of deaths in period 1 to 13% (8543/65,472) in period 2, even as the sex-specific trajectories diverged by intent (Figure 1). Across the full 25-year window, overall crude firearm mortality declined steadily from 1999 through 2013, reversed direction in 2014, and accelerated sharply during the pandemic period, rising 28.9% between 2019 and 2020 (from 7.69 to 9.91 per 100,000). Rates remained elevated through 2022 before partially stabilizing in 2023 and declining in 2024 (Figure 1). The recent 2024 overall rate of 8.24 per 100,000 represented a 12% decrease from 2023, the largest single-year decline observed during this 25-year period, though overall youth firearm mortality remained above the 2019 prepandemic baseline.

When examined by intent, firearm homicide and firearm suicide followed divergent trajectories within this aggregate trend. Among Black non-Hispanic male youth, homicide rates rose from 33.44 per 100,000 in period 1 to 44.03 per 100,000 in period 2, and Black non-Hispanic female homicide rates nearly doubled from 3.46 to 6.03 per 100,000 (Tables 1 and 2). Firearm suicide rates among AI/AN non-Hispanic male youth rose from 8.47 to 11.74 per 100,000 across the 2 periods, and Black non-Hispanic male suicide rates rose from 3.51 to 6.33 per 100,000. The 2 periods nearly coincide across the 2018 to 2020 overlap in Figure 1. The small residual differences reflect changes in population denominators between the bridged-race and single-race files rather than true changes in mortality; therefore, rates were not directly compared across periods.

XSL•FO
RenderX

Figure 1. Crude firearm mortality rates among US youth and young adults aged 0 to 24 years, by injury intent and sex, 1999‑2024. Homicide, suicide, and unintentional intent are shown as separate trend lines. The shaded region indicates the elevated mortality period of 2020‑2024. The vertical dashed line at 2018 marks the CDC transition from bridged to single-race classification; dotted lines across 2018‑2020 indicate overlapping years and do not represent a true mortality trend. Across the 2018 to 2020 overlap, the 2 periods nearly coincide, and any residual difference reflects the change in population denominators between the bridged-race and single-race files rather than a true change in mortality, so rates from the two periods are not directly compared. Source: CDC WONDER [25]. CDC: Centers for Disease Control and Prevention.



XSL•FO
RenderX

Table . Crude firearm mortality rates per 100,000 by race, Hispanic origin, sex, and injury intent, US youth ages 0 - 24 years, and years 1999 - 2020 (bridged-race classification). CIs were calculated using the Byar approximation.

Race/ethnicity group	Sex	Homicide rate (95% CI)	Suicide rate (95% CI)	Unintentional rate (95% CI)
Black or African American				
Non-Hispanic	Female	3.46 (3.37 - 3.55)	0.36 (0.33 - 0.39)	0.09 (0.07 - 0.10)
Non-Hispanic	Male	33.44 (33.16 - 33.72)	3.51 (3.42 - 3.60)	0.73 (0.69 - 0.77)
Hispanic	Female	0.70 (0.56 - 0.86)	Unreliable[a]	Suppressed[b]
Hispanic	Male	4.49 (4.14 - 4.85)	0.92 (0.76 - 1.08)	Unreliable
White				
Non-Hispanic	Female	0.58 (0.56 - 0.60)	0.67 (0.65 - 0.69)	0.04 (0.04 - 0.05)
Non-Hispanic	Male	1.59 (1.56 - 1.62)	5.33 (5.28 - 5.39)	0.35 (0.34 - 0.37)
Hispanic	Female	0.95 (0.91 - 0.99)	0.28 (0.26 - 0.30)	0.03 (0.02 - 0.03)
Hispanic	Male	8.21 (8.09 - 8.33)	2.39 (2.33 - 2.46)	0.25 (0.23 - 0.27)
American Indian or Alaska Native				
Non-Hispanic	Female	1.28 (1.07 - 1.49)	1.08 (0.89 - 1.27)	Unreliable
Non-Hispanic	Male	6.19 (5.74 - 6.65)	8.47 (7.93 - 9.01)	0.87 (0.70 - 1.04)
Hispanic	Female	Unreliable	Unreliable	Suppressed
Hispanic	Male	Suppressed	Suppressed	Suppressed
Asian or Pacific Islander				
Non-Hispanic	Female	0.39 (0.34 - 0.44)	0.29 (0.25 - 0.33)	Suppressed
Non-Hispanic	Male	2.03 (1.91 - 2.14)	1.74 (1.63 - 1.84)	0.09 (0.07 - 0.12)
Hispanic	Female	Suppressed	Unreliable	Suppressed
Hispanic	Male	1.59 (1.24 - 2.01)	0.94 (0.68 - 1.27)	Suppressed

[a]Unreliable: fewer than 20 deaths; interpret with caution. Source: CDC (Centers for Disease Control and Prevention) WONDER [25].

[b]Suppressed: fewer than 10 deaths. Source: CDC (Centers for Disease Control and Prevention) WONDER [25].

XSL•FO
RenderX

Table . Crude firearm mortality rates per 100,000 population pooled 2018-2024, 6-category single-race classification per 1997 OMB[a] standards. CIs were calculated using the Byar approximation.

Race/ethnicity group	Sex	Homicide rate (95% CI)	Suicide rate (95% CI)	Unintentional rate (95% CI)
Black or African American				
Non-Hispanic	Female	6.03 (5.82 - 6.25)	0.87 (0.79 - 0.95)	0.14 (0.10 - 0.17)
Non-Hispanic	Male	44.03 (43.46 - 44.61)	6.33 (6.11 - 6.55)	0.86 (0.78 - 0.94)
Hispanic	Female	1.03 (0.77 - 1.35)	0.35 (0.21 - 0.55)	Suppressed[b]
Hispanic	Male	6.71 (6.01 - 7.40)	2.30 (1.89 - 2.70)	0.21 (0.10 - 0.37)
White				
Non-Hispanic	Female	0.65 (0.61 - 0.69)	0.95 (0.90 - 0.99)	0.04 (0.03 - 0.05)
Non-Hispanic	Male	1.85 (1.79 - 1.92)	7.01 (6.89 - 7.13)	0.30 (0.27 - 0.32)
Hispanic	Female	1.20 (1.13 - 1.28)	0.50 (0.45 - 0.55)	0.02 (0.01 - 0.03)
Hispanic	Male	7.33 (7.14 - 7.52)	3.70 (3.57 - 3.84)	0.22 (0.19 - 0.26)
American Indian or Alaska Native				
Non-Hispanic	Female	2.08 (1.59 - 2.68)	1.77 (1.32 - 2.33)	Suppressed
Non-Hispanic	Male	9.29 (8.20 - 10.38)	11.74 (10.51 - 12.97)	1.14 (0.79 - 1.59)
Hispanic	Female	0.40 (0.21 - 0.71)	0.40 (0.21 - 0.71)	Suppressed
Hispanic	Male	1.93 (1.47 - 2.49)	1.61 (1.19 - 2.12)	Suppressed
Asian				
Non-Hispanic	Female	0.26 (0.19 - 0.34)	0.43 (0.35 - 0.53)	Suppressed
Non-Hispanic	Male	1.11 (0.97 - 1.26)	2.41 (2.20 - 2.63)	0.06 (0.03 - 0.11)
Hispanic	Female	Suppressed	Suppressed	Suppressed
Hispanic	Male	1.63 (0.96 - 2.57)	2.17 (1.39 - 3.23)	Suppressed
Native Hawaiian or Other Pacific Islander				
Non-Hispanic	Female	Suppressed	1.35 (0.65 - 2.48)	Suppressed
Non-Hispanic	Male	8.34 (6.44 - 10.64)	5.13 (3.67 - 6.99)	Suppressed
Hispanic	Female	Suppressed	Suppressed	Suppressed
Hispanic	Male	2.62 (1.25 - 4.81)	3.14 (1.62 - 5.48)	Suppressed
More than 1 race				
Non-Hispanic	Female	1.02 (0.86 - 1.18)	0.50 (0.39 - 0.63)	Suppressed
Non-Hispanic	Male	3.40 (3.12 - 3.69)	3.28 (3.00 - 3.57)	0.15 (0.09 - 0.22)
Hispanic	Female	0.73 (0.48 - 1.05)	0.49 (0.30 - 0.77)	Suppressed
Hispanic	Male	3.22 (2.66 - 3.77)	2.25 (1.81 - 2.76)	Suppressed

[a]OMB: Office of Management and Budget.

[b]Suppressed: fewer than 10 deaths. Source: CDC (Centers for Disease Control and Prevention) WONDER [25].

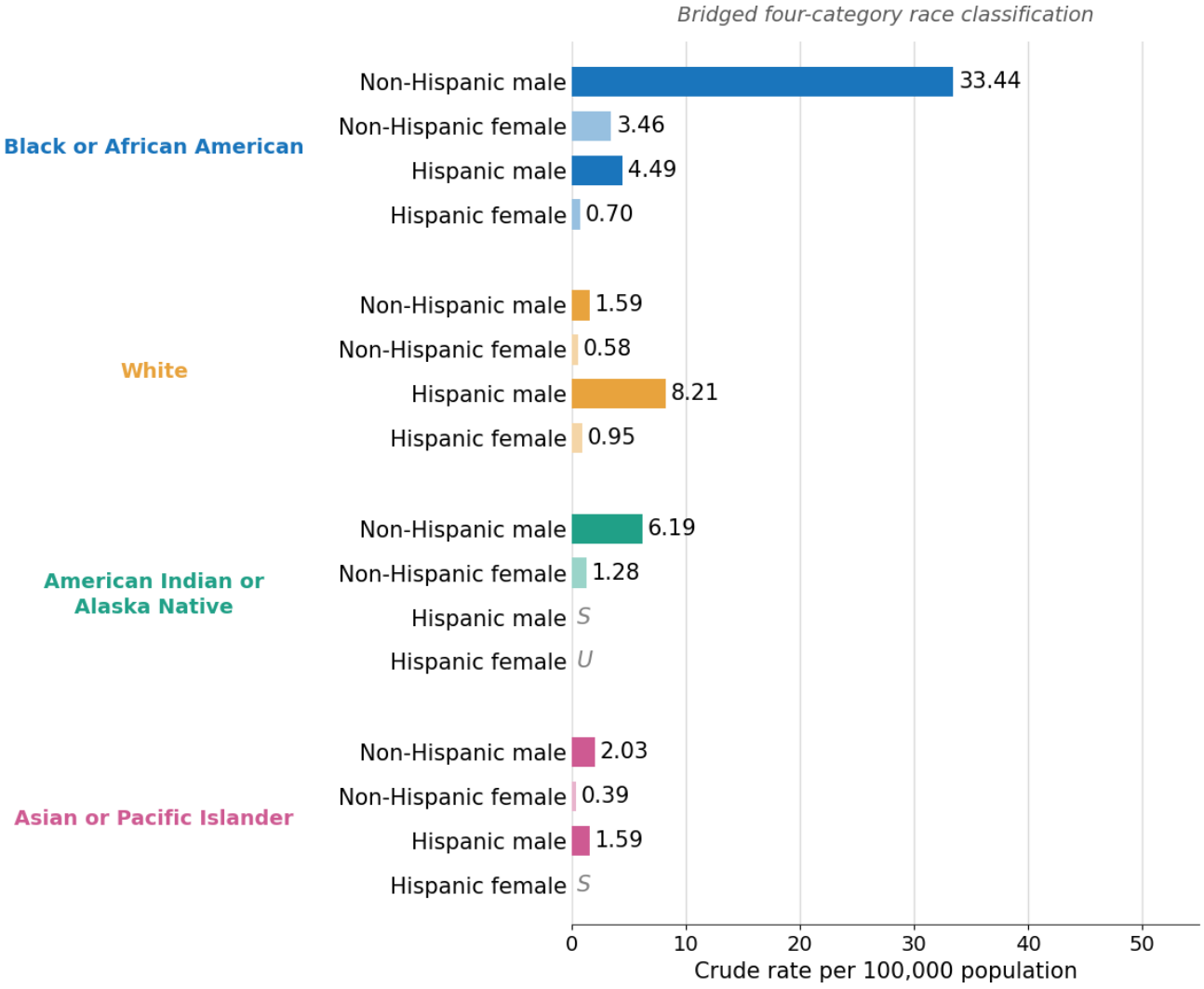
Firearm Homicide Rates by Race or Ethnicity and Sex, 2018-2024

Male Homicides

During period 2, Black non-Hispanic male youth experienced the highest firearm homicide rate of any group examined (44.03 per 100,000; 95% CI 43.46 - 44.61), which was 23.8 times the rate among White non-Hispanic male youth (1.85 per 100,000; 95% CI 1.79 - 1.92; Figures 2-5; Table 2). AI/AN non-Hispanic male youth exhibited the second-highest rate (9.29 per 100,000; 95% CI 8.20 - 10.38), followed by NH/OPI non-Hispanic males (8.34 per 100,000; 95% CI 6.44 - 10.64). Among Hispanic-origin male strata, White Hispanic males (7.33 per 100,000; 95% CI 7.14 - 7.52) and Black Hispanic males (6.71 per 100,000; 95% CI 6.01 - 7.40) carried the highest homicide burdens.

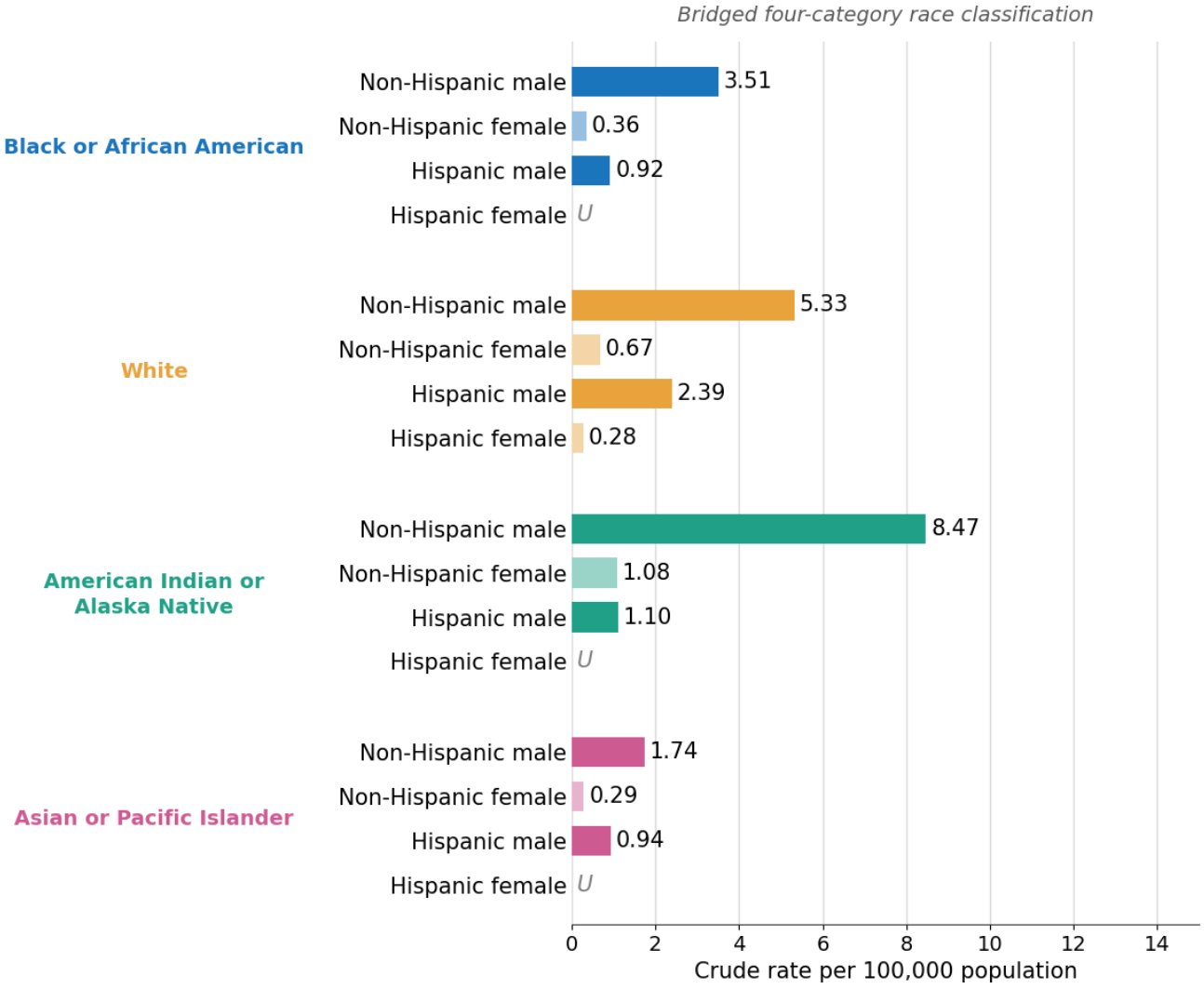
XSL•FO
RenderX

Figure 2. Crude firearm homicide rates among US youth and young adults aged 0 to 24 years, by race, Hispanic ethnicity, and sex, using the bridged 4-category race classification for 1999 to 2020. S indicates suppressed data because the cell contained fewer than 10 deaths; U indicates unreliable data because the estimate was based on fewer than 20 deaths. Source: CDC WONDER [25]. CDEC: Centers for Disease Control and Prevention.



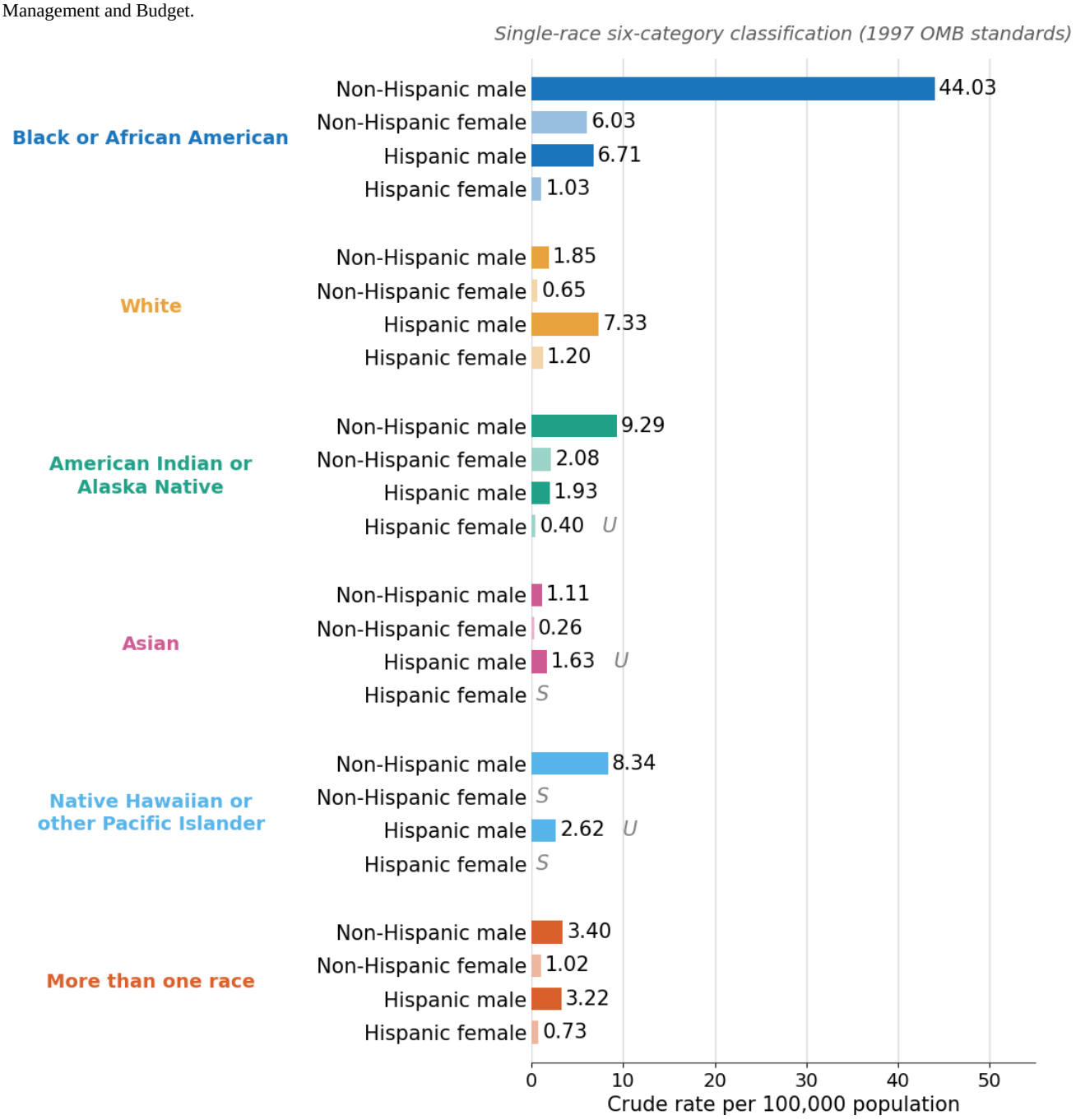
XSL•FO
RenderX

Figure 3. Crude firearm suicide rates among US youth and young adults aged 0 to 24 years, by race, Hispanic ethnicity, and sex, using the bridged 4-category race classification for 1999 to 2020. S indicates suppressed data because the cell contained fewer than 10 deaths; U indicates unreliable data because the estimate was based on fewer than 20 deaths. Source: CDC WONDER [25]. CDC: Centers for Disease Control and Prevention.



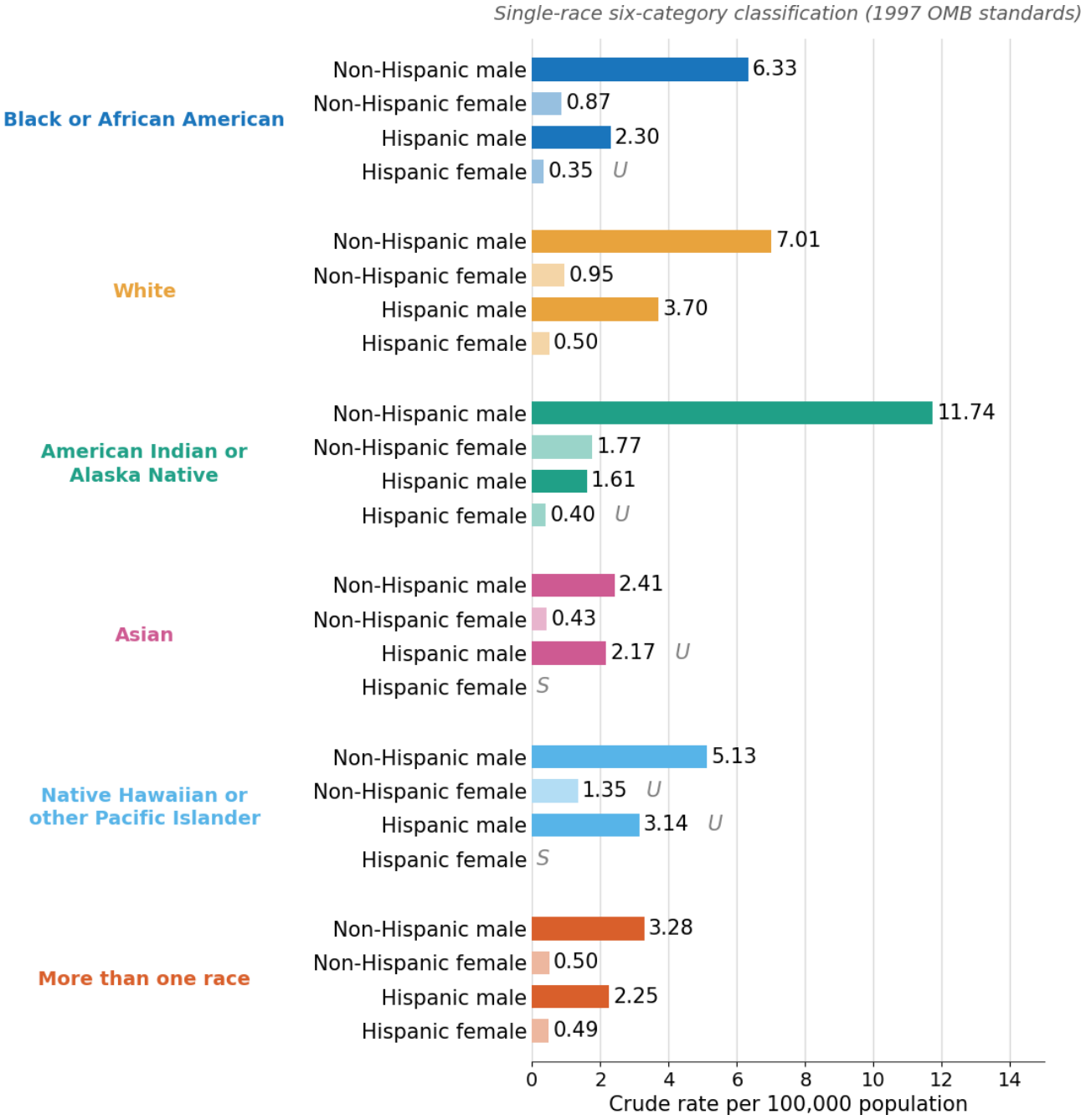
XSL•FO
RenderX

Figure 4. Crude firearm homicide rates among US youth and young adults aged 0 to 24 years, by race, Hispanic ethnicity, and sex, using the single-race 6-category classification for 2018 to 2024. S indicates suppressed data because the cell contained fewer than 10 deaths; U indicates unreliable data because the estimate was based on fewer than 20 deaths. Source: CDC WONDER [25]. CDC: Centers for Disease Control and Prevention; OMB: Office of Management and Budget.



XSL•FO
RenderX

Figure 5. Crude firearm suicide rates among US youth and young adults aged 0 to 24 years, by race, Hispanic ethnicity, and sex, using the single-race 6-category classification for 2018 to 2024. S indicates suppressed data because the cell contained fewer than 10 deaths; U indicates unreliable data because the estimate was based on fewer than 20 deaths. Source: CDC WONDER [25]. CDC : Centers for Disease Control and Prevention.



Female Homicides

Black non-Hispanic female youth had the highest firearm homicide rate of any female stratum (6.03 per 100,000; 95% CI 5.82 - 6.25), 9.3 times the rate among White non-Hispanic female youth (0.65 per 100,000; 95% CI 0.61 - 0.69). Notably, this rate exceeded the homicide rates observed in male youth across 7 of the 12 other male groups examined, a pattern typically obscured by sex-aggregated reporting or by race-only stratifications that pool data for Black non-Hispanic girls with other female populations.

Firearm Suicide Rates by Race or Ethnicity and Sex, 2018-2024

Male Suicides

The demographic distribution of firearm suicide differed substantively from that of homicide. AI/AN non-Hispanic male youth experienced the highest firearm suicide rate of any group (11.74 per 100,000; 95% CI 10.51 - 12.97), 1.7 times the White non-Hispanic male rate (7.01 per 100,000; 95% CI 6.89 - 7.13; Table 2). Black non-Hispanic male youth also showed an elevated rate (6.33 per 100,000; 95% CI 6.11 - 6.55), placing them above NH/OPI non-Hispanic males (5.13 per 100,000; 95% CI 3.67 - 6.99) and below the AI/AN and White non-Hispanic strata.

XSL•FO
RenderX

Female Suicides

Among female youth, AI/AN non-Hispanic females had the highest firearm suicide rate (1.77 per 100,000; 95% CI 1.32 - 2.33), followed by NH/OPI non-Hispanic females (1.35 per 100,000; 95% CI 0.65 - 2.49, interpret with caution) and White non-Hispanic females (0.95 per 100,000; 95% CI 0.90 - 0.99). Black non-Hispanic female firearm suicide rates (0.87 per 100,000; 95% CI 0.79 - 0.95), while lower in absolute terms than rates for AI/AN and White non-Hispanic females, showed the steepest within-period relative increase of any female stratum examined (see AAPC analysis in the section below titled "AAPCs in Firearm Homicide and Suicide Rates, 2018-2024").

AAPCs in Firearm Homicide and Suicide Rates, 2018-2024

Overview

AAPCs were estimated for all non-Hispanic race-sex strata with at least 3 nonsuppressed annual observations between 2018 and 2024 (Table 3), separately for firearm homicide and suicide.

XSL•FO
RenderX

Table . AAPCs[a,b] in crude firearm homicide and suicide mortality rates among US youth and young adults aged 0 to 24 years, by race or ethnicity and sex, non-Hispanic strata, 2018 - 2024.

Race/ethnicity group	Sex	Years observed	AAPC % (95% CI)	*P* value	Annual crude rates per 100,000 across years (2018 - 2024)
Firearm homicide					
Black NH[c]	Male	7	+2.4% (−7.2 to +13.0)	.56	33.28; 37.10; 52.34; 53.75; 49.74; 45.08; 37.06
Black NH	Female	7	+4.7% (−6.2 to +16.8)	.33	4.35; 4.60; 6.55; 8.10; 7.11; 6.21; 5.31
White NH	Male	7	−1.5% (−8.3 to +5.8)	.61	1.72; 1.68; 2.33; 1.99; 1.93; 1.80; 1.52
White NH	Female	7	+0.1% (−5.3 to +5.8)	.96	0.61; 0.58; 0.77; 0.64; 0.72; 0.64; 0.59
AI/AN[d] NH	Male	7	+6.4% (−1.0 to +14.3)	.08	7.65; 8.59; 9.07; 6.96; 11.47; 10.70; 10.87
AI/AN NH	Female	4	+10.2% (−9.9 to +34.7)	.17	2.40; 2.40; 2.48; 3.28
Asian NH	Male	7	−2.7% (−14.5 to +10.7)	.61	0.94; 1.17; 1.29; 1.28; 1.01; 1.44; 0.69
Asian NH	Female	__[e]	—	—	Insufficient data
NH/OPI[f] NH	Male	5	−3.0% (−8.3 to +2.7)	.19	10.37; 9.46; 10.74; 8.87; 8.41
NH/OPI NH	Female	—	—	—	No data
Multiracial NH	Male	7	+5.8% (−2.9 to +15.2)	.15	2.38; 2.92; 3.21; 4.44; 3.89; 3.58; 3.30
Multiracial NH	Female	7	+5.8% (−6.0 to +19.1)	.27	0.84; 0.68; 0.94; 1.51; 0.99; 1.20; 0.96
Firearm suicide					
Black NH	Male	7	+9.2% (+3.2 to +15.4)	.01*	4.68; 4.51; 6.05; 7.28; 7.25; 7.24; 7.28
Black NH	Female	7	+11.7% (−1.7 to +27.0)	.08	0.62; 0.51; 0.68; 1.22; 1.17; 0.81; 1.07
White NH	Male	7	−0.1% (−3.4 to +3.2)	.91	6.98; 6.54; 7.20; 7.92; 6.79; 6.86; 6.80
White NH	Female	7	−0.1% (−3.4 to +3.4)	.95	1.02; 0.88; 0.95; 0.98; 0.86; 1.00; 0.96
AI/AN NH	Male	7	+5.6% (−5.2 to +17.6)	.25	9.01; 8.36; 14.41; 13.23; 14.10; 9.73; 13.59
AI/AN NH	Female	—	—	—	Insufficient data
Asian NH	Male	7	+2.4% (−4.2 to +9.6)	.40	2.02; 2.26; 2.27; 2.98; 2.43; 2.75; 2.17
Asian NH	Female	6	+0.5% (−10.0 to +12.3)	.90	0.48; 0.43; 0.43; 0.42; 0.35; 0.59
NH/OPI NH	Male	—	—	—	No data
NH/OPI NH	Female	—	—	—	No data
Multiracial NH	Male	7	+2.6% (−2.1 to +7.5)	.22	2.86; 3.53; 2.94; 3.52; 3.02; 3.36; 3.72
Multiracial NH	Female	5	+4.0% (−2.5 to +10.9)	.15	0.49; 0.51; 0.46; 0.58; 0.61

[a]AAPC: average annual percentage change.

XSL•FO
RenderX

[b]Average annual percentage changes were estimated using log-linear regression of the natural logarithm of the annual crude rate on calendar year.

[c]NH: non-Hispanic.

[d]AI/AN: American Indian or Alaska Native.

[e]Not available. Insufficient nonsuppressed annual observations for estimation.

[f]NH/OPI: Native Hawaiian or Other Pacific Islander.

Firearm Homicides

No non-Hispanic race-sex stratum showed a statistically significant linear trend in firearm homicide across 2018 - 2024, reflecting the nonmonotonic trajectories across this window. Among Black non-Hispanic male youth, annual rates rose from 33.28 per 100,000 in 2018 to a peak of 53.75 in 2021, then declined progressively to 37.06 in 2024 (AAPC +2.4%; 95% CI −7.2 to +13.0; *P*=.56). Similar nonmonotonic trajectories appeared among Black non-Hispanic female youth (AAPC +4.7%; *P*=.33) and AI/AN non-Hispanic male youth (AAPC +6.4%; *P*=.08), with each group returning toward prepandemic levels by 2024. White non-Hispanic male rates remained low and stable (AAPC −1.5%; *P*=.61). Despite the relative reversion among the burdened groups, the Black-to-White non-Hispanic male homicide rate ratio increased modestly from 19.3 in 2018 to 24.4 in 2024, indicating that the partial reversion in relative rates was concentrated in the reference group rather than in the Black non-Hispanic male group, and that the relative disparity widened rather than narrowed across the period (Table 3).

Firearm Suicides

In contrast to homicide, firearm suicide showed sustained directional trends concentrated in Black non-Hispanic strata. Rates among Black non-Hispanic male youth increased significantly at an AAPC of +9.2% per year (95% CI +3.2 to +15.4; *P*=.01), shifting the Black-to-White non-Hispanic male suicide rate ratio from 0.67 in 2018 to 1.07 in 2024, marking the first year in the 2018 to 2024 series in which the firearm suicide rate among Black non-Hispanic male youth exceeded that among White non-Hispanic male youth. Black non-Hispanic female youth showed a comparable rise in magnitude (AAPC +11.7%; 95% CI −1.7 to +27.0; *P*=.08), falling just outside the conventional threshold for statistical significance and constituting a notable trend warranting close epidemiological surveillance despite the wide CI. White non-Hispanic male and female rates were stable across the same period (AAPC −0.1%; *P*=.91 and *P*=.95, respectively). AI/AN non-Hispanic male suicide rates showed a nonsignificant increasing trend (AAPC +5.6%; 95% CI −5.2 to +17.6; *P*=.25), with substantial year-to-year variability reflecting small population denominators (Table 3).

Age-Specific Firearm Mortality Patterns

Within period 2, age-specific analysis revealed steep developmental gradients concentrated in 2 intersectional locations (Multimedia Appendix 1). Among Black non-Hispanic male youth, firearm homicide rates increased from 6.63 per 100,000 at ages 10 to 14 years to 86.06 at ages 15 to 19 years and 114.44 at ages 20 to 24 years. The Black-to-White non-Hispanic male homicide rate ratio widened across consecutive age bands, from 26.1 at ages 15 to 19 years to 27 at ages 20 to 24 years, indicating that the greatest absolute firearm homicide burden occurred during late adolescence and young adulthood. This pattern challenges prevention frameworks that define youth as ending at age 17 or 18 because they may exclude young adults aged 18 to 24, including the age group with the highest observed firearm homicide rate.

Among AI/AN non-Hispanic male youth, firearm suicide rates increased from 2.92 per 100,000 at ages 10 to 14 years to 20.16 at ages 15 to 19 years and 31.71 at ages 20 to 24 years. In the oldest age band, the AI/AN non-Hispanic male firearm suicide rate was 1.6 times the corresponding White non-Hispanic male rate (19.61 per 100,000). The parallel adolescent-to-young-adult escalations observed for AI/AN male firearm suicide and Black non-Hispanic male firearm homicide underscore the need for prevention frameworks that extend beyond adolescence into the 18-to-24-year developmental window.

Among female youth, age gradients were less pronounced for homicide but evident for suicide. White non-Hispanic female firearm suicide rates increased from 0.47 per 100,000 at ages 10 to 14 years to 2.40 at ages 20 to 24 years. Black non-Hispanic female firearm suicide rates showed a similar but shallower gradient, increasing from 0.52 to 2.44 per 100,000 across the same age bands. In contrast to the pattern among males, Black and White non-Hispanic female firearm suicide rates across age groups were statistically indistinguishable, with overlapping CIs at ages 10 to 14 and at ages 20 to 24 years, warranting continued monitoring in future surveillance cycles as an indicator of the shifting racial distribution of firearm suicide risk.

Discussion

Principal Findings

This 25-year national analysis advances 3 critical findings that aggregate surveillance routinely conceals. Primarily, firearm homicide remains concentrated with extreme disproportionality among Black non-Hispanic male youth and, to a degree not captured by race-only or sex-aggregated reporting, among Black non-Hispanic girls. Annual firearm homicide rates among Black non-Hispanic male youth rose sharply during 2020 and 2021 and substantially reverted by 2024, with the 2024 rate approaching the 2018 level. However, the disparity ratio relative to White non-Hispanic male youth did not narrow across this period. The Black-to-White non-Hispanic male homicide rate ratio increased from 19.3 in 2018 to 24.4 in 2024, indicating that improvement in the absolute homicide rate among Black non-Hispanic males was not accompanied by a reduction in relative inequality.

In addition, firearm suicide burden was highest among AI/AN non-Hispanic male youth and intensified with age, peaking at ages 20 to 24 years. Furthermore, firearm suicide rates among Black non-Hispanic male youth increased at a statistically significant rate of +9.2% per year, with a comparable-magnitude

XSL•FO
RenderX

rise among Black non-Hispanic female youth and stable rates among White non-Hispanic male and female youth. Consequently, the Black-to-White non-Hispanic male suicide rate ratio shifted from 0.67 in 2018 to 1.07 in 2024, marking the first year in recent surveillance history that Black non-Hispanic male rates exceeded those of White non-Hispanic male youth. Our findings show that firearm mortality among US youth represents 2 epidemiologically distinct phenomena that require intent-specific, intersectionally disaggregated surveillance and prevention paradigms.

Comparison With Prior Work

Firearm Homicide

The firearm homicide rate among Black non-Hispanic male youth (44.03 per 100,000), 23.8 times that of White non-Hispanic male youth, is consistent with prior evidence linking elevated community violence exposure to structural conditions, not limited to residential segregation, concentrated economic disinvestment, and limited access to opportunity in the particular neighborhoods where these conditions co-occur, rather than to membership in any racial or ethnic group [5,26]. Crucially, this disparity ratio widened further with age, reaching 27.0 per 100,000 at ages 20 to 24 years, indicating that the transition from adolescence to young adulthood represents a critical period of escalating rather than declining risk. This age gradient has direct programmatic implications, as prevention frameworks that treat individuals aged 18 years and older as outside the scope of youth-focused intervention may be systematically misaligned and disconnected from the life stage at which absolute homicide risk peaks. Community violence intervention programs, hospital-based violence interruption, and structural investments in housing stability and economic opportunity represent vital avenues for reducing this burden [27]. However, young adults aged 18 to 24 may therefore fall into a prevention and service gap, aging out of pediatric and adolescent systems without a corresponding pathway into adult violence-prevention services unless they experience a violent injury, encounter the criminal legal system, or live in a community served by a dedicated intervention program.

Notably, the firearm homicide rate among Black non-Hispanic female youth (6.03 per 100,000) exceeded that of male youth in 7 of the 12 other racial and ethnic groups. This pattern may be obscured in surveillance reports that aggregate firearm mortality by race and ethnicity without simultaneous stratification by sex [6,8], underscoring the need to explicitly integrate intersectional metrics into clinical and community risk assessment, prevention programming, and policy.

Firearm Suicide

AI/AN non-Hispanic male youth had the highest firearm suicide rate of any group (11.74 per 100,000), 1.7 times that of White non-Hispanic males. The absolute burden was greatest among those aged 20 to 24 years, whose firearm suicide rate reached 31.71 per 100,000. Lethal-means safety counseling and secure storage legislation are associated with reduced firearm suicide risk [28], but universal approaches are likely insufficient for AI/AN youth without culturally grounded multitiered strategies developed with (and led by) AI/AN communities. AI/AN youth live across a wide range of contexts, including reservations, rural nonreservations, urban settings, among others, and effective approaches must reflect these distinct community settings. Tribally led prevention partnerships and collaborations with Urban Indian Health Program partnerships offer important infrastructure for addressing community-defined priorities, including limited behavioral health access, the legacy of historical and ongoing structural trauma, and unequal exposure to firearm-injury risk environments [29]. The high year-to-year rate volatility observed in the AI/AN series further indicates that standard surveillance infrastructure is poorly suited to monitoring risk in numerically small populations with elevated burden, and supports the use of pooled estimates, confidence intervals, and surveillance approaches designed for small populations.

The convergence of Black and White non-Hispanic male firearm suicide rates between 2018 and 2024 represents an emerging disparity signal with direct relevance to prevention planning. As firearm homicide and suicide follow distinct demographic concentration patterns, disaggregating by intent is a prerequisite for prevention systems that can match intervention resources to need [2-5,8,9]. The 2024 overall decline in firearm mortality, the largest single-year decline in the current study period, is encouraging, although rates remained elevated relative to pre-2014 levels, and the trajectories among several subgroups warrant continued monitoring.

Intersectional Interpretation

Three findings illustrate why an intersectional multi-year lens, rather than race-only or sex-only stratification, is necessary to characterize the firearm mortality burden in this age group. Primarily, the firearm homicide rate among Black non-Hispanic female youth (6.03 per 100,000) exceeded male homicide rates in 7 of 12 other racial and ethnic male subgroups examined, a result that is invisible in sex-aggregated reporting and in race-only reporting that pools Black non-Hispanic girls with girls of all other groups. This is the empirical pattern that intersectional theory predicts, noting that risk concentrates at specific social locations defined by the simultaneous operation of multiple structural axes [14-16]. Moreover, the elevated firearm suicide rate among AI/AN non-Hispanic young adult males (31.71 per 100,000 at ages 20 to 24 y) similarly emerges only when race or ethnicity, sex, age, and intent are considered together, reflecting the joint operation of historical, geographic, and behavioral health system exposures rather than any single attribute [29].

In addition, the convergence of Black and White non-Hispanic male firearm suicide rates between 2018 and 2024 demonstrates that intersectional locations are not static. The rate ratio shifted from 0.67 to 1.07 over 6 years, indicating that the demographic distribution of firearm suicide risk is itself in motion. This shift is reinforced by a parallel rise among Black non-Hispanic female youth (AAPC +11.7%; 95% CI −1.7 to +27.0), which approaches but does not reach the conventional threshold for statistical significance and which is similarly invisible in sex-aggregated or race-only reporting. The sex-symmetric character of the Black non-Hispanic firearm suicide rise is precisely the kind of dynamic that intersectional surveillance

XSL•FO
RenderX

can detect and that aggregated reporting cannot. For prevention planning, these findings imply that resources designed around the historical White male suicide profile or the historical Black male homicide profile cannot be assumed to reach the youth populations now bearing the greatest and most rapidly changing burden. Intersectional surveillance must be matched with intersectional intervention design, developed with the affected communities [15,16].

Surveillance Implications

Several populations with elevated burden in this analysis, including Black non-Hispanic girls, AI/AN youth, and NH/OPI adolescents, remain difficult to characterize reliably in national datasets because of small population denominators, historical racial category aggregation, misclassification of injury intent, and suppression thresholds that remove cells with fewer than 10 deaths [13,30]. Of the 72 possible intersectional cells in period 2, only 42 (58.3%) yielded stable estimates. NH/OPI youth were the most severely affected by suppression, with 10 of 12 (83.3%) cells either suppressed or unreliable, meaning that a population group with documented elevated homicide burden is effectively invisible in standard surveillance outputs. This gap reflects a systemic shortcoming in national mortality surveillance, particularly for populations facing the highest levels of risk.

Addressing these gaps requires action at multiple levels. Vital statistics systems should improve the completeness and accuracy of race and ethnicity classification, particularly for AI/AN, Hispanic, Asian, and multiracial decedents. In addition, supporting tribally led and tribally governed data-linkage initiatives, developed through sovereign-to-sovereign partnerships among Tribal Nations, Tribal Epidemiology Centers, Urban Indian Organizations, the Indian Health Service, and state vital statistics agencies, may help reduce racial misclassification in mortality records and improve identification of AI/AN deaths. Such initiatives must proceed under tribal authority, community-defined governance, formal data-use agreements, and appropriate tribal review, consistent with Indigenous Data Sovereignty and CARE principles (Collective benefit, Authority to control, Responsibility, Ethics). Because linkage primarily improves numerator accuracy rather than the size or stability of the underlying population denominator, it should be paired with multiyear estimates, transparent uncertainty reporting, and improved population-estimation methods for small and geographically diverse AI/AN populations. Linkage cannot, however, resolve the small-denominator problem itself, as the instability of rates in numerically small populations is a function of population size rather than of counting error.

Tribal data sovereignty frameworks that enable AI/AN communities to generate, govern, and analyze their own public health data could also improve the capacity to detect and respond to concentrated burden [30]. Finally, the 2018 race-classification transition, while methodologically necessary, created a discontinuity that has constrained intersectional analyses for nearly a decade. Prospective planning for future classification updates should include explicit transition protocols that preserve analytic continuity (e.g., bridging algorithms), rather than requiring investigators to construct post hoc bridging designs.

National vital statistics surveillance is not the only level at which firearm injury can be counted or addressed. Community-based surveillance systems, including hospital trauma registries, school-based injury reporting, faith-based and neighborhood-level violence interruption networks, and tribal injury surveillance partnerships, may generate higher-resolution and more contextually grounded data than national mortality files in populations where suppression and small denominators limit the utility of vital statistics; the White Mountain Apache surveillance and prevention system offers one such model [29]. Pairing these systems with community-based interventions [27], such as hospital-linked violence interruption programs, school-based behavioral health, and lethal-means counseling [28], could allow detection and response to be matched at the scale where families and providers actually encounter risk, moving prevention science well beyond universal screening and counseling [27,29]. These findings indicate which populations bear concentrated burden, and do not posit that descriptive mortality data can adjudicate these issues; that work requires evaluation designs this analysis does not provide.

Limitations

This analysis has several limitations. CDC WONDER suppresses strata with fewer than 10 deaths, which limits estimation for some intersectional subgroups. Small denominators among AI/AN, NH/OPI, and multiracial youth increase rate instability. The 2018 race-classification transition prevents direct cross-period comparisons. Additionally, crude rather than age-adjusted rates were used to preserve interpretability within fine-grained strata but may limit comparability with age-adjusted estimates reported elsewhere. AAPC estimates rest on a 7-year annual series (2018‑2024); the short window precludes reliable detection of trend inflection points, and log-linear regression may understate nonmonotonic dynamics such as the pandemic-era homicide surge and subsequent partial reversion. Our use of AAPC assumes a linear trajectory across the 2018‑2024 period. However, firearm homicides in heavily burdened strata followed a distinctly nonmonotonic path (a sharp pandemic-era spike followed by a post-2022 decline). The resulting nonsignificant *P* values reflect this U-shaped deviation rather than a lack of meaningful year-to-year shifts. Future analyses with longer longitudinal windows should use nonlinear modeling to capture these specific inflection points. AAPCs should be interpreted as directional rather than definitive. Our descriptive design does not account for socioeconomic conditions, urbanicity, or state-level firearm violence policies and precludes causal inference. Sex was determined from death certificate records, which did not include gender identity data. Consequently, the analysis could not examine firearm mortality among transgender, nonbinary, and other gender-diverse populations at the intersections of race and ethnicity, age, and gender. Intent and race or ethnicity classification on death certificates are similarly imperfect, particularly for AI/AN, Hispanic, and Asian American or Pacific Islander decedents [13,30]. Geographic variation was not examined, though prior work demonstrates substantial heterogeneity [2,10].

XSL•FO
RenderX

Conclusions

Twenty-five years of national mortality data show that firearm mortality among US youth follows distinct, intent-specific patterns across race or ethnicity, sex, and age that aggregate surveillance routinely obscures. The post-2020 surge in firearm homicide among Black non-Hispanic male youth has reverted by 2024, with the 2024 rate approaching the 2018 baseline. However, the Black-to-White non-Hispanic male homicide rate ratio widened modestly across the period (from 19.3 in 2018 to 24.4 in 2024), indicating that the partial reversion was concentrated in the reference group rather than in the burdened group, and that the structural drivers producing this disparity did not narrow. Frameworks that interpret the 2024 decline as evidence of progress should attend to this distinction. Concurrently, firearm suicide rates among Black non-Hispanic male and female youth rose at a comparable magnitude while White non-Hispanic rates remained stable, producing the first year in this dataset in which Black non-Hispanic male suicide rates exceeded those of White non-Hispanic males. This study advances 3 contributions to the surveillance literature: (1) a coherent bridging design that links bridged-race and single-race surveillance eras within a single intersectional framework, enabling continued long-term monitoring across the 2018 classification transition; (2) finalized 2024 estimates with explicit attention to nonmonotonic postpandemic dynamics, including parallel computation of homicide and suicide AAPCs that prior partially stratified analyses have not provided; and (3) an empirical case for intersectional surveillance as the unit of analysis appropriate to populations whose burden and trajectory cannot be characterized by race alone or sex alone. Intent-specific, intersectionally disaggregated surveillance and prevention strategies, developed with the affected communities, continue to be the appropriate response [27,29].

Interpretation Notes for AAPCs

For firearm homicide, the log-linear AAPC understates nonmonotonic dynamics. Annual homicide rates among Black non-Hispanic male youth illustrate this pattern: the rate rose from 33.28 in 2018 to a peak of 53.75 in 2021 (a 61.5% relative increase over 3 years), then declined progressively to 37.06 in 2024 (a 31% relative decline over 3 years). Consequently, the log-linear fit across the overall 7-year window did not reach statistical significance because the pandemic-era increase was partially offset by subsequent declines. The pattern is interpreted as pandemic-era elevation with partial reversal rather than as a sustained directional trend.

For firearm suicide, 2 strata show sustained directional trends consistent with the intersectional pattern described in the main text: Black non-Hispanic male youth (AAPC +9.2%, P=.01) and Black non-Hispanic female youth (AAPC +11.7%, P=.08, approaching but not reaching the conventional threshold). White non-Hispanic male and female rates are stable, producing the convergence in the Black-to-White non-Hispanic male suicide rate ratio reported in the main text.

Acknowledgments

The authors thank the National Center for Health Statistics for maintaining the CDC WONDER Underlying Cause of Death public-use database. Generative AI tools (Claude, Anthropic PBC) were used during preparation of this paper for language editing only, including grammar checking, sentence-level clarity, chart displays, and proofreading of author-drafted content. The authors verified all output produced with the assistance of these tools, take full responsibility for the accuracy, integrity, and originality of this paper.

Funding

The authors declared no financial support was received for this work.

Data Availability

The data analyzed during this study are publicly available through the CDC (Centers for Disease Control and Prevention) WONDER Underlying Cause of Death database maintained by the National Center for Health Statistics. Bridged-race mortality data for 1999 to 2020 are accessible online [31], and single-race mortality data for 2018 to 2024 are accessible online [32]. No registration or data use agreement is required for access.

Authors' Contributions

Conceptualization: CE, TU
Data curation: CE
Formal analysis: CE
Investigation: CE, TU
Methodology: CE
Project administration: CE
Visualization: CE
Writing – original draft: CE
Writing – review & editing: CE, TU

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Age-specific firearm homicide and suicide mortality rates and distribution of firearm deaths among US youth and young adults aged 0 to 24 years, 1999 to 2020 and 2018 to 2024.
[DOCX File, 22 KB - publichealth_v12i1e96931_app1.docx]

Checklist 1
STROBE checklist.
[PDF File, 160 KB - publichealth_v12i1e96931_app2.pdf]

References

1. Cunningham RM, Walton MA, Carter PM. The major causes of death in children and adolescents in the United States. N Engl J Med 2018 Dec 20;379(25):2468-2475. [doi: 10.1056/NEJMsr1804754] [Medline: 30575483]
2. Degli Esposti M, Pulcini CD, Fleegler EW, et al. Youth firearm mortality in the Americas from 2015 to 2022. JAMA Netw Open 2024 Oct 1;7(10):e2437395. [doi: 10.1001/jamanetworkopen.2024.37395] [Medline: 39356508]
3. Goldstick JE, Zeoli A, Mair C, Cunningham RM. US firearm-related mortality: national, state, and population trends, 1999-2017. Health Aff (Millwood) 2019 Oct;38(10):1646-1652. [doi: 10.1377/hlthaff.2019.00258] [Medline: 31589525]
4. Fowler KA, Dahlberg LL, Haileyesus T, Gutierrez C, Bacon S. Childhood firearm injuries in the United States. Pediatrics 2017 Jul;140(1):e20163486. [doi: 10.1542/peds.2016-3486] [Medline: 28630118]
5. Cheon C, Lin Y, Harding DJ, Wang W, Small DS. Neighborhood racial composition and gun homicides. JAMA Netw Open 2020 Nov 2;3(11):e2027591. [doi: 10.1001/jamanetworkopen.2020.27591] [Medline: 33252687]
6. Betz ME, Boudreaux ED. Managing suicidal patients in the emergency department. Ann Emerg Med 2016 Feb;67(2):276-282. [doi: 10.1016/j.annemergmed.2015.09.001] [Medline: 26443554]
7. Roberts BK, Nofi CP, Cornell E, Kapoor S, Harrison L, Sathya C. Trends and disparities in firearm deaths among children. Pediatrics 2023 Sep 1;152(3):e2023061296. [doi: 10.1542/peds.2023-061296] [Medline: 37599647]
8. Price JH, Khubchandani J, Foh EP. Unintentional firearm mortality in African-American youths, 2010-2019. J Natl Med Assoc 2021 Oct;113(5):580-586. [doi: 10.1016/j.jnma.2021.05.009] [Medline: 34172295]
9. Mariño-Ramírez L, Jordan IK, Nápoles AM, Pérez-Stable EJ. Comparison of US firearm-related deaths among children and adolescents by race and ethnicity, 1999-2020. JAMA 2022 Dec 20;328(23):2359-2360. [doi: 10.1001/jama.2022.19508] [Medline: 36538317]
10. Lee LK, Tang SFS, Cull WL, Fleegler EW, Olson LM. Firearm and motor vehicle pediatric deaths-intersections of age, sex, race, and ethnicity. JAMA Pediatr 2025 Mar 1;179(3):310-320. [doi: 10.1001/jamapediatrics.2024.5089] [Medline: 39714814]
11. Weybright EH, Terral HF, Conrick KM, Carter PM, Rowhani-Rahbar A. Trends in firearm death among middle and high-school aged rural and urban adolescents from 2001 to 2022. Inj Epidemiol 2025 Sep 1;12(1):53. [doi: 10.1186/s40621-025-00613-w] [Medline: 40890889]
12. Arias E, Heron M, National Center for Health Statistics, Hakes J, US Census Bureau. The validity of race and Hispanic-origin reporting on death certificates in the United States: an update. Vital Health Stat 2 2016 Aug 1(172):1-21. [Medline: 28436642]
13. Klein-Cloud R, Roberts B, Cornell E, Nofi C, Sathya C. Racial and ethnic disparities in pediatric firearm deaths persist in 2022 and 2023. Inj Epidemiol 2025 Mar 24;12(1):19. [doi: 10.1186/s40621-025-00571-3] [Medline: 40128885]
14. Crenshaw K. Mapping the margins: intersectionality, identity politics, and violence against women of color. Stanford Law Rev 1991 Jul;43(6):1241. [doi: 10.2307/1229039]
15. Bowleg L. The problem with the phrase women and minorities: intersectionality-an important theoretical framework for public health. Am J Public Health 2012 Jul;102(7):1267-1273. [doi: 10.2105/AJPH.2012.300750] [Medline: 22594719]
16. Bauer GR. Incorporating intersectionality theory into population health research methodology: challenges and the potential to advance health equity. Soc Sci Med 2014 Jun;110:10-17. [doi: 10.1016/j.socscimed.2014.03.022] [Medline: 24704889]
17. Krieger N. Methods for the scientific study of discrimination and health: an ecosocial approach. Am J Public Health 2012 May;102(5):936-944. [doi: 10.2105/AJPH.2011.300544] [Medline: 22420803]
18. Xu JQ, Murphy SL, Kochanek KD, Arias E. Mortality in the United States, 2024. NCHS Data Brief 2026 Jan(548). [doi: 10.15620/cdc/174641] [Medline: 41678830]
19. Arnett JJ. Emerging adulthood. A theory of development from the late teens through the twenties. Am Psychol 2000 May;55(5):469-480. [doi: 10.1037/0003-066X.55.5.469] [Medline: 10842426]
20. National Academies of Sciences, Engineering, and Medicine. The promise of adolescence: realizing opportunity for all youth. : National Academies Press; 2019. [doi: 10.17226/25388]
21. Rees CA, Monuteaux MC, Steidley I, et al. Trends and disparities in firearm fatalities in the United States, 1990-2021. JAMA Netw Open 2022 Nov 1;5(11):e2244221. [doi: 10.1001/jamanetworkopen.2022.44221] [Medline: 36445703]

XSL•FO
RenderX

22. Ingram DD, Parker JD, Schenker N, et al. United States Census 2000 population with bridged race categories. Vital Health Stat 2 2003 Sep(135):1-55. [Medline: 14556588]
23. Klein RJ, Schoenborn CA. Healthy people 2010 statistical notes: age adjustment using the 2000 projected US population. : Centers for Disease Control and Prevention; 2001 Jan URL: https://www.cdc.gov/nchs/data/statnt/statnt20.pdf [accessed 2026-07-10]
24. Rothman KJ, Greenland S, Lash TL. Modern Epidemiology, 3rd edition: Lippincott Williams & Wilkins; 2008.
25. About underlying cause of death, 1999-2020. CDC WONDER. URL: https://wonder.cdc.gov/ucd-icd10.html [accessed 2026-07-01]
26. Krieger N. Structural racism, health inequities, and the two-edged sword of data: structural problems require structural solutions. Front Public Health 2021;9:655447. [doi: 10.3389/fpubh.2021.655447] [Medline: 33937178]
27. Butts JA, Roman CG, Bostwick L, Porter JR. Cure violence: a public health model to reduce gun violence. Annu Rev Public Health 2015 Mar 18;36:39-53. [doi: 10.1146/annurev-publhealth-031914-122509] [Medline: 25581151]
28. Runyan CW, Becker A, Brandspigel S, Barber C, Trudeau A, Novins D. Lethal means counseling for parents of youth seeking emergency care for suicidality. West J Emerg Med 2016 Jan;17(1):8-14. [doi: 10.5811/westjem.2015.11.28590] [Medline: 26823923]
29. Cwik MF, Tingey L, Maschino A, et al. Decreases in suicide deaths and attempts linked to the White Mountain Apache Suicide Surveillance and Prevention System, 2001-2012. Am J Public Health 2016 Dec;106(12):2183-2189. [doi: 10.2105/AJPH.2016.303453] [Medline: 27736202]
30. Espey DK, Jim MA, Richards TB, Begay C, Haverkamp D, Roberts D. Methods for improving the quality and completeness of mortality data for American Indians and Alaska Natives. Am J Public Health 2014 Jun;104 Suppl 3(Suppl 3):S286-S294. [doi: 10.2105/AJPH.2013.301716] [Medline: 24754557]
31. About multiple cause of death, 1999-2020. CDC WONDER. URL: https://wonder.cdc.gov/mcd-icd10.html [accessed 2026-07-10]
32. About multiple cause of death, 2018-2024, single race. CDC WONDER. URL: https://wonder.cdc.gov/mcd-icd10-expanded.html [accessed 2026-07-10]

Abbreviations

AAPC: average annual percentage change
AI/AN: American Indian or Alaska Native
CDC: Centers for Disease Control and Prevention
CFR: Code of Federal Regulations
***ICD-10*:** *International Statistical Classification of Diseases, Tenth Revision*
NCHS: National Center for Health Statistics
NH/OPI: Native Hawaiian or Other Pacific Islander
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Mavragani, T Sanchez; submitted 02.Apr.2026; peer-reviewed by AA Villa-Rueda, GMA Malik, K Gelberg; revised version received 25.Jun.2026; accepted 30.Jun.2026; published 31.Jul.2026.

Please cite as:
Emezue C, Udmuangpia T
Twenty-Five Years of Firearm Homicides and Suicides Among US Children and Young Adults, 1999 to 2024: Intersectional Surveillance Analysis
JMIR Public Health Surveill 2026;12:e96931
URL: https://publichealth.jmir.org/2026/1/e96931
doi:10.2196/96931

© Chuka Emezue, Tipparat Udmuangpia. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 31.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Temporal Trends in Type 2 Diabetes Burden and Care Outcomes in Taiwan From 2012 to 2021: Nationwide Population-Based Cohort Study

Ming-Yen Lin[1], PhD; Yi-Ling Wu[2], DrPH; Hsiaoting Wu[2], MS; Chih-Cheng Hsu[2,3,4,5], MD, DrPH; iH3 Research Group[6]

[1]Precision Sports Medicine and Health Promotion Center, Kaohsiung Medical University, Kaohsiung, Taiwan
[2]Institute of Population Health Sciences, National Health Research Institutes, Miaoli County, Taiwan
[3]Department of Health Services Administration, China Medical University, Taichung, Taiwan
[4]Department of Family Medicine, Min-Sheng General Hospital, Taoyuan, Taiwan
[5]National Center for Geriatrics and Welfare Research, National Health Research Institutes, Yunlin, Taiwan
[6]See Acknowledgments

Corresponding Author:
Chih-Cheng Hsu, MD, DrPH
Institute of Population Health Sciences
National Health Research Institutes
No. 35, Keyan Road
Miaoli County, 35053
Taiwan
Phone: 886 37 206 166 ext 3633
Fax: 886 37 586 261
Email: cch@nhri.edu.tw

Abstract

Background: Understanding temporal changes in noncommunicable diseases such as diabetes is crucial for informing health policy precisely and improving population-level outcomes.

Objective: This study aimed to evaluate temporal trends in key health metrics of type 2 diabetes using nationwide surveillance data to inform future prevention and care strategies in a resource-rich country.

Methods: We conducted a retrospective, population-based cohort study to analyze temporal trends in annual incidence, prevalence, mortality, and life expectancy from 2012 to 2021; the ABC (hemoglobin A_{1c} [HbA_{1c}], blood pressure, and cholesterol) control metrics from 2015 to 2021; and complication-specific incidence rates from 2017 to 2021. The selected health metrics were quantified overall and stratified by age and sex and standardized to understand disparities in trends. An autoregressive model was used to evaluate temporal change after appropriately accounting for serial correlation in the data.

Results: The standardized incidence and mortality rates of type 2 diabetes slightly declined, by an average of 0.029 (95% CI 0.011-0.047) and 0.043 (95% CI 0.011-0.076) per 100 patient-years annually, respectively (P=.006 and P=.02). These declines coincided with a significant average annual increase of 0.17 (95% CI 0.15-0.21) in the standardized prevalence proportion (P<.001). While glycemic and lipid control improved steadily, blood pressure control remained suboptimal, reaching only 35.9% in 2021. All examined complications, except nephropathy, showed age-specific declines over time. The disparity in life expectancy between people with type 2 diabetes and the general population narrowed over time, but the gap was still pronounced among younger age groups.

Conclusions: This nationwide analysis highlights critical care gaps, particularly in blood pressure management, nephropathy prevention, and care for younger individuals with type 2 diabetes. Addressing these deficiencies is essential for achieving more equitable and effective diabetes care.

(JMIR Public Health Surveill 2026;12:e86356) doi:10.2196/86356

KEYWORDS

type 2 diabetes mellitus; incidence; prevalence; complications; life expectancy; mortality; trend

XSL•FO
RenderX

Introduction

The global diabetes population is estimated to rise by more than 50%, from 463 million in 2019 to 700 million by 2045 [1]. Early detection of blood glucose abnormalities through the fasting plasma glucose test is essential for initiating further interventions and management [2]. However, despite increasing awareness of diabetes, the International Diabetes Federation recently estimated that nearly 45% of adults with diabetes worldwide remain unaware of their condition [3]. Consequently, delayed blood glucose control interventions can result in the onset of associated complications before the diabetes diagnosis. One large US cohort study demonstrated that more than 10% of patients had developed nephropathy, and 3.3% of patients were comorbid with macrovascular disease at type 2 diabetes diagnosis [4]. In addition, 1% to 3% of the cohort developed cardiovascular disease, nephropathy, and peripheral neuropathy annually. Therefore, early detection and timely delivery of comprehensive care play a key role in preventing the development of diabetes complications.

Improving type 2 diabetes prevention and care can rely on long-term monitoring of multiple health metrics across the disease spectrum. Incidence, prevalence, and mortality are commonly used to assess the overall status of diabetes prevention and care; however, these metrics often involve heterogeneous data collection methods, making comparisons across countries and over time challenging [5]. Moreover, relying solely on these measures provides limited guidance for improving clinical care and disease management. Therefore, incorporating additional process and prognosis metrics, such as glycemic control rates and complication incidence, is essential to provide actionable insights and enable decision-makers to allocate resources more precisely [6].

Comprehensive diabetes management requires both medication and patient self-care, including dietary modifications, lifestyle improvements, and home monitoring of blood pressure and blood glucose. To deliver holistic care, Taiwan launched pay-for-performance (P4P) reimbursement for diabetes to enhance the health care system's support for patient self-care in 2001. In short, this program encourages physicians to deliver more comprehensive care through a multidisciplinary team (physician, dietitian, and nurse educator) based on international clinical practice guidelines. Evidence from the past decade has shown that these programs substantially reduce the risks of complications and mortality, thereby lowering medical expenditures [7-9]. However, updated evidence on type 2 diabetes care from the surveillance data system is useful in re-examining the care framework and process, identifying care gaps and disadvantages, and redesigning an inclusive care program to address patient needs throughout their life.

Disease surveillance through a health data system can help us understand disease dynamics, medical services use, and care quality monitoring, also offering essential insights into future disease trend prediction and resource arrangements [10]. Our previous research has demonstrated how a health surveillance system using administrative health databases helps improve the effectiveness of diabetes and kidney disease care and management evaluations [11-13]. This study aimed to explore the updated trends in incidence, prevalence, mortality, life expectancy, care quality, and relevant complications to sketch out preventive and care challenges, as well as corresponding strategic actions, for type 2 diabetes in the coming decade.

Methods

Study Design and Database

We conducted a retrospective, population-based cohort study and performed time-series analyses using data systematically integrated from multiple databases described in the Introduction section [11]. In brief, the National Health Insurance Program in Taiwan is a single-payer, government-administered system that covers most of the population. The Taiwan National Health Insurance Research Database comprises several claims databases that enable linking with various governmental registry databases, including those for cancer, catastrophic illness, and death, through unique pseudo-identifiers under specific regulations. They have been widely applied to detect disease dynamics, assess clinical efficacy, and evaluate health policies [14]. In addition, the study also applied the P4P database, 2015 to 2021, to understand key care elements of diabetes care quality, including blood pressure, blood glucose, and blood lipids [15,16]. The research databases were applied to the Health and Welfare Data Science Center and the Health Research Data Integration Service from the National Health Insurance Administration, Ministry of Health and Welfare, from 2011 to 2022.

Ethical Considerations

The institutional review boards of the National Health Research Institutes (EC1131006-E) reviewed and approved the study. As all patients' IDs in the study databases are scrambled, the institutional review boards of the National Health Research Institutes waived informed consent. All study procedures were conducted in accordance with the principles of the Declaration of Helsinki (2008 revision) and the regulations of the Health and Welfare Data Science Center.

Diabetes Mellitus; Hemoglobin A1c, Blood Pressure, and Cholesterol Control; and Relevant Complications

Diabetes cases were identified by a series of diagnosis codes (Table S1 in Multimedia Appendix 1) that appeared 3 or more times in outpatient records or in 1 or more inpatient records within 1 year. The first diagnosis date between 2011 and 2022 was considered the incident date. As patients with type 1 diabetes were registered in the catastrophic illness registry, the type 2 diabetes population can be confirmed by excluding type 1 diabetes from the overall diabetes population. The incidence, prevalence, mortality, and average life expectancy statistics for type 2 diabetes were presented from 2012 to 2021 due to the unreliable data in 2011 and 2022, which resulted from prevalent case contamination and incomplete end-of-year data uploads. An incident case was defined as the absence of diabetes diagnosis codes prior to the first recorded diagnosis date during 2011 to 2022. As incident cases were identified from 2012 onward, this approach ensured that each case had at least a 1-year period without a diabetes diagnosis before classification

XSL•FO
RenderX

as incident. Prevalent cases were identified each year based on the surviving population at the start of the given year.

The diabetes P4P care program asks health care providers to upload patients' essential laboratory data seasonally to ensure care quality. ABC (hemoglobin A_{1c} [HbA_{1c}], blood pressure, and cholesterol) control trends were obtained for the patients enrolled in the care program from 2015 to 2021. The proportions of individuals achieving HbA_{1c} <53 mmol/mol (7%), systolic blood pressure <130 mm Hg, and low-density lipoprotein (LDL) <2.59 mmol/L (100 mg/100 mL) reflected good control of glycemic status, blood pressure, and lipid profile, respectively. The values of each selected laboratory item within 1 calendar year were averaged to reflect long-term control, with the proportion that met these 3 control criteria considered to constitute optimal control of all 3 risk factors.

The studied type 2 diabetes complications were detected from 2017 to 2021. Stroke, major adverse cardiovascular events (MACEs), retinopathy, diabetic foot, peripheral vascular disease, and nephropathy were determined by the relevant diagnosis appearing ≥2 times in outpatient or once in inpatient records within 1 year and relevant procedure codes in the diabetes population (Tables S1 and S2 in Multimedia Appendix 1). MACEs included death caused by cardiovascular diseases, myocardial infarction, heart failure, and stroke. In addition, heart failure occurrence was identified by the appropriate diagnosis codes appearing in one-time inpatient records, as suggested by the study committee. Mortality was confirmed by linking data to the National Death Registry database.

Statistical Analysis

The annual incident rate was calculated by dividing the number of newly diagnosed type 2 diabetes cases in a given year by the number of insured individuals at risk at midyear, excluding those with prevalent type 2 diabetes, and multiplying the result by 100. The annual prevalent proportion was calculated by dividing the year's type 2 diabetes prevalent cases by the total number of insurance cases and multiplying the result by 100. The annual mortality rate was calculated by dividing the number of all-cause deaths among individuals with prevalent type 2 diabetes in a given year by the total patient-years of prevalent type 2 diabetes in that year and multiplying the result by 100. For the convenience of international comparison, the overall and sex-specific incident rates, prevalence proportions, and mortality rates were standardized by 5-year age groups (0-4, 5-9...95-99, and >100 years) using the World Health Organization (WHO) 2000 to 2025 standard population as reference [17]. The demographic life table approach [18] was applied to calculate the average life expectancy, and differences were compared with those of the Taiwan general population in the 20- to 24-year, 40- to 44-year, and 65- to 69-year age groups for both sexes. Temporal trends of the studied health metrics were evaluated using autoregressive models to account for serial correlation in the data. The presence and direction of trends were assessed by including a trend variable, defined as a transformed calendar year coded sequentially as (1, 2,...). For each outcome, residual autocorrelation was assessed using the Durbin-Watson test, and autoregressive terms were incorporated as needed to account for serial dependence. The trend effect for each outcome was then estimated using the Prais-Winsten method after adjustment for autocorrelation. Estimates of the trend effect and their 95% CIs are reported. A 2-sided *P* value of <.05 was deemed statistically significant. All data management and analysis were performed using SAS (version 9.4; SAS Institute). Figures were plotted using GraphPad Prism (version 10.4.0; GraphPad Software).

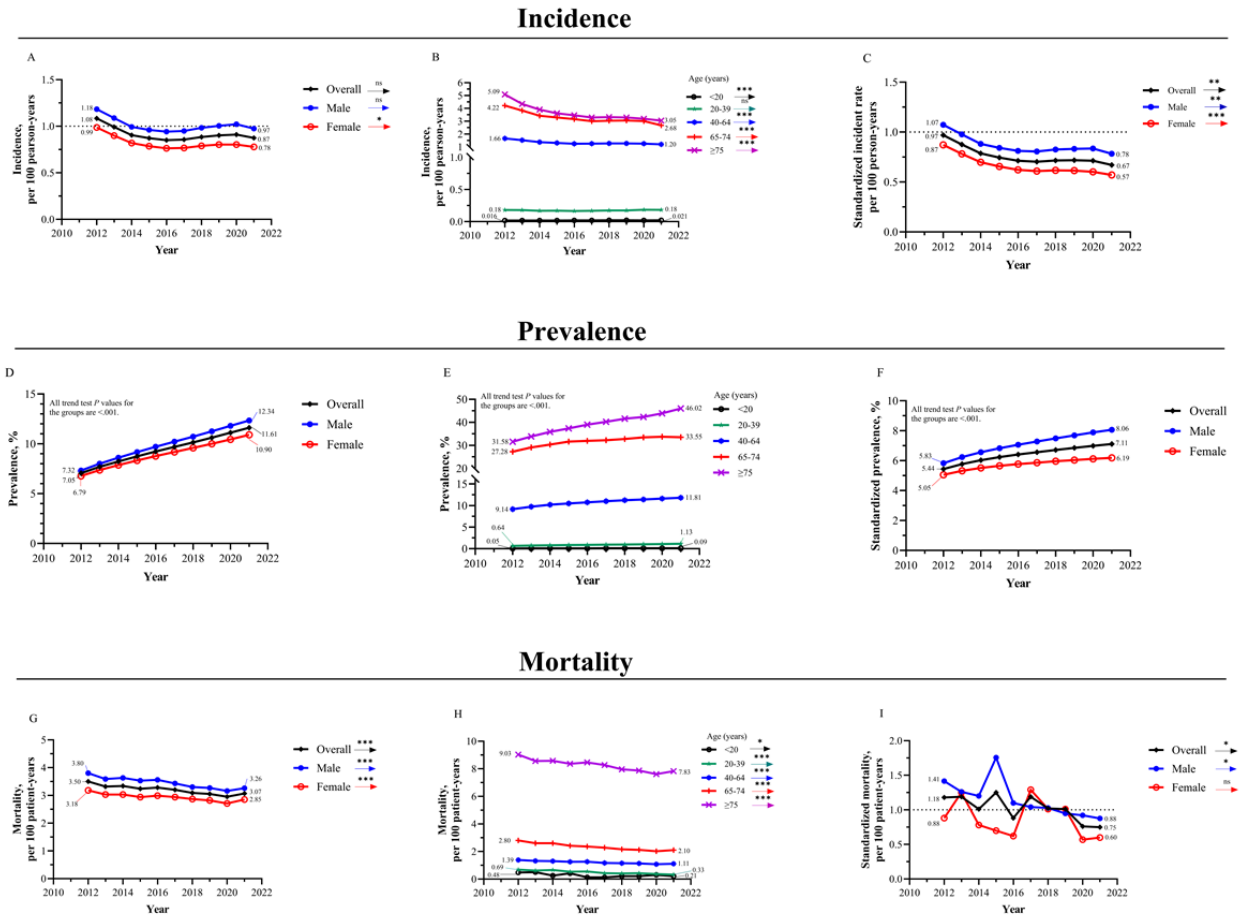
Results

Trends in Incidence, Prevalence, and Mortality

During 2012 to 2022, the average annual numbers of new diagnosis cases, existing cases, and deaths due to type 2 diabetes in Taiwan were 187,119, 2,310,140, and 63,694, respectively. In 2012, Taiwan's total population was 23.31 million, and the population at risk of developing type 2 diabetes was 20.75 million. The overall and sex-specific type 2 diabetes annual incidence showed similar patterns, with a slight reduction during the study period, which was more apparent in the early phase (2012-2014) and tended to be moderated tempered in recent years (Figure 1A). After adjusting for autocorrelation, female participants showed a significant annual reduction of .02 in the incidence rate (*P*<.04; Table S3 in Multimedia Appendix 1). It must be noted that male participants had a higher incidence than female participants, with the older age groups (≥40 years) being the main contributors to the reduction in incidence rates (Figure 1B). The standardized incidence of type 2 diabetes slightly declined by an average of 0.029 (95% CI 0.011-0.047) per 100 patient-years annually (*P*=.006; Figure 1C).

XSL•FO
RenderX

Figure 1. Trends of annual incidence, prevalence, and mortality for type 2 diabetes in Taiwan from 2012 to 2021. (A) Annual incidence of type 2 diabetes for the overall and sex-specific populations. (B) Annual incidence by age. (C) Annual standardized incidence in the overall and sex-specific populations. (D) Annual prevalence in the overall and sex-specific populations. (E) Annual prevalence by age. (F) Annual standardized prevalence in the overall and sex-specific populations. (G) Annual mortality rate in overall and sex-specific type 2 diabetes populations. (H) Annual mortality rate by age. (I) Annual standardized mortality rate in overall and sex-specific type 2 diabetes populations. For A, C, D, F, G, and I, overall populations are represented by black lines and solid diamonds, male populations are represented by blue lines and solid circles, and female populations are represented by red lines and open circles. For B, E, and H, populations <20 years are represented by black lines and open circles, 20-39 years by green lines and solid stars, 40-64 years by blue lines and solid asterisks, 65-74 years by red lines and plus signs, and ≥75 years by purple lines and X's. The World Health Organization (WHO) 2000 to 2025 standard population was used for age standardization. We used an autoregressive model to estimate temporal change after accounting for serial correlation. Serial correlations were evaluated using the Durbin-Watson test. *P* values for temporal trends are denoted as follows: **P*<.05, ***P*<.01, and ****P*<.001. ns: not significant (P≥.05).



Unlike the observed decline in incidence, the annual type 2 diabetes prevalence in the overall population significantly increased, by approximately 65%, from 7.05% in 2012 to 11.61% in 2021 (Figure 1D). Similarly, male participants had higher annual prevalence proportions than female participants, and the discrepancy in annual prevalence between the sexes tended to be larger in more recent years. The older age groups experienced greater increases in prevalence during the follow-up years (Table S4 in Multimedia Appendix 1), with the most notable increase observed among individuals aged >75 years (Figure 1E). The age-standardized prevalence proportions show similar upward trends for overall and various sexes (Figure 1F). The overall standardized prevalence proportion shows a significant average annual increase of 0.17 (95% CI 0.15-0.21, *P*<.001; Table S4 in Multimedia Appendix 1).

As expected, the annual mortality rate decreased slightly from 3.5 per 100 patient-years in 2012 to 3.1 per 100 patient-years in 2021 (Figure 1G). The standardized mortality rates of type 2 diabetes declined by an average of 0.043 (95% CI 0.011-0.076, *P*=.02) per 100 patient-years annually (Table S5 in Multimedia Appendix 1). Although the mortality rate in 2021 remained higher in male participants and older age groups in general (Figures 1G, 1H, and 1I), the mortality trends suggest that those groups have more declining changes than the others (the estimated annual difference is 0.063 reduction in male participants and 0.136 in those aged ≥75 years; Table S5 in Multimedia Appendix 1).

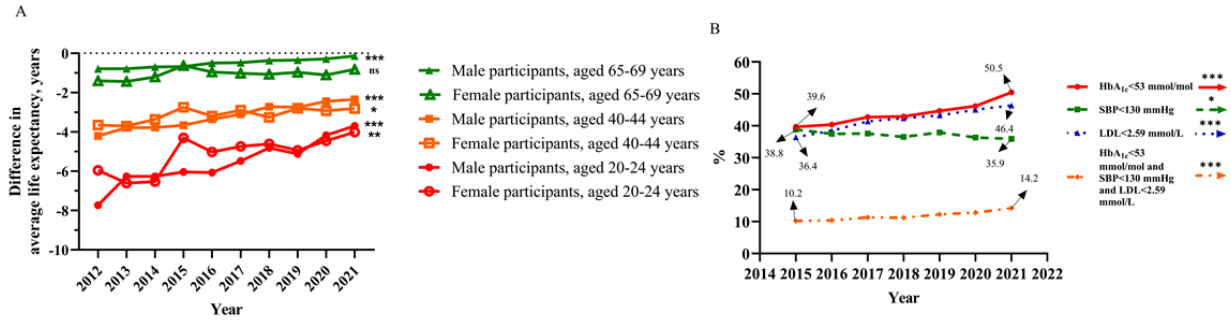
Difference in Average Life Expectancy

The average life expectancy of type 2 diabetes populations was shorter than that of the general population, with differences ranging from 0.13 to 7.7 years in the selected groups (Figure 2A). Notably, younger age groups exhibited larger life expectancy gaps between the 2 groups, irrespective of sex. Over the study period, these differences gradually narrowed across all age groups, particularly in male participants. Male participants with type 2 diabetes approached the average life

XSL•FO
RenderX

expectancy of the general population more rapidly than female participants over time (estimates and 95% CIs were significantly larger in male participants than in female participants at the same age level; Table S6 in Multimedia Appendix 1).

Figure 2. Trends in average life expectancy differences and ABC (hemoglobin A1c, blood pressure, and cholesterol) control in patients with type 2 diabetes in Taiwan. (A) Differences in average life expectancy by sex and age. Differences in average life expectancy were calculated by subtracting the average life expectancy of the general population from that of the type 2 diabetes population. For ages 65-69 years, green lines and solid triangles represent men and green lines and open triangles represent women; for ages 40-44 years, orange lines and solid squares represent men and orange lines and open squares represent women; for ages 20-24 years, red lines and solid circles represent men and red lines and open circles represent women. (B) Annual proportion of patients with type 2 diabetes enrolled in the pay-for-performance (P4P) program who achieved ABC care targets. Red solid lines and solid circles represent HbA1c < 53 mmol/mol, green dashed lines and solid squares represent SBP < 130 mmHg, blue dotted lines and solid triangles represent LDL < 2.59 mmol/L, and orange dashed dotted lines and solid diamonds represent achievement of all 3 targets. We used an autoregressive model to estimate temporal change after accounting for serial correlation. Serial correlations were evaluated using the Durbin-Watson test. We used an autoregressive model to estimate temporal change after accounting for serial correlation. Serial correlations were evaluated using the Durbin-Watson test. **P*<.05, ***P*<.01, and ****P*<.001. HbA1c: hemoglobin A1c; LDL: low-density lipoprotein cholesterol; ns: not significant (P≥.05); SBP: systolic blood pressure.



Trends of ABC Controls in the National P4P Care

From 2015 to 2021, an average of 670,981 existing patients annually participated in the national P4P care program. During the study period, the annual proportion of participants with HbA_{1c} levels <53 mmol/mol (<7%) increased by 10.9% (Figure 2B). A similarly significant improvement was observed in cholesterol control (the annual proportion of LDL <2.59 mmol/L [100 mg/100 mL] increased by 10.0 percentage points, from 36.4% in 2015 to 46.4% in 2021; *P*<.001). However, less than 40% of patients achieved optimal blood pressure control (systolic blood pressure <130 mm Hg) in 2021, even under the P4P care program. The annual proportion of patients achieving optimal blood pressure control was even significantly lower recently (estimate: −0.383 per year, 95% CI −0.700 to −0.065; *P*=.03; Table S7 in Multimedia Appendix 1). The yearly proportion achieving optimal control in the ABC index (above 3 indexes) only increased by less than 1% per year from 2015 to 2021. The results of the autoregressive models confirmed that the annual percentages of blood sugar and lipid control increased significantly by 1.6, whereas blood pressure control showed a significant declining trend (Table S7 in Multimedia Appendix 1).

Trends in the Incidence of Associated Complications

The annual occurrence of stroke (10.1% in 2017 to 9.4% in 2021; *P*=.048) and MACEs (16.2% in 2017 to 14.9% in 2021, *P*=.002) in the overall type 2 diabetes population significantly decreased from 2017 to 2021 (Table S8 in Multimedia Appendix 1). However, the reductions were not homogeneous after stratification by sex and age, with no meaningful changes observed in stroke incidence among female participants aged <40 years or in MACEs among male participants aged 40 to 59 years and female participants aged <40 years.

Similarly, the annual heart failure proportion (1.7% in 2017 to 1.4% in 2021; *P*=.04) and the occurrence proportions of peripheral vascular disease (1.8% in 2017 to 1.8% in 2021; *P*=.03) generally reduced (Table S8 in Multimedia Appendix 1). However, several subgroups, including overall male participants, overall female participants, and individuals aged <40 and 40 to 59 years in both sexes, did not show a noticeable improvement in the annual heart failure occurrence. The nonobvious reduction in peripheral vascular disease occurrence was also observed in male participants aged <40 and female participants aged <40, 40 to 59, and ≥80 years.

Except for male participants (3.1% in 2017 to 2.8% in 2021; *P*=.09; Table 1), the type 2 diabetes population showed significant improvement in diabetic foot development across the selected groups. However, retinopathy showed only slight improvement among male participants overall and male participants aged ≥80 years (overall male participants: −14% from 2017; *P*=.04; male participants aged ≥80 years: −14.2% from 2017; *P*=.04; Table 1).

In contrast, annual nephropathy occurrence across time tended to be stable (36.5% in 2017 to 37.6% in 2021; *P*=.18; Table 1). The Prais-Winsten analysis confirmed significant annual declines in the incidence of stroke, MACEs, heart failure, peripheral vascular disease, diabetes-related foot complications, and retinopathy (Tables S9-S14 in Multimedia Appendix 1).

XSL•FO
RenderX

Table 1. Trends in overall and complication-specific annual incidence of diabetic foot, retinopathy, and nephropathy among individuals with type 2 diabetes from 2017–2021 .

	Year, (%)					*P* value[a]
	2017	2018	2019	2020	2021	
Diabetic foot						
Overall	3.1	3	3	2.9	2.8	.01
Male (age in years)	3.1	3.1	3	2.9	2.8	.09
<40	3.2	3.1	2.9	2.9	2.7	.007
40-59	2.8	2.7	2.7	2.6	2.6	.02
60-79	3.0	3	2.9	2.8	2.7	.001
≥80	4.7	4.7	4.5	4.4	4.2	.003
Female (age in years)	3	3	2.9	2.8	2.8	.01
<40	2.5	2.5	2.4	2.4	2.3	.04
40-59	2.3	2.3	2.2	2.2	2.2	.001
60-79	2.9	2.8	2.7	2.6	2.5	.001
≥80	4.4	4.4	4.4	4.2	4	.06
Retinopathy						
Overall	4.3	4.2	4.2	3.9	3.6	.16
Male (age in years)	4.1	4.1	4	3.8	3.6	.04
<40	2.2	2.3	2.3	2.2	2	.35
40-59	3.3	3.3	3.2	3.1	3	.06
60-79	4.9	4.9	4.7	4.5	4.1	.10
≥80	3.6	3.6	3.6	3.3	3.1	.04
Female (age in years)	4.5	4.4	4.3	4.0	3.7	.16
<40	2.8	2.9	2.8	2.8	2.7	.05
40-59	3.8	3.7	3.7	3.5	3.3	.10
60-79	5.2	5.1	5	4.6	4.2	.14
≥80	3.3	3.3	3.3	3	2.9	.20
Nephropathy						
Overall	36.5	37.2	37.6	37.5	37.6	.18
Male (age in years)	37.1	37.9	38.3	38.2	38.3	.16
<40	29.7	30.4	30.6	30.7	31.2	.009
40-59	31.9	32.7	33	33.2	33.3	.03
60-79	39	39.6	39.9	39.7	39.7	.25
≥80	46.9	47.7	48.3	47.7	47.5	.35
Female (age in years)	35.8	36.5	36.9	36.8	36.8	.20
<40	27.4	27.9	28.1	28	28.6	.03
40-59	30.2	30.9	31.2	31.4	31.4	.03
60-79	36.3	36.8	37.1	36.8	36.7	.55
≥80	42.7	43.8	44.4	44.2	44	.31

[a]We used an autoregressive model to estimate temporal change after accounting for serial correlation. Serial correlations were evaluated using the Durbin-Watson test. The *P* value for the temporal change was <.05, considered statistically significant.

Interestingly, the annual nephropathy occurrence noticeably increased in younger age groups (<40 and 40-59 years) of both male participants and females participants (<40 years: increase of 0.33 per year, 95% CI 0.16-0.51 in male participants; *P*=.009 and increase of 0.24 per year, 95% CI 0.03-0.44 in female participants; *P*=.03; 40-59 years: increase of 0.32 per year, 95%

XSL•FO
RenderX

CI 0.06-0.58 in male participants; *P*=.03 and increase of 0.28 per year, 95% CI 0.05-0.50 in female participants; *P*=.03; Table 1 and Table S15 in Multimedia Appendix 1).

Discussion

Principal Findings

The study identified several important features by systematically quantifying and integrating time-series analyses from different universal databases. We found that the annual incidence of type 2 diabetes remained stable, whereas mortality declined over time, contributing to an approximately 65% increase in type 2 diabetes prevalence from 2012 to 2021. The older age groups were the major contributors to reductions in incidence and mortality. Although the differences in average life expectancy in the type 2 diabetes population gradually decreased between the sex and age groups, the younger age groups still had substantial gaps in their average life expectancy compared to the general population. Analysis of the annual proportion of optimal ABC control trends indicated that the annual proportion of individuals achieving optimal blood pressure control did not improve as expected. The time-series analysis in the occurrence of associated complications demonstrated that the proportions were generally stable or reduced, except for nephropathy in younger patients (<40 or 40-59 years).

A reduction in the global incidence of type 2 diabetes was supported by a recent study using high-quality databases from resource-rich countries [19]. Studies in the United States also reported a declining trend in type 2 diabetes incidence and a lower lifetime risk of type 2 diabetes in 2014 to 2015 than in 2009 to 2010 for both sexes [19-21]. However, some argue that these reductions may be a result of a change in type 2 diabetes diagnosis criteria in 2011 or using survey-based data with different approaches [22]. Notably, the Global Burden of Disease Study 2017 also reported a decrease in the age-standardized incidence of diabetes from 1990 to 2017 in Australasia and Tropical Latin America [23]. Previous evidence combining the current findings suggests that controlling air pollutant concentration rather than changing personal lifestyles in Taiwan may contribute to the reduced incidence of type 2 diabetes [24], but more research is required to clarify the causal effect.

The differences in the analyzed trends between age- and sex-specific groups raise concerns about the social determinants of health, the need to promote precise preventive actions, and health equity. Youth-onset type 2 diabetes is a global public health challenge, particularly in disadvantaged populations [25]. Modern societal environments may limit access to healthy foods and opportunities for physical activity, potentially contributing to the development of metabolic conditions that increase the risk of type 2 diabetes among younger individuals. In addition, our findings identified another vulnerable population, middle-aged adults, as evidenced by a less reduced incidence of type 2 diabetes and heart failure development. Most of these individuals are working, raising concerns about whether health care and working environments offer sufficient resources to support medical access and self-care for diabetes management. Further studies are needed to clarify the underlying mechanisms.

The WHO launched a global diabetes initiative to promote early diabetes diagnosis and control: 80% of people with diabetes are diagnosed, 80% of the diagnosed patients have good glycemic control (HbA_{1c} <64 mmol/mol [8%]), 80% of the diagnosed patients have good blood pressure control (systolic blood pressure/diastolic blood pressure of 140/90 mm Hg [18.67/12.00 kPa]), and 60% of the diagnosed patients aged >40 years use statins. Our findings on ABC control trends align with the WHO targets, demonstrating gradual improvements in glycemic and lipid control, whereas improvements in blood pressure control were less pronounced. It is reasonable to speculate that the stable occurrence of nephropathy is likely related to long-term, uncontrolled high blood pressure, which accelerates kidney dysfunction to end-stage kidney disease more quickly in Taiwan than in other countries [26]. Nonetheless, this insight may help governments allocate resources for the subsequent preventive actions more precisely. For example, the government can develop strategies to detect key unmet chronic conditions earlier and implement the necessary procedures in line with the WHO's targets.

Life expectancy in the type 2 diabetes population is not always quantified and compared with the overall population in previous studies. Similar to the recent observations that display years of life lost due to type 2 diabetes from 23 developed jurisdictions [19], our results reflect that those with type 2 diabetes have a lower average life expectancy than the general population. Additionally, younger individuals with type 2 diabetes tend to experience a substantially greater loss in life expectancy than older individuals, with nearly 5 years of life lost among those aged 20 to 24 years versus less than 2 years among those aged 65 to 69 years. Encouragingly, both the overall gap in years of life lost between individuals with type 2 diabetes and the general population and the age-related disparity in life years lost have narrowed. This trend may be attributed to the widespread adoption of more effective interventions, such as sodium-glucose cotransporter 2 inhibitors and glucagon-like peptide 1 receptor agonists, which remarkably reduce the risks of complications and mortality [27].

The study has several advantages. First, high-quality databases with strict definitions were used to ensure the findings' accuracy and reliability. Next, integrating several meaningful health metrics across the lifetime of patients with type 2 diabetes offers more comprehensive health and care insights for stakeholders, allowing decision-makers to optimize their health decisions and explore possible mechanisms continuously.

However, some limitations need to be disclosed. First, the strict definitions to ensure accurate results may compromise generalizability to some patients with mild conditions, both for diabetes and its relevant complications. In addition, the ABC control time-series analysis from the P4P care program may be subject to optimistic bias and may not be representative of the entire population with type 2 diabetes in Taiwan. Factors influencing ABC control are likely multifactorial, involving patient and physician behaviors as well as broader social determinants, which are not captured within this surveillance system. Finally, the study was conducted in a high-income country with a single-payer health insurance system, so caution

XSL•FO
RenderX

should be exercised when generalizing the findings to other countries, particularly those in low- and middle-income settings or with different health care systems.

Conclusions

In conclusion, the comprehensive analyses revealed substantial progress across several type 2 diabetes metrics but highlighted warning signs, particularly in blood pressure management, nephropathy prevention, and care for younger individuals. The surveillance results underscore the need for timely, targeted policy actions, including strengthening blood pressure control, enhancing early kidney protection strategies (eg, the use of sodium-glucose cotransporter 2 inhibitors), and delivering more accessible care for younger age groups to increase health parities.

Acknowledgments

The authors thank the Taiwan Instrument Research Institute, National Applied Research Laboratories, for its assistance with project coordination. In addition, the authors thank Miss Chi-Ting Huang for her assistance with data management. Portions of this work were presented orally at the Annual Meeting of the Taiwan Society of Nephrology in December 2024. The authors used ChatGPT to review selected author-written sentences for language fluency and grammatical accuracy. All suggested revisions were reviewed and approved by the authors, who retained full responsibility for the final content of the manuscript.

The authors would like to thank the collaborators in the iH3 Research Group: Jeng-Huei Chen, Jia-Sin Liu, Lii-Jia Yang, Cheng-Yin Chung, Ping-Hsun Wu, Yihuang Kang, Jia-Ying Song, Yi-Wen Chiu, Hsing Luh, Shang-Jyh Hwang, Yu-Hsuan Lin.

Data Availability

All data generated or analyzed during this study are included in this paper and its supplementary information files. The raw data can be obtained from the Taiwan Ministry of Health and Welfare upon reasonable request [28,29].

Funding

This work was primarily supported by the National Health Research Institutes, Taiwan (grant NHRI-EX115-11208PI). Additional support was provided by Kaohsiung Medical University Hospital, Kaohsiung Medical University (grants KMUH112-2R20, KMUH113-3R17, and KMUH114-4R21), and the Precision Sports Medicine and Health Promotion Center, Kaohsiung Medical University, Taiwan (grant KMU-TC114A06-1), awarded to MYL.

Authors' Contributions

Conceptualization: CCH (lead), MYL (equal)
Methodology: MYL (lead), YLW (equal), HW (supporting), CCH (supporting)
Software: YLW (lead), MYL (equal), CCH (supporting)
Formal analysis: YLW (lead), HW (supporting)
Validation: MYL (lead), CCH (equal)
Writing—original draft preparation: MYL (lead), YLW (equal), HW (supporting)
Writing—review and editing: CCH
Supervision: CC
Project administration: MYL (lead), YLW (equal), CCH (supporting)
Funding acquisition: MYL (lead), CCH (equal).
All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

MYL received a consultancy fee from WS Far IR Medical Technology Co, Ltd, New Taipei City, Taiwan, for designing a customer report system. This financial relationship did not influence this paper's adherence to scientific rigor and objectivity. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
Relevant diagnostic and procedure codes, modeling results, and standard reporting checklists used in the study.
[DOCX File , 105 KB - publichealth_v12i1e86356_app1.docx]

References

1. Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N, et al. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: results from the International Diabetes Federation Diabetes Atlas, 9th edition. Diabetes Res Clin Pract 2019 Nov;157:107843. [doi: 10.1016/j.diabres.2019.107843] [Medline: 31518657]
2. American Diabetes Association Professional Practice Committee. 2. Diagnosis and classification of diabetes: standards of care in diabetes-2025. Diabetes Care 2025 Jan 01;48(1 Suppl 1):S27-S49. [doi: 10.2337/dc25-S002] [Medline: 39651986]

XSL•FO
RenderX

3. Ogurtsova K, Guariguata L, Barengo NC, Ruiz PL, Sacre JW, Karuranga S, et al. IDF Diabetes Atlas: global estimates of undiagnosed diabetes in adults for 2021. Diabetes Res Clin Pract 2022 Jan;183:109118. [doi: 10.1016/j.diabres.2021.109118] [Medline: 34883189]
4. An J, Nichols GA, Qian L, Munis MA, Harrison TN, Li Z, et al. Prevalence and incidence of microvascular and macrovascular complications over 15 years among patients with incident type 2 diabetes. BMJ Open Diabetes Res Care 2021 Jan;9(1):e001847 [FREE Full text] [doi: 10.1136/bmjdrc-2020-001847] [Medline: 33397671]
5. Khan MA, Hashim MJ, King JK, Govender RD, Mustafa H, Al Kaabi J. Epidemiology of type 2 diabetes - Global Burden of Disease and forecasted trends. J Epidemiol Glob Health 2020 Mar;10(1):107-111 [FREE Full text] [doi: 10.2991/jegh.k.191028.001] [Medline: 32175717]
6. Wens J, Dirven K, Mathieu C, Paulus D, Van Royen P, Belgian Diabetes Project Group. Quality indicators for type-2 diabetes care in practice guidelines: an example from six European countries. Prim Care Diabetes 2007 Feb;1(1):17-23. [doi: 10.1016/j.pcd.2006.07.001] [Medline: 18632015]
7. Chen YC, Lee CT, Lin BJ, Chang YY, Shi HY. Impact of pay-for-performance on mortality in diabetes patients in Taiwan: a population-based study. Medicine (Baltimore) 2016 Jul;95(27):e4197 [FREE Full text] [doi: 10.1097/MD.0000000000004197] [Medline: 27399144]
8. Kornelius E, Chiou JY, Yang YS, Lu YL, Peng CH, Huang CN. The diabetes shared care program and risks of cardiovascular events in type 2 diabetes. Am J Med 2015 Sep;128(9):977-85.e3. [doi: 10.1016/j.amjmed.2015.03.025] [Medline: 25908394]
9. Hsieh HM, Lin MY, Chiu YW, Wu PH, Cheng LJ, Jian FS, et al. Economic evaluation of a pre-ESRD pay-for-performance programme in advanced chronic kidney disease patients. Nephrol Dial Transplant 2017 Jul 01;32(7):1184-1194 [FREE Full text] [doi: 10.1093/ndt/gfw372] [Medline: 28486670]
10. Virnig BA, McBean M. Administrative data for public health surveillance and planning. Annu Rev Public Health 2001;22:213-230. [doi: 10.1146/annurev.publhealth.22.1.213] [Medline: 11274519]
11. Lin MY, Liao CT, Lin MH, Wu HT, Wu MS, Hsu YH, et al. Taiwan renal data report system: framework, function, and development. Nephrology (Carlton) 2023 Nov;28(11):581-587. [doi: 10.1111/nep.14228] [Medline: 37549919]
12. Sheen YJ, Hsu CC, Jiang YD, Huang CN, Liu JS, Sheu WH. Trends in prevalence and incidence of diabetes mellitus from 2005 to 2014 in Taiwan. J Formos Med Assoc 2019 Nov;118 Suppl 2:S66-S73 [FREE Full text] [doi: 10.1016/j.jfma.2019.06.016] [Medline: 31300322]
13. Taiwan Society of Nephrology. 2020 annual report on kidney disease in Taiwan. Journal of the Formosan Medical Association. 2022. URL: https://www.sciencedirect.com/journal/journal-of-the-formosan-medical-association/vol/121/suppl/S1 [accessed 2026-07-24]
14. Hsing AW, Ioannidis JP. Nationwide population science: lessons from the Taiwan National Health Insurance Research Database. JAMA Intern Med 2015 Sep;175(9):1527-1529. [doi: 10.1001/jamainternmed.2015.3540] [Medline: 26192815]
15. Lin TY, Chen CY, Huang YT, Ting MK, Huang JC, Hsu KH. The effectiveness of a pay for performance program on diabetes care in Taiwan: a nationwide population-based longitudinal study. Health Policy 2016 Nov;120(11):1313-1321. [doi: 10.1016/j.healthpol.2016.09.014] [Medline: 27780591]
16. Lee IT, Hsu CC, Sheu WH, Su SL, Wu YL, Lin SY. Pay-for-performance for shared care of diabetes in Taiwan. J Formos Med Assoc 2019 Nov;118 Suppl 2:S122-S129 [FREE Full text] [doi: 10.1016/j.jfma.2019.08.011] [Medline: 31471222]
17. Ahmad OB, Boschi-Pinto C, Lopez AD, Murray CJ, Lozano R, Inoue M. Age standardization of rates: a new WHO standard. World Health Organization. 2001. URL: https://cdn.who.int/media/docs/default-source/gho-documents/global-health-estimates/gpe_discussion_paper_series_paper31_2001_age_standardization_rates.pdf [accessed 2026-07-20]
18. Silcocks PB, Jenner DA, Reza R. Life expectancy as a summary of mortality in a population: statistical considerations and suitability for use by health authorities. J Epidemiol Community Health 2001 Jan;55(1):38-43 [FREE Full text] [doi: 10.1136/jech.55.1.38] [Medline: 11112949]
19. Tomic D, Morton JI, Chen L, Salim A, Gregg EW, Pavkov ME, et al. Lifetime risk, life expectancy, and years of life lost to type 2 diabetes in 23 high-income jurisdictions: a multinational, population-based study. Lancet Diabetes Endocrinol 2022 Nov;10(11):795-803 [FREE Full text] [doi: 10.1016/S2213-8587(22)00252-2] [Medline: 36183736]
20. National Diabetes Statistics Report. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/diabetes/php/data-research/index.html [accessed 2025-05-07]
21. Benoit SR, Hora I, Albright AL, Gregg EW. New directions in incidence and prevalence of diagnosed diabetes in the USA. BMJ Open Diabetes Res Care 2019 May 28;7(1):e000657 [FREE Full text] [doi: 10.1136/bmjdrc-2019-000657] [Medline: 31245008]
22. Selvin E, Ali MK. Declines in the incidence of diabetes in the U.S.-real progress or artifact? Diabetes Care 2017 Sep;40(9):1139-1143 [FREE Full text] [doi: 10.2337/dc16-2442] [Medline: 28830954]
23. Liu J, Ren ZH, Qiang H, Wu J, Shen M, Zhang L, et al. Trends in the incidence of diabetes mellitus: results from the Global Burden of Disease Study 2017 and implications for diabetes mellitus prevention. BMC Public Health 2020 Sep 17;20(1):1415 [FREE Full text] [doi: 10.1186/s12889-020-09502-x] [Medline: 32943028]
24. Chen CC, Wang YR, Liu JS, Chang HY, Chen PC. Associations between long-term ambient PM2.5 exposure and the incidence of cardiopulmonary diseases and diabetes, attributable years lived with disability, and policy implication. Ecotoxicol Environ Saf 2024 Sep 01;282:116688 [FREE Full text] [doi: 10.1016/j.ecoenv.2024.116688] [Medline: 38971102]

XSL•FO
RenderX

25. Bjornstad P, Chao LC, Cree-Green M, Dart AB, King M, Looker HC, et al. Youth-onset type 2 diabetes mellitus: an urgent challenge. Nat Rev Nephrol 2023 Mar;19(3):168-184 [FREE Full text] [doi: 10.1038/s41581-022-00645-1] [Medline: 36316388]
26. Hsu CC, Hsu YH, Wu MS, Hwang SJ. Achievements and challenges in chronic kidney disease care in Taiwan. J Formos Med Assoc 2022 Feb;121 Suppl 1:S3-S4 [FREE Full text] [doi: 10.1016/j.jfma.2021.12.016] [Medline: 35016834]
27. Xie Y, Bowe B, Xian H, Loux T, McGill JB, Al-Aly Z. Comparative effectiveness of SGLT2 inhibitors, GLP-1 receptor agonists, DPP-4 inhibitors, and sulfonylureas on risk of major adverse cardiovascular events: emulation of a randomised target trial using electronic health records. Lancet Diabetes Endocrinol 2023 Sep;11(9):644-656. [doi: 10.1016/S2213-8587(23)00171-7] [Medline: 37499675]
28. Regarding the Health and Welfare Information Science Center [Article in Chinese]. Ministry of Health and Welfare Taiwan. URL: https://dep.mohw.gov.tw/DOS/cp-5119-59201-113.html [accessed 2026-07-27]
29. National health insurance insurer information integration and application service center. National Health Insurance Administration, Ministry of Health and Welfare Taiwan. URL: https://www.nhi.gov.tw/ch/np-3134-1.html [accessed 2026-07-27]

Abbreviations

LDL: low-density lipoprotein
MACE: major adverse cardiovascular event
P4P: pay-for-performance
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 23.Oct.2025; peer-reviewed by SJ Kim, JC da Mota; comments to author 06.Apr.2026; revised version received 03.May.2026; accepted 11.Jun.2026; published 24.Aug.2026.

Please cite as:
Lin MY, Wu YL, Wu H, Hsu CC, iH3 Research Group
Temporal Trends in Type 2 Diabetes Burden and Care Outcomes in Taiwan From 2012 to 2021: Nationwide Population-Based Cohort Study
JMIR Public Health Surveill 2026;12:e86356
URL: https://publichealth.jmir.org/2026/1/e86356
doi:10.2196/86356
PMID:

©Ming-Yen Lin, Yi-Ling Wu, Hsiaoting Wu, Chih-Cheng Hsu, iH3 Research Group. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 24.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Adapting the Tobacco Pack Surveillance System Protocol to Assess Electronic Cigarette Packaging: Protocol for a Content Analysis

Reiley Hartmuller[1], BS; Jennifer L Brown[1], MPH, PhD; Hannah Elise Barker[1], MPH; Qinghua Nian[1], MA, PhD; Kevin Welding[1], MA, PhD; Joanna E Cohen[1], MHSc, PhD; Katherine Clegg Smith[1], MA, PhD

Johns Hopkins Bloomberg School of Public Health, Department of Health, Behavior and Society, Institute for Global Tobacco Control, Baltimore, MD, United States

Corresponding Author:
Reiley Hartmuller, BS
Johns Hopkins Bloomberg School of Public Health
Department of Health, Behavior and Society
Institute for Global Tobacco Control
2213 McElderry St., Fourth Floor
Baltimore, MD, 21205
United States
Phone: 1 4436213079
Email: rhartmu2@jhu.edu

Abstract

Background: In 2021, a total of 82 million people used electronic cigarettes (e-cigarettes) globally. E-cigarette regulations around the globe vary widely from the product being banned in some jurisdictions to being completely unregulated in others. The Tobacco Pack Surveillance System (TPackSS) was initiated in 2012 to monitor tobacco packs available in 14 low- and middle-income countries with the greatest number of people who smoke. The aim of TPackSS is to assess compliance with country-specific tobacco packaging and labeling requirements and identify marketing features and appeals used on tobacco packaging.

Objective: The objective of this study was to adapt and expand the previous TPackSS protocol to also include disposable e-cigarette devices and their accompanying products: e-cigarette liquids and pods or cartridges.

Methods: E-cigarettes were added to TPackSS data collection in Indonesia in 2022 and in China in 2023. Collection took place in Jakarta, Medan, and Surabaya in Indonesia and Shanghai, Beijing, Chongqing, Guangzhou, Kunming, and Shenzhen in China. A total of 15 neighborhoods per city in Indonesia and 12 per city in China were visited. The TPackSS protocol developed for tobacco products in 2012 to 2013 was used as the foundation for the adapted protocol for e-cigarettes. The adaptation of the original TPackSS protocol followed an iterative process involving extensive discussions among the TPackSS study team, comprising researchers with expertise in tobacco control and 13 years of experience in tobacco product packaging surveillance. The study adapted the sampling frame and sampling strategy, pack collection procedures, photography guide, shipment and translation procedures, and codebook to account for the unique country contexts for e-cigarettes (eg, regulations, market for e-cigarettes, and retailer landscape) and the heterogeneity of e-cigarette packaging observed. These adaptations were based on the review of peer-reviewed literature, white papers, Euromonitor data, and consultation with researchers with expertise in e-cigarette marketing and with in-country public health professionals. Using Indonesia as a pilot country, the protocol was evaluated and further refined post implementation for use in China.

Results: The TPackSS study began in 2012 and was funded by Bloomberg Philanthropies. Data collection took place from September to October 2022 in Indonesia and April to May 2023 in China. Across Indonesia and China, 968 unique e-cigarette products were collected. TPackSS data collection, coding procedures, and study findings are publicly available for use on the TPackSS website.

Conclusions: The protocols presented here, which were used in 2 countries with contrasting e-cigarette markets and regulatory requirements, can be adapted for use in other countries. Findings from use of this adapted protocol can inform policy by providing insights into the design features and marketing appeals of e-cigarette products available on the market, as well as compliance with health warning label requirements where applicable.

XSL•FO
RenderX

(JMIR Public Health Surveill 2026;12:e89325) doi:10.2196/89325

KEYWORDS

electronic nicotine delivery devices; packaging and labeling; global health; low- or middle-income country; health warning label; electronic cigarettes; electronic cigarette marketing; tobacco and nicotine products; electronic cigarette packaging

Introduction

An estimated 82 million people globally used electronic cigarettes (e-cigarettes) in 2021 [1]. In many countries, including Indonesia and China, rates of e-cigarette use among youth exceed those of adults [2]. Many countries do not regulate flavors, nicotine strength, or liquid volumes, nor do they require graphic health warning labels (HWLs) on e-cigarette products [3-6].

E-Cigarette Packaging as Advertising

Tobacco and nicotine product packaging serves as communication real estate for both tobacco company marketing and health communication through required government warnings [7,8]. E-cigarette packages can be designed to resemble other objects (eg, Tic Tac container and technology devices such as mobile phones) [9-11]. They also feature a wide variety of marketing appeals such as themes of fantasy or violence, graphic elements such as cartoon characters, and heavily stylized brand names [10]. Among youth, there is an increased interest in trying fully branded e-cigarette packaging when compared to e-cigarette products with plain packaging [12-14]. E-cigarette package attractiveness is also enhanced by a variety of flavors, vibrant colors, and other branding that targets specific subgroups [12,15].

In addition to company branding (names, colors, textual elements, imagery, shapes, and sizes), e-cigarette packaging can also include HWLs, nicotine concentration, and ingredient lists. HWLs on e-cigarette packaging are associated with reduced e-cigarette use initiation and higher intentions to quit vaping among adults and youth [15-19].

Study Overview

The Tobacco Pack Surveillance System (TPackSS) study was initiated in 2012 with the goal to collect and analyze tobacco products (cigarettes, kreteks, bidis, and straw cigarettes) and their packaging in low- and middle-income countries with the greatest number of people who smoke. Data collection started in 2013 [20]. Roll your own and smokeless tobacco were added to tobacco product collection in 2015 and 2016, respectively, and heated tobacco products and nicotine pouches were added to the types of products to be collected in 2021 and 2022, respectively. The data collection protocol aims to collect 1 of every unique package available for sale in each country. The data generated from TPackSS data collection can inform policies that reduce appeal and attractiveness of tobacco product packaging and more effective HWL implementation and enforcement.

Beginning in 2022, e-cigarettes were added to the TPackSS data collection protocol. In the context of this paper, the term “e-cigarettes” refers to disposable e-cigarette devices and/or their accompanying products: e-cigarette liquids and pods or cartridges. It also can be used to refer to the packaging of these products when discussing them together as a whole. It does not include reusable e-cigarette devices such as tank systems or e-cigarette accessories and parts. In the work described here, e-cigarettes are not included in our definition of tobacco products.

E-cigarettes were an important product type to add to surveillance as the market has grown significantly since their debut in 2007 [21]. As of 2025, a total of 12 million youth aged 13 to 15 years across 110 countries reported current use of e-cigarettes [22]. Surveillance is critical in a market where regulations vary across the globe, including many jurisdictions with no or partial regulations [22]. The TPackSS protocol needed to be adapted for e-cigarette products because the retailer landscape often differs between e-cigarette and tobacco vendors (eg, e-cigarettes are commonly sold at vape shops). In addition, they are often shaped differently than tobacco products, can be more heterogeneous, and consist of different components (eg, liquids compared to sticks). Indonesia was the first country in which e-cigarettes were collected as part of the TPackSS study in 2022. Since then, e-cigarettes were also collected in China in 2023, with other country collections planned.

This paper presents an adaptation of the existing TPackSS protocol [20], originally developed for tobacco products, to optimize it for collection of e-cigarettes in parallel with tobacco products and the process for adaptation, including implementation evaluation and protocol refinements. We describe the specific protocol changes made based on product and market differences unique to e-cigarettes and reflect on our experience implementing and refining the adapted protocol in Indonesia and China. Through these case studies, we highlight key methodological considerations for conducting similar research in jurisdictions with varied e-cigarette landscapes and illustrate how these considerations may be incorporated into the protocol.

Methods

Overview

The original TPackSS protocol was adapted for e-cigarettes using an iterative process to enhance validity and to ensure revisions aligned with the study objectives while maintaining methodological rigor. Initial adaptations were made based on our knowledge of the e-cigarette landscape, country-specific regulations, and e-cigarette products and marketing, informed by the peer-reviewed literature, white papers, Euromonitor data, and consultation with experts in e-cigarette marketing research and in-country public health professionals. All adaptations to the protocol were extensively discussed and agreed upon by the TPackSS study team, comprising researchers with expertise in tobacco advertising, tobacco package HWLs, and packaging design, who have conducted research on tobacco packaging for

XSL•FO
RenderX

13 years. Using Indonesia as a pilot country, implementation of each phase of the adapted protocol was monitored, and the research team held regular debriefings to identify challenges, assess adherence to the protocol, and address any needed clarifications. Reflections made during implementation of the protocol in Indonesia also informed further adaptations to the China protocol. This iterative process of revisions, observation, group discussion and reflection, and refinement enhanced feasibility and data quality while preserving the study's core objectives.

The original TPackSS protocol included a comprehensive overview of each phase of the research from data collection to dissemination of findings [20]. Here, we present how each step was adapted for e-cigarettes and the factors considered when making adaptations. Descriptions of how the protocol differed in Indonesia and China provide examples of how the protocol can be customized to reflect jurisdiction-specific factors while still adhering to the key steps of the protocol. Table 1 provides a summary of the key differences in the original TPackSS protocol and the protocol developed for e-cigarettes.

Table 1. Similarities and differences between the original tobacco product protocol and electronic cigarette (e-cigarette) protocol.

Steps or procedures included in the protocol	Original tobacco product protocol	E-Cigarette protocol
Products collected	Cigarettes, kreteks, bidis, straw cigarettes, roll your own, smokeless, and heated tobacco product sticks	Disposable e-cigarettes, pods or cartridges, and e-cigarette liquids (must contain nicotine)
Vendor types	Various tobacco vendor types based on country context	Various tobacco vendor types based on country context and vape stores
Selecting venues for product purchases	Walking protocol	Preselected vape stores in Indonesia and walking protocol in China
Photography guide	Standard photography guide	In addition to the standard photography guide, steps were added to account for extra packaging layers and variety of e-cigarette product packaging
Shipment and disposal	Shipped product and packaging back to office in the United States	Shipped empty packaging back to office in the United States
Translation	Used professional translation services	Used Chinese-speaking coders (China) or Google Translate (Indonesia)
Codebook	Separate intake, HWL[a] compliance, and feature and appeals codebooks	Combined intake, HWL compliance (when applicable), and features and appeals into a single codebook and added a nicotine-related section
Features highlighted on the TPackSS[b] website	HWL noncompliance, flavor appeal, brand family, product type, country, city, collection date, price, less harm appeal, lipstick pack, and stick count	HWL noncompliance, flavor appeal, brand family, product type, country, city, collection date, and price

[a]HWL: health warning label.

[b]TPackSS: Tobacco Pack Surveillance System.

Study Design and Sampling Framework

Defining Inclusion and Exclusion Criteria

Only e-cigarette products that were prepackaged with liquid containing nicotine were purchased, namely, disposable e-cigarette devices and e-cigarette refills (cartridges or pods or premixed liquids).

Determining the Geographic Scope

The methods used for selection of cities remained the same as the original protocol, apart from adding 1 city in China [20]. In Indonesia, data were collected in Jakarta, Medan, and Surabaya. In China, data were collected in Shanghai, Beijing, Chongqing, Guangzhou, Kunming, and Shenzhen. Shenzhen was added because of its central role in global e-cigarette production.

In the original protocol, 12 neighborhoods per city, stratified by high, middle, and low socioeconomic status (SES), were selected based on diversity in terms of geographic and residential composition [20]. Neighborhood SES in Indonesia was determined by the number of households living in poverty, as reported in the government's integrated social welfare database. Neighborhood SES in China was determined by land area, population, and property prices obtained primarily from government websites. To factor in a low concentration of vape stores in Indonesia, 3 additional neighborhoods in each city, 1 per SES strata, were visited, dependent on where vape stores were located; a total of 45 neighborhoods were visited. In China, where vape stores are relatively concentrated, in addition to the original criteria used for selection of neighborhoods, each neighborhood also had to contain a cluster of vape stores; a total of 72 neighborhoods were visited in China.

Defining the Sampling Frame

In the original protocol, vendor types were selected based on Euromonitor data on where tobacco products are most frequently sold [20]. To adapt the protocol for e-cigarettes, vape stores were added to the list of vendor types based on Euromonitor data [23]. In China, 2 types of vape stores were identified: (1) branded vape stores that sell primarily 1 name brand (eg, RELX, YOOZ, Snowplus, and MOTI) and (2) nonbranded vape stores that sell many different e-cigarette brands.

XSL•FO
RenderX

Developing the Sampling Strategy

In the original protocol, at the first tobacco vendor in each city, every unique tobacco product package for sale was purchased in what was referred to as the index purchase [20]. Similar to the original protocol, every unique e-cigarette package was purchased at the first vape store in each city. The definition of a "unique package" was expanded for e-cigarettes to account for e-cigarette–specific characteristics, defined as any package with at least 1 difference in an exterior feature of the pack, including volume, flavor, nicotine strength, brand name presentation, descriptor, or colors.

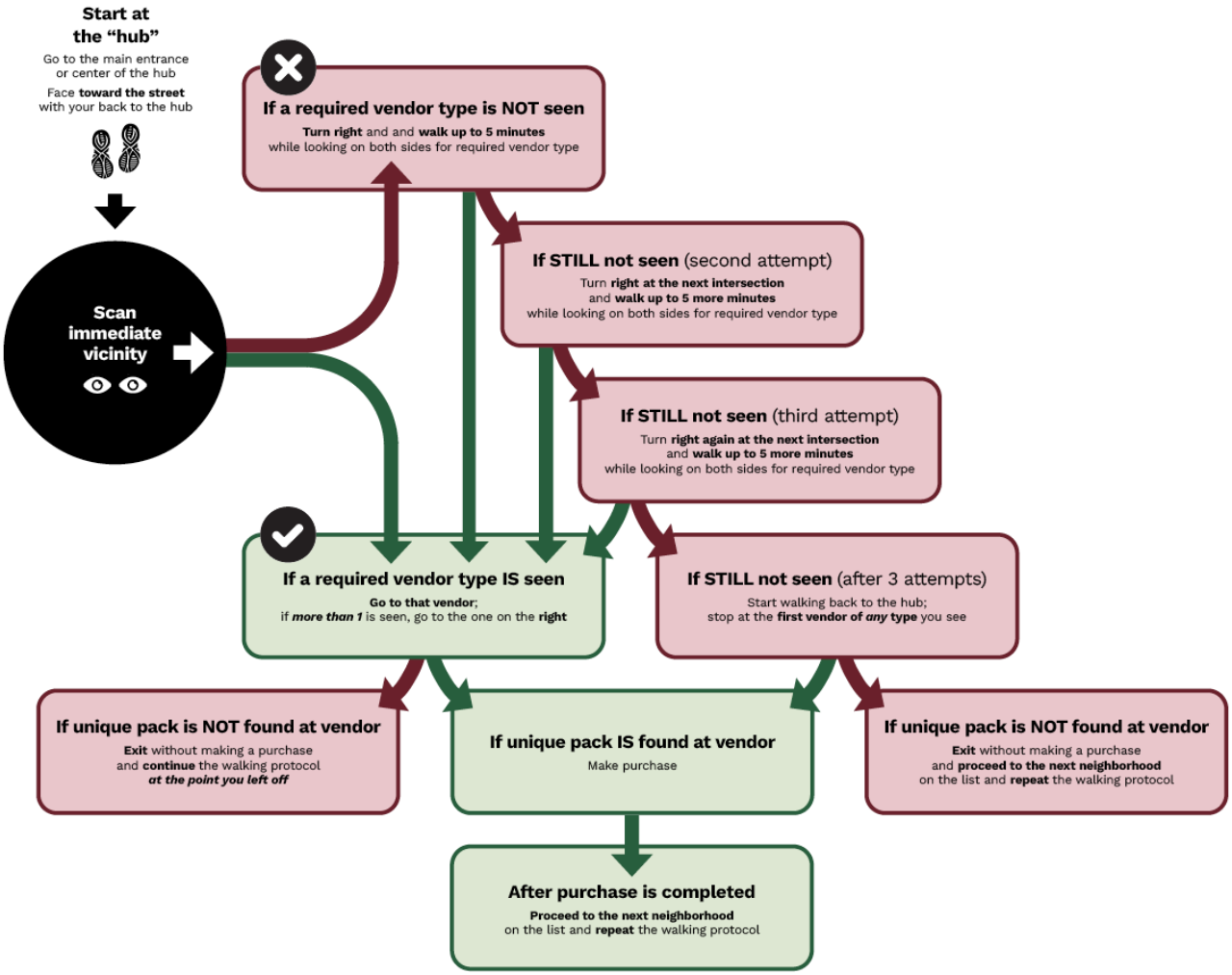
Identification and selection of vape stores was intertwined with the selection of neighborhoods. Being less concentrated in Indonesia, the 3 vape stores selected in each city were those with the greatest variety of products, determined based on online descriptions and pictures, with the requirement that they not belong to the same brand chain or owner [24].

After the 3 preselected vape stores were visited in each city in Indonesia, the original walking protocol was followed by data collectors in the remaining 12 neighborhoods in each city. Data collectors purchased unique e-cigarette packages that had not been previously collected from vape stores if they were identified at subsequently visited tobacco vendors.

In China, with a relatively high concentration of vape stores, the original walking protocol was adapted to include identification of branded and nonbranded vape stores. Following the adapted walking protocol, data collectors identified and visited vape stores and tobacco vendors simultaneously.

When following the walking protocol, in each neighborhood, data collectors started at a hub, defined as a central area in the neighborhood. The hub could be a major transit point (eg, bus terminal and train station), historical landmark, tourist attraction, major shopping center, large plaza, square, or major intersection and source of commerce. In the first city visited in China, data collectors started in a neighborhood classified as middle SES and purchased all unique e-cigarette packages from the first vape store they visited. If unique packages were not found at the first vendor visited, data collectors visited up to 2 additional e-cigarette vendors in a given neighborhood until at least 1 unique package was found (Figure 1). At the second and third neighborhood, similar procedures were followed, except either a brand or nonbrand vape store had to be visited depending on the type of vape stores visited previously to ensure both types were visited and variety was achieved. In the remaining 8 neighborhoods, the same procedures for the third neighborhood were used. The entire method was repeated for the remaining 5 cities in China [24].

Figure 1. Main walking protocol steps in China. This walking protocol was also followed in Indonesia for tobacco product collection. While not the primary focus, unique electronic cigarette packages were purchased if they were sold at tobacco vendors identified following the protocol.



XSL•FO
RenderX

Package Collection

The training for data collectors was adapted to contain e-cigarette–specific content and procedures. The specific procedures followed by data collectors to identify unique e-cigarette packages, label e-cigarette packages with unique IDs, create an e-cigarette image archive, document the details of each e-cigarette purchase at the neighborhood level, and create an inventory of e-cigarette packages purchased mirrored the procedures used for tobacco products, as described in the original protocol [20].

E-Cigarette Photography Guide

The photography guide for e-cigarettes followed the original guide [20], with instructions added to capture additional e-cigarette pack images as needed to account for multiple layers of packaging not commonly found for tobacco products and the greater variety of e-cigarette packaging shapes relative to tobacco products. All layers of e-cigarette packaging were captured, and for e-cigarette packages with unique shapes, all sides that contained descriptors, imagery, or HWLs were captured.

Shipping and Receiving the Packages

We received updated approval from the US Food and Drug Administration and the US Federal Trade Commission to import and store empty e-cigarette packages for research purposes. E-cigarette products were disposed of prior to shipping; only empty packaging was imported due to safety concerns with shipping batteries and liquid, an issue not encountered with tobacco products. As with tobacco products, packages are stored in our office space in labeled and cataloged boxes. Unlike the intake procedures described previously for tobacco products, intake for e-cigarettes was completed at the same time as coding [20].

Translation

Since publication of the original protocol, professional translation services were substituted with Google Translate to translate all non-English text on packaging, including tobacco and e-cigarette packaging, into English. Research assistants now use the image translation feature available via the Google Translate smartphone app to record English translations in Microsoft Excel spreadsheets for reference during coding [24]. For the coding of Chinese e-cigarette packages, all team members were able to read Chinese; therefore, translation was not required.

Creating an E-Cigarette Codebook

The e-cigarette codebook was developed through an iterative process involving incorporation of existing measures, consultation with subject matter experts, in-person review of e-cigarettes collected in Indonesia, and pilot testing followed by refinement. First, we combined our previously separate tobacco product codebooks on intake, warning labels, design features, and marketing appeals into a single codebook to be adapted for e-cigarettes. We then added questions on external packaging, nicotine type and presentation, and cessation claims and expanded flavor and youth appeal response options based on the literature on e-cigarettes and personal observation of e-cigarette packaging. Next, we conducted a literature review to identify existing content analyses of e-cigarette advertising and packaging, identifying 15 relevant peer-reviewed publications [9,25-38]. We conducted informational meetings with research teams that authored 5 of these publications, which used data from Greece, England, and the United States. Prior to the meetings, research teams received a list of proposed codebook constructs (ie, product information, nicotine content, flavors, packaging features, warning labels and messages, marketing messages and images, and advertising tactics). During the meetings, researchers provided feedback on these constructs and responded to targeted questions about their experiences in codebook development, features they would have liked to have coded, coding challenges, and unexpected findings. Detailed notes were taken to document lessons learned, suggestions for improvement, and recommendations for additional codebook constructs. Insights gathered were incorporated into the first full draft of the codebook to enhance its practicality and relevance.

To aid codebook development, we also viewed the physical e-cigarette packages collected in Indonesia in person. We found that compared to cigarettes, e-cigarettes from Indonesia came in a greater variety of shapes and displayed unique and varied design elements. The design also seemed more tailored to a younger audience, consistent with the literature on cigarette pack design elements that make cigarettes more attractive to youth and young adults (Figure 2) [39,40]. On the basis of these observations, we added questions and response options to the codebook to capture specific packaging elements likely to appeal to youth (eg, cartoons and fun fonts).

XSL•FO
RenderX

Figure 2. Comparison between cigarette and electronic cigarette (e-cigarette) packages collected in Indonesia, 2022. All e-cigarette images featured are the external packaging of e-cigarette liquids.

Next, we pilot-tested the codebook with coders and further revised the codebook and accompanying coding guide. A random subsample of 125 packages from Indonesia were used to pilot the first full version of the e-cigarette codebook. Revisions were made for clarity to address package characteristics not previously addressed and to eliminate items that had low interrater reliability. For example, we found the units used to communicate nicotine concentration varied (eg, percentage, “mg,” and “mg/mL”) and added response options accordingly. We also found that “masculine” and “feminine” appeals and brand names could not be determined reliably; these codebook questions were eliminated or edited.

A high level of intercoder agreement was reached for most variables (80%); however, for marketing appeals, there was a trade-off between simplifying the codebook to achieve higher agreement and including appeals that could provide greater context to the marketing being used on the packaging, resulting in lower levels of initial agreement (50%-70%) on some variables.

The codebooks for Indonesia and China e-cigarettes also varied due to different packaging and labeling regulations. At the time of data collection, e-cigarettes were unregulated in Indonesia; therefore, no questions applicable to compliance with existing policies were included in the Indonesia e-cigarette codebook, and only the external packaging layer visible to the consumer at point of sale was coded. Given the existing e-cigarette packaging and labeling policies in China at the time of data collection, the China e-cigarette codebook included related items, and coders were instructed to observe all layers of the packaging and the product. Further revisions were made to the China codebook based on packaging elements specific to Chinese e-cigarettes, additional pilot testing of 10 e-cigarette packages from China, and decisions to simplify the codebook after determining that certain items were not of high priority.

The base codebook structure used in both countries contained 4 sections: product information, nicotine, warning labels, and marketing appeals (Table 2), with some codes only included in the Indonesia or China codebook based on lessons learned and country-specific regulations and context.

XSL•FO
RenderX

Table 2. Sections of the codebook and variables used in Indonesia and China (full codebooks are available on the Tobacco Pack Surveillance System (TPackSS) website [24])[a].

Section	Variables
Basic product information	• Duplicate pack status[b] • Purchase date • Price • Product type • External packaging layers • Number of puffs as stated on the pack (for disposables)[b] • Number of pods or cartridges[b] • Brand names • Manufacturer name[b] • Manufacturer location[b] • Presence of English[c] • Web presence[c] • Insert or onsert
Nicotine	• Mention of nicotine • Type of nicotine composition • Total amount of e-cigarette[d] liquid • Nicotine concentration units • Nicotine concentration amount
Health warning and labeling–related items[e]	• Presence of text and/or image warnings[c] • HWL[f] location (for each layer: product, internal packaging, and external packaging)[b] • Observance of HWL through the internal packaging layer[b] • HWL obstruction (for each layer)[b] • HWL in Chinese language (for each layer)[b,g] • HWL text (for each layer)[b,g] • HWL located on bottom of the pack face (for product and external packaging)[b,g] • Branding beneath HWL (for product and external packaging)[b,g] • HWL color (for product and external packaging)[b,g] • HWL text contrast (for product and external packaging)[b,g] • Part of the HWL not present (for product and external packaging)[b,g] • HWL area (for external packaging)[b,g] • Misleading descriptors (for any layer)[b,h] • Health-promoting descriptors (for any layer)[b,h] • Minor protection descriptors (for any layer)[b,h] • Minor protection imagery (for each layer)[b,h]

XSL•FO
RenderX

Section	Variables
Features and appeals (coded for on any layer in China and only for the most external visible packaging layer in Indonesia)	• Branding information on insert or onsert[b] • Presence of branding information on each packaging layer[b] • Web presence (for each layer)[b] • Presence of company phone number (for each layer)[b] • Marketing claims • Claim of cessation aid[c] • Mention of recognition by award or standard[b] • Environmentalism and civic responsibility lexical and imagery appeals[b] • Technology appeal[b] • Taste or sensation terminology • Flavor terminology or imagery (for external packaging) • Full flavor name[c] • Description of tobacco flavor (for external packaging)[b] • Feminine terminology or imagery[c] • Masculine terminology or imagery[c] • Presence of youth appeal[c] • Specific youth appeals[c] • Youth appeals by text, imagery, and shape[b] • China-specific cultural appeals[b] • Global or luxury or quality appeals[b] • Other notable appeals

[a]Web presence appears in separate sections of the Indonesia (basic product information) and China (features and appeals) codebooks.

[b]Variable appears only in the China codebook.

[c]Variable appears only in the Indonesia codebook.

[d]E-cigarette: electronic cigarette.

[e]All items included in the China codebook are government mandated.

[f]HWL: health warning label.

[g]Repeated for each HWL.

[h]Variables were coded for presence on any layer of the packaging, instead of focusing solely on the HWL.

Coding

Coding of e-cigarette packages followed the same procedures as previously described for tobacco products [20]. Packages were double-coded independently, and discrepancies were resolved by a third reviewer. Since 2021, we have transitioned to coding all products online using pack images generated using the photography guide rather than in person using the physical packages.

Coders used an extensive coding guide with questions, variable definitions, and package examples as a reference during coding. During a virtual training, coders were introduced to the codebook, coding guide, and coding platform Research Electronic Data Capture (REDCap; Vanderbilt University) and practiced coding packages as a group [41]. As with tobacco products, weekly coding review meetings were held to discuss any unusual packages encountered during coding and resolve coding discrepancies.

For Indonesia, separate qualitative coding of cultural appeals was conducted because the original coders were not all native to Indonesia or closely familiar with its cultural context and, therefore, were not able to adequately assess the meaning of some package elements. Two native Bahasa Indonesian speakers who were raised in Indonesia examined all packages and made an initial assessment of whether text or imagery referred to anything related specifically to Indonesia (eg, a popular character in pop culture, food, or a slang phrase); discrepancies were discussed and resolved by a third party when needed. Next, one of these initial coders reviewed the images of the e-cigarette packages again and wrote a short qualitative description (2-4 sentences) of the cultural appeal. A third study member, who was born in Indonesia, conducted a thematic analysis of the cultural appeals across packages and summarized the results. A fourth study member, who was born in Indonesia, reviewed all packages to determine if cultural appeal overlapped with youth appeal and, when appropriate, categorized and described the culturally specific youth appeal. Two additional Indonesians reviewed these youth appeal categories and descriptions and provided additional insight on whether the appeals were mostly known and liked by younger consumers or if they also appealed to a wider audience. For China, cultural appeals were coded quantitatively as part of the standard codebook because all members of the study team were native to or closely familiar with the Chinese cultural context [24].

XSL•FO
RenderX

Results

Packages Collected

The TPackSS study began in 2012 with funding from Bloomberg Philanthropies' Bloomberg Initiative to Reduce Tobacco Use. Between September and October 2022, we collected 825 unique packages in Indonesia, comprising 782 e-cigarette liquid bottles, 26 disposable e-cigarette devices, and 17 disposable pods or cartridges. Of the 825 unique packages collected in Indonesia, 483 were collected in Jakarta, 203 in Surabaya, and 139 in Medan. Between April and May 2023, we collected 143 unique packages in China, comprising 5 disposable e-cigarette devices and 138 disposable pods or cartridges. No e-cigarette liquid bottles were collected in China. This is likely because the law prohibits the sale of e-cigarette products that allow people to add their own e-liquid to a device [42]. Of the 143 unique packages collected in China, 41 were collected in Shanghai, 31 in Beijing, 6 in Chongqing, 25 in Guangzhou, 15 in Kunming, and 25 in Shenzhen. Data analyses for Indonesia are completed, and secondary analyses are being completed for China. Detailed findings from these collections are being disseminated through separate publications scheduled for 2026 [43].

Presentation on the TPackSS Website

As with our tobacco product collections, we upload images of the e-cigarette packages to the TPackSS searchable online database [20]. Similar to tobacco products, information about a specific pack is provided along with its photos. Individual e-cigarette packages have the same description categories as tobacco products: brand family, product type, country, city, collection date, and price. Additionally, similar to tobacco products, these packages can be filtered and searched by product type, country, brand family, collection year, and specific features. Some of the current features available for tobacco products (ie, HWL noncompliance and flavor appeal) are applicable to e-cigarettes.

The TPackSS website also features a "Share A Pack" section where users can upload and share photos of tobacco packs from their country. These user submissions contain similar description categories and searchable features to the official uploaded TPackSS collections. This user submission feature has been expanded to encourage uploads of e-cigarette products.

As with previous tobacco product collections, in addition to uploading photos of the packages collected to the project website, we are continuing to create and disseminate country-specific fact sheets. These fact sheets are available for download from the website [24]. The field procedures and codebooks used for these collections and our publications on TPackSS findings are also available via the website.

Discussion

This protocol provides procedures for systematically purchasing unique e-cigarette products; assessing the products and their packaging for HWL compliance, nicotine information, and design features and marketing appeals; and disseminating the findings. This protocol can be adapted to other countries by taking into consideration the unique policy and market context. The protocol is applicable and adaptable to multiple policy and market contexts. It draws strength from being adapted from the TPackSS project's years of experience conducting similar work with tobacco products. The protocol has also been informed by the experience of other research teams that have conducted research on e-cigarette packaging.

The adapted protocol was successfully used to obtain or purchase a wide collection of unique e-cigarette packages sold in Indonesia and China. E-cigarette coding resulted in policy-relevant findings that address the study's main objectives to assess e-cigarette compliance with country-specific packaging and labeling regulations and identify marketing tactics being used on packaging. Findings have been disseminated through fact sheets, peer-reviewed journal publications, academic presentations, and communications with public health professionals in Indonesia and China.

The TPackSS e-cigarette protocol does have some limitations. While the data collection protocol is designed to ideally capture every unique product on the market, it is possible that some unique products are available at vendors we did not visit. In addition, this approach does not allow for the capture of the actual market share of collected products as information on sales is not collected. Products exclusively sold through online retailers or vendors are not captured. For countries where not all coders are culturally tied to the study country, it is possible that specific cultural nuances are missed during coding. Furthermore, coding of appeals such as youth and cultural references involves subjective interpretation. Despite these limitations, this protocol provides a valuable structure for monitoring e-cigarette products and packaging as well as compliance with HWL regulations. Beyond narrative findings, the protocol also ensures the production of visual data that can be incorporated into dissemination materials for diverse stakeholders and aids clear and accessible communication of study findings.

The e-cigarette market is changing rapidly, and continual monitoring is required to inform regulations to protect the public's health. TPackSS data collection of e-cigarettes will be repeated in the future based on warning labels and packaging requirement changes. The findings from these collections can inform policy by providing insights into the design features and marketing appeals of e-cigarette products available on the market, as well as compliance with HWL requirements where applicable.

XSL•FO
RenderX

Acknowledgments

The authors would like to acknowledge the entire TPackSS team, including the work of our many coders: Chunlin (Eileen) Ren, Emily Xing, Natalie Nguyen, Jaydon Martinez, Raniyan Zaman, Rozanne Lim, Xinao (Lorraine) Luo, Yohana Revi Imanita, and Zhehan Wang. The authors would also like to thank those who assisted with the qualitative coding in Indonesia: Dr Beladenta Amalia, Joshua Sinamo, Adhiyatma Fuaddy, Yohana Revi Imanita, Nabilah Kusuma Wardhani, and Helen Natalia. In addition, the authors acknowledge the invaluable help of the researchers they consulted during codebook development and their in-country partners who assisted with data collection, translation, and disposal of the electronic cigarette products: Tobacco Control Support Center, Dr Abdillah Ahsan and his research team at the University of Indonesia, Rising Sun, and Dr Pinpin Zheng at Fudan University. Furthermore, the authors acknowledge Dr Beladenta Amalia for her contributions in leading the 2022 Tobacco Pack Surveillance System Indonesia data collection. The authors would also like to thank Jared Earley for his contribution to the design of the flowchart figure.

Funding

This work was supported by the Bloomberg Philanthropies' Bloomberg Initiative to Reduce Tobacco Use.

Data Availability

The data supporting this research are available from the authors on reasonable request.

Authors' Contributions

JEC and KCS developed the original Tobacco Pack Surveillance System (TPackSS) protocol. JLB managed the project, and all authors contributed to the development of the modified TPackSS protocol and codebook for electronic cigarettes. QN managed data collection and coding of the China data. HEB and RH oversaw coding of the Indonesia data. RH and QN performed the statistical analyses. RH drafted the initial manuscript. All authors provided critical feedback on the manuscript and approved the final version.

Conflicts of Interest

JEC was a paid consultant in litigation against a tobacco company.

References

1. Jerzyński T, Stimson G. Estimation of the global number of vapers: 82 million worldwide in 2021. Drugs Habits Soc Policy 2023 Apr 17;24(2):91-103 [FREE Full text] [doi: 10.1108/DHS-07-2022-0028]
2. WHO global report on trends in prevalence of tobacco use 2000–2024 and projections 2025–2030. World Health Organization. 2025 Oct 06. URL: https://www.who.int/publications/i/item/9789240116276 [accessed 2025-10-21]
3. Grilo G, Iacobelli M, Welding K. International regulatory landscape for e-cigarettes and heated tobacco products. Institute for Global Tobacco Control, Johns Hopkins Bloomberg School of Public Health. 2021 Mar 2. URL: https://www.gov.br/anvisa/pt-br/assuntos/regulamentacao/participacao-social-antigo/tomada-publica-de-subsidios/arquivos/tomada-publica-de-subsidios-no-6-de-11-04-2021/anvisa_report_jhu_igtc.pdf [accessed 2025-10-09]
4. Atella L, Duren M, Cohen J, Welding K, Kennedy RD. Nature and extent of country-level policies regulating flavors in electronic nicotine delivery systems (ENDS) [Poster]. In: Proceedings of the Society for Research on Nicotine & Tobacco Annual Meeting. 2023 Presented at: SRNT 2023; March 1-4, 2023; San Antonio, TX URL: https://www.globaltobaccocontrol.org/sites/default/files/2023-02/SRNT_2023_LA_Policy_Flavors_ENDS.pdf?utm_source=schedule&utm_medium=web&utm_campaign=Laraflavors&utm_id=SRNT [doi: 10.1093/ntr/nts151]
5. Kennedy RD, Atella L, Duren M, Welding K, Cohen J. Country-level policies regulating nicotine strength and liquid volumes for electronic nicotine delivery systems (ENDS) [Poster]. In: Proceedings of the Society for Research on Nicotine & Tobacco Annual Meeting. 2023 Presented at: SRNT 2023; March 1-4, 2023; San Antonio, TX URL: https://www.globaltobaccocontrol.org/sites/default/files/2023-02/SRNT_2023_RDK_Policy_Strength_Liquid_ENDS.pdf [doi: 10.1093/ntr/nts151]
6. Atella L, Duren M, Cohen J, Welding K, Kennedy RD. Prevalence of policies requiring health warning labels on electronic nicotine delivery systems and heated tobacco products around the globe. In: Proceedings of the Society for Research on Nicotine & Tobacco Annual Meeting. 2023 Presented at: SRNT 2023; March 1-4, 2023; San Antonio, TX URL: https://www.globaltobaccocontrol.org/sites/default/files/2023-03/SRNT_2023_LA_HWL.pdf
7. Wakefield M, Morley C, Horan JK, Cummings KM. The cigarette pack as image: new evidence from tobacco industry documents. Tob Control 2002 Mar;11 Suppl 1(Suppl 1):I73-I80 [FREE Full text] [doi: 10.1136/tc.11.suppl_1.i73] [Medline: 11893817]

XSL•FO
RenderX

8. State of evidence review: health warning labels on tobacco. Global Tobacco Control. 2013 Oct. URL: http://globaltobaccocontrol.org/sites/default/files/HealthWarnings_state_of_evidence_final_11_18_2013_web_0.pdf [accessed 2025-10-09]
9. Kong AY, Derrick JC, Abrantes AS, Williams RS. What is included with your online e-cigarette order? An analysis of e-cigarette shipping, product and packaging features. Tob Control 2018 Nov;27(6):699-702 [FREE Full text] [doi: 10.1136/tobaccocontrol-2016-053061] [Medline: 27357936]
10. Nottage MK, Taylor EV, East KA, McNeill A, Thrasher JF, Reid JL, et al. Packaging of disposable vaping products and e-liquids in England, Canada and the United States: a content analysis. Addiction 2025 Mar;120(3):483-495. [doi: 10.1111/add.16611] [Medline: 38970447]
11. FDA warns retailers and manufacturers over e-cigarettes resembling smart devices. Tobacco Twenty One. 2024 Nov 2. URL: https://tobacco21.org/fda-warning-letters-e-cigarettes-targeting-youth/ [accessed 2025-10-09]
12. Gomes MN, Reid JL, Hammond D. The effect of branded versus standardized e-cigarette packaging and device designs: an experimental study of youth interest in vaping products. Public Health 2024 May;230:223-230. [doi: 10.1016/j.puhe.2024.02.001] [Medline: 38429123]
13. Taylor E, Arnott D, Cheeseman H, Hammond D, Reid JL, McNeill A, et al. Association of fully branded and standardized e-cigarette packaging with interest in trying products among youths and adults in Great Britain. JAMA Netw Open 2023 Mar 01;6(3):e231799 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.1799] [Medline: 36917111]
14. Simonavičius E, East K, Taylor E, Nottage M, Reid JL, Arnott D, et al. Impact of e-liquid packaging on vaping product perceptions among youth in England, Canada, and the United States: a randomized online experiment. Nicotine Tob Res 2024 Feb 22;26(3):370-379 [FREE Full text] [doi: 10.1093/ntr/ntad144] [Medline: 37542732]
15. Shang C, Huang J, Chaloupka FJ, Emery SL. The impact of flavour, device type and warning messages on youth preferences for electronic nicotine delivery systems: evidence from an online discrete choice experiment. Tob Control 2018 Oct;27(e2):e152-e159 [FREE Full text] [doi: 10.1136/tobaccocontrol-2017-053754] [Medline: 29097588]
16. Avery RJ, Kalaji M, Niederdeppe J, Mathios A, Dorf M, Byrne S, et al. Perceived threat and fear responses to e-cigarette warning label messages: results from 16 focus groups with U.S. youth and adults. PLoS One 2023 Jun 23;18(6):e0286806 [FREE Full text] [doi: 10.1371/journal.pone.0286806] [Medline: 37352255]
17. Brewer NT, Jeong M, Hall MG, Baig SA, Mendel JR, Lazard AJ, et al. Impact of e-cigarette health warnings on motivation to vape and smoke. Tob Control 2019 Jul 10;28(e1):e64-e70 [FREE Full text] [doi: 10.1136/tobaccocontrol-2018-054878] [Medline: 31292169]
18. Lee HY, Lin HC, Seo DC, Lohrmann DK. The effect of e-cigarette warning labels on college students' perception of e-cigarettes and intention to use e-cigarettes. Addict Behav 2018 Jan;76:106-112. [doi: 10.1016/j.addbeh.2017.07.033] [Medline: 28779651]
19. Nguyen T, Shah G, Barefield AC. The influence of e-cigarette warning labels on youths' use intentions - a mediation analysis of role of perceived harm. Subst Use Misuse 2023;58(5):709-716. [doi: 10.1080/10826084.2023.2184205] [Medline: 36861946]
20. Smith K, Washington C, Brown J, Vadnais A, Kroart L, Ferguson J, et al. The Tobacco Pack Surveillance System: a protocol for assessing health warning compliance, design features, and appeals of tobacco packs sold in low- and middle-income countries. JMIR Public Health Surveill 2015 Aug 12;1(2):e8 [FREE Full text] [doi: 10.2196/publichealth.4616] [Medline: 27227142]
21. Noel JK, Rees VW, Connolly GN. Electronic cigarettes: a new 'tobacco' industry? Tob Control 2011 Jan;20(1):81. [doi: 10.1136/tc.2010.038562] [Medline: 20930060]
22. WHO report on the global tobacco epidemic, 2025: warning about the dangers of tobacco. World Health Organization. 2025 Jun 23. URL: https://www.who.int/publications/i/item/9789240116276 [accessed 2026-03-03]
23. Smokeless tobacco, e-vapour products and heated tobacco in China. Euromonitor International. 2021 Jun. URL: https://www.euromonitor.com/smokeless-tobacco-e-vapour-products-and-heated-tobacco-in-china/report [accessed 2023-02-01]
24. Resources. TPackSS: Tobacco Pack Surveillance System. URL: https://tpackss.globaltobaccocontrol.org/resources [accessed 2025-10-09]
25. Chacon V, Arriaza A, Cavazos-Rehg P, Barnoya J. Availability, price, and packaging of electronic cigarettes and e-liquids in Guatemala city retailers. Nicotine Tob Res 2018 Jan 05;20(2):253-257. [doi: 10.1093/ntr/ntx071] [Medline: 28371843]
26. Buonocore F, Marques Gomes AC, Nabhani-Gebara S, Barton SJ, Calabrese G. Labelling of electronic cigarettes: regulations and current practice. Tob Control 2017 Jan;26(1):46-52 [FREE Full text] [doi: 10.1136/tobaccocontrol-2015-052683] [Medline: 26790924]
27. Buettner-Schmidt K, Miller DR, Balasubramanian N. Electronic cigarette refill liquids: child-resistant packaging, nicotine content, and sales to minors. J Pediatr Nurs 2016;31(4):373-379 [FREE Full text] [doi: 10.1016/j.pedn.2016.03.019] [Medline: 27079973]
28. Buettner-Schmidt K, Miller DR, Orr M, Balasubramanian N, Rykal K, Steward KF, et al. Electronic cigarette refill liquids: nicotine content, presence of child-resistant packaging, and in-shop compounding. J Pediatr Nurs 2021;59:45-54 [FREE Full text] [doi: 10.1016/j.pedn.2020.12.016] [Medline: 33460879]

XSL•FO
RenderX

29. Seitz CM, Orsini MM, Jung G, Butler K. Cartoon images on e-juice labels: a descriptive analysis. Nicotine Tob Res 2020 Oct 08;22(10):1909-1911. [doi: 10.1093/ntr/ntaa029] [Medline: 31996912]
30. Dube SR, Pathak S, Nyman AL, Eriksen MP. Electronic cigarette and electronic hookah: a pilot study comparing two vaping products. Prev Med Rep 2015;2:953-958 [FREE Full text] [doi: 10.1016/j.pmedr.2015.10.012] [Medline: 26740911]
31. Nikitin D, Timberlake DS, Williams RS. Is the e-liquid industry regulating itself? A look at e-liquid internet vendors in the United States. Nicotine Tob Res 2016 Oct;18(10):1967-1972 [FREE Full text] [doi: 10.1093/ntr/ntw091] [Medline: 26995792]
32. Fagan P, Pokhrel P, Herzog TA, Guy MC, Sakuma KL, Trinidad DR, et al. Warning statements and safety practices among manufacturers and distributors of electronic cigarette liquids in the United States. Nicotine Tob Res 2018 Jul 09;20(8):970-976 [FREE Full text] [doi: 10.1093/ntr/ntx101] [Medline: 28520985]
33. D'Mello K, Hammond D, Mahamad S, Wiggers D, East K. Nicotine content, labelling and flavours of e-liquids in Canada in 2020: a scan of the online retail market [Article in English, French]. Health Promot Chronic Dis Prev Can 2022 Jan;42(1):4-11 [FREE Full text] [doi: 10.24095/hpcdp.42.1.02] [Medline: 35044139]
34. Berg CJ, Barker DC, Meyers C, Weber A, Park AJ, Patterson A, et al. Exploring the point-of-sale among vape shops across the United States: audits integrating a mystery shopper approach. Nicotine Tob Res 2021 Feb 16;23(3):495-504 [FREE Full text] [doi: 10.1093/ntr/ntaa041] [Medline: 32149340]
35. Kim S, Goniewicz ML, Yu S, Kim B, Gupta R. Variations in label information and nicotine levels in electronic cigarette refill liquids in South Korea: regulation challenges. Int J Environ Res Public Health 2015 May 05;12(5):4859-4868 [FREE Full text] [doi: 10.3390/ijerph120504859] [Medline: 25950652]
36. Trtchounian A, Talbot P. Electronic nicotine delivery systems: is there a need for regulation? Tob Control 2011 Jan;20(1):47-52. [doi: 10.1136/tc.2010.037259] [Medline: 21139013]
37. Ribisl K, Eaddy J, Cavender S. Standardized tobacco assessment for retail settings: vape shop: vSTARS training manual. State and Community Tobacco Control Research. 2016 Aug 16. URL: http://countertobacco.wpengine.com/wp-content/uploads/2016/06/Final-Training-Manual-for-vSTARS-Beta-Release.pdf [accessed 2023-01-05]
38. Staal Y, Havermans A, van Nierop L, Visser W, Wijnhoven S, Bil W, et al. Conceptual model for the evaluation of attractiveness, addictiveness and toxicity of tobacco and related products: the example of JUUL e-cigarettes. Regul Toxicol Pharmacol 2021 Dec;127:105077 [FREE Full text] [doi: 10.1016/j.yrtph.2021.105077] [Medline: 34748878]
39. Grilo G, Lagasse LP, Cohen JE, Moran MB, Reynales-Shigematsu LM, Smith KC. "It's all about the colors:" how do Mexico City youth perceive cigarette pack design. Int J Public Health 2021 Mar 10;66:585434 [FREE Full text] [doi: 10.3389/ijph.2021.585434] [Medline: 34744557]
40. Borland R, Savvas S, Sharkie F, Moore K. The impact of structural packaging design on young adult smokers' perceptions of tobacco products. Tob Control 2013 Mar;22(2):97-102. [doi: 10.1136/tobaccocontrol-2011-050078] [Medline: 22166265]
41. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
42. Main policies. Tobacco Control Laws. 2022. URL: https://www.tobaccocontrollaws.org/legislation/china/e-cigarettes [accessed 2025-10-09]
43. Barker HE, Brown JL, Hartmuller R, Amalia B, Luo X, Zaman R, et al. Content analysis of e-cigarette packaging in Indonesia: youthful, fun and flavorful. Tob Control 2026 May 20:1 [FREE Full text] [doi: 10.1136/tc-2026-060081] [Medline: 42161505]

Abbreviations

e-cigarette: electronic cigarette
HWL: health warning label
REDCap: Research Electronic Data Capture
SES: socioeconomic status
TPackSS: Tobacco Pack Surveillance System

Edited by B Biswas; submitted 11.Dec.2025; peer-reviewed by R Mejia, O Ruokolainen; comments to author 27.Jan.2026; accepted 06.May.2026; published 02.Jul.2026.

Please cite as:
Hartmuller R, Brown JL, Barker HE, Nian Q, Welding K, Cohen JE, Smith KC
Adapting the Tobacco Pack Surveillance System Protocol to Assess Electronic Cigarette Packaging: Protocol for a Content Analysis
JMIR Public Health Surveill 2026;12:e89325
URL: https://publichealth.jmir.org/2026/1/e89325
doi:10.2196/89325
PMID:42390918

XSL•FO
RenderX

©Reiley Hartmuller, Jennifer L Brown, Hannah Elise Barker, Qinghua Nian, Kevin Welding, Joanna E Cohen, Katherine Clegg Smith. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 02.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Integration of VectorCam, an AI-Enabled Mosquito Surveillance Tool, Into Uganda's National Malaria Elimination Program: Protocol for a Mixed Methods Hybrid Type III Implementation Study

Mei-Li Hey[1*], MSc; Marina Rincon Torroella[2,3*], MPH; Diwakar Mohan[4], MBBS, MPH, MD, DrPH; Sunny Patel[2,3], MSE; Winnifred Kansiime[5], MSc; Catherine Maiteki-Sebuguzi[6], MBChB, MSc; Aryaman Shodhan[2], MSE; Youseph Yazdi[2], MBA, PhD; Peter Waiswa[7], MPH, MD, PhD; Soumyadipta Acharya[2,3], MBBS, MS, PhD

[1]Department of Environmental Health and Engineering, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States
[2]Department of Biomedical Engineering, Whiting School of Engineering, Johns Hopkins University, Baltimore, MD, United States
[3]Vector Control Innovations, Inc, Ellicott City, MD, United States
[4]Department of International Health, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States
[5]Department of Disease Control and Environmental Health, School of Public Health, Makerere University, Kampala, Central Region, Uganda
[6]Uganda National Malaria Elimination Division, Ministry of Health, Kampala, Central Region, Uganda
[7]Department of Health Policy, Planning, and Management, School of Public Health, Makerere University, Kampala, Central Region, Uganda
[*]these authors contributed equally

Corresponding Author:
Mei-Li Hey, MSc
Department of Environmental Health and Engineering
Bloomberg School of Public Health
Johns Hopkins University
615 N Wolfe St
Baltimore, MD, 21205
United States
Phone: 1 (410) 955 3543
Email: meilihey1@gmail.com

Abstract

Background: Malaria remains one of the most significant global health threats, and its control depends on timely, accurate entomological surveillance. In Uganda, routine vector surveillance relies heavily on microscope-based mosquito identification by a limited number of trained entomologists, which can delay reporting. VectorCam is an AI-enabled digital tool to support real-time mosquito identification and can be operated by community health workers, supporting rapid data reporting. After co-designing the implementation package using a human-centered design approach with Uganda's Ministry of Health, we are implementing a surveillance strategy that incorporates VectorCam into routine work.

Objective: This study aims to evaluate acceptability, fidelity, feasibility, cost, and preliminary effectiveness of VectorCam, the intervention, along with supporting implementation strategies as it is integrated into Uganda's national malaria control program.

Methods: This mixed methods hybrid type III implementation-effectiveness study uses district-level restricted randomization to assign 12 districts to four sequential rollout waves following a 2-district pilot phase. The implementation package will be introduced over a one-year period across malaria-endemic districts in Northern and Eastern Uganda. Implementation outcomes, including acceptability, feasibility, fidelity, and cost, will be assessed using qualitative interviews, system usage logs, field observations, and program documentation. Effectiveness outcomes include technical performance of mosquito species identification, time efficiency, and operational use of surveillance data for decision-making. Accuracy will be evaluated through comparison with molecular identification methods, alongside additional classification performance metrics. Data will be analyzed using a mixed methods approach, with triangulation of qualitative and quantitative data and longitudinal analyses across districts and rollout waves. An adaptive learning approach will guide iterative refinement of the implementation package.

Results: Fourteen districts were selected by the Ministry of Health in September 2025 for VectorCam introduction. A pilot phase conducted between December 2025 and January 2026 in two districts was used to refine implementation strategies and research tools through stakeholder workshops. Ethical approval was obtained from Johns Hopkins in January 2026 and Ugandan

XSL•FO
RenderX

ethics committees in September 2025. Implementation began in February 2026 and will continue through March 2027 with concurrent data collection and analysis throughout the implementation timeline.

Conclusions: This study will evaluate a task-shifting approach to entomological surveillance by integrating an AI-enabled tool into routine workflows of community health workers through a co-designed implementation package. By examining how the VectorCam intervention and supporting implementation strategies function together within real-world health system contexts, this study will generate evidence on the sustainability of digitally enabled surveillance. Findings will provide insight into how task shifting and digital tools can expand surveillance capacity, improve timeliness of data reporting, and strengthen the use of entomological data for decision-making. These results will inform national scale-up efforts in Uganda and offer a model for other malaria-endemic settings seeking to modernize vector surveillance.

(JMIR Public Health Surveill 2026;12:e99717) doi:10.2196/99717

KEYWORDS

vector surveillance; malaria; digital technology; implementation research; AI

Introduction

Background

Malaria is a major public health challenge globally, with an estimated 263 million cases and nearly 600,000 deaths recorded in 2023 [1]. Africa is disproportionately affected, as 94% of all malaria cases and 95% of deaths occur on this continent [2]. In Uganda, malaria is endemic in nearly all of the 146 districts and accounts for approximately 30% to 50% of outpatient visits and hospital admissions [3,4]. Despite progress in malaria control over the past 2 decades, recent years have seen a plateau in health gains, highlighting the need for more adaptive, data driven, and locally tailored strategies to work toward elimination [1,5].

A cornerstone of effective malaria control and elimination is timely, accurate entomological surveillance [6,7]. Vector surveillance provides essential data on mosquito species composition, abundance, behavior, and insecticide resistance, all of which are critical for informing targeted interventions such as long-lasting insecticide-treated nets, indoor residual spraying, and larval source management [8-10]. For example, the most common malaria vectors in Uganda are *Anopheles gambiae s.l.* and *Anopheles funestus s.l.* [4]. Previous studies found that certain interventions are more effective in targeting these mosquito groups than others, a fact that underscores the importance of improving the efficiency of surveillance data collection and usage when choosing the correct intervention [11].

However, Uganda and similar locales have a major challenge: the existing vector surveillance systems are limited by their reliance on manual, microscope-based identification of mosquitoes by trained entomologists [12,13]. Uganda's Ministry of Health (MOH) conducts routine monthly mosquito surveillance through district vector control officers (VCOs). The method currently used is a manual process in which mosquitoes are collected, primarily through pyrethrum spray catches (PSCs), human-landing catches (HLCs), or Centers for Disease Control light trap catches (CDC-LTC), and then identified individually under a microscope to determine the genus and species [14-16]. The data are recorded, aggregated, and manually uploaded to centralized health informatics systems such as District Health Information System 2 (DHIS2). The Vector Control Division then uses these data for decision-making and mosquito population mapping [17].

The current vector surveillance methods are labor- and time-intensive and require special expertise. Therefore, generating actionable insights at the pace needed for effective malaria control is challenging [16]. These limitations delay decision-making and reduce the ability to respond rapidly to changes in mosquito populations and malaria transmission dynamics. To address these surveillance gaps, VectorCam was developed at the Johns Hopkins University, Center for Bioengineering Innovation and Design as a low-cost, AI-enabled tool that allows village health teams (VHTs) and other community health workers to perform mosquito identification tasks without prior entomological training [18].

VectorCam, a Field Tool for Rapid Morphological Identification of Mosquitoes

AI-enabled tools are quickly gaining recognition as some of the most effective ways to improve data quality and dissemination [19-23]. The VectorCam platform uses a novel combination of a low-cost, 3D-printed device, smartphone-based microscopy, AI, and a mobile app to enable real-time mosquito identification. The devices use a convolutional neural network (CNN), running locally on a low-cost smartphone to photographically identify species, sex, and abdominal status of each mosquito specimen [18]. Then, the data are uploaded to a central server, which generates reports that are automatically sent to malaria control teams. VectorCam is designed to work in offline and low-bandwidth environments and integrates easily into existing workflows. A feasibility study by Couret et al [24] applied a similar CNN to images of mosquitoes for classification of sex, genus, species, and strains and found accuracy rates of up to 98%. This type of technology has implications in low-resource settings where AI-assisted technologies could close critical capacity gaps, such as when trained entomologists and surveillance infrastructure are limited.

In 2023-2024, a randomized controlled trial (RCT) conducted in 2 high malaria-burden districts of Uganda (Mayuge and Adjumani) demonstrated VectorCam's efficacy and feasibility [25]. During the RCT's study period, VectorCam generated predictions for primary malaria vectors including *Anopheles gambiae s.l.* (AG) and *Anopheles funestus s.l.* (AF), a heterogeneous grouping labeled Anopheles Other (AO)

XSL•FO
RenderX

representing locally relevant secondary vectors, Culicine genera including Culex spp. (CU), Mansonia spp. (MAN), Aedes spp. (AED), and a nonmosquito class to screen out other arthropods commonly captured during PSCs and CDC-LTCs. Across all taxonomic categories, VectorCam achieved an overall identification accuracy of 93% when compared with the gold-standard entomological identification methods [25]. The VectorCam technology also scored above the 90th percentile on the System Usability Scale, indicating that perceived usability is excellent [25]. Additionally, over 77,000 mosquitoes were identified by 24 VHTs over the 12-month study, with an average processing time of 17 seconds per specimen. This rate is an approximately two-fold improvement over the routine, microscope-based method of mosquito identification, which averages 34.5 seconds per specimen. Additionally, the VectorCam arm of the study accomplished a 98% data completion rate (29,785 entries completed out of 30,360 entries possible), compared to 45% (13,667/30,360) in the paper-based, routine surveillance control arm. These percentages reflect the number of mosquito data entries at day 14 (ie, the standard lag time between data collection and submission in Uganda's routine surveillance workflow) over an 11-month period and across 240 households (120 per arm). By addressing 2 persistent bottlenecks of limited entomological expertise and the slow pace of manual data entry, VectorCam has demonstrated an increase in data reporting efficiency, offering potential for national scale up, ultimately leading to better malaria control.

The technical performance of VectorCam and its feasibility have been established. However, widespread adoption of such technology requires an understanding of how to successfully implement, integrate, sustain, and scale the intervention in real-world health system contexts [4,26]. In Uganda and similar settings, implementation of novel technologies frequently fails or lacks sustainability, with common challenges stemming from technical (eg, infrastructure and expertise) and social (eg, culture, societal context, and human factors) dimensions [27,28]. This study uses an implementation science approach using the Consolidated Framework for Implementation Research (CFIR) and Expert Recommendations for Implementing Change (ERIC) strategies to assess and inform the VectorCam rollout, aiming to improve both uptake and sustainability of its implementation [29-32].

Objectives

The main objective of this study is to evaluate how VectorCam can be sustainably integrated into Uganda's national vector surveillance infrastructure across 12 districts, with a goal of informing future scale-up efforts across the country and in similar settings. The three aims of the study are to (1) identify contextual factors that enable the implementation of VectorCam in Uganda and evaluate the mechanisms through which they affect the implementation process; (2) assess the acceptability, feasibility, fidelity, and cost of VectorCam's implementation model; and (3) evaluate the effectiveness of VectorCam in providing timely, accurate vector surveillance data and influencing malaria control decision-making.

Methods

Study Design Overview

This study uses a mixed methods, hybrid type III implementation-effectiveness design [33]. The primary aim is to evaluate implementation outcomes, including acceptability, feasibility, fidelity, and cost, while the secondary aim is to assess preliminary effectiveness outcomes, including technical accuracy of mosquito identification and operational use of VectorCam data in decision-making. Table 1 outlines these outcomes along with assessment methodology for each, demonstrating our multiple methods of data collection that will generate predefined key measures and are analyzed using both qualitative and quantitative approaches.

XSL•FO
RenderX

Table 1. Implementation and effectiveness outcomes, methods, and metrics by aim.

Outcome	Methods of data collection	Key measures	Planned analysis
Aim 1: assessment of contextual determinants	• CFIR[a]-determinants evaluation survey with implementers; ERIC[b] strategy matching tool • Semistructured pre- and post-KIIs[c] evaluating context with users[d]	• CFIR barriers and facilitators by domain	• Descriptive summaries of survey notes; quantitative summaries of Likert Scale ratings • Thematic analysis of CFIR constructs; cross-district & cross-wave comparison matrices
Aim 2: acceptability, feasibility, fidelity (qualitative)	• Semistructured pre- and post-KIIs evaluating implementation with users[d] • Midpoint IDIs[e] with Master Trainers, VCOs[f], and VHTs[g] • Quarterly stakeholder engagement meeting • FRAME-IS[h] documentation tool	• Perceived usefulness, ease of use, workflow fit, perceived protocol adherence, sustained use, pain points, adaptation tracking	• Thematic analysis; descriptive statistics; longitudinal summaries by district and rollout wave
Aim 2: fidelity (quantitative)	• App usage and completeness metrics • Field observation checklist and supervisory toolkit	• Completeness of reporting; protocol adherence, district-level sustained use • Protocol adherence rate	• Proportions with 95% CIs; ordinal or binary mixed models where appropriate; district quarter trend plots
Aim 2: cost	• Costing database, budgets, procurement and supervision records	• Total and unit costs, budget impact, scale-up costs	• Budget impact analysis modeling for 1-year, 3-year, and 5-year scenarios; ingredients-based costing; annualization of capital costs; one-way sensitivity analysis
Aim 3: technical effectiveness	• Validation subset with VectorCam, routine method, and PCR[i] (or equivalent)	• Sensitivity, specificity, PPV[j], NPV[k], F1-score, agreement	• Diagnostic accuracy analysis with 95% CIs; paired comparison when both methods are available; descriptive comparison of classification performance
Aim 3: operational effectiveness and data use	• Quarterly NMED[l] meeting attendance checklist • KIIs to measure time allocation with users • KIIs with district and national decision-makers	• Use of data in planning; reporting timeliness • Time allocation; task-shifting model metrics • Data usage by decision-makers; impact on malaria control actions	• Descriptive comparisons of time allocation metrics and data-use indicators across districts and rollout phases; mixed effects regression for time metrics; qualitative thematic analysis of data-use pathways

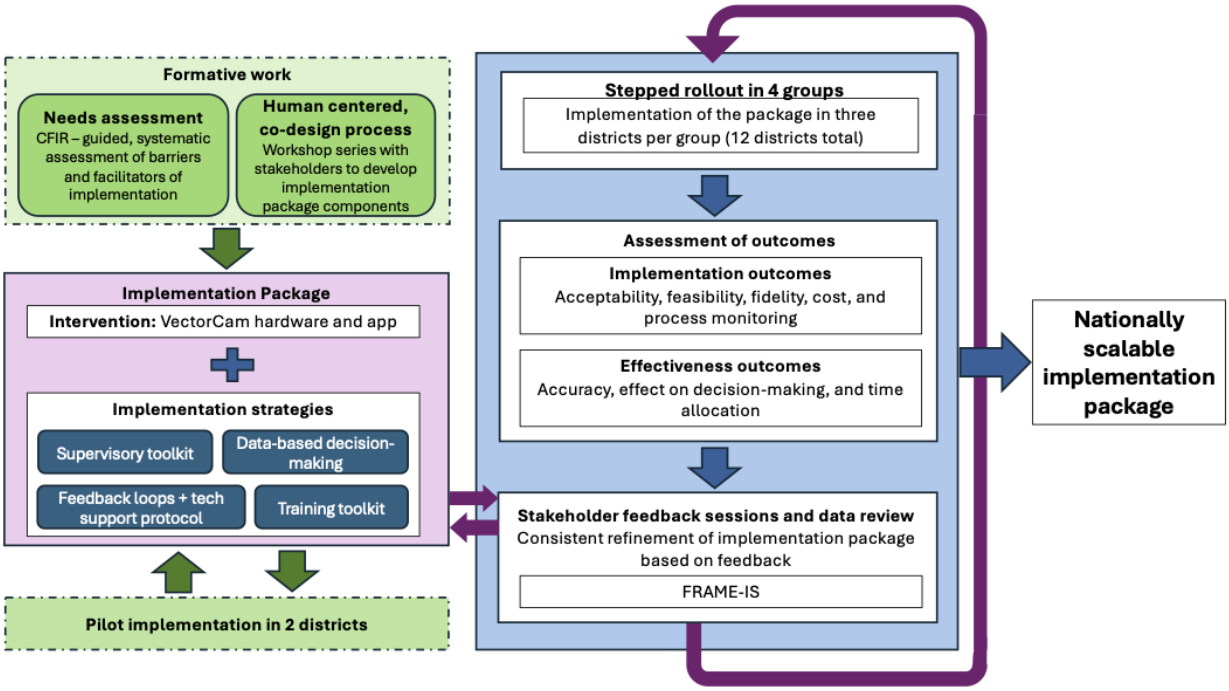
[a]CFIR: Consolidated Framework for Implementation Research.
[b]ERIC: Expert Recommendations for Implementing Change.
[c]KII: key informant interview.
[d]Pre and postinterviews are conducted 1-week after training and 3 months after implementation.
[e]IDI: in-depth interview.
[f]VCO: vector control officer.
[g]VHT: village health team.
[h]FRAME-IS: Framework for Reporting Adaptations and Modifications–Enhanced for Implementation Strategies.
[i]PCR: polymerase chain reaction.
[j]PPV: positive predictive value.
[k]NPV: negative predictive value.
[l]NMED: National Malaria Elimination Division.

The study includes a randomized rollout of VectorCam across a group of 3 districts every 3 months, supporting real-time learning and adaptations. The staggered rollout enables longitudinal assessment of implementation and effectiveness outcomes across districts and rollout waves. A midline review will synthesize this feedback to inform data-driven adaptations to the implementation package and technology design, subject to review and approval by Uganda's MOH officials before rollout in the subsequent districts. Adaptations may include modifications to the intervention (eg, app updates) or to the implementation strategies (revised training toolkit), depending on identified needs. We apply a combination of complementary

XSL•FO
RenderX

implementation science frameworks to guide design, evaluation, and adaptive refinement. Figure 1 illustrates the structure of this study, showing how implementation science theory will be used and the cyclic, stepped rollout that will allow us to systematically gather feedback and improve the intervention prior to each phase. This protocol aligns with the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) guidelines, and the intervention description follows the Template for Intervention Description and Replication (TIDieR) checklist (Multimedia Appendices 1 and 2). A protocol timeline is provided in Multimedia Appendix 3.

Figure 1. Conceptual overview of the hybrid type III implementation-effectiveness study design. CFIR: Consolidated Framework for Implementation Research; FRAME-IS: Framework for Reporting Adaptations and Modifications–Enhanced for Implementation Strategies.



Study Setting and Participants

The MOH selected 14 districts in Uganda to receive VectorCam based on the following inclusion criteria: (1) Presence of diverse mosquito species relevant to malaria transmission; (2) workforce capacity - availability of trained VCOs and VHTs to participate in data collection; (3) district–level commitment to adopting digital vector surveillance tools; (4) willingness of local health authorities to support and sustain implementation; and (5) logistical feasibility to operate within the region (eg, security concerns, seasonal accessibility, and surveillance supply availability). A total of 2 districts, Koboko and Mayuge, were selected as pilot implementation districts to provide an opportunity to test and adapt the initial implementation package, methods of data collection, and key measures based on their high level of implementation readiness. Findings from the pilot will be reviewed jointly by the study team and the MOH to inform the refinement of research tools and implementation strategies. The selected districts represent a range of malaria transmission settings and implementation contexts while maintaining sufficient operational readiness to support the initial phase of national rollout. Although this study does not specifically target the most resource-constrained districts, a key objective is to identify the contextual determinants, implementation strategies, and support structures required for successful adoption of VectorCam. Findings from this study will inform future adaptation and scale-up efforts in settings with more limited surveillance infrastructure and workforce capacity.

The remaining 12 districts will implement the package in 4 sequential rollout phases, with 3 districts added every 3 months following completion of the pilot phase. This design enables evaluation of implementation processes across diverse contexts and over time, while allowing for iterative refinement of strategies between phases. Districts are assigned to rollout waves using restricted randomization at the district level via a computer-generated random sequence created by a study programmer prior to implementation. To ensure geographic balance, the allocation sequence will be constrained such that no rollout group includes more than one district from the same region. This approach was developed collaboratively by the study team and Uganda's MOH to support evaluation across different districts while promoting equitable distribution of the intervention across regions during rollout. The rollout waves are intended to support implementation sequencing and adaptive learning rather than serve as comparison groups for causal inference. Multimedia Appendices 4 and 5 show information on the districts in each rollout group. To facilitate interpretation of our outcomes, contextual district characteristics, including malaria burden and workforce composition (eg number of VHTs and VCOs per district), will be documented during implementation activities and added descriptively to the table shown in Multimedia Appendix 5.

XSL•FO
RenderX

Key participants include VHTs, VCOs, master trainers (ie, district-level supervisors who are trained in the VectorCam system), District Health Officers, and staff from the National Malaria Elimination Division (NMED), representing multiple levels of the health system engaged in malaria surveillance and control.

Intervention Description

The proposed task-shifting model for vector surveillance includes the VectorCam hardware and software (intervention) along with the strategies supporting its integration into Uganda's routine vector surveillance program (implementation package). We used a human-centered design approach to developing the implementation package, with a particular focus on determining the details of the implementation strategies and materials. The implementation package was co-designed in consultation with national, district, and community-level stakeholders during a collaborative workshop held in August 2025. The participants included officials from the Ministry of Health, master (lead) VCOs, district VCOs, and members of the research team.

The intervention includes VectorCam imaging devices, smartphones preloaded with the VectorCam application, and charging accessories provided to each district and shown in Multimedia Appendix 6. The package includes 4 core implementation strategies: (1) a training toolkit and capacity building, (2) data-based decision making, (3) a supervisory toolkit and quality assurance, and (4) learning loops and procedures for ongoing technical support. These strategies are designed to address known barriers to implementation by improving user knowledge and skills, strengthening feedback and accountability systems, reinforcing perceived usefulness, and supporting integration into routine workflows. Because VectorCam is designed to support task shifting of surveillance activities to community-level personnel, the study will also evaluate the training intensity qualitatively and by number of retraining sessions completed, supervision requirements, technical support needs, and workflow adaptations necessary to maintain high-quality surveillance data. Figure 1 depicts the implementation package components conceptually.

Standardized training materials and the supervisory toolkit, including demonstrations, user manuals, inventory tracking procedures, and troubleshooting checklists, were codeveloped by the VectorCam implementation team, the researchers, and the MOH. Training will follow a cascade model in which Master Trainers from NMED will receive training on VectorCam operations, troubleshooting, supervision procedures, and supply management. These Master Trainers will then conduct district-level, hands-on training for VCOs and VHTs during the first week of rollout in the district. The training also includes pre- and posttraining knowledge assessments for VHTs that will identify any remaining gaps prior to VectorCam use in the district. Additionally, the supervisory toolkit includes a structured audit to assess protocol adherence, data quality, and equipment functionality. These toolkits are complemented by the data dashboards and monthly review meetings for regular feedback cadence. Technical support includes a tiered escalation pathway. First-line support is provided by VCOs, with second-line support from the VectorCam technical team for software issues or hardware failures. Procedures for reporting and resolving technical issues are documented using an "Implementation Issues Log" and reinforced during training. These strategies in combination are designed to support fidelity and acceptability of the intervention by improving consistency, knowledge, and quality of data collection.

After each implementation cycle, we will conduct structured stakeholder engagement meetings with district health teams, VCOs, and NMED representatives. This meeting cadence is built into the implementation package specifically to collect feedback from users, disseminate results, and share lessons learned. The Framework for Reporting Adaptations and Modifications–Enhanced for Implementation Strategies (FRAME-IS) will be used to document any significant changes to the implementation package, and these will be shared with stakeholders during the quarterly meeting [34]. Notes from these sessions will also help us identify facilitators and barriers to implementing and sustaining VectorCam, including factors related to resource availability, leadership support, and integration into routine workflows.

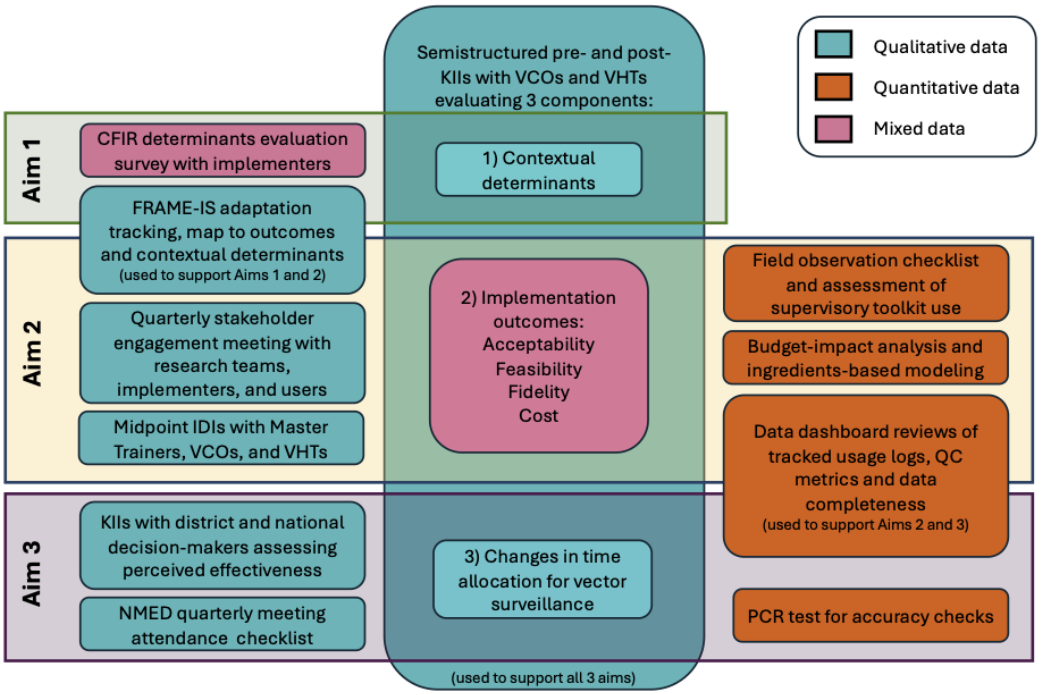
Outcomes Assessment

Table 1 shows a breakdown of outcomes, methods for data collection, and evaluation metrics aligned with each aim. The table outlines how implementation and effectiveness will be assessed through mixed methods, ensuring that outcomes are aligned with the study's overall evaluation framework. All surveys and interviews have been developed with feedback from the study's in-country partners and will be administered with printed forms in either English or local languages as appropriate. The interviews will be administered by a researcher who is fluent in both English and the local language and trained to obtain consent before interviews.

Figure 2 illustrates all the methods of data collection used across the aims of this study, correlating with Table 1. An overview of each aim and a description of the methods used for data collection within them is detailed in the following sections.

XSL•FO
RenderX

Figure 2. Overview of the mixed methods research tools used by aim. CFIR: Consolidated Framework for Implementation Research; IDI: in depth interview; KII: key informant interviews; NMED: National Malaria Elimination Division; PCR: polymerase chain reaction; QC: quality control; VCO: Vector Control Officers; VHT: village health teams.



Aim 1: Evaluation of Contextual Determinants

We will assess the contextual determinants that influence implementation of the VectorCam platform by using an approach grounded in CFIR methodology and guided by established implementation mapping procedures [32]. This aim combines semistructured surveys and in-depth, pre- and post-key informant interviews (KIIs) to generate an understanding of the barriers and facilitators to successful implementation and evaluate the mechanisms through which they influence the implementation and effectiveness outcomes of interest.

Methods for data collection for Aim 1 include (1) CFIR-determinants evaluation survey with implementers, and (2) semistructured, pre- and post-KIIs to evaluate contextual characteristics with users. Firstly, for the CFIR-determinants evaluation surveys, we will administer a semistructured survey to select implementers: 4 MOH representatives and 6 VCOs or District Health Officials. This process will yield a total of 10 interviews conducted between months 4 and 6 of the staggered rollout. This survey will collect information on the context, probing for barriers and facilitators to implementation success. The survey instrument will primarily include Likert-scale questions to understand how the participant ranks the importance of specific determinants, supplemented by 5 open-ended questions to probe in-depth explanations for their choice of ranking. All questions are adapted from validated CFIR measurement tools to assess the 5 domains, including perceptions of intervention characteristics (eg, relative advantage and adaptability), inner setting (eg, organizational culture and implementation climate), outer setting (eg, external policies and community support), characteristics of individuals (eg, self-efficacy and knowledge), and process (eg, leadership engagement and planning).

Secondly, in each district, pre-KIIs will be conducted during the first week of intervention implementation (after training), and post-KIIs will then be administered 3 months later, focusing on the evaluation of contextual characteristics with users. Four VHTs and 2 VCOs will be purposely selected per district to ensure a balance of perspectives and varying levels of prior experience with entomological surveillance. The same participants will be contacted again for postimplementation interviews, enabling us to make paired comparisons of changes in attitudes and reported practices over time. Whenever possible, the same participants will be recontacted for follow-up interviews. If a participant is unavailable due to transfer, withdrawal, or other forms of attrition, a replacement participant occupying a similar implementation role will be purposively selected. Data collected through these sessions will inform

XSL•FO
RenderX

iterative refinements to the implementation package and materials for subsequent adaptive cycles. By collecting rich qualitative data, the study team can examine what contextual determinants are most important during the initial phases of VectorCam rollout and, more precisely, how these factors are associated with changes in implementation outcomes.

Aim 2: Acceptability, Feasibility, Fidelity, and Cost

The primary purpose of this aim is to ascertain the implementation outcomes of VectorCam use as defined by the Implementation-Outcome Framework [35]. Acceptability, fidelity, and feasibility of the implementation package will be assessed by using an integrated mixed methods approach that combines qualitative and quantitative data. For qualitative data, we will deploy pre- and postimplementation semistructured interviews, quarterly stakeholder engagement meeting notes, midpoint in-depth interviews with implementers and users, and FRAME-IS guided survey. Quantitative data collection tools for Aim 2 include data dashboard metrics, such as usage logs and percent completeness, and the field observation checklist. Additionally, the cost of VectorCam integration into Uganda's vector surveillance system will be assessed with a combination of prospective costing methods, budget impact modeling, routine monitoring data, and structured feedback from district and national stakeholders. This aim will generate a comprehensive understanding of user experiences, protocol adherence, ease of use, fit with existing workflows, and evidence on the long-term financial feasibility of maintaining the implementation package at scale. Data from this aim can be used by Uganda's MOH and its implementing partners to plan for a scalable and sustainable nationwide rollout.

Methods of data collection for Aim 2 include:

1. Semistructured pre- and post-KIIs to evaluate implementation outcomes with users
2. Field observation checklist to evaluate fidelity and supervisory toolkit
3. VectorCam app-system usage logs
4. Costing database, budgets, and procurement records
5. FRAME-IS documentation tool

The first data collection tool, pre- and post-KIIs, includes interview guides that are designed to be used at the same time points as the semistructured pre- and post-KII to evaluate context and administered to the same participants (ie, 4 VHTs and 2 VCOs) introduced in Aim 1 data collection methods. The purpose of this interview guide is to collect information about implementation outcomes. It will include structured questions for assessing satisfaction with the technology, perceived ease of use, alignment with existing workflows, confidence in the tool, and reflections on the frequency and consistency of use.

Each month, a VCO will use the second research tool and the field observation checklist to directly observe VHTs using the VectorCam system in the field to assess fidelity and the deployment of the supervisory toolkit. We will review these logs to verify the proportion of eligible VHTs trained in each cycle, equipment distribution, and posttraining support provided. The supervisory tool also contains a standardized field observation checklist and a set of structured questions designed to identify factors that may affect supervision, these questions will explore domains such as availability of logistical resources (eg, transport and devices), communication and coordination mechanisms, and context and barriers that will be used to assess how well VectorCam users are adhering to the protocol, including preparation and calibration of the device, correct specimen handling, adherence to identification procedures, and successful data upload to DHIS2. A VCO and a researcher will complete this checklist separately so that we will have 2 observations with which to assess district-level fidelity. Observers will rate protocol adherence as high, medium, or low and will provide narrative field notes describing deviations and potential contributing factors.

Quantitative data collected from VectorCam's app will be used to assess percent of data completeness, and fidelity will be collected through system-generated usage logs exported monthly for analysis. Completeness of reporting and sustained use metrics will be summarized descriptively. Mixed-effects logistic or ordinal regression models may be used, where appropriate, to explore temporal patterns in completeness of reporting and implementation outcomes related to sustained use while accounting for clustering at the district level.

Throughout the study period, we will use a cost-tracking system to capture direct and indirect costs associated with VectorCam implementation in each district. Cost data will be collected prospectively and will include expenses related to procurement of hardware and consumables, training and supervision activities, technical support, and system maintenance. Financial records, procurement invoices, and project budgets will be reviewed to ensure comprehensive accounting of expenditures. Health economists on the study team will use these data to develop annual and multiyear budget projections, estimate total annual implementation costs, and calculate operational cost savings that result from replacement of manual surveillance methods. These models will generate 1-year, 3-year, and 5-year budget projections to inform national malaria control planning.

Lastly, to ensure transparent reporting of changes to the package over time, we will systematically document all modifications to the intervention and implementation strategies using the FRAME-IS [34]. Throughout implementation, project staff and implementing partners will record proposed or observed changes during routine monitoring and bi-weekly coordination calls. Each adaptation will be documented on a structured form aligned to FRAME-IS domains: (1) the nature of the modification (content, context, personnel, or delivery mode); (2) the level at which the change occurred (individual provider, facility, district, or system level); (3) whether the modification was planned or reactive; (4) the goal of the change, including improving feasibility, acceptability, fidelity, reach, or equity; (5) the decision-making process and actors involved; and (6) the extent to which the change was fidelity-consistent with core program components. We will then categorize adaptations into higher-level themes to enable comparison and identify patterns in how the implementation strategies evolved over time. This process will ensure transparent reporting of adaptations and allow linking of modifications to implementation outcomes and contextual drivers.

XSL•FO
RenderX

Aim 3: Evaluation of Effectiveness

Effectiveness of the VectorCam intervention (as delivered with implementation strategies) in maintaining accuracy, improving time allocation metrics, and supporting operational use of surveillance data will be assessed with a combination of KIIs, quantitative analyses of system data logs, and qualitative interviews with national and district-level decision-makers. This aim is designed to generate preliminary evidence regarding VectorCam's performance under ideal implementation conditions and whether the task-shifting model is associated with improvements in data quality and timelier, evidence-based decision-making relative to microscopy-based surveillance. Routine microscope-based entomological surveillance conducted by district entomology teams serves as the comparator condition, as it represents the standard of practice within Uganda's national malaria surveillance system.

Methods of data collection for Aim 3 include:

1. Pre- and post-KIIs to evaluate time allocation with users
2. App-recorded data logging time spent per mosquito identification event
3. Molecular analysis comparison
4. KIIs with national and district-level decision-makers on operational effectiveness and data use
5. NMED quarterly meeting attendance and observational checklist

Pre- and postimplementation interviews with VHTs and VCOs will be used to assess the time required for mosquito identification and data submission. This tool is designed to be utilized at the same time point as the pre- and post-KIIs for data collection within Aims 1 and 2. Participants will be asked to describe the average time spent on identification tasks before and after the introduction of VectorCam, including steps related to specimen preparation, processing, and reporting. Questions on the KII guides differ for VHTs and VCOs in order to capture changes to workflow based on the task-shifting model.

Quantitative data on actual time spent per mosquito identification event will be extracted from VectorCam usage logs over the study period. Logs will include timestamps marking initiation of identification, confirmation of results, and data upload.

For the molecular analysis comparison, accuracy of mosquito species identification using VectorCam will be evaluated by comparing system-generated identifications to a gold-standard reference. Molecular identification will serve as the primary gold standard, using polymerase chain reaction (PCR) when species-specific reagents are available and sequencing when they are not. A random subset of 500 specimens collected during routine surveillance will be selected for molecular validation. This sample size was selected to provide sufficient precision for estimating identification accuracy under routine implementation conditions. Based on the 93% species identification accuracy observed in the previous RCT, a validation sample of 500 specimens yields an approximate 95% CI with a margin of error of +/- 2.2 percentage points around the overall accuracy estimate. These estimates will be used to characterize VectorCam performance descriptively where sample sizes are sufficient, recognizing that precision will vary by the number of specimens available for each species. These specimens will be analyzed independently by an entomologist blinded to the VectorCam results and PCR findings. Accuracy will be defined as the proportion of VectorCam identifications that match PCR results when available, or morphologic identification otherwise. In addition, model performance will be evaluated using standard classification metrics from an AI perspective. In addition to this primary accuracy measure, performance will also be evaluated using standard metrics, including precision, recall, and the F_1-score, which together provide a balanced assessment of how well the model identifies mosquito species.

Next, we will explore VectorCam's operational effectiveness and data use on malaria control decision-making by interviewing 8-10 MOH-NMED staff and District Health Officers. An interview guide was developed to facilitate the KIIs with national and district-level decision makers to collect qualitative information regarding operational effectiveness and VectorCam data use. The NMED groups that will be interviewed will either have routine interactions with the entomology database, responsibilities for improving the database, and/or responsibilities that involve using the database to make decisions about malaria control programs, including roles such as the Director of the National Malaria Program, Information Technology Senior Developer, lead entomologists, and coordinators of routine surveillance activities. Interviews will explore perceptions of the reliability and utility of VectorCam data, comparisons to previous surveillance approaches, and examples of how data informed vector control interventions or planning. Interviews will be conducted a minimum of 6 months after implementation in each district cohort to allow sufficient time for data review and use.

Lastly, a researcher from either Makerere University or Johns Hopkins University will attend NMED's quarterly meetings and use the NMED quarterly meeting attendance and observational checklist to document how the VectorCam data are reviewed and used during routine decision-making processes. The checklist will capture whether VectorCam data were presented, discussed, or referenced in relation to surveillance planning, vector control activities, resource prioritization, or other operational decisions. Findings from these observations will be triangulated with key informant interviews to assess the perceived utility and operational use of VectorCam data.

Data Management and Analysis

This study uses a convergent mixed methods analytic approach in which qualitative and quantitative data are collected and analyzed concurrently across all study aims. Data collection and analysis will occur iteratively across the staggered rollout cycles, allowing findings from earlier phases to inform subsequent data collection, refinement of tools, and interpretation of results. This approach is intended to capture the dynamic and context-dependent nature of implementation and to support adaptive learning throughout the study. Across all aims, data collected through multiple methods will be used to generate and validate key measures, and analyses will be

XSL•FO
RenderX

conducted in alignment with the "Planned Analysis" column in Table 1.

All qualitative data, including semistructured key informant interviews, in-depth interviews, open-ended survey responses, and stakeholder engagement discussions, will be collected by trained research staff and transcribed verbatim when audio recording is permitted or expanded into detailed notes immediately following data collection. Research staff involved in observational data collection will receive standardized training on data collection procedures, use of observational tools, and documentation practices to promote consistency across observers and study sites. Data will be managed and analyzed using Atlas.TI. A single, unified codebook will be applied across all qualitative data sources to ensure consistency and enable comparison across stakeholder groups and time points. This approach is particularly important given that the same interview streams (ie, the pre and post-KIIs) contribute to multiple study aims and are intended to capture changes over time in determinants, implementation processes, and operational effectiveness.

Codebook development will follow a hybrid deductive and inductive approach (see Multimedia Appendix 7 for deductive codebook). Deductive codes will be derived from established implementation science frameworks, including the CFIR, as well as the Implementation Outcomes Framework (IOF) to capture constructs such as acceptability, feasibility, fidelity, and cost. After the first implementation rollout, the codebook will be iteratively refined to incorporate inductive codes that emerge from the data, including unanticipated barriers, contextual dynamics, adaptations, and potential district-specific codes. Coding will initially be conducted by at least 2 coders, with regular meetings to reconcile discrepancies, refine code definitions, and ensure consistent application of the coding framework. Prior to independent coding of the full dataset, coders will establish interrater reliability using a subset of transcripts. Interrater reliability will be assessed using percent agreement, with a minimum threshold of 0.75 required before proceeding to full coding. If agreement falls below this threshold, discrepancies will be reviewed jointly, code definitions will be refined, and additional training and calibration exercises will be conducted before reassessment. During coding, disagreements that cannot be resolved through discussion will be adjudicated by a senior member of the research team. Ongoing reflexive discussions will be used to examine assumptions and strengthen interpretive validity within the research teams. Framework analysis will be used to identify themes within and across stakeholder groups and over time, with particular attention to how contextual determinants influence implementation outcomes and how these relationships evolve across rollout cycles.

Quantitative data will be derived from structured surveys, field observation checklists, costing databases, and the data dashboard, which tracks key metrics such as data completeness. Descriptive analyses will be conducted to summarize key variables, including measures of data completeness, usage frequency, and accuracy. Implementation outcomes will be assessed using multiple complementary indicators and, where appropriate, proportions will be calculated with 95% CIs. Fidelity will be evaluated through both system-generated metrics, such as completeness of required data fields, and observational assessments of protocol adherence. Patterns of system use will be examined longitudinally to assess trends over time and variation across districts. Descriptive summaries will be generated by district and rollout wave, including district-quarter trend plots to visualize changes over time.

Cost analysis will be conducted from the program perspective, with scenario analyses for broader health system scale-up. An ingredients-based costing approach will be used to estimate startup and recurrent costs, including hardware, consumables, training, supervision, maintenance, connectivity, and technical support. Capital items will be annualized over their useful life using a standard discount rate. Unit costs will be reported per district covered, per user trained, and per identification event completed. Budget impact models will be developed to project costs under different scale-up scenarios, incorporating both observed expenditures and anticipated changes in resource requirements. Sensitivity analyses will test uncertainty in key assumptions.

Effectiveness outcomes will be assessed using both quantitative and qualitative data. Time allocation will be evaluated through both self-reported data (collected via the pre- and post- KIIs) and system-recorded measures of time spent on surveillance activities. To evaluate the task-shifting model that VectorCam used, time allocation comparisons will be made across pre- and postimplementation periods. Accuracy of mosquito identification will be assessed by comparing VectorCam outputs to the gold-standard molecular identification (PCR or sequencing for species for which PCR reagents are not available). Qualitative data from interviews with decision-makers will be analyzed thematically to assess the perceived utility of surveillance data for decision-making using the same codebook as other qualitative data collection tools.

Integration of qualitative and quantitative data will occur throughout the analytic process to enable triangulation and strengthen the validity of findings. This will occur through joint displays, triangulation matrices, and iterative comparison of quantitative and qualitative findings. Quantitative results will be interpreted alongside qualitative insights to assess convergence and to provide context for observed patterns. Qualitative data will also be used to explain unexpected quantitative findings and to identify mechanisms underlying observed relationships. Integration will be further supported using an Implementation Research Logic Model, which will link contextual determinants, implementation strategies, and outcomes across data sources.

Documentation of adaptations using FRAME-IS will be analyzed alongside implementation and effectiveness metrics to examine whether changes to the implementation approach are associated with improvements in the implementation outcomes we are analyzing. Comparisons across stakeholder groups will also be conducted to assess alignment between frontline experiences, implementation processes, and decision-making.

Through this integrated analytic approach, the study will generate a comprehensive understanding of how VectorCam is

XSL•FO
RenderX

implemented within routine surveillance systems, how implementation processes influence outcomes, and how improvements in data quality and timeliness translate into actionable use for malaria control.

Sample Size and Power

Across all aims, sample size determination is guided by the descriptive and implementation-focused nature of the study rather than formal hypothesis testing. In each of the 12 districts, 4 VHTs and 2 VCOs will participate in pre- and postimplementation qualitative assessments to collect data across all 3 aims. These KIIs will yield a total of 72 interviews at baseline and 72 at follow-up. The sample size was selected to ensure representation of the primary implementation actors across all participating districts while remaining feasible within a national implementation study. Participants will be purposively selected to capture perspectives across diverse implementation contexts, and the anticipated number of interviews is expected to provide sufficient breadth and depth to identify recurring themes related to implementation processes and outcomes. In addition, approximately 10 KIIs will be conducted with district-level and national decision-makers to assess effectiveness for Aim 3. We again expect this number will be sufficient to achieve thematic saturation for our (preliminary) effectiveness outcomes of interest. Quantitative system data logs will include all identification records generated by trained VHTs during the study period, and for accuracy validation, at least 500 specimens will be randomly selected and assessed across all districts. Because the study primarily employs descriptive rather than hypothesis-driven testing, no formal power calculations were conducted. The overall sample is expected to be sufficient to generate estimates of implementation outcomes and detect meaningful differences in time efficiency, identification accuracy, and data usage.

Ethical Considerations

This study has received ethical approval from the Johns Hopkins Medicine Institutional Review Board (protocol number IRB00505343) and, in Uganda, from the AIDS Support Organization (TASO) Research Ethics Committee (TASO-2025-879) and the Makerere University School of Public Health Research Review Committee. The VectorCam device was reviewed and determined to be a nonsignificant risk (NSR) device meeting criteria for an abbreviated investigational device exemption. No significant harms are anticipated due to the minimal-risk nature of the study.

This study is embedded within ongoing Ministry of Health routine entomological surveillance activities. For components involving human participants (eg, interviews, surveys, and observational assessments), written informed consent is obtained prior to participation. Consent forms are available in English and in relevant local languages (including Lusoga, Ngakarimojong, Luo, and Lugbara), and trained research staff conduct consent discussions in the participant's preferred language, using certified translations. Participants who complete interviews or surveys will receive 20,000 Ugandan Shillings (UGX) as compensation for their time. The compensation amount was reviewed and approved by the relevant ethics committees and is not considered coercive.

At the start of the consent process, participants will be asked to indicate what language they are comfortable with, and the appropriate consent forms will be selected. All sections of the consent document will be read verbatim in the selected language and explained to ensure comprehension. All participants will be informed that participation in the study is entirely voluntary and that they may withdraw from the study at any time. Signed consent forms are retained in secure study files in accordance with institutional and national regulatory requirements.

All identifiable data will be stored on secure, password-protected servers at Johns Hopkins University and Makerere University, accessible only to authorized study personnel. Deidentified data will be used for analysis, and no identifying information will be shared or published.

Results

In collaboration with Uganda's Ministry of Health, the study team identified 22 districts that met the inclusion criteria, from which 14 districts were selected to receive the VectorCam intervention in September 2025. The pilot phase was conducted from December 2025 to January 2026, during which research tools were tested during rollout in two districts. Stakeholder workshops in the pilot phase were used to refine the implementation package and research instruments. Ethics approval was granted in January 2026. Data collection for the first rollout group began in February 2026, initiating a year-long data collection period. As of March 2026, the second rollout group of 3 districts is scheduled to begin data collection in June 2026. Data collection and subsequent analysis will be concurrent in districts where the intervention has been deployed and are expected to continue through March 2027.

Discussion

Anticipated Findings

This protocol represents one of the first hybrid type III studies to evaluate how an AI-enabled surveillance tool can be scaled and sustained within a national vector control program in low- and middle-income countries (LMICs). Building on strong feasibility evidence from VectorCam's prior RCT, we will assess how contextual determinants influence uptake, identify and refine implementation strategies, and generate robust evidence on both implementation outcomes and effectiveness across diverse malaria-endemic districts in Uganda. The study is designed to contribute to the evidence base on implementation of AI-enabled digital health innovations in routine public health systems of LMICs.

Beyond evaluating digital surveillance technology, this study examines a task-shifting model in which community-level personnel conduct activities traditionally performed by trained entomologists. Understanding the contextual determinants, implementation strategies, training requirements, and supervision structures associated with successful implementation will be critical for future expansion of surveillance activities into districts with more limited workforce capacity and infrastructure. Although the current rollout focuses on districts selected for implementation readiness, findings from this study

XSL•FO
RenderX

are intended to inform future deployment in underserved settings where entomological surveillance remains limited or unavailable.

Several features strengthen the rigor and relevance of this study. First, our implementation strategy development is grounded in the CFIR and the ERIC taxonomy, ensuring a theory-driven and context-sensitive approach to scale-up. Use of the FRAME-IS allows systematic documentation of adaptations to the implementation model, enabling us to address the growing demand for transparency and reproducibility in implementation research. Additionally, we integrate learning processes across all phases of the study to support iterative adaptation and stakeholder co-ownership, which bolsters the probability of long-term sustainability.

Challenges and Limitations

The reliance on system usage logs and structured observations as proxies for implementation quality may underestimate informal use or undocumented fidelity. However, triangulation with qualitative interviews, cost tracking, and FRAME-IS documentation is expected to provide a more complete picture.

Another limitation is that data from VHTs and VCOs will be collected through self-report and field observations, both of which may be influenced by social desirability bias or observer effects. To mitigate this limitation, researchers will undergo structured training, and all the survey tools will be codeveloped, pilot tested, or both. Additionally, this implementation represents the inaugural national rollout of VectorCam into Uganda's routine vector surveillance program, no prior equivalent program exists against which a concurrent or historical control comparison could be meaningfully constructed. Thus, effectiveness outcomes should be interpreted as preliminary and descriptive rather than causal.

Although VectorCam has demonstrated strong technical performance in prior trials, scale-up may introduce new technological or bandwidth challenges not previously encountered. The monitoring, evaluation, and learning framework and staggered implementation schedule will help mitigate this risk through continuous monitoring and rapid-cycle problem solving. Finally, although the budget impact analysis will model cost scenarios at 1- and 5-year horizons, longer-term sustainability will also depend on shifts in MOH budget allocations and digital infrastructure, which are factors outside the scope of this study but relevant for future planning.

Conclusions

This study represents an important contribution to implementation science, digital public health, and malaria vector control, as we aim to evaluate the implementation and effectiveness of a co-designed, AI-enabled platform integrated within Uganda’s national malaria surveillance system. Through the examination of contextual determinants, implementation strategies and outcomes, and preliminary effectiveness outcomes, this study seeks to generate evidence on the feasibility of task-shifting routine entomological surveillance activities to community-level personnel. The insights generated from utilizing this protocol for Uganda’s VectorCam rollout can inform the scale-up of similar surveillance technologies across LMICs and help accelerate progress toward malaria elimination.

Acknowledgments

We would like to thank Justine Inhensiko, Patience Oputan, Samara Toussaint, Kigongo Siriman, James Kawesa, Alex Musiime, Mike Kato, Jovan Bate, David Onanyang, and Jimmy Opigo for their assistance with in-field coordination support, and Neil Lobo, Atul Zacharias, Sai Reddy, Parthvi Mehta, Deming (Remus) Li, Shreya Raman, Christina Heal Kowal, Santiago Sanchez, and Kyle Cooper for their contributions to the design of the VectorCam technology.

Funding

This work was supported by the Bill & Melinda Gates Foundation [Grant Number: INV-079162]. The sponsor can be contacted at the Bill & Melinda Gates Foundation, Seattle, WA, USA. The funder had no role in the design of the study; collection, management, analysis, or interpretation of data; or decision to submit the manuscript for publication.

Data Availability

No datasets have yet been generated, as this manuscript reports a study protocol. All research-related datasets, study instruments, implementation materials, codebooks, data dictionaries, and analytic code used to generate published findings will be made publicly available through an open-access repository (eg, OSF) upon publication of primary study results. Where permitted by ethical, regulatory, and Ministry of Health requirements, deidentified datasets or aggregate data supporting published findings may be made available through a controlled-access request process.

Authors' Contributions

MH: methodological development, writing – original and final drafts, development of the research tools, writing – review & editing; MRT: conceptualization, writing – original draft, project administration, writing – review & editing; DM: methodological development, writing – review & editing; SP: conceptualization, writing – review & editing; WKK: writing – original draft, writing – review & editing; PW: project administration, writing – review & editing; AS: writing – review & editing; CMS: review & editing; YY: review & editing; SA: conceptualization, project administration, writing – review & editing.

XSL•FO
RenderX

Conflicts of Interest

A product described in this protocol publication is manufactured by Vector Control Innovations, Inc (VCI), which is a collaborator for this study. SP, MRT, and SA are cofounders and unpaid members of the board of directors of VCI. This arrangement has been reviewed and approved by the Johns Hopkins University in accordance with its conflict-of-interest policies.

Multimedia Appendix 1
SPIRIT checklist 2025.
[PDF File (Adobe PDF File), 131 KB - publichealth_v12i1e99717_app1.pdf]

Multimedia Appendix 2
TIDieR checklist for Protocols.
[PDF File (Adobe PDF File), 129 KB - publichealth_v12i1e99717_app2.pdf]

Multimedia Appendix 3
Protocol timeline describing research and implementation activities and the staggered rollout design.
[PDF File (Adobe PDF File), 43 KB - publichealth_v12i1e99717_app3.pdf]

Multimedia Appendix 4
Map with Ugandan districts and rollout groups depicted.
[PDF File (Adobe PDF File), 105 KB - publichealth_v12i1e99717_app4.pdf]

Multimedia Appendix 5
This table indicates the districts receiving the VectorCam implementation package and sorts them by rollout group.
[PDF File (Adobe PDF File), 69 KB - publichealth_v12i1e99717_app5.pdf]

Multimedia Appendix 6
Depiction of the novel VectorCam intervention device and smart phone application.
[PNG File , 1186 KB - publichealth_v12i1e99717_app6.png]

Multimedia Appendix 7
Preliminary deductive codebook used for the protocol's qualitative analyses.
[XLSX File (Microsoft Excel File), 40 KB - publichealth_v12i1e99717_app7.xlsx]

References

1. World malaria report 2024. World Health Organization. 2024. URL: https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2024 [accessed 2025-08-20]
2. Malaria. World Health Organization. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/malaria [accessed 2025-04-15]
3. Uganda. Target Malaria. URL: https://targetmalaria.org/about-us/where-we-operate/uganda/ [accessed 2025-08-20]
4. Yeka A, Gasasira A, Mpimbaza A, Achan J, Nankabirwa J, Nsobya S, et al. Malaria in Uganda: challenges to control on the long road to elimination: I. epidemiology and current control efforts. Acta Trop 2012 Mar;121(3):184-195 [FREE Full text] [doi: 10.1016/j.actatropica.2011.03.004] [Medline: 21420377]
5. Rosenthal PJ, John CC, Rabinovich NR. Malaria: how are we doing and how can we do better? Am J Trop Med Hyg 2019;100(2):239-241 [FREE Full text] [doi: 10.4269/ajtmh.18-0997] [Medline: 30628572]
6. Bhatt S, Weiss DJ, Cameron E, Bisanzio D, Mappin B, Dalrymple U, et al. The effect of malaria control on Plasmodium falciparum in Africa between 2000 and 2015. Nature 2015 Oct 08;526(7572):207-211 [FREE Full text] [doi: 10.1038/nature15535] [Medline: 26375008]
7. Shretta R, Liu J, Cotter C, Cohen J, Dolenz C, Makomva K. Malaria elimination and eradication. In: Major Infectious Diseases. Washington (DC): The World Bank; 2017.
8. CDC. Larval source management and other vector control interventions. CDC Malaria. 2024. URL: https://www.cdc.gov/malaria/php/public-health-strategy/larval-management.html [accessed 2025-08-20]
9. Essential programme on immunization. World Health Organization. URL: https://www.who.int/teams/immunization-vaccines-and-biologicals/essential-programme-on-immunization/integration/linking-with-other-health-interventions/long-lasting-insecticidal-nets-(llins) [accessed 2025-08-20]
10. Pluess B, Tanser FC, Lengeler C, Sharp BL. Indoor residual spraying for preventing malaria. Cochrane Database Syst Rev 2010;2010(4):CD006657 [FREE Full text] [doi: 10.1002/14651858.CD006657.pub2] [Medline: 20393950]

XSL•FO
RenderX

11. Kaindoa EW, Matowo NS, Ngowo HS, Mkandawile G, Mmbando A, Finda M, et al. Interventions that effectively target Anopheles funestus mosquitoes could significantly improve control of persistent malaria transmission in south-eastern Tanzania. PLoS One 2017;12(5):e0177807 [FREE Full text] [doi: 10.1371/journal.pone.0177807] [Medline: 28542335]
12. Russell TL, Farlow R, Min M, Espino E, Mnzava A, Burkot TR. Capacity of national malaria control programmes to implement vector surveillance: a global analysis. Malar J 2020;19(1):422 [FREE Full text] [doi: 10.1186/s12936-020-03493-1] [Medline: 33228725]
13. Kumar G, Baharia R, Singh K, Gupta SK, Joy S, Sharma A. Addressing challenges in vector control: a review of current strategies and the imperative for novel tools in india's combat against vector-borne diseases. BMJ Public Health 2024;2(1):17.
14. Gillies MT, Meillon BA. The Anophelinae of Africa south of the Sahara (Ethiopian Zoogeographical Region). Johannesburg: The South African Institute for Medical Research; 1968.
15. Kilama M, Smith DL, Hutchinson R, Kigozi R, Yeka A, Lavoy G, et al. Estimating the annual entomological inoculation rate for Plasmodium falciparum transmitted by Anopheles gambiae s.l. using three sampling methods in three sites in Uganda. Malar J 2014 Mar 21;13:111 [FREE Full text] [doi: 10.1186/1475-2875-13-111] [Medline: 24656206]
16. Rueda LM. Pictorial keys for the identification of mosquitoes (Diptera: Culicidae) associated with dengue virus transmission. Zootaxa 2004;589(1):1-60. [doi: 10.11646/zootaxa.589.1.1]
17. Welcome. Vector Control Division. URL: https://vectorcontroldivisionuganda.wordpress.com/ [accessed 2026-08-06]
18. Dasari S, Gopinath B, Gaulke CJ, Patel SM, Merali KK, Sunil Kumar A, et al. A handheld tool for the rapid morphological identification of mosquito species (VectorCam) for community-based malaria vector surveillance: summative usability study. JMIR Hum Factors 2024 Aug 16;11:e56605 [FREE Full text] [doi: 10.2196/56605] [Medline: 39150762]
19. Li D, Hegde S, Sunil Kumar A, Zacharias A, Mehta P, Mukthineni V, et al. Towards transforming malaria vector surveillance using VectorBrain: a novel convolutional neural network for mosquito species, sex, and abdomen status identifications. Sci Rep 2024 Oct 10;14(1):23647 [FREE Full text] [doi: 10.1038/s41598-024-71856-8] [Medline: 39384771]
20. Kostkova P, Saigí-Rubió F, Eguia H, Borbolla D, Verschuuren M, Hamilton C, et al. Data and digital solutions to support surveillance strategies in the context of the COVID-19 pandemic. Front Digit Health 2021;3:707902 [FREE Full text] [doi: 10.3389/fdgth.2021.707902] [Medline: 34713179]
21. Joshi A, Miller C. Review of machine learning techniques for mosquito control in urban environments. Ecol Inform 2021;61:101241. [doi: 10.1016/j.ecoinf.2021.101241]
22. Motta D, Santos A, Winkler I, Machado BAS, Pereira DADI, Cavalcanti AM, et al. Application of convolutional neural networks for classification of adult mosquitoes in the field. PLoS One 2019;14(1):e0210829 [FREE Full text] [doi: 10.1371/journal.pone.0210829] [Medline: 30640961]
23. Goodwin A, Padmanabhan S, Hira S, Glancey M, Slinowsky M, Immidisetti R, et al. Mosquito species identification using convolutional neural networks with a multitiered ensemble model for novel species detection. Sci Rep 2021 Jul 01;11(1):13656 [FREE Full text] [doi: 10.1038/s41598-021-92891-9] [Medline: 34211009]
24. Couret J, Moreira DC, Bernier D, Loberti AM, Dotson EM, Alvarez M. Delimiting cryptic morphological variation among human malaria vector species using convolutional neural networks. PLoS Negl Trop Dis 2020;14(12):e0008904 [FREE Full text] [doi: 10.1371/journal.pntd.0008904] [Medline: 33332415]
25. Patel S, Rincon TM, Li D. VectorCam: A novel AI-powered digital tool for automated morphological identification of mosquito species, sex, and abdominal status by village health teams in Uganda: A randomized controlled trial. Am J Trop Med Hyg 2024;112(6) [FREE Full text]
26. Barry S, Davies C. Creating system-wide change in medicine: the role of implementation science in achieving scale and adoption. Future Healthc J 2025;12(3):100452 [FREE Full text] [doi: 10.1016/j.fhj.2025.100452] [Medline: 40896305]
27. Agbeyangi A, Suleman H. Advances and challenges in low-resource-environment software systems: a survey. Informatics 2024;11(4):90. [doi: 10.3390/informatics11040090]
28. Kiberu VM, Mars M, Scott RE. Barriers and opportunities to implementation of sustainable e-health programmes in Uganda: a literature review. Afr J Prim Health Care Fam Med 2017;9(1):e1-e10 [FREE Full text] [doi: 10.4102/phcfm.v9i1.1277] [Medline: 28582996]
29. Kaboré SS, Ngangue P, Soubeiga D, Barro A, Pilabré AH, Bationo N, et al. Barriers and facilitators for the sustainability of digital health interventions in low and middle-income countries: A systematic review. Front Digit Health 2022;4:1014375 [FREE Full text] [doi: 10.3389/fdgth.2022.1014375] [Medline: 36518563]
30. van Gemert-Pijnen JL. Implementation of health technology: directions for research and practice. Front Digit Health 2022;4:1030194 [FREE Full text] [doi: 10.3389/fdgth.2022.1030194] [Medline: 36440459]
31. Damschroder LJ, Reardon CM, Opra Widerquist MA, Lowery J. Conceptualizing outcomes for use with the consolidated framework for implementation research (CFIR): the CFIR outcomes addendum. Implement Sci 2022;17(1):7 [FREE Full text] [doi: 10.1186/s13012-021-01181-5] [Medline: 35065675]
32. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated consolidated framework for implementation research based on user feedback. Implement Sci 2022;17(1):75 [FREE Full text] [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]
33. Peters DH, Adam T, Alonge O, Agyepong IA, Tran N. Implementation research: what it is and how to do it. BMJ 2013;347:f6753. [doi: 10.1136/bmj.f6753] [Medline: 24259324]

XSL•FO
RenderX

34. Miller CJ, Barnett ML, Baumann AA, Gutner CA, Wiltsey-Stirman S. The FRAME-IS: a framework for documenting modifications to implementation strategies in healthcare. Implement Sci 2021;16(1):36 [FREE Full text] [doi: 10.1186/s13012-021-01105-3] [Medline: 33827716]
35. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011 Mar;38(2):65-76. [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]

Abbreviations

AED: Aedes spp.
AF: Anopheles funestus s.l.
AG: Anopheles gambiae s.l.
AO: Anopheles Other
CDC-LTC: Centers for Disease Control light trap catches
CFIR: Consolidated Framework for Implementation Research
CNN: convolutional neural network
CU: Culex spp.
DHIS2: District Health Information System 2
ERIC: Expert Recommendations for Implementing Change
FRAME-IS: Framework for Reporting Adaptations and Modifications–Enhanced for Implementation Strategies
HLC: human-landing catch
IOF: Implementation Outcomes Framework
KII: key informant interview
LMIC: low- and middle-income country
MAN: Mansonia spp.
MOH: Ministry of Health
NMED: National Malaria Elimination Division
NSR: nonsignificant risk
PCR: polymerase chain reaction
PSCs: pyrethrum spray catches
RCT: randomized controlled trial
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
TASO: AIDS Support Organization
TIDieR: Template for Intervention Description and Replication
VCOs: vector control officers
VHT: village health team

Edited by M Kopka; submitted 28.Apr.2026; peer-reviewed by M Opiyo, A Patel; comments to author 23.May.2026; revised version received 12.Jun.2026; accepted 20.Jun.2026; published 10.Sep.2026.

Please cite as:
Hey ML, Rincon Torroella M, Mohan D, Patel S, Kansiime W, Maiteki-Sebuguzi C, Shodhan A, Yazdi Y, Waiswa P, Acharya S Integration of VectorCam, an AI-Enabled Mosquito Surveillance Tool, Into Uganda's National Malaria Elimination Program: Protocol for a Mixed Methods Hybrid Type III Implementation Study
JMIR Public Health Surveill 2026;12:e99717
URL: https://publichealth.jmir.org/2026/1/e99717
doi:10.2196/99717
PMID:

©Mei-Li Hey, Marina Rincon Torroella, Diwakar Mohan, Sunny Patel, Winnifred Kansiime, Catherine Maiteki-Sebuguzi, Aryaman Shodhan, Youseph Yazdi, Peter Waiswa, Soumyadipta Acharya. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 10.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Broadband Access and Ophthalmologist Density in the United States: Cross-Sectional Questionnaire Study

Patrick Akarapimand[1], BS; John C Lin[1], BS; Ingrid U Scott[2,3], MD, MPH; Sally L Baxter[4,5], MD, MSc

[1]Department of Ophthalmology, Scheie Eye Institute, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

[2]Department of Ophthalmology, Pennsylvania State College of Medicine, Pennsylvania State University, Hershey, PA, United States

[3]Department of Public Health Sciences, Pennsylvania State College of Medicine, Pennsylvania State University, Hershey, PA, United States

[4]Hamilton Glaucoma Center and Division of Ophthalmology Informatics and Data Science, Viterbi Family Department of Ophthalmology and Shiley Eye Institute, University of California San Diego, 9415 Campus Point Drive, MC 0946, La Jolla, CA, United States

[5]Division of Biomedical Informatics, Department of Medicine, University of California San Diego

Corresponding Author:
Sally L Baxter, MD, MSc
Hamilton Glaucoma Center and Division of Ophthalmology Informatics and Data Science, Viterbi Family Department of Ophthalmology and Shiley Eye Institute, University of California San Diego, 9415 Campus Point Drive, MC 0946, La Jolla, CA, United States

Abstract

Background: Rural US communities experience disproportionately high rates of visual disability yet have limited access to ophthalmologists. Teleophthalmology may help address these gaps, but its effectiveness depends on broadband connectivity. The relationship between broadband access and ophthalmologist density has not been well characterized.

Objective: The aim of this study is to quantify the association between household broadband access—defined as subscription rates or connection prevalence—and county-level ophthalmologist density and to identify sociodemographic predictors of access.

Methods: We conducted an ecological study of all 3141 US counties using 2019 data from the American Community Survey, Area Health Resources File, and National Center for Health Statistics (NCHS). Broadband access was the primary exposure; ophthalmologist count with county population as an offset was the outcome. The primary analysis used negative binomial regression, adjusting for urbanicity, income, education, age, sex, race/ethnicity, unemployment, and insurance status. Sensitivity analyses included population-weighted linear regression and state fixed effects models. County-level heatmaps illustrated geographic patterns.

Results: Median household broadband access was 56.6%, ranging from 72.2% in the most urban counties (NCHS category 1) to 49.1% in the most rural (NCHS category 6). In unadjusted negative binomial regression, each 10-percentage-point increase in broadband access was associated with a 68% higher ophthalmologist rate (incidence rate ratio=1.68, 95% CI 1.61 - 1.76; $P<.001$). After adjustment, each 10-percentage-point increase was associated with a 46% higher rate (incidence rate ratio=1.46, 95% CI 1.37 - 1.56; $P<.001$). Sensitivity analyses were consistent with primary analysis. Regions with both low broadband access and zero ophthalmologist density were concentrated in the South, Mountain West region, and Alaska.

Conclusions: Broadband access is strongly associated with ophthalmologist availability across US counties, independent of sociodemographic factors. Areas lacking ophthalmologists also tend to lack broadband adoption, creating compounded barriers to both in-person and teleophthalmic care. Efforts to expand broadband may support more equitable access to vision services in underserved regions.

(JMIR Public Health Surveill 2026;12:e88473) doi:10.2196/88473

KEYWORDS

broadband; teleophthalmology; ophthalmologist density; rural urban divide; public health

Introduction

Roughly 20% of the US population lives in a rural setting; those living in rural settings are more likely than those living in urban setting to experience disability [1]. Of those experiencing disability, visual disability in particular is highly prevalent among adults aged 18 years or older, with recent estimates approximating that 7 million people in the United States live with visual impairment as defined by visual acuity loss [2,3]. Although 90% of people with visual impairment have preventable or treatable conditions [4], areas with limited or no access to local eye care providers pose a significant concern in the United States [5,6]. This disparity is particularly concentrated in rural areas, where there are lower densities of

XSL•FO
RenderX

ophthalmologists and patients are unable to travel the distance required to attend appointments [5,6].

Improving access to eye care in rural areas remains a key health policy; however, many obstacles exist for increasing provider density, access, and funding in these areas. Currently, only 11% of physicians practice in rural areas [7]. This lack of physicians is exacerbated by estimates that predict a nearly 25% reduction in the rural physician workforce as many rural physicians near retirement [8]. The effort to draw younger physicians to rural settings is complicated by fewer opportunities for spousal employment, more limited school resources for their children, and the financial disincentive of a lower income in the face of mounting student debt [7,9,10]. The idea of specialization in ophthalmology is also in particular discordance with the notion that rural physicians must practice more broadly due to fewer health care resources [11,12].

During the height of the COVID-19 pandemic, ambulatory outpatient visits declined by 79% [13], with studies showing lasting reductions in ophthalmic care utilization [14]. As an answer to the physical limitations of pandemic lockdown orders and national policies, teleophthalmology gained unprecedented attention and adoption on a large scale [15]. This growth of interest in teleophthalmology as a potential solution for otherwise geographically limited vision services also highlights the importance of broadband internet access, which is crucial for many components of teleophthalmology visits such as synchronous video with audio and asynchronous messaging and sharing of vital health information between patient and provider [16]. Beyond the primary connection that broadband access enables, it also connects patients to online resources containing vision health information [17].

Given the linkage of broadband access and teleophthalmology services, broadband access may be a useful target to improve access to vision care, particularly for rural populations who otherwise lack access to providers. However, to our knowledge, broadband access has not been studied in association with access to physical ophthalmic facilities, personnel, and services. Additionally, the intersection of broadband access across geographical distributions and other social determinants of health is poorly understood. Exploring this association may help inform public interventions to increase equity in broadband access and access to ophthalmic services. Therefore, we utilized data from the Area Health Resource Files, National Center for Health Statistics (NCHS), and American Community Survey (ACS) to assess for associations between disparities in broadband internet access and access to ophthalmology specialists.

Methods

Ethical Considerations

This study was considered exempt from full review per the University of Pennsylvania Institutional Review Board guidelines [18], as it used only deidentified, publicly available, aggregate county-level data and did not involve human subjects research requiring informed consent. The study adhered to the principles of the Declaration of Helsinki. No individual-level patient data were accessed or collected. All data were obtained from 2019 versions of each dataset.

Study Design

The number of practicing ophthalmologists per county and unemployment rates were obtained from the Area Health Resource Files. Urbanicity was categorized according to the NCHS 6-level urban–rural classification scheme (1=large central metro, 2=large fringe metro, 3=medium metro, 4=small metro, 5=micropolitan, and 6=noncore). This approach avoids the assumption of a linear step relationship between categories that would be imposed by treating the codes as continuous. All other demographic and socioeconomic variables were obtained from ACS, including median household income, percentage of the population who were female, Black, Hispanic, aged ≥65 years, with less than a high school education, and uninsured. Case complete analysis was performed with 91 counties omitted for missing data.

The primary exposure was broadband access, defined as the percentage of households in a county with broadband internet access such as cable, fiber optic, or DSL. This measure reflects the proportion of households with functional internet connectivity, rather than direct availability of broadband infrastructure, speed, or affordability. Broadband access was treated as a continuous variable and adjusted per 10% for interpretability. The primary outcome was the number of ophthalmologists per county. County population size was incorporated as an offset term to model ophthalmologist availability relative to county population.

Descriptive statistics were used to summarize broadband access across all counties. Bivariate associations between broadband access and ophthalmologist count were initially assessed using simple linear regression. Because the outcome variable represented a count of ophthalmologists and demonstrated overdispersion, the primary analysis used negative binomial regression to estimate the association between broadband access and the number of ophthalmologists per county. County population size was included as an offset term to model ophthalmologist supply relative to population. The multivariable model adjusted for potential county-level confounders including urbanicity, median household income, and percentage of the population who were female, Black, Hispanic, aged ≥65 years, with less than a high school education, unemployed, and uninsured. Incidence rate ratios (IRR) were then calculated. To assess potential multicollinearity among predictors, variance inflation factors were calculated for all independent variables.

Sensitivity analyses were conducted to evaluate the robustness of the findings. First, a population-weighted multiple linear regression model was fitted with ophthalmologist density (ophthalmologists per 100,000 population) as the outcome, weighting counties by population size to reduce the influence of sparsely populated counties. Second, models incorporating state fixed effects were estimated to control for unobserved state-level factors, such as regulatory environments, telehealth policy, and workforce training infrastructure.

County-level geographic boundaries were obtained from the US Census Bureau's 2019 TIGER/Line Shapefiles [19].

XSL•FO
RenderX

Broadband access and ophthalmologist density, adjusted into unique deciles for interpretability, were then mapped to US counties. Additional heatmaps identified counties with low broadband access (defined as <1st quartile) and low ophthalmologist density (defined as 0 ophthalmologists per 100,000 people). We also conducted further exploratory analysis to investigate whether household broadband access varies systematically across county socioeconomic characteristics. All analyses and visualizations were conducted using RStudio (version 4.4.0; Posit).

Results

Data from 3141 US counties were included. Median broadband access across all counties was 56.6%. In the most urban counties (NCHS category 1), 72.2% of households had broadband, and there were 9.1 ophthalmologists per 100,000 people, whereas the most rural counties (NCHS category 6) had 49.1% of households with broadband and 0.43 ophthalmologists per 100,000.

Bivariate analysis showed a significant association between broadband access and ophthalmologist counts, where each 10-percentage-point increase in household broadband access was associated with a 68% higher rate of ophthalmologists per county (IRR=1.68, 95% CI 1.61 - 1.76; *P*<.001). After adjusting for urbanicity, income, education, age, sex, race/ethnicity, unemployment, and insurance status, each 10-percentage-point increase in broadband access was associated with a 46% higher ophthalmologist supply per capita (IRR 1.46, 95% CI 1.37 - 1.56; *P*<.001, Table 1).

Table . Negative binomial regression of ophthalmologist counts on broadband access and covariates.

Variable	Incidence rate ratio	95% CI	*P* value
Household broadband access (per 10%)	1.46	1.37 - 1.56	<.001
NCHS category 2 (large fringe metro)	0.47	0.39 - 0.55	<.001
NCHS category 3 (medium metro)	0.82	0.69 - 0.98	.03
NCHS category 4 (small metro)	0.99	0.82 - 1.19	.93
NCHS category 5 (micropolitan)	0.85	0.7 - 1.04	.12
NCHS category 6 (noncore)	0.32	0.25 - 0.42	<.001
Median household income (per $10,000)	0.99	0.95 - 1.03	.54
Less than high school education (%)	0.96	0.94 - 0.97	<.001
Age ≥65 (%)	1.01	0.99 - 1.02	.35
Female (%)	1.21	1.16 - 1.26	<.001
Black (%)	1.02	1.01 - 1.02	<.001
Hispanic (%)	1.01	1.01 - 1.02	<.001
Unemployment rate (%)	0.94	0.90 - 0.98	.005
Uninsured (%)	0.99	0.97 - 1.00	.08

In multiple linear regression, we assessed ophthalmologist density and broadband access after adjusting for urbanicity (NCHS Urban-Rural Classification), median household income (per $10,000), and percentage of the population who were female, Black, Hispanic, aged 65 years and older, with less than a high school education, unemployed, and uninsured.

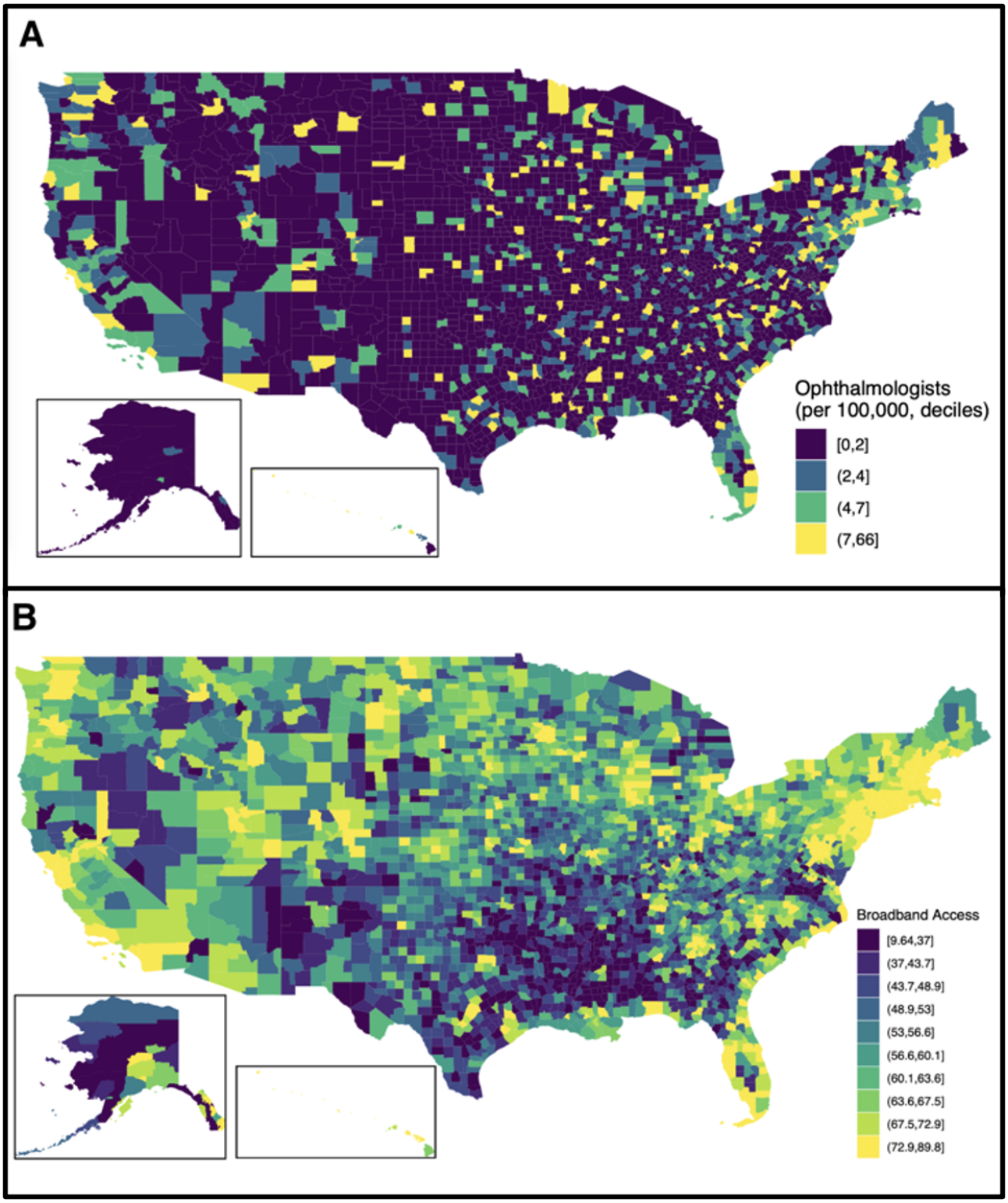
Compared with large central metropolitan counties, noncore counties had a 68% lower ophthalmologist availability (IRR 0.32, 95% CI 0.25 - 0.42; *P*<.001). Variance inflation factors for all predictors were below 2, indicating minimal multicollinearity.

Population-weighted linear regression produced consistent findings. Each 10-percentage-point increase in broadband access was associated with 0.43 additional ophthalmologists per 100,000 population ((β=.43, 95% CI 0.18 - 0.68; *P*<.001). Furthermore, models including state fixed effects demonstrated a persistent association (IRR 1.63, 95% CI 1.51 - 1.76; *P*<.001), suggesting that the relationship was present even accounting for unobserved state-level factors.

Figure 1 illustrates geographic variation in ophthalmologist density and broadband access by US counties. Figure 2 illustrates that counties with both low broadband access and ophthalmologist density were concentrated in the South, Mountain West, and Alaska regions.

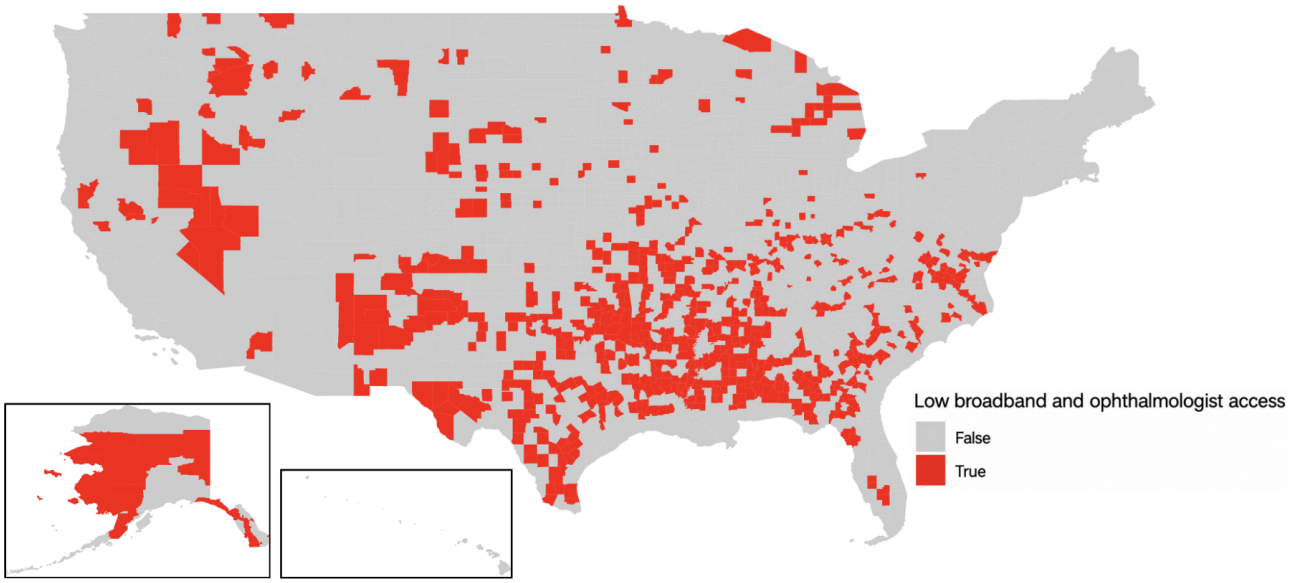
XSL•FO
RenderX

Figure 1. Heatmaps of ophthalmologist density (Panel A) and broadband access (Panel B) by US county, 2019.



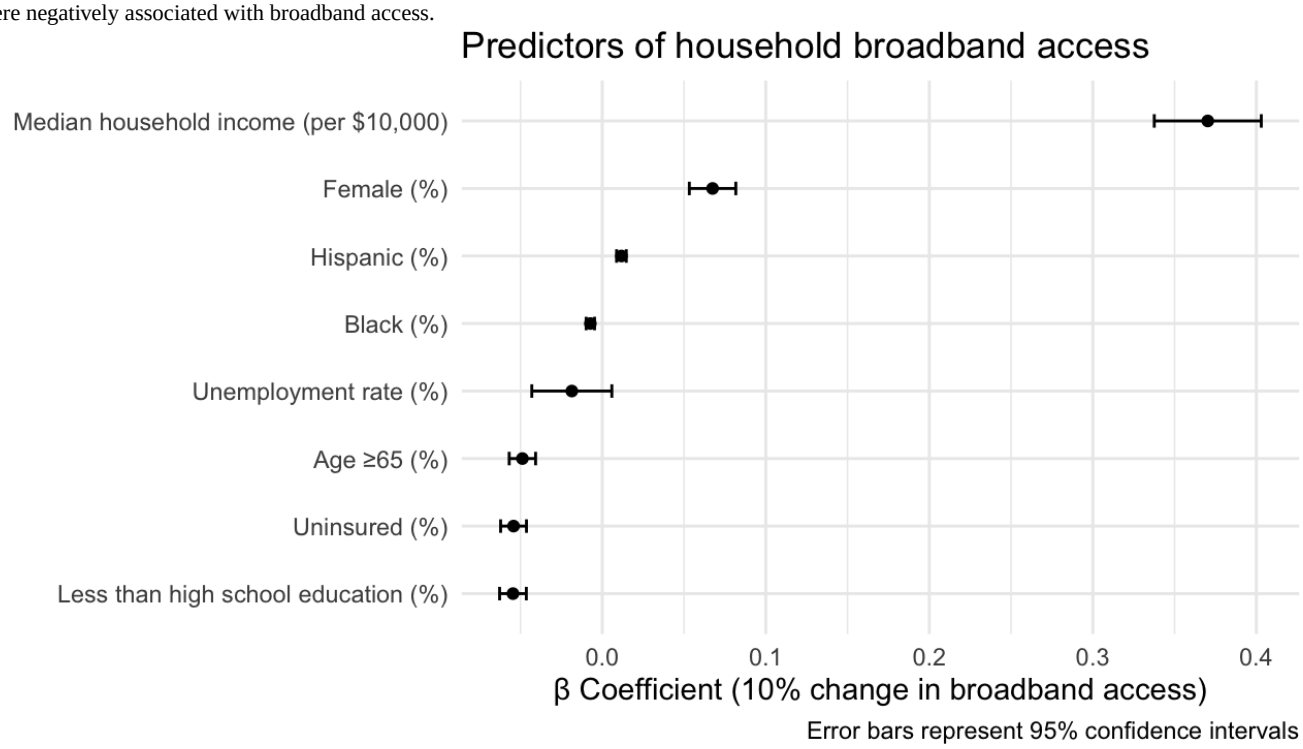
XSL•FO
RenderX

Figure 2. Heatmap of US counties with both low broadband access and low ophthalmologist access, 2019. Defined by broadband access <1st quartile (46.27%) and ophthalmologist density of 0 per 100,000 people. In total, 700 counties met these criteria, encompassing approximately 11.4 million residents. The heatmap demonstrates concentration in the South, Mountain West, and Alaska regions.



In multivariable linear regression, several sociodemographic characteristics were associated with household broadband access across the United States. Lower educational attainment (β=−0.05, 95% CI −0.06 to −0.05), higher proportions of adults ≥65 years (β=−0.05, 95% CI −0.06 to −0.05), higher percentages of Black residents (β=−0.007, 95% CI −0.010 to −0.005), and higher percentages of uninsured residents (β=−0.05, 95% CI −0.06 to −0.05) were associated with lower broadband access. Conversely, higher median household income (β=0.37, 95% CI 0.34 to 0.40), greater proportions of female residents (β=0.07, 95% CI 0.05 to 0.08), and higher percentages of Hispanic residents (β=0.01, 95% CI 0.009 to 0.015) were positively associated with access. Unemployment showed a marginal negative association that did not reach significance (β=−0.02, 95% CI −0.04 to 0.006). Although several predictors reached statistical significance, the effect sizes were modest. Results are visualized in Figure 3.

Figure 3. Forest plot showing β coefficients and 95% CIs from multiple linear regression of household broadband access among socioeconomic covariates. Results show that higher proportions of females and Hispanic people, as well as higher median household income, are positively associated with broadband access, whereas high proportions of Black people, adults aged 65 or older, uninsured people, and those with no high school education were negatively associated with broadband access.



XSL•FO
RenderX

Discussion

Teleophthalmology has been suggested as an innovative solution to improve access to ophthalmic services in communities with fewer vision care resources. However, broadband access and adoption to enable the full functionalities of teleophthalmology in these areas has not been well studied. In this study assessing broadband access with ophthalmologist density in 3141 US counties, we found that US counties with reduced broadband access also have lower average counts and densities of ophthalmologists. Importantly, the association between broadband access and ophthalmologist supply remained robust across multiple sensitivity analyses designed to address potential methodological concerns. Results were consistent after population-weighted modeling and incorporating state fixed effects, suggesting that the observed relationship is not driven by small population counties or unmeasured state-level factors, strengthening the inference that broadband access and ophthalmologist distribution are closely linked structural features of health care access. This finding underscores that while teleophthalmology may mitigate key barriers to in-person vision care, lack of broadband access may present an additional, underexplored barrier to care for individuals across the spectrum of urbanicity. These findings may be particularly concentrated among rural regions in the South, Mountain West region, and Alaska. Furthermore, in secondary analysis, we found that broadband access tends to be independently lower in communities that are more rural, have lower levels of educational attainment, and higher proportions of older adults over 65, Black residents, and uninsured individuals, although with modest effect sizes.

An important interpretive consideration is that the ACS broadband measure used in this study captures household subscription rates—reflecting whether households have an active broadband connection—rather than physical infrastructure availability, service speeds, or affordability. A county may have fiber infrastructure yet show low subscription rates due to cost barriers, limited digital literacy, or a lack of devices. Conversely, a county may have high adoption rates despite suboptimal infrastructure due to demographic and socioeconomic characteristics. Therefore, our findings should be interpreted as demonstrating an association between low broadband adoption and low ophthalmologist density, rather than between infrastructure deficits and provider scarcity per se. Efforts to improve broadband equity must therefore address both availability (deployment of physical infrastructure) and adoption (affordability subsidies, device access programs, and digital literacy training), as low adoption can persist even after infrastructure gaps are closed.

Our findings align with current literature demonstrating sociodemographic disparities in broadband access. For instance, several studies suggest that areas with larger proportions of Black and American Indian/Alaska Native populations, lower economic status, and lower educational attainment tend to have less broadband access [20,21]. Our results also further support several studies demonstrating disproportionately limited access to ophthalmic care in rural areas and communities with lower socioeconomic status [22-24]. Our study adds to the current literature that the association between household broadband access and ophthalmologist density persists across the urban/rural divide. This means that even in metropolitan counties where in-person care is relatively accessible, low broadband subscription may independently limit the reach of supplemental teleophthalmology services. This is exacerbated in nonmetropolitan areas, where even fewer ophthalmologists are available and there is poorer broadband access. This study thus expands on the current literature by directly contextualizing broadband access within the broader scope of ophthalmology access, showing that areas void of physical providers also lack the connectivity necessary for remote solutions, creating compounded access barriers.

The clear association between broadband access and socioeconomic factors raises the concern that broadband expansion alone may not lead to meaningful increases in health care utilization. Digital literacy gaps, especially among older adults, disabled people, or low-income populations, may persist independent of connectivity. Indeed, studies have found that older adults and individuals with lower education levels may be less likely to engage with online health tools or complete telehealth visits due to unfamiliarity with technology or distrust in virtual care models [25-29]. Others have argued that teleophthalmology is limited in clinical scope for procedures requiring in-person physical exams, diagnostic imaging, or surgical interventions [30-33]. Another critique is based on cost-effectiveness: broadband infrastructure costs are particularly high in geographically isolated regions, where population density may be too low to justify investment on health care grounds alone [34-36].

However, these concerns, while valid, do not negate the importance of broadband as a foundational infrastructure for equitable health care delivery. First, digital literacy gaps are surmountable with targeted interventions. Programs that pair broadband rollouts with community-based digital navigator training have successfully improved technology adoption in older adults, racial/ethnic minorities, and low-income groups [37-40]. Second, teleophthalmology has evolved beyond synchronous video visits to include asynchronous image-based screening, such as store-and-forward models for diabetic retinopathy and glaucoma, which require minimal digital interaction by the patient and are well-suited for underserved communities [30,41-45]. For instance, Google/Verily's ARDA partnership with Aravind Eye Care in India demonstrates that clinic-based broadband can mitigate household limitations [46]. Nevertheless, widespread household connectivity may enhance patient participation in remote screening and education.

Federal efforts such as the Broadband Equity, Access, and Deployment program, part of the 2021 Infrastructure Investment and Jobs Act, already allocate billions of dollars toward rural broadband with explicit equity goals, aligning policy momentum with the needs demonstrated in this study [47]. Further interventions include the expansion of federal programs such as the US Department of Agriculture Rural eConnectivity Pilot Program [48], including the removal of eligibility restrictions to allow all rural areas with inadequate broadband access to apply for funding. Furthermore, telehealth policy reform, such as reimbursement parity and interstate licensure compacts, will

XSL•FO
RenderX

further support teleophthalmology and eye care access in underserved regions [34,49]. Our results provide empirical evidence that leveraging policy funding to support prioritizing broadband investments, particularly in the rural South, Mountain West region, and Alaska and in regions with lower educational attainment and higher proportions of older adults, Black people, and uninsured people, may be particularly high yield for both technological and health system gains.

This study contains a few key limitations. First, the ACS broadband measure captures household subscription rates, not infrastructure availability, speed, latency, or affordability, and therefore cannot distinguish between cost barriers, digital literacy gaps, device limitations, and physical infrastructure deficits as drivers of low broadband access. Furthermore, our data only consider the percentage of households that lack access to broadband internet. As a result, access to broadband outside the household in facilities such as workplaces, schools, or public facilities remains unclear. However, transportation, lack of convenience, and privacy concerns arising from utilizing broadband access outside of one's home may present as barriers; therefore, accessibility to household broadband internet is still relevant. In addition, our data were obtained from 2019, which was before the massive expansion in broadband internet access in response to the COVID-19 pandemic. Lastly, this ecological, cross-sectional study cannot establish causation or temporal sequence, and unmeasured county-level variables may confound associations, as is the case with all cross-sectional studies.

To our knowledge, this study represents the first national analysis examining the relationship between broadband access and ophthalmologist workforce distribution across all US counties. These two factors appear to be intertwined and may reinforce each other in determining rural and underserved communities' connectivity to vision care. As teleophthalmology becomes integral to future eye care delivery, advancing broadband equity should be positioned as a public health strategy that simultaneously bolsters virtual and in-person specialist access.

Funding

This research was supported by the National Institutes of Health (NIH) grants T35 EY033704, DP5OD029610, OT2OD032644, and P30EY022589 and an unrestricted departmental grant from Research to Prevent Blindness. The sponsor or funding organization had no role in the design or conduct of this research.

Data Availability

The dataset created for this study is available in Multimedia Appendix 1.

Authors' Contributions

PA: Conceptualization, formal analysis, investigation, methodology, software, visualization, writing – original draft, writing – review & editing
JCL: Conceptualization, formal analysis, methodology, software, writing – review & editing
IUS: Conceptualization, methodology, validation, writing – review & editing
SLB: Conceptualization, funding acquisition, methodology, project administration, validation, writing – review & editing

Conflicts of Interest

None declared.

Multimedia Appendix 1
Dataset.
[ZIP File, 4748 KB - publichealth_v12i1e88473_app1.zip]

References

1. The South had highest disability rate among regions in 2021. United States Census Bureau. 2023. URL: https://www.census.gov/library/stories/2023/06/disability-rates-higher-in-rural-areas-than-urban-areas.html [accessed 2024-09-13]
2. Fast facts: vision loss. Centers for Disease Control and Prevention. 2024 May 15. URL: https://www.cdc.gov/vision-health/data-research/vision-loss-facts/index.html [accessed 2024-09-13]
3. Flaxman AD, Wittenborn JS, Robalik T, et al. Prevalence of visual acuity loss or blindness in the US: a Bayesian meta-analysis. JAMA Ophthalmol 2021 Jul 1;139(7):717-723. [doi: 10.1001/jamaophthalmol.2021.0527] [Medline: 33983373]
4. Burton MJ, Ramke J, Marques AP, et al. The Lancet Global Health Commission on Global Eye Health: vision beyond 2020. Lancet Glob Health 2021 Apr;9(4):e489-e551. [doi: 10.1016/S2214-109X(20)30488-5] [Medline: 33607016]
5. Lundeen EA, Flaxman AD, Wittenborn JS, et al. County-level variation in the prevalence of visual acuity loss or blindness in the US. JAMA Ophthalmol 2022 Aug 1;140(8):831-832. [doi: 10.1001/jamaophthalmol.2022.2405] [Medline: 35797011]

XSL•FO
RenderX

6. Ervin AM, Solomon SD, Shoge RY. Access to eye care in the United States: evidence-informed decision-making is key to improving access for underserved populations. Ophthalmology 2022 Oct;129(10):1079-1080. [doi: 10.1016/j.ophtha.2022.07.011] [Medline: 36058738]
7. Jaret P. Attracting the next generation of physicians to rural medicine. AAMC. 2020. URL: https://www.aamc.org/news/attracting-next-generation-physicians-rural-medicine [accessed 2024-09-13]
8. Skinner L, Staiger DO, Auerbach DI, Buerhaus PI. Implications of an aging rural physician workforce. N Engl J Med 2019 Jul 25;381(4):299-301. [doi: 10.1056/NEJMp1900808] [Medline: 31340091]
9. Medical student debt. AMSA. URL: https://www.amsa.org/action-committee/wellness-student-life-action-committee/medical-student-debt/ [accessed 2024-09-13]
10. Burr WH, Everitt JG, Johnson JM. “The debt is suffocating to be honest”: student loan debt, prospective sensemaking, and the social psychology of precarity in an allopathic medical school. SSM - Qualitative Research in Health 2023 Dec;4:100304. [doi: 10.1016/j.ssmqr.2023.100304]
11. Kaneko M, Higuchi T, Ohta R. Primary care physicians working in rural areas provide a broader scope of practice: a cross-sectional study. BMC Prim Care 2024 Jan 2;25(1):9. [doi: 10.1186/s12875-023-02250-y] [Medline: 38166745]
12. Peterson LE, Fang B. Rural family physicians have a broader scope of practice than urban family physicians. : Rural and Underserved Health Research Center; 2018 URL: https://ruhrc.uky.edu/assets/ruhrc-publication-2018-02-23-rural-family-physicians-have-a-broader-scope-of-practice-than-urban-family-physicians.pdf [accessed 2026-05-14]
13. Mehrotra A, Chernew ME, Linetsky D, Hatch H, Cutler DM. The impact of the COVID-19 pandemic on outpatient visits: a rebound emerges. : Commonwealth Fund; 2020. [doi: 10.26099/ds9e-jm36]
14. Li C, Lum F, Chen EM, et al. Shifts in ophthalmic care utilization during the COVID-19 pandemic in the US. Commun Med (Lond) 2023 Dec 14;3(1):181. [doi: 10.1038/s43856-023-00416-4] [Medline: 38097811]
15. Ricur G, Reyes J, Alfonso E, Marino RG. Surfing the COVID-19 tsunami with teleophthalmology: the advent of new models of eye care. Curr Ophthalmol Rep 2023;11(1):1-12. [doi: 10.1007/s40135-023-00308-9] [Medline: 36743397]
16. Introduction to broadband and high speed internet. BroadbandUSA, National Telecommunications and Information Administration. 2022. URL: https://broadbandusa.ntia.gov/sites/default/files/2022-12/Introduction_to_Broadband_and_High_Speed_Internet_FINAL_0.pdf [accessed 2026-05-14]
17. Benda NC, Veinot TC, Sieck CJ, Ancker JS. Broadband internet access is a social determinant of health!. Am J Public Health 2020 Aug;110(8):1123-1125. [doi: 10.2105/AJPH.2020.305784] [Medline: 32639914]
18. Standard operating policies (SOP) (Version 14.0). University of Pennsylvania Institutional Review Board. URL: https://irb.upenn.edu/documents/standard-operating-policies/ [accessed 2026-05-14]
19. TIGER/line with selected demographic and economic data. United States Census Bureau. URL: https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-data.html [accessed 2025-07-20]
20. Zahnd WE, Bell N, Larson AE. Geographic, racial/ethnic, and socioeconomic inequities in broadband access. J Rural Health 2022 Jun;38(3):519-526. [doi: 10.1111/jrh.12635] [Medline: 34792815]
21. Reddick CG, Enriquez R, Harris RJ, Sharma B. Determinants of broadband access and affordability: an analysis of a community survey on the digital divide. Cities 2020 Nov;106:102904. [doi: 10.1016/j.cities.2020.102904] [Medline: 32921864]
22. Soares RR, Mokhashi N, Sharpe J, et al. Patient accessibility to eye care in the United States. Ophthalmology 2023 Apr;130(4):354-360. [doi: 10.1016/j.ophtha.2022.11.017] [Medline: 36427602]
23. Hemmerich C, Jones G, Staggs J, Anderson RM, Bacani R, Vassar M. Inequities and research gaps in ophthalmology: a scoping review. JAMA Ophthalmol 2023 Jan 1;141(1):63-70. [doi: 10.1001/jamaophthalmol.2022.5237] [Medline: 36480183]
24. Yadlapalli N, Hollinger R, Berzack S, Spies D, Patel A, Sridhar J. Potential gaps in eye care based on evaluation of federally qualified health centers. JAMA Ophthalmol 2024 Sep 19;142(11):1018-1026. [doi: 10.1001/jamaophthalmol.2024.3569] [Medline: 39298148]
25. Masterson Creber R, Dodson JA, Bidwell J, et al. Telehealth and health equity in older adults with heart failure: a scientific statement from the American Heart Association. Circ Cardiovasc Qual Outcomes 2023 Nov;16(11):e000123. [doi: 10.1161/HCQ.0000000000000123] [Medline: 37909212]
26. Li W, Ornstein KA, Li Y, Liu B. Barriers to learning a new technology to go online among older adults during the COVID-19 pandemic. J Am Geriatr Soc 2021 Nov;69(11):3051-3057. [doi: 10.1111/jgs.17433] [Medline: 34409589]
27. Choxi H, VanDerSchaaf H, Li Y, Morgan E. Telehealth and the digital divide: identifying potential care gaps in video visit use. J Med Syst 2022 Jul 30;46(9):58. [doi: 10.1007/s10916-022-01843-x] [Medline: 35906432]
28. Airola E. Learning and use of eHealth among older adults living at home in rural and nonrural settings: systematic review. J Med Internet Res 2021 Dec 2;23(12):e23804. [doi: 10.2196/23804] [Medline: 34860664]
29. Ladin K, Porteny T, Perugini JM, et al. Perceptions of telehealth vs in-person visits among older adults with advanced kidney disease, care partners, and clinicians. JAMA Netw Open 2021 Dec 1;4(12):e2137193. [doi: 10.1001/jamanetworkopen.2021.37193] [Medline: 34870680]

XSL•FO
RenderX

30. Mosenia A, Li P, Seefeldt R, Seitzman GD, Sun CQ, Kim TN. Longitudinal use of telehealth during the COVID-19 pandemic and utility of asynchronous testing for subspecialty-level ophthalmic care. JAMA Ophthalmol 2023 Jan 1;141(1):56-61. [doi: 10.1001/jamaophthalmol.2022.4984] [Medline: 36454548]
31. Aziz K, Moon JY, Parikh R, et al. Association of patient characteristics with delivery of ophthalmic telemedicine during the COVID-19 pandemic. JAMA Ophthalmol 2021 Nov 1;139(11):1174-1182. [doi: 10.1001/jamaophthalmol.2021.3728] [Medline: 34554212]
32. Fatehi F, Jahedi F, Tay-Kearney ML, Kanagasingam Y. Teleophthalmology for the elderly population: a review of the literature. Int J Med Inform 2020 Apr;136:104089. [doi: 10.1016/j.ijmedinf.2020.104089] [Medline: 32044698]
33. Chen EM, Andoh JE, Nwanyanwu K. Socioeconomic and demographic disparities in the use of telemedicine for ophthalmic care during the COVID-19 pandemic. Ophthalmology 2022 Jan;129(1):15-25. [doi: 10.1016/j.ophtha.2021.07.003] [Medline: 34245753]
34. Cortelyou-Ward K, Atkins DN, Noblin A, Rotarius T, White P, Carey C. Navigating the digital divide: barriers to telehealth in rural areas. J Health Care Poor Underserved 2020;31(4):1546-1556. [doi: 10.1353/hpu.2020.0116] [Medline: 33416736]
35. Li X, Huang L, Zhang H, Liang Z. Enabling telemedicine from the system-level perspective: scoping review. J Med Internet Res 2025;27:e65932. [doi: 10.2196/65932]
36. Loccoh EC, Nguyen A, Kim G, Warraich HJ. Geospatial analysis of access to health care and internet services in the US. JAMA Netw Open 2022 Nov 1;5(11):e2243792. [doi: 10.1001/jamanetworkopen.2022.43792] [Medline: 36416826]
37. Neumann AV, Gonzalez A, Walter LC, Rivera J. Improving older adults' telehealth through a novel community-academic partnership: preliminary data. J Am Geriatr Soc 2023 Dec;71(12):3886-3895. [doi: 10.1111/jgs.18556] [Medline: 37607098]
38. Jezewski E, Miller A, Eusebio M, Potter J. Targeted telehealth education increases interest in using telehealth among a diverse group of low-income older adults. Int J Environ Res Public Health 2022 Oct 16;19(20):13349. [doi: 10.3390/ijerph192013349] [Medline: 36293929]
39. Bailey JE, Gurgol C, Pan E, et al. Early patient-centered outcomes research experience with the use of telehealth to address disparities: scoping review. J Med Internet Res 2021 Dec 7;23(12):e28503. [doi: 10.2196/28503] [Medline: 34878986]
40. Adepoju OE, Dang P, Fuentes C, Liaw W. Examining the affordable connectivity program and telehealth use: a pilot survey of the affordable connectivity program, telehealth, video and audio visits in a racially diverse, lower-income population. Sci Rep 2025 Jan 17;15(1):2321. [doi: 10.1038/s41598-025-86728-y] [Medline: 39833310]
41. Rathi S, Tsui E, Mehta N, Zahid S, Schuman JS. The current state of teleophthalmology in the United States. Ophthalmology 2017 Dec;124(12):1729-1734. [doi: 10.1016/j.ophtha.2017.05.026] [Medline: 28647202]
42. Bai P, Burt SS, Woodward MA, et al. Federally qualified health centers as a model to improve vision health: a systematic review. JAMA Ophthalmol 2025 Mar 1;143(3):242-251. [doi: 10.1001/jamaophthalmol.2024.6264] [Medline: 39946139]
43. Bonilla-Escobar FJ, Ghobrial AI, Gallagher DS, Eller A, Waxman EL. Comprehensive insights into a decade-long journey: the evolution, impact, and human factors of an asynchronous telemedicine program for diabetic retinopathy screening in Pennsylvania, United States. PLOS One 2024;19(7):e0305586. [doi: 10.1371/journal.pone.0305586] [Medline: 38995899]
44. Horton MB, Brady CJ, Cavallerano J, et al. Practice guidelines for ocular telehealth-diabetic retinopathy, third edition. Telemed J E Health 2020 Apr;26(4):495-543. [doi: 10.1089/tmj.2020.0006] [Medline: 32209018]
45. Jo JJ, Pasquale LR. Recent developments of telemedicine in glaucoma. Curr Opin Ophthalmol 2024 Mar 1;35(2):116-123. [doi: 10.1097/ICU.0000000000001019] [Medline: 38295153]
46. Brant A, Singh P, Yin X, et al. Performance of a deep learning diabetic retinopathy algorithm in India. JAMA Netw Open 2025 Mar 3;8(3):e250984. [doi: 10.1001/jamanetworkopen.2025.0984] [Medline: 40105843]
47. Broadband Equity, Access, and Deployment (BEAD) Program. National Telecommunications and Information Administration. URL: https://www.ntia.gov/funding-programs/high-speed-internet-programs/broadband-equity-access-and-deployment-bead-program [accessed 2025-07-20]
48. ReConnect loan and grant program. USDA. 2025 Nov 13. URL: https://www.usda.gov/sustainability/infrastructure/broadband/reconnect-loan-and-grant-program [accessed 2025-11-16]
49. Brady CJ, D'Amico S, Withers N, Kim BY. Using public datasets to identify priority areas for ocular telehealth. Telemed J E Health 2021 Nov;27(11):1293-1298. [doi: 10.1089/tmj.2020.0433] [Medline: 33600257]

Abbreviations

ACS: American Community Survey
IRR: incidence rate ratio
NCHS: National Center for Health Statistics

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 25.Nov.2025; peer-reviewed by S Wu, Y Yang; revised version received 08.Apr.2026; accepted 10.Apr.2026; published 19.May.2026.

Please cite as:
Akarapimand P, Lin JC, Scott IU, Baxter SL
Broadband Access and Ophthalmologist Density in the United States: Cross-Sectional Questionnaire Study
JMIR Public Health Surveill 2026;12:e88473
URL: https://publichealth.jmir.org/2026/1/e88473
doi:10.2196/88473

© Patrick Akarapimand, John C Lin, Ingrid U Scott, Sally L Baxter. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 19.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Digital Health Center App for Community and National Malaria Surveillance in Cambodia: Implementation Case Study

Pengby Ngor[1,2,3], BSc, MPhil; Huy Rekol[1], MD, MPH, DTMH, EFTP; Eng Thyda[1], BSc; Ry Rith[1], BSc, MSc; Sok Kimleng[1], BSc; Yoem Rattana[1], BSc; Ou Vunsokserey[1], BSc; Hem Vanna[1], BSc; Chang Pumeoung[1], BSc; Pascal Ringwald[4], MD, PhD; Richard J Maude[2,3,5], BSc, MBChB, DTM&H, DPhil, MD; Lisa White[6], BSc, PhD; Siv Sovannaroth[1], MAGMed, DipAPE

[1]National Center for Parasitology, Entomology and Malaria Control, No. 477, Betong (St.), corner of St. 92, Trapeang Svay Village, Phnom Penh, Cambodia

[2]Mahidol Oxford Tropical Medicine Research Unit, Bangkok, Thailand

[3]The Open University, Milton Keynes, United Kingdom

[4]WHO Mekong Malaria Elimination Programme, Phnom Penh, Cambodia

[5]Department of Medicine, Centre for Tropical Medicine and Global Health, University of Oxford, Oxford, United Kingdom

[6]Model Health Ltd, Brixham, United Kingdom

Corresponding Author:
Pengby Ngor, BSc, MPhil
National Center for Parasitology, Entomology and Malaria Control, No. 477, Betong (St.), corner of St. 92, Trapeang Svay Village, Phnom Penh, Cambodia

Abstract

Background: By 2015, the emergence and dissemination of multidrug-resistant *Plasmodium falciparum* in the Greater Mekong Subregion threatened regional and global malaria control efforts. In response, Greater Mekong Subregion countries committed to malaria elimination by 2030, with strengthened surveillance as a strategic pillar. In 2017, Cambodia introduced an elimination-oriented digital Malaria Information System (MIS). Its health center app enables real-time, geo-located, case-based malaria reporting across primary health centers, and is fully integrated with the MIS.

Objective: This study aimed to evaluate the real-world national implementation of Cambodia's Android-based health center app, considering coverage, fidelity, timeliness, and data use, and their effects on malaria surveillance performance, case management, programmatic response, and public health outcomes.

Methods: System performance and public health use were assessed using system-generated metadata, national surveillance data, and user surveys. Operational indicators included technical performance, data completeness, and reporting timeliness, alongside surveillance outcomes such as case notification, classification, reactive case detection, and foci investigation. Nationwide user experience was measured via a survey of 761 health centers across 21 provinces, with in-depth structured surveys at 9 health centers in 3 provinces. Descriptive analyses evaluated system functionality, contribution to malaria surveillance and response, and usability among frontline health workers.

Results: The health center app demonstrated strong technical performance, with rapid loading and resilient data transmission under low-bandwidth conditions, supporting reliable reporting in resource-constrained settings. Integrated real-time dashboards provided analytics for case management, surveillance monitoring, risk stratification, and targeted public health interventions. Data completeness remained high (99%, 89/90 fields in 2024), demonstrating consistent routine use even as case incidence declined. Between January 1, 2025, and July 31, 2025, 69 malaria cases were reported nationally (23 locally acquired, 7 domestically imported, and 39 internationally imported). Of these, 95.7% (66/69) were notified and classified within 1 day. Reactive case detection was completed within 3 days for all 21 eligible cases, and 16 of 19 eligible foci received a response within 7 days, indicating strong operational responsiveness. User surveys showed 96.3% (733/761) of health centers were satisfied or very satisfied, 90.1% (686/761) reported rare or no technical issues, and 91.7% (698/761) found the app easy to navigate. Operational challenges included limited internet connectivity, transport to remote areas, and electricity interruptions. In-depth surveys confirmed high uptake, confidence in reporting, and routine use of surveillance data, although gaps in local analytical capacity were identified.

Conclusions: Developed and managed locally to enhance sustainability, the MIS drove significant reductions in malaria case incidence, with the health center app contributing timely, complete, structured reporting at the point of care. Public health responses were facilitated by real-time analysis, targeted interventions, and decentralized decision-making. User engagement was sustained as malaria cases declined, and further enhancements are planned to ensure seamless transition to postelimination surveillance, reducing the risk of malaria reestablishment in Cambodia.

XSL•FO
RenderX

(JMIR Public Health Surveill 2026;12:e85881) doi:10.2196/85881

KEYWORDS

malaria; Cambodia; surveillance; elimination; information systems

Introduction

Malaria is a life-threatening parasitic disease caused by *Plasmodium spp*., transmitted to humans through the bites of infected *Anopheles* mosquitoes. Malaria continues to pose a major global health challenge, causing significant morbidity and mortality [1].

The Greater Mekong Subregion (GMS) has been a focus for the development of parasite resistance to antimalarial drugs. In the late 1950s, chloroquine-resistant *Plasmodium falciparum* emerged in the GMS, and over the following 2 decades disseminated worldwide [2]. This led to extensive treatment failures and significantly undermined malaria control programs globally [3,4]. Similarly, the GMS was the source of pyrimethamine-resistant parasites which quickly spread to Africa [5]. The introduction of artemisinin-based combination therapy (ACT) in 2001 marked a turning point in malaria control, restoring treatment efficacy and becoming the cornerstone of global treatment strategies. Consequently, reports in 2008 of the emergence of *P. falciparum* artemisinin partial resistance on the Thai-Cambodia border generated considerable concern. The subsequent evolution of multiresistant parasites and their rapid proliferation across the GMS complicated efforts to treat and contain the disease [6-8]. There was a considerable risk that these multiresistant parasites would spread globally, causing malaria to become untreatable [9]. This led to urgent and coordinated regional efforts to contain malaria [10], which by 2015 had evolved into a commitment to attain malaria elimination in the entire GMS by 2030 [11].

Malaria elimination is the interruption of local transmission in a defined geographical area [12]. It is an ambitious objective that is complex to achieve. Elimination is pursued in phases [11,12]. In the transmission-reduction phase, broad interventions are applied to reduce disease burden, such as widespread vector control, including the distribution of insecticide bed nets, and strengthening malaria services to deliver effective medicines. Once case incidence is low (typically below 1 case/1000 people/year), the elimination phase begins, and surveillance becomes the primary intervention [11,12]. The aim is to systematically drain the human transmission reservoir by finding and eliminating all parasites and interrupting transmission. To achieve this, a strong and responsive surveillance system is needed, capable of promptly detecting and treating every case, determining the origin of each case, triggering a timely and appropriate public health response, and detecting any changes in therapeutic drug efficacy or transmission patterns. This is generally achieved via one or more of three main surveillance strategies:

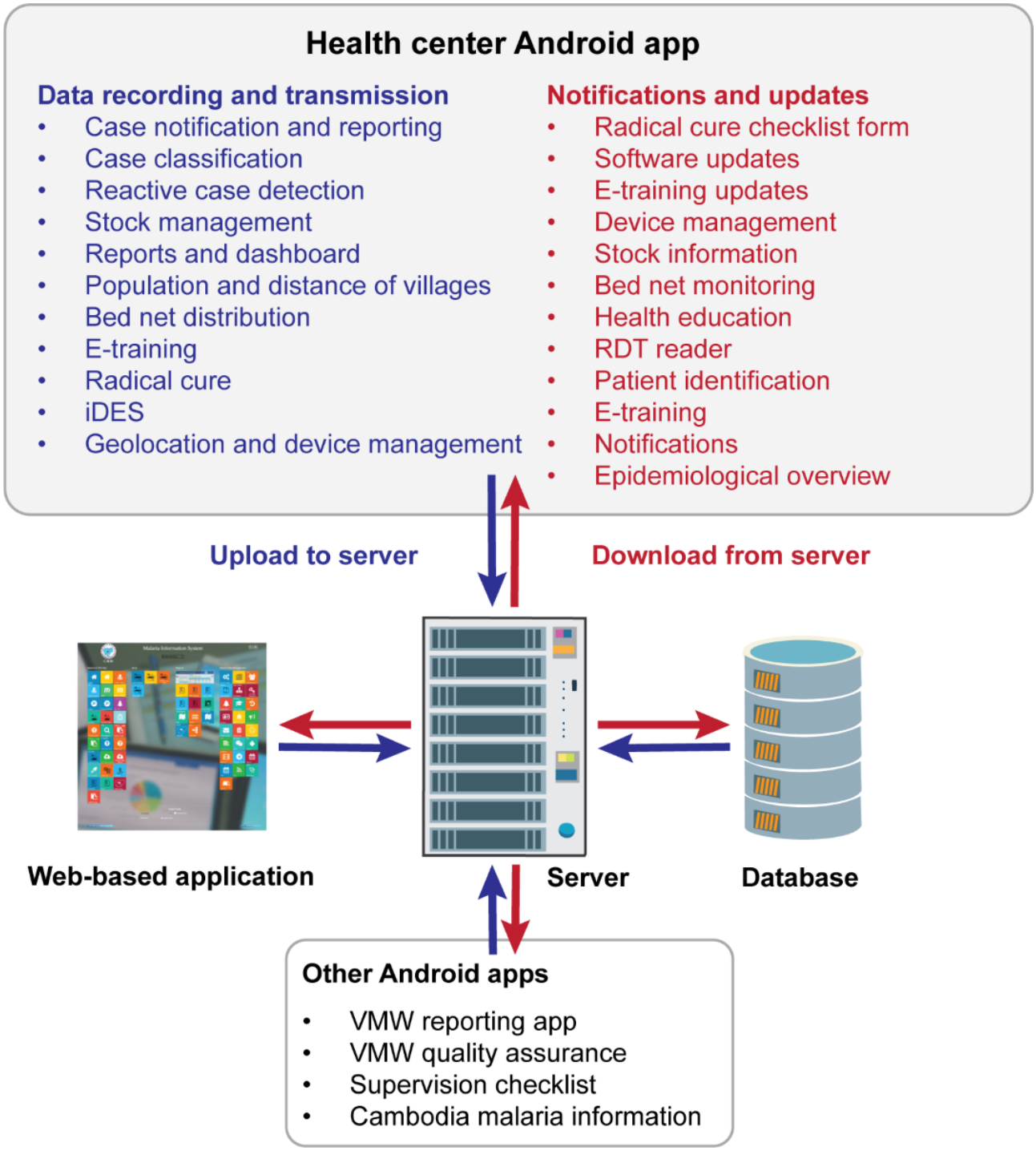
1. The 1-3-7 strategy: supporting malaria elimination in China and elsewhere, each case is reported within 1 day, investigated within 3 days, and followed by a targeted response within 7 days to prevent onward transmission [13-15].
2. Integrated drug efficacy surveillance (iDES): an early warning of drug resistance to ensure effective treatment; follow-up is extended to Day 42 for *P. falciparum* and Day 90 for *P. vivax* cases to identify treatment failures; blood spot samples are also collected for molecular analysis of known drug resistance markers [16-19].
3. Stratification of malaria risk: detailed epidemiological and ecological data at a high level of granularity are used for targeting to ensure effective deployment of essential resources.

Although these strategies can be implemented independently, their integration strengthens malaria surveillance and response systems by linking timely case detection, investigation, targeted intervention, and drug efficacy monitoring. This ensures good patient outcomes, reduces transmission potential, and maximizes the public health benefit. A coordinated approach also reduces duplication of effort, improves the efficiency of resource use, and enhances the sustainability of malaria elimination programs.

In Cambodia in 2008, malaria surveillance was severely limited and ineffective malaria treatments were commonly prescribed, deepening the public health impact and fueling transmission. Initial efforts to contain artemisinin resistance included banning artemisinin monotherapies in favor of more effective ACTs, phasing out the private sector, deploying insecticide-treated bed nets, facilitating information sharing across borders, and introducing malaria surveillance using a Microsoft Access Database. As the regional response pivoted to target malaria elimination, Cambodia implemented a National Strategic Plan for Malaria Elimination (2011 - 2025). This included substantial investments in surveillance infrastructure by the Cambodian Ministry of Health, led by the National Center for Parasitology, Entomology and Malaria Control (CNM) [20,21]. These efforts expanded surveillance activities via the Malaria Information System (MIS), which was migrated to a web-based platform in 2015 [22,23]. However, data were still collected on paper forms, with aggregation at the operational district level and monthly reporting. This process was slow, data were often incomplete, and outputs lacked the required granularity to support elimination efforts within an increasingly heterogeneous and fragmented transmission landscape [22-24]. To overcome these limitations, a fully digital system was launched in 2017, combining web-based analytics on the MIS, with 2 mobile apps enabling digital real-time reporting from health workers in the field, one for village malaria workers (VMWs) and another for staff at health centers (Figure 1) [24,25]. The MIS thereby supports the 1-3-7 strategy, iDES, and malaria risk stratification within one fully integrated platform extending throughout the public health system and the community.

XSL•FO
RenderX

Figure 1. Overview of health center app within the Cambodian malaria surveillance system. iDES: integrated drug efficacy surveillance; RDT: rapid diagnostic test; VMW: village malaria worker.



Cambodia's national malaria surveillance system has been widely recognized internationally as a model for digital transformation in elimination settings [26-31]. Efficient surveillance has supported an extraordinary decline in malaria cases from over 66,000 in 2018 to 355 in 2024, with no deaths from malaria recorded over this period [1,25,32]. Given that malaria elimination remains a priority for the World Health Organization and the Global Fund to Fight AIDS, Tuberculosis and Malaria, documentation of a fully implemented, nationally integrated digital surveillance platform has substantial global relevance. The broader surveillance architecture and its evolution have been described in recent publications [24,25]. This study extends that work by documenting the integration of the health center-level digital surveillance application into the country's malaria surveillance architecture. Operationalization of the system at the primary health care level is essential, as real-time case reporting, treatment verification, and response coordination directly influence elimination outcomes. This implementation case study focuses on a snapshot of recent data to highlight both the strengths and limitations of the health center app as Cambodia strives to complete the "last mile" to malaria elimination. The architecture and implementation principles are directly transferable to other malaria-endemic countries transitioning from control to

XSL•FO
RenderX

elimination, focusing on practical insights for national-level digital health surveillance planning and sustainability in preelimination settings.

Methods

Ethical Considerations

This work was conducted as part of CNM's ongoing efforts to strengthen routine surveillance, drawing on existing program data and surveys with VMWs about their professional use of a government-supported malaria reporting system. All respondents were public health personnel acting in their official roles, and participation in the surveys was voluntary and anonymous. No information that could identify individual patients or reveal personal health details was collected or assessed. Under national guidelines, the activities did not require formal ethics committee review or written informed consent. All procedures complied with relevant institutional, national, and international standards for data protection. No formal ethics committee approval was sought for this work because the project constituted a service implementation and evaluation activity using routinely collected, fully anonymized patient data, with no direct patient contact or intervention. Staff participants contributed only in their professional capacities as employees of the digital health service. Guidance from the University of Cambridge Research Ethics framework states that ethical review is not normally required for the secondary use of robustly anonymized data where individuals are not identifiable to the research team. Relevant guidance is available from the University of Cambridge Research Ethics Handbook [33] and associated ethics review guidance: University of Cambridge Ethics Review Flowchart [34]. Cambodian national health research ethics oversight information is available from the Cambodian National Ethics Committee for Health Research [35].

Malaria Surveillance in Cambodian Health Centers

Health centers serve as the primary formal health care contact for rural populations, offering essential diagnosis, treatment, and referral services for common illnesses. Each center covers a defined catchment area, typically including 8000 - 12,000 persons, and operates under Ministry of Health guidelines. For malaria, health center staff use rapid diagnostic tests (RDTs), prescribe ACTs, and record clinical and epidemiological details of each case. Health centers also provide support for VMWs, compile routine reports, and conduct follow-up investigations to direct local malaria responses.

Since 2017, health centers have conducted real-time case-based reporting via a dedicated digital application (Figure 1). Health center participation is mandatory in endemic districts, with malaria designated as a notifiable disease, supporting broad coverage and adherence to reporting requirements. Consequently, health center data form a core component of Cambodia's malaria surveillance infrastructure, contributing to individual case management, targeted local responses, and national-level monitoring and elimination planning.

App Development

The health center mobile app was developed in-house by CNM, with support from the Global Fund, to strengthen malaria elimination activities. Its primary aim is to enable timely, accurate, standardized malaria case reporting from public health centers, aligning with national strategy [20,21].

Built as a native Android app using Java, the app was designed for frontline health staff, considering feasibility, usability in low-connectivity settings, and compatibility with Ministry of Health workflows. Supporting Android versions 5.0+, this platform was selected over iOS and web-based alternatives due to its affordability, flexibility, offline functionality, and compatibility with the low-cost, older smartphones and tablets available within the public health system.

Recognizing the limited network access in rural areas, the app features offline-first functionality with local SQLite data storage, synchronizing automatically to the central MIS (Microsoft SQL Server) when internet becomes available (Figure 1). This integration supports near real-time data consolidation with other sources (eg, VMW app). The MIS also automatically shares data with the National Health Information System and the Malaria Reference Laboratory, with population information maintained through the Ministry of Planning Commune Database.

A user-centered design approach guided development. The interface uses icons to enhance usability and organizes functions into distinct modules (Figure 2). The system brings together a comprehensive suite of malaria surveillance and management functions, beginning with patient reporting, which captures essential epidemiological case data at the point of care. Additionally, the system supports detailed case classification, distinguishing local from imported cases and further categorizing indigenous, induced, recrudescent, and relapsing infections. It supports timely reactive case detection to actively identify additional infections around confirmed cases that might otherwise be missed and enables effective management of radical cure for *P. vivax* malaria, including tracking treatment eligibility and implementation across health facilities. Demographic insights are strengthened through population data that help determine accurate incidence rates, complemented by bed-net status features that record distribution dates, coverage, and expiry alerts. An iDES module is included to provide early warning signals of emerging drug resistance through extended follow-up of *P. falciparum* and *P. vivax* cases. Operational efficiency is supported by stock management tools that monitor commodity levels, expiry dates, and re-ordering needs, while an offline mode ensures continuous data capture in low-connectivity settings. The platform also offers robust reporting capabilities through automated and customizable analytics and facilitates oversight of VMWs by tracking their locations, training, reporting performance, and supervision. A centralized dashboard synthesizes key indicators to present a clear picture of the local malaria situation with precise village location mapping assigned using standardized GPS-based codes.

Prototypes underwent scenario-based testing at CNM with iterative refinement involving health staff, CNM surveillance officers, and district managers, ensuring usability and relevance. Updates continue based on input from subnational teams and evolving program needs. A dedicated 7-member CNM team manages maintenance, updates, training, and technical support.

XSL•FO
RenderX

A custom device management module enables remote software updates, monitors device use, and helps resolve hardware issues. Health centers can also report lost or damaged devices directly to CNM.

Figure 2. Health center app graphical user interface home screen. Top right: E-training – online resources and training materials and courses. (1) Patient reporting; (2) Integrated drug efficacy surveillance (iDES); (3) Reactive case detection; (4) Stock management; (5) Offline mode – management of data transfer offline and online; (6) Reporting; (7) Village malaria workers; (8) Population; (9) Bed-net status; (10) Dashboard; (11) Village location; (12) Case classification; (13) *P. vivax* radical cure. The figure was derived with permission from CNM's online malaria information system available via their website [36].

Data Processing

To facilitate adoption and reduce training needs, the app replicates the national paper-based malaria reporting forms used at health centers, aligning with national surveillance standards. Structured digital forms feature drop-down menus, prepopulated fields, and built-in logic to minimize errors and ensure consistency. This familiar design supports harmonized data collection, standardized case classification, and seamless analytics for trend monitoring and outbreak detection. Automated data quality checks verify completeness and internal consistency. In-app alerts notify users of missing fields or late submissions, while recent entries can be reviewed for prompt correction, reinforcing accountability. Entries are time-stamped and geotagged using built-in GPS functionality, enabling fine-scale spatial mapping, real-time risk stratification, and targeted response activities.

Once uploaded, case data feeds into the MIS, which applies validation checks and flags cases requiring follow-up. The MIS generates visual outputs, including time-series graphs, incidence maps, and summary tables, accessible to national and subnational teams via a secure online interface to track surveillance quality, treatment timelines, and emerging hotspots. To enhance performance in low-bandwidth areas, the MIS uses pregenerated dashboards and analytics that refresh hourly. Visuals are cached and compressed for fast access on both desktop and mobile devices, even with limited connectivity.

Data Security

Security and data protection were integrated from the design stage. Data transmitted between the app and MIS server is encrypted via HTTPS, with user access controlled through unique logins authenticated against CNM's central system. To safeguard patient confidentiality, personally identifiable information is minimized and retained only at the local level for follow-up care. Each case is assigned a unique identifier, and data access is role-based. CNM uses basic pseudonymization for national reporting, aligning with international standards for data minimization and secure storage.

Implementation

In 2017, CNM procured 800 tablets for distribution to health facilities across 21 provinces, completing rollout by early 2018. Since 2021, the app has been running on general-use health center devices, improving sustainability. As of June 2025, there are 977 devices active across 871 health centers, recognizing that most health centers now have more than one device (Figure 3).

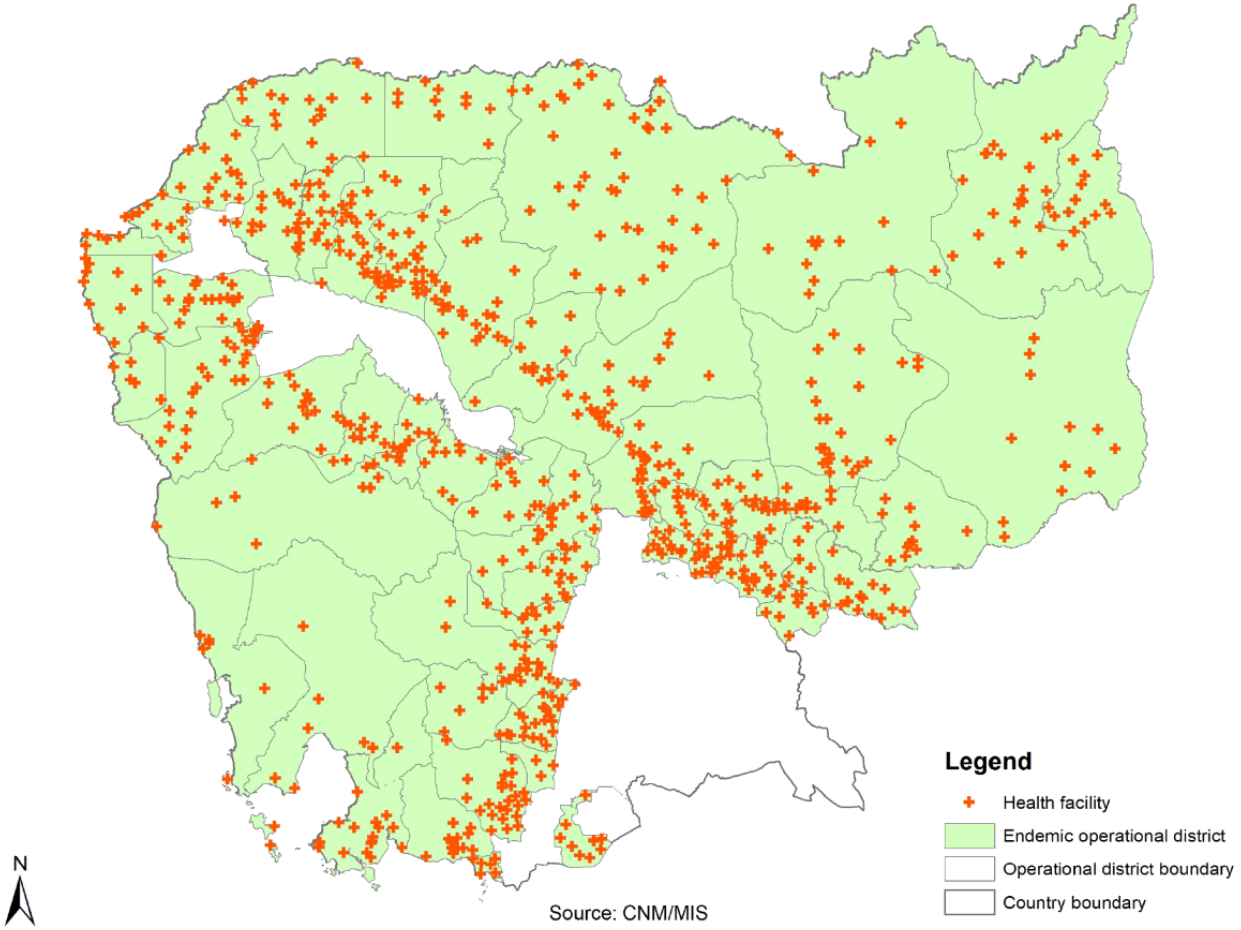
Initial deployment included in-person training, followed by field visits for reinforcement and troubleshooting. New staff are onboarded through a similar process, with training typically

XSL•FO
RenderX

requiring 2 hours. Annual refresher sessions are held to maintain user skills. CNM also provides ongoing support via phone or Telegram, with on-site visits for unresolved issues.

To enhance continuous learning, the app includes an embedded e-training module offering videos, slides, checklists, and quizzes. This enables staff to troubleshoot issues independently and assess their knowledge. CNM can remotely update training materials and send reminders to ensure compliance. For example, during the COVID-19 pandemic, the e-training module allowed uninterrupted training while supporting infection control. In particular, advice could be updated on how to continue to deliver malaria services effectively during the pandemic and differential diagnosis of fever. Additionally, an in-app messaging feature enables health workers to submit questions directly to CNM, which also helps identify broader knowledge gaps.

Figure 3. Distribution of health centers using the mobile app in 2024. The figure was derived with permission from CNM's online malaria information system available via their website [36].



Monitoring and Evaluation

A comprehensive evaluation and monitoring framework assesses the performance, usability, and programmatic use of the health center app. This framework combines continuous system-generated monitoring, routine surveillance validation, user feedback mechanisms, and periodic in-depth assessments to ensure the system remains aligned with malaria elimination requirements and responsive to user needs.

Technical performance is monitored continuously using automated metadata collected through the app and the MIS. Key performance indicators include app launch time, data upload and download success rates, synchronization times under varying network conditions, and crash frequency. A third-party monitoring service (Firebase Crashlytics) generates real-time alerts enabling rapid identification and rectification of technical issues. Timing logs capture request–response durations, supporting evaluation of system resilience under low-bandwidth conditions. These real-time indicators allow the CNM to assess system reliability, detect emerging technical constraints, and prioritize software improvements.

Routine monitoring includes tracking the completeness, accuracy, and timeliness of case-based reporting. Data submitted through the app are automatically transferred to the MIS, where automatic validation checks assess internal consistency, missing data, and unusual entries. Unusual or inconsistent patterns are flagged and followed up with field visits. Health center participation rates, frequency of app use, and completeness of mandatory fields are also monitored. Evaluation of case form content includes verification of demographic, diagnostic, and treatment variables. Malaria elimination surveillance performance is assessed through real-time longitudinal monitoring of the 1-3-7 framework, iDES, and foci clearance.

XSL•FO
RenderX

The app also supports risk stratification at the village level, a key programmatic tool.

Ongoing evaluation of the health center app's usability and contextual performance is essential. Real-time feedback is routinely collected, and user surveys are conducted periodically to check alignment between the app functionality, user experiences, and programmatic objectives. More in-depth surveys are conducted to obtain fine-grained information on potential app improvements and training needs.

Results

Data Collection for Case Management

Data submitted through the health center app are automatically transmitted to the MIS, where they undergo routine validation and analysis. These data underpin surveillance outputs guiding malaria elimination across all health system levels. In 2024, 99% (871/882) of endemic-area health centers used the app, logging an average of 4 visits per month. Data quality was high, with 99% (89/90) of submitted records complete and accurate.

The malaria case form captures all essential data, including demographics, diagnostic methods and results, treatment regimens, and epidemiological risk factors such as travel, forest exposure, or occupation to inform accurate case classification (Figure 4A). Each patient has a unique code, confirmed with fingerprint identification, allowing linked interactions for follow-up and retreatment (Figure 4A). Accurate time and location information allows cases to be tracked in real-time and supports the deployment of rapid responses to prevent or contain outbreaks.

XSL•FO
RenderX

Figure 4. Examples of reporting workflow on the health center app. A: Case reporting workflow on the health center app. Starting from the health center app main menu health center form (1), the menu for data entry is selected (2), the month is selected and patient details are entered, including demographic and clinical characteristics, diagnostic test results, treatment and referral details (3); data can be linked to previous records using the patient code, name, sex, age, or contact details (3.1) and patient details can be viewed (3.2) with any additional patient details completed (4), and further details on treatment, referrals, and pregnancy, hemoglobin, and glucose-6-phosphate dehydrogenase test results as well as drug doses recorded (5). B: Follow-up schedule for *P. vivax* radical cure on the health center app. Starting from the health center app home menu health center form (1), the data entry menu is accessed, including *P. vivax* radical cure follow-up (2). Year and month are used to search for patient information (3) and follow-up is scheduled depending on treatment regimen (3.1). At follow-up, the date and time are noted and the patient asked about possible signs of hemolysis (paleness, yellow skin and eyes, shortness of breath, etc) (4). Treatment follow-up by the village malaria worker is detailed in section (5), providing information on adherence, recovery, and adverse events. The figure was derived with permission from CNM's online malaria information system available via their website [36].

As malaria elimination has progressed in Cambodia, the epidemiological context has shifted from a high burden of *P. falciparum* malaria to most cases now caused by *P. vivax*. The treatment protocol for *P. vivax* is complex, requiring prior testing to determine glucose-6-phosphate dehydrogenase (G6PD) activity before treatment with an ACT for 3 days plus daily primaquine for 7 days for patients with normal or intermediate G6PD activity, or weekly primaquine for 8 weeks for patients with G6PD deficiency. Therefore, a specific workflow is incorporated to support efficient case management of *P. vivax* malaria cases, including follow-up and adherence (Figure 4B).

The app also provides individualized patient tracking, highlighting follow-up needs and adherence checks to support quality case management and programmatic objectives. Automatic reminders for follow-up interventions (phone calls

XSL•FO
RenderX

or in-person visits) are generated, for example, to track patients enrolled in iDES or to support *P. vivax* radical cure adherence.

Data Analytics for Malaria Elimination

To support immediate data use by frontline staff, the app includes an integrated dashboard showing key indicators by catchment area. Performance against the 1-3-7 framework is tracked with visualizations of cases disaggregated by *Plasmodium spp*., age, and sex using time-series graphs. Comparative summaries with peer facilities promote data use and performance monitoring. Automated flags identify cases for investigation. For example, delayed treatment suggesting access issues, retreatment possibly due to drug resistance, or cases in elimination zones triggering outbreak responses.

In addition to case data, the dashboard tracks preventive measures. Bed-net distribution status, including maps and expiry dates, supports redeployment planning. Also, procurement is guided by commodity stock level monitoring, with analytics identifying overstocked, adequate, or depleted sites. By delivering timely, accessible, and actionable insights, the dashboard transforms surveillance into a 2-way process, embedding data-driven decision-making into routine care and elimination activities.

System Performance

In terms of technical operation, the system is robust, with key technical performance indicators shown in Table 1. The app opens rapidly, typically within 3 seconds. This ensures that the app is available at the time of patient consultation and is quickly accessible given the workload and time pressure that health staff often face. Data uploads and downloads are resilient, even on older 2G and 3G networks. Success rates for uploads (95%) and downloads (98%) are high, providing reassurance that the design of the system has been optimized for a low-bandwidth setting.

Table . Key indicators of health enter app technical operation.

Indicators	Outcomes
Time between app data submission and updated dashboard analytics availability, median (IQR)[a]	3 (3 - 20)
Time taken for the app to launch and become usable after being opened, median (IQR)[a]	3 (3 - 20)
Data uploads and downloads that were successful without errors or retries (%)	
Upload success rate	95
Download success rate	98
Proportion of attempts that required retries due to network failure, timeouts, or other issues (retry rate per 1000 operations)	50
Interactions that occurred in offline mode and later synchronized successfully (%)	10
Number of crashes (per 1000 active user-days)	20 - 30
Time taken for synchronization segmented by network, median (IQR)[a]	
2G	13 (13 - 30)
3G	10 (7 - 30)
4G	3 (4 - 20)
Wi-Fi	3 (4 - 20)

[a]Measured in seconds.

Surveillance Performance

The MIS aggregates case data for all reported malaria cases in Cambodia. Malaria cases continued to decline in 2025 compared to previous years (Figure 5A), with no malaria deaths since 2018. Between January 1, 2025, and July 31, 2025, 69 cases were reported: 23 locally acquired, 7 domestically imported (from other districts), and 39 internationally imported (Figure 5B). Of these, 16 were reported via the health center app, 14 via the VMW app, and 39 internationally imported cases were added by MIS staff as these cases occurred in non-endemic areas.

Among the 69 cases, 95.7% (66/69) were notified and classified within 1 day. Reactive case detection was completed within 3 days for 21 of 21 eligible cases. Of 19 eligible foci, a foci response was initiated within 7 days for 16 (Figure 5C), reflecting strong adherence to the 1-3-7 framework.

National roll-out of iDES was completed in 2020. In 2024, 91.3% (173/186) of eligible cases were enrolled. Follow-up was completed to Day 28 for 99.4% (172/173) of cases and for *P. vivax* cases 92.9% (158/170) completed Day 90 follow-up. In 2024, dried blood spots were collected on Day 0 for 80.9% (140/173) of cases and on Day 28 for 77.5% (134/173) of cases permitting later analysis for molecular resistance markers. Declining case numbers reduce the sensitivity of clinical outcome data for detecting resistance, increasing the importance of molecular resistance markers [37].

Real-time foci identification classifies villages as active (≥1 indigenous case in the current year), residual (cases within the

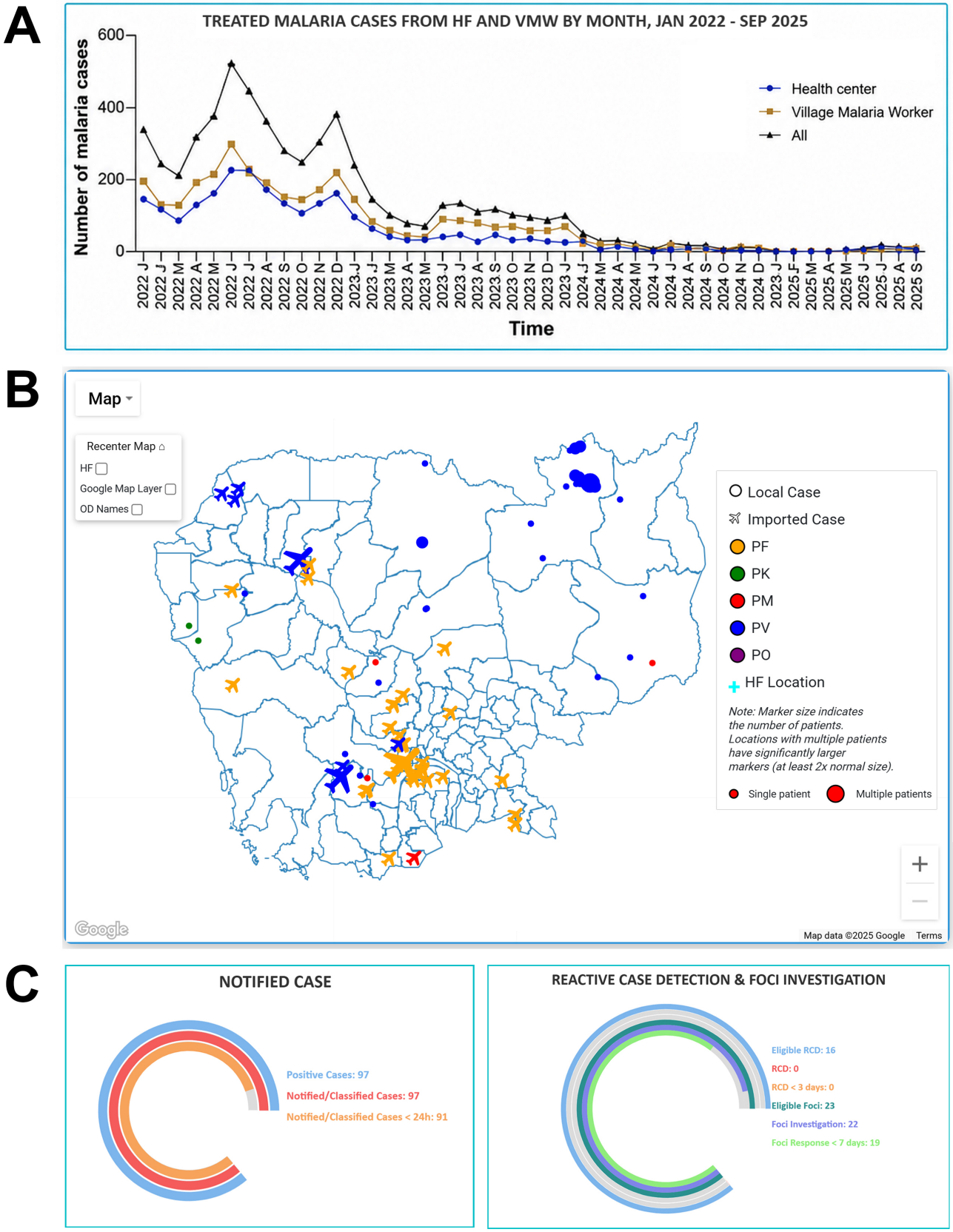
XSL•FO
RenderX

past 1 - 3 years), or cleared. As of August 2025, there were 27 active and 70 residual foci. Foci are prioritized for intensified elimination efforts, including increased testing, behavior change interventions, bed-net distribution, and forest pack allocation (eg, hammock nets, repellents, and educational materials) [32].

Risk stratification ensures efficient resource allocation. In 2025, of 9.8 million residents in endemic areas, 22,119 were high risk, 147,602 medium risk, 2.8 million low risk, and 6.8 million at no risk. Stratification relies on the annual blood examination rate (ABER) and annual parasite index (API), both requiring comprehensive reporting and accurate population data. These metrics are automatically generated on the dashboard to the village level (Figure 6).

XSL•FO
RenderX

Figure 5. Snapshot of indicators of surveillance performance from the Malaria Information System dashboard. (**A**) Number of malaria cases and reporting route (January 2022 to September 2025). (**B**) Status of malaria cases in Cambodia (January to September 2025). (**C**) Performance against 1-3-7 targets (January to September 2025). These outputs are automatically generated from data submitted via the health center and VMW apps and included on the dashboard. The figure was derived with permission from CNM's online malaria information system available via their website [36].



XSL•FO
RenderX

Figure 6. Core surveillance indicators for malaria stratification from the Malaria Information System dashboard. (**A**) Annual blood examination rate (ABER) (January to September 2025), defined as the number of blood samples tested, divided by the total population, then multiplying by 100. (**B**) Annual parasite index (API) (January to September 2025), defined as the number of confirmed malaria cases per 1000 people within a specific geographic area and time period. These outputs are automatically generated from data submitted via the health center and VMW apps and included on the dashboard. The figure was derived with permission from CNM's online malaria information system available via their website [36].

XSL•FO
RenderX

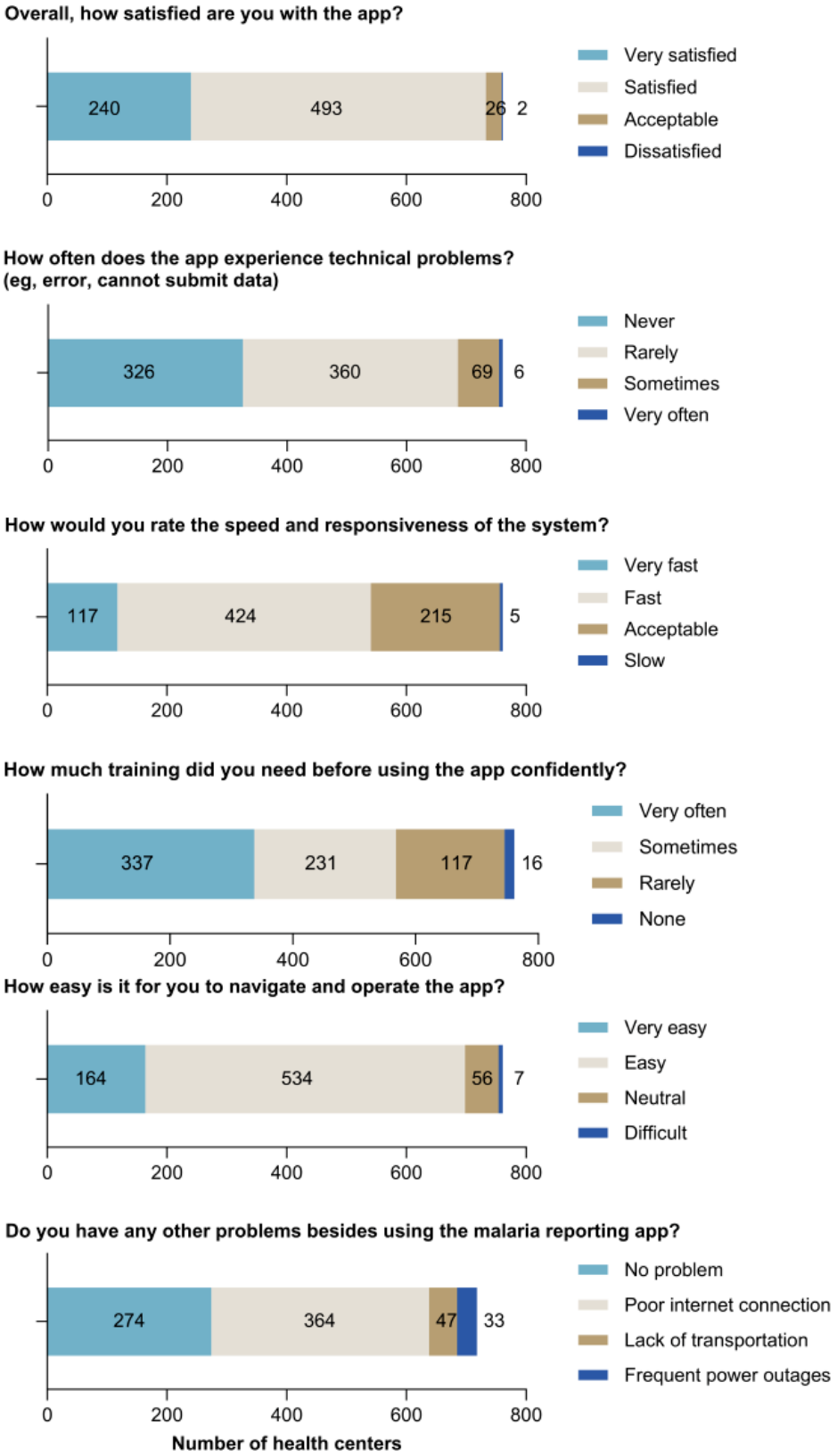
For ABER (Figure 6A), the proportion of the population tested annually indicates surveillance coverage. While ≥10% is recommended in control settings, elimination contexts require more targeted approaches, such as reactive testing and screening high-risk groups. Therefore, high ABER is expected in high-risk areas and lower in areas where risk is minimal. For API (Figure 6B), confirmed cases per 1000 people; measures transmission intensity and guides intervention focus. In elimination settings with low incidence, uneven surveillance can distort API, so village-level granularity is critical. Cambodia's low API supports readiness for national elimination.

User Feedback

A usability and satisfaction survey conducted in July 2025 of all health centers received responses from 86.3% (761/882) of facilities within 1 month (Figure 7). Overall, 96.3% (733/761) of health centers were satisfied or very satisfied with the app, 90.1% (686/761) found that technical issues were rare or absent, and 71.1% (541/761) thought that system speed was fast or very fast. Although 44.3% (337/761) of centers confirmed that they needed regular training on the app, 91.7% (698/761) found navigating the app easy or very easy. Other common issues that impacted surveillance activities were poor internet connection, a lack of transportation to reach remote locations, and frequent power outages (Figure 7). This highlights the infrastructure constraints that had to be considered during the MIS system design.

XSL•FO
RenderX

Figure 7. Health center user feedback on app technical performance and usability. The figure was derived with permission from CNM's online malaria information system available via their website [36].



The most recent in-depth survey was conducted in 9 randomly selected health centers across Battambang, Kampong Speu, and Mondul Kiri provinces (August-November 2024). Responses were obtained from the designated malaria focal person at each health center, all of whom had received standardized training on the use of the health center app and surveillance.

All 9 centers were aware of malaria reporting mandates, the requirement for case-based data submission, and national surveillance guidelines, demonstrating strong procedural awareness. All used both the health center app and CNM-issued paper forms (not processed, retained for audit purposes). All health centers confirmed immediate case reporting and

XSL•FO
RenderX

compliance with WHO recommendations, including reporting zero cases, demonstrating high uptake and process adherence. Their understanding of surveillance objectives showed alignment with programmatic goals (Textbox 1) [20,21].

Textbox 1. Summary of health center perceptions of the purposes of surveillance.

- The system is designed to monitor malaria trends and track the number of tests conducted, regardless of whether they are few or many.
- It serves as a reminder and alert to keep health workers aware of the current malaria situation, helping them understand fluctuations over time.
- By tracking malaria cases, it enables identification of areas at risk and provides real-time data, making it easier to generate timely reports.
- Reports can be accurately managed at the health center level and efficiently shared with higher levels of the health system.
- The data are valuable for understanding where malaria cases are concentrated, whether in villages with many or few cases, and supports planning for the following month or year.

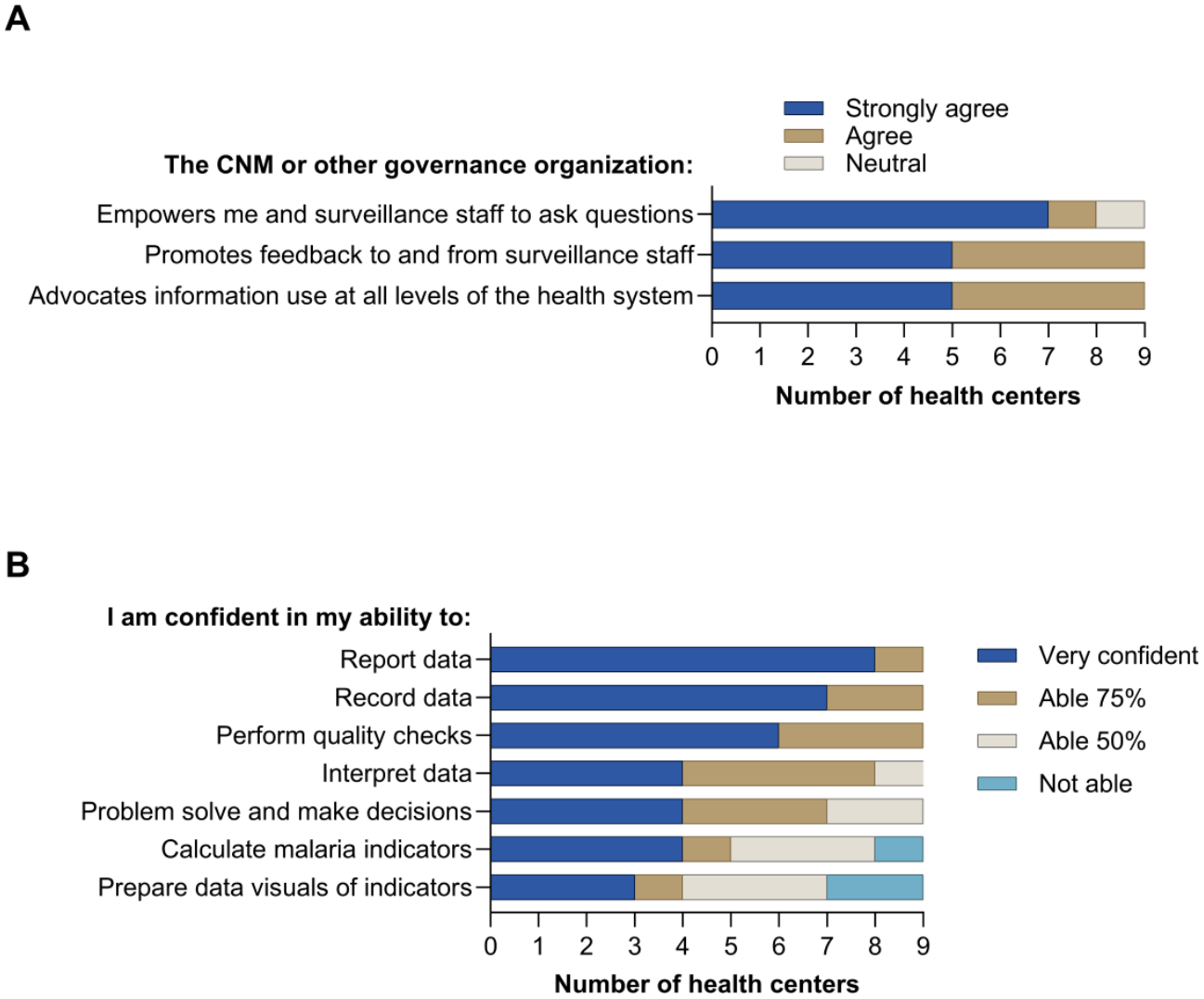
Self-reported comprehension of surveillance data was high: 100% at 3 centers, 80% at 6 centers. All found job-aids and SOPs clear, although one center noted complexity requiring further training. All centers knew where to seek support, indicating adequate system support. Eight of 9 centers conducted monthly data quality checks, all had supervision records; recent supervisory visits included data quality assessments, corrective actions, and written feedback, with standardized checklists used in 8/9 cases.

Staffing was adequate, with 6 centers reporting the correct number of staff (between 1 and 5 persons), 2 centers needed an additional staff member each, and 1 center had 1 more member of staff versus requirements. All centers had patient record books and tablets, with all tablets operational over the previous 30 days. This probably reflects the app's device management module which supports a high level of functionality by quickly identifying technical issues. Three centers had computers and 5 had printers, with 1 computer and 2 printers fully functional for the last 30 days. Printers and computers are maintained locally, often with limited technical support. Infrastructure challenges persist: 4 centers experienced power issues, 3 had internet disruptions, and 2 had temporary system failures in the past 4 weeks. However, all centers reported access to troubleshooting support. Stock-outs pose a risk to elimination efforts and one Mondul Kiri center experienced a stock-out of RDTs and ACTs, highlighting the need for responsive surveillance systems in the preelimination phase.

All centers accessed dashboard visualizations and 4 downloaded data for further analysis. All held data review meetings, with bulletins produced in 6 centers. However, only 5/9 could calculate key indicators, with gaps in training and data availability reported in several centers. Despite this, 8 centers accessed data daily or weekly, and all at least monthly, and feedback on CNM and partner support was positive (Figure 8A). Consistent with the annual training cascade, all health centers received surveillance training in the previous 12 months on case notification, supervision, case data review, case report compilation, and recording of patient information. Six centers also received training on report preparation and 8 on data analysis and use. While confidence was high for data entry and quality checks, staff were less confident in calculating indicators and preparing visual outputs (Figure 8B).

XSL•FO
RenderX

Figure 8. Health center user feedback on surveillance activities supported by the health center app. (**A**) Feedback on the supportive supervision of CNM (the National Center for Parasitology, Entomology and Malaria Control) regarding surveillance. (**B**) Confidence of users in their ability to conduct core surveillance activities, analyze data and interpret and use data. The figure was derived with permission from CNM's online malaria information system available via their website [36].



Discussion

Study Findings and Generalizability

The Cambodian health center app constitutes a pivotal component of the national malaria surveillance system, enabling real-time, case-based data capture from endemic districts. By facilitating structured digital reporting at the point of care, the system improves the timeliness and completeness of malaria case notification, supports treatment adherence, and permits granular analysis of clinical and epidemiological data, enhancing case classification accuracy and enabling faster public health responses.

Effective surveillance prevents delayed treatment and onward transmission. The key strength of case-based data are alignment with the 1-3-7 surveillance framework [13,14]. This strategy contributed to malaria elimination in China, and similar approaches have been adopted in Cambodia, Thailand, Lao PDR, Myanmar, Viet Nam, India, Tanzania, and Zambia [13,14,38]. In Cambodia, digital surveillance at the point-of-care enables timely case management and follow-up, meeting 1-3-7 targets while integrating drug efficacy surveillance to ensure that treatments remain effective. At the same time, the Cambodian MIS fulfills WHO requirements for malaria elimination certification, ie, the surveillance system must be able to detect, investigate, and classify cases as locally acquired or imported [39]. Thus, no additional systems are required, avoiding duplication of effort in preparing for malaria certification.

While grounded in a specific country context, the challenges addressed by the health facility app are common across malaria-eliminating countries globally; malaria becomes increasingly spatially and temporally heterogeneous, focused in hard-to-reach communities. In Cambodia, geolocated case data have revealed persistent transmission in border and forested regions despite low overall case numbers, allowing tailored vector control and community engagement strategies [32,40]. Without sufficiently sensitive and sustainable surveillance, hard-won gains can be quickly lost following the withdrawal of control measures, as outbreaks in at-risk communities spread unchecked [41]. For example, in Sri Lanka in the 1960s, the Huai River basin in China in the 1960s and 1970s, and more

XSL•FO
RenderX

recently in Thailand in 2017 [42-44]. Sub-national tailoring of interventions based on risk stratification informed by real-time surveillance allows targeting of residual foci and prompt identification and control of outbreaks, preventing resurgence [45].

As Cambodia moves toward a postelimination phase, emphasis will shift from local transmission detection to identifying imported cases, especially across porous international borders. Vigilance must be maintained and malaria should still be perceived as a threat both in terms of the risk of resurgence and for individual patients who may be at risk of poor outcomes if presentation, diagnosis and treatment are delayed [43,46]. The health facility app is integrated with complementary tools such as the VMW app, allowing detection in both formal health facilities and community settings. This digitally integrated approach allows the surveillance system to transition seamlessly from elimination to postelimination activities.

Sustainability is essential. Global funding is contracting, and as malaria cases decline, there is a risk of de-prioritization within national health budgets. The sustainable design principles applied to the Cambodian MIS and mobile reporting apps have broad applicability: national ownership, low recurrent costs, offline-first functionality, integration with routine workflows, efficient training requirements, and the ability to adapt rapidly to evolving epidemiological and programmatic needs while avoiding duplication. By contrast, a mobile case-based surveillance system in Myanmar struggled with long-term sustainability due to heavy reliance on external support, duplicated systems, and dependence on internet access [47-49].

In a broader context, the Cambodian MIS is an example of a successful digital health system implemented in a low-resource setting. Digital health tools show potential to widen care access and improve service quality, but in many low-resource countries progress has stalled at brief pilot stages with limited scale-up and unclear long-term viability [50]. A systematic review of digital health interventions in low-and middle-income countries identified barriers relating to infrastructure, equipment, internet access, electricity, and the nature of the intervention [51]. Facilitators were associated with stakeholder commitment and participation [51]. The experience of Cambodia illustrates that meaningful local involvement in design, deployment, and ongoing development is fundamental to context-specific problem solving and ensuring that systems remain flexible and robust in the face of evolving operational demands [52].

Limitations

Despite the system's strengths, challenges persist. While reporting and data entry have been widely adopted, gaps remain in health center staff's capacity to analyze, interpret, and act on data. Although the dashboard offers summary indicators, users often lack the skills to conduct more complex analyses, even when data are exportable to Excel. Limited digital literacy and a lack of tailored analytic tools at the local level restrict the full use of surveillance data for decision-making.

Connectivity remains another constraint, particularly in remote areas. Although the system supports offline data entry, delayed synchronization still occurs, potentially affecting real-time response and 1-3-7 targets. Despite the expansion of internet access in Cambodia, this issue is clearly still impacting day-to-day surveillance, with most health centers reporting poor internet connectivity. Similar challenges are a common issue for digital health interventions in low-resource settings [51].

Future Directions

Expanding the analytical capacity of the app to address evolving user needs represents a key area for development. As malaria transmission becomes increasingly isolated, analytics must become more specialized. Centralized generation of customized reports or automated analytics tailored to specific settings could enhance data use at health centers. Additionally, broadening the training curriculum to emphasize data interpretation and visualization skills would strengthen local decision-making.

Future system upgrades could include integration with other data sources, such as population mobility, environmental data, and socioeconomic indicators. Linking health and geospatial data with climate datasets may allow more accurate prediction of transmission risk and better outbreak preparedness [53].

Post-elimination, the ability to rapidly stratify malaria risk in terms of environmental receptivity, historical transmission, and importation rate is required for the prevention of reestablishment of malaria transmission. Incorporating machine learning and predictive analytics could enhance early warning capabilities, automate anomaly detection, and support more targeted public health interventions [54]. Given the existing integration of the MIS with the National Malaria Laboratory, there is also potential for further molecular surveillance to discriminate between local and imported infections [55].

Finally, expanding the platform to support surveillance for other diseases or syndromic conditions typically encountered at health centers may increase long-term sustainability, reduce duplication within the health system, and build broader public health resilience in Cambodia.

Conclusions

Without urgent commitments from the GMS countries, resistant *P. falciparum* could have spread globally, leading to a resurgence in the disease. The considerable investment in surveillance and its contribution toward malaria elimination in the GMS requires greater examination and recognition. Also, the experience of Cambodia and other GMS countries provides valuable learnings and templates for designing flexible surveillance systems which are applicable in other malaria endemic settings transitioning from burden reduction to targeting malaria elimination. This is particularly relevant given the emergence of artemisinin partial resistance in Africa [1].

This report showed that strengthening the functionality and reach of Cambodia's health center malaria surveillance app enhanced the efficiency and responsiveness of the national malaria elimination program. The app's integration of case-based reporting, interactive dashboards, and training modules supports data timeliness, completeness, and use, transforming surveillance from a passive reporting task into an active, data-driven tool for decision-making.

XSL•FO
RenderX

The system's national coverage and strong uptake among frontline health workers aim to capture every malaria case. Its open-source, in-house development model ensures adaptability to evolving operational needs, including new features for treatment adherence monitoring and foci response. Real-time data collection has facilitated high adherence to the 1-3-7 framework and informed targeted responses in residual transmission zones.

To ensure continued success and sustainability, future efforts should focus on strengthening analytic capacity among health center staff, addressing infrastructure limitations such as connectivity, developing predictive capabilities, and further integrating the app with other health and environmental information systems. Long-term sustainability will depend on continued political commitment, integration with surveillance for other health priorities, and active engagement of internal and external stakeholders. As Cambodia moves into the postelimination phase, a robust digital surveillance infrastructure is essential for protecting the population from malaria resurgence.

Acknowledgments

Technical support was provided by the Mahidol Oxford Tropical Medicine Research Unit. Naomi Richardson (Magenta Communications Ltd) provided writing, editorial, and graphical services.

Funding

Funding was provided by the Global Fund to Fight AIDS, Tuberculosis and Malaria, the World Health Organization, the U.S. President's Malaria Initiative, the United Nations Office for Project Services and the Bill and Melinda Gates Foundation (OPP1193472_2018). This publication was funded in whole, or in part, by the Wellcome Trust [220211/Z/20/Z]. For the purpose of Open Access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission.

Data Availability

All relevant data are available in the manuscript.

Authors' Contributions

PN wrote the initial report for further paper development. PN, SS, ET, RY, SK, YR, OV, HM, and CP contributed data and analysis. PR contributed to data interpretation. PN, SS, RJM, and LJW contributed to the concept of the paper. All authors critically reviewed the manuscript, approved the final version and take responsibility for its publication.

Conflicts of Interest

PR is a staff member of the World Health Organization. The author alone is responsible for the views expressed in this publication, which do not necessarily represent the decisions, policies, or views of World Health Organization. All other authors declare no competing interests.

References

1. World malaria report 2025. World Health Organization. 2025. URL: https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2025 [accessed 2026-05-06]
2. Harinasuta T, Suntharasamai P, Viravan C. Chloroquine-resistant falciparum malaria in Thailand. Lancet 1965 Oct 2;2(7414):657-660. [doi: 10.1016/s0140-6736(65)90395-8] [Medline: 4158213]
3. Trape JF. The public health impact of chloroquine resistance in Africa. Am J Trop Med Hyg 2001;64(1-2 Suppl):12-17. [doi: 10.4269/ajtmh.2001.64.12] [Medline: 11425173]
4. Wellems TE, Plowe CV. Chloroquine-resistant malaria. J Infect Dis 2001 Sep 15;184(6):770-776. [doi: 10.1086/322858] [Medline: 11517439]
5. Vinayak S, Alam MT, Mixson-Hayden T, et al. Origin and evolution of sulfadoxine resistant Plasmodium falciparum. PLoS Pathog 2010 Mar 26;6(3):e1000830. [doi: 10.1371/journal.ppat.1000830] [Medline: 20360965]
6. Wongsrichanalai C, Meshnick SR. Declining artesunate-mefloquine efficacy against falciparum malaria on the Cambodia-Thailand border. Emerg Infect Dis 2008 May;14(5):716-719. [doi: 10.3201/eid1405.071601] [Medline: 18439351]
7. Noedl H, Se Y, Schaecher K, et al. Evidence of artemisinin-resistant malaria in western Cambodia. N Engl J Med 2008 Dec 11;359(24):2619-2620. [doi: 10.1056/NEJMc0805011] [Medline: 19064625]
8. Imwong M, Suwannasin K, Kunasol C, et al. The spread of artemisinin-resistant Plasmodium falciparum in the Greater Mekong Subregion: a molecular epidemiology observational study. Lancet Infect Dis 2017 May;17(5):491-497. [doi: 10.1016/S1473-3099(17)30048-8] [Medline: 28161569]
9. Canavati SE, Lawford HLS, Fatunmbi BS, et al. Establishing research priorities for malaria elimination in the context of the emergency response to artemisinin resistance framework-the Cambodian approach. Malar J 2016 Feb 25;15:120. [doi: 10.1186/s12936-016-1117-9] [Medline: 26916933]

XSL•FO
RenderX

10. Emergency response to artemisinin resistance in the Greater Mekong Subregion: regional framework for action 2013-2015. World Health Organization. 2013. URL: https://apps.who.int/iris/handle/10665/79940 [accessed 2026-05-06]
11. Strategy for malaria elimination in the Greater Mekong Subregion: 2015-2030. World Health Organization. 2015. URL: https://apps.who.int/iris/handle/10665/208203 [accessed 2026-05-06]
12. A framework for malaria elimination. World Health Organization. 2017. URL: https://www.who.int/publications/i/item/9789241511988 [accessed 2026-05-06]
13. Sogandji N, Stevenson A, Luo MY, Qi G, Maude RJ. Systematic review of evidence for the impact and effectiveness of the 1-3-7 strategy for malaria elimination. Malar J 2024 Dec 18;23(1):371. [doi: 10.1186/s12936-024-05200-w] [Medline: 39695680]
14. Yi B, Zhang L, Yin J, Zhou S, Xia Z. 1-3-7 surveillance and response approach in malaria elimination: China's practice and global adaptions. Malar J 2023 May 9;22(1):152. [doi: 10.1186/s12936-023-04580-9] [Medline: 37161379]
15. Kheang ST, Sovannaroth S, Barat LM, et al. Malaria elimination using the 1-3-7 approach: lessons from Sampov Loun, Cambodia. BMC Public Health 2020 Apr 22;20(1):544. [doi: 10.1186/s12889-020-08634-4] [Medline: 32321475]
16. Sudathip P, Saejeng A, Khantikul N, et al. Progress and challenges of integrated drug efficacy surveillance for uncomplicated malaria in Thailand. Malar J 2021 Jun 9;20(1):261. [doi: 10.1186/s12936-021-03791-2] [Medline: 34107955]
17. Liu H, Xu JW, Deng DW, et al. Dihydroartemisinin-piperaquine efficacy in Plasmodium falciparum treatment and prevalence of drug-resistant molecular markers along China-Myanmar border in 2014-2023. J Glob Antimicrob Resist 2023 Dec;35:271-278. [doi: 10.1016/j.jgar.2023.10.001] [Medline: 37816434]
18. Somé AF, Somé A, Sougué E, et al. Safety and efficacy of repeat ivermectin mass drug administrations for malaria control (RIMDAMAL II): a phase 3, double-blind, placebo-controlled, cluster-randomised, parallel-group trial. Lancet Infect Dis 2025 Jul;25(7):737-750. [doi: 10.1016/S1473-3099(24)00751-5] [Medline: 39919778]
19. Wasakul V, Verschuuren TD, Thuy-Nhien N, et al. Genetic surveillance of Plasmodium falciparum populations following treatment policy revisions in the Greater Mekong Subregion. Nat Commun 2025 May 20;16(1):4689. [doi: 10.1038/s41467-025-59946-1] [Medline: 40394107]
20. The national strategic plan for elimination of malaria in the Kingdom of Cambodia 2011-2025. Kingdom of Cambodia Nation Religion King. 2011. URL: https://www.cnm.gov.kh/userfiles/file/N%20Strategy%20Plan/National%20Strtegyin%20English.pdf [accessed 2026-05-07]
21. Cambodia malaria elimination action framework 2016-2020. Kingdom of Cambodia Ministry of Health.: Royal Government of Cambodia; 2021 URL: https://mesamalaria.org/wp-content/uploads/2016/03/Cambodia-Malaria-Elimination-Action-Framework-.pdf [accessed 2026-05-18]
22. Fu C, Lopes S, Mellor S, Aryal S, Sovannaroth S, Roca-Feltrer A. Experiences from developing and upgrading a web-based surveillance system for malaria elimination in Cambodia. JMIR Public Health Surveill 2017 Jun 14;3(2):e30. [doi: 10.2196/publichealth.6942] [Medline: 28615155]
23. Mellor S. Moving towards malaria elimination: developing innovative tools for malaria surveillance in Cambodia. : Malaria Consortium; 2013 URL: https://www.malariaconsortium.org/media-library/media-files/201312100515-lp2013-asia-malaria-surveillance.pdf [accessed 2026-05-18]
24. Ngor P, Rekol H, Sovannaroth S, et al. A bespoke, real-time surveillance system for malaria elimination in Cambodia: architecture, applications and impact. Malar J 2026;25(1). [doi: 10.1186/s12936-025-05772-1]
25. Sovannaroth S, Kugler AM, Ngor P, et al. Key findings and recommendations from assessment of Cambodia's national malaria surveillance system. PLOS Glob Public Health 2025;5(12):e0005165. [doi: 10.1371/journal.pgph.0005165] [Medline: 41348915]
26. Expanding Cambodia's malaria information system to advance elimination goals. PMI Evolve. URL: https://pmievolve.org/expanding-cambodias-malaria-information-system-advance-elimination-goals [accessed 2026-05-07]
27. Cambodia's final push: innovation, community, and policy driving malaria elimination. APMEN Asia Pacific Malaria Elimination Network.: APMEN URL: https://apmen.org/blog/cambodias-final-push-innovation-community-and-policy-driving-malaria-elimination [accessed 2026-05-07]
28. Moving towards malaria elimination: developing innovative tools for malaria surveillance in Cambodia. Malaria Knowledge Hub.: MESA; 2013. URL: https://mesamalaria.org/resource-hub/moving-towards-malaria-elimination-developing-innovative-tools-malaria-surveillance [accessed 2026-05-07]
29. Community engagement is key to malaria elimination in Cambodia. World Health Organization.: WHO; 2023. URL: https://www.who.int/westernpacific/newsroom/feature-stories/item/community-engagement-is-key-to-malaria-elimination-in-cambodia [accessed 2026-05-07]
30. Photo story: sharing lessons from Cambodia on the malaria elimination journey. Malaria Consortium. 2024. URL: https://www.malariaconsortium.org/photo-story-sharing-lessons-from-cambodia-on-the-malaria-elimination-journey [accessed 2026-05-07]
31. Workshop on malaria surveillance and malaria elimination in the Greater Mekong Subregion, Siem Reap, Cambodia, 15-17 November 2023. World Health Organization. URL: https://www.who.int/publications/i/item/RS-2023-GE-46 [accessed 2026-05-07]

XSL•FO
RenderX

32. Sovannaroth S, Ngor P, Khy V, et al. Accelerating malaria elimination in Cambodia: an intensified approach for targeting at-risk populations. Malar J 2022 Jul 2;21(1):209. [doi: 10.1186/s12936-022-04234-2] [Medline: 35780142]
33. Research ethics handbook for research ethics committee. : University of Cambridge; 2025 URL: https://www.research-services.admin.cam.ac.uk/files/urec-research-ethics-handbook.pdf?utm_source=chatgpt.com [accessed 2026-05-18]
34. Is independent ethical review required? (flowchart). University of Cambridge. URL: https://www.research-services.admin.cam.ac.uk/manage-risk/research-ethics/independent-ethical-review-required?utm_source=chatgpt.com [accessed 2026-05-18]
35. NECHR official website (homepage). National Election Committee for Human Rights (NECHR). URL: https://www.demo.nechr.org.kh/?utm_source=chatgpt.com [accessed 2026-05-18]
36. Malaria dashboard. URL: https://mis.cnm.gov.kh/Dashboard [accessed 2026-06-10]
37. Martin RE, Shafik SH, Richards SN. Mechanisms of resistance to the partner drugs of artemisinin in the malaria parasite. Curr Opin Pharmacol 2018 Oct;42(71-80):71-80. [doi: 10.1016/j.coph.2018.07.010] [Medline: 30142480]
38. Martin AC, Matoba J, Sing'anga C, et al. Implementation outcomes of 1-3-7 focus investigation for malaria in a low transmission setting in Southern Province, Zambia: a mixed methods study. PLOS Glob Public Health 2025;5(1):e0004179. [doi: 10.1371/journal.pgph.0004179] [Medline: 39823475]
39. Preparing for certification of malaria elimination, second edition. World Health Organization.: WHO; 2022. URL: https://www.who.int/publications/i/item/9789240062009 [accessed 2026-05-07]
40. Stratil AS, Vernaeve L, Lopes S, et al. Eliminating Plasmodium falciparum malaria: results from tailoring active case detection approaches to remote populations in forested border areas in north-eastern Cambodia. Malar J 2021 Feb 22;20(1):108. [doi: 10.1186/s12936-021-03622-4] [Medline: 33618718]
41. Cohen JM, Smith DL, Cotter C, et al. Malaria resurgence: a systematic review and assessment of its causes. Malar J 2012 Apr 24;11(122):122. [doi: 10.1186/1475-2875-11-122] [Medline: 22531245]
42. Roh ME, Lausatianragit K, Chaitaveep N, et al. Civilian-military malaria outbreak response in Thailand: an example of multi-stakeholder engagement for malaria elimination. Malar J 2021 Dec 7;20(1):458. [doi: 10.1186/s12936-021-03995-6] [Medline: 34876133]
43. Wickremasinghe AR. Elimination of malaria from Sri Lanka and beyond; lessons for other countries in elimination phase. Ceylon Med J 2023 Aug 24;68(S1):39-45. [Medline: 37610920]
44. Liu X, Guo P, Liang Y, et al. Lessons from failure to success on malaria elimination in the Huai River Basin in China. BMJ 2025 Apr 22;389:e080658. [doi: 10.1136/bmj-2024-080658] [Medline: 40262830]
45. Reithinger R. The road to malaria elimination: why a robust surveillance system is imperative. Lancet Glob Health 2025 Sep;13(9):e1498-e1499. [doi: 10.1016/S2214-109X(25)00282-7] [Medline: 40845872]
46. Seneviratne S, Fernando D, Chulasiri P, et al. A malaria death due to an imported Plasmodium falciparum infection in Sri Lanka during the prevention of re-establishment phase of malaria. Malar J 2023 Aug 24;22(1):243. [doi: 10.1186/s12936-023-04681-5] [Medline: 37620890]
47. Win Han O, Kaung Myat T, Cutts JC, et al. Sustainability of a mobile phone application-based data reporting system in Myanmar's malaria elimination program: a qualitative study. BMC Med Inform Decis Mak 2021 Oct 18;21(1):285. [doi: 10.1186/s12911-021-01646-z] [Medline: 34663306]
48. Win H, Nay Y, Kyawt Mon W, et al. Reactive surveillance and response strategies for malaria elimination in Myanmar: a literature review. Malar J 2023;22(1). [doi: 10.1186/s12936-023-04567-6] [Medline: 37814267]
49. Win Han O, Win H, Cutts JC, et al. A mobile phone application for malaria case-based reporting to advance malaria surveillance in Myanmar: a mixed methods evaluation. Malar J 2021 Mar 26;20(1):167. [doi: 10.1186/s12936-021-03701-6] [Medline: 33771144]
50. Ebenso B, Namisango E, Abejirinde IO, Allsop MJ. Editorial: the scale-up and sustainability of digital health interventions in low- and middle-income settings. Front Digit Health 2025;7:1634223. [doi: 10.3389/fdgth.2025.1634223] [Medline: 41058973]
51. Kaboré SS, Ngangue P, Soubeiga D, et al. Barriers and facilitators for the sustainability of digital health interventions in low and middle-income countries: a systematic review. Front Digit Health 2022;4:1014375. [doi: 10.3389/fdgth.2022.1014375] [Medline: 36518563]
52. Sylla B, Ismaila O, Diallo G. 25 years of digital health toward universal health coverage in low- and middle-income countries: rapid systematic review. J Med Internet Res 2025 May 29;27:e59042. [doi: 10.2196/59042] [Medline: 40440696]
53. Keita IM, Diouf M, Ndiop M, et al. Integration of sentinel surveillance and climate factors to accelerate malaria elimination in a changing climate of Senegal. Sci One Health 2025;4:100112. [doi: 10.1016/j.soh.2025.100112] [Medline: 40510868]
54. Yang S, Li RY, Yan SN, et al. Risk assessment of imported malaria in China: a machine learning perspective. BMC Public Health 2024 Mar 20;24(1):865. [doi: 10.1186/s12889-024-17929-9] [Medline: 38509529]
55. Safarpour M, Cabrera-Sosa L, Gamboa D, Van Geertruyden JP, Delgado-Ratto C. Detecting imported malaria infections in endemic settings using molecular surveillance: current state and challenges. Front Epidemiol 2025;5:1490141. [doi: 10.3389/fepid.2025.1490141] [Medline: 40078574]

XSL•FO
RenderX

Abbreviations

ABER: annual blood examination rate
ACT: artemisinin-based combination therapy
API: annual parasite index
CNM: National Center for Parasitology, Entomology, and Malaria Control
G6PD: glucose-6-phosphate dehydrogenase
iDES: integrated drug efficacy surveillance
MIS : Malaria Information System
RDT: rapid diagnostic test
VMW: village malaria worker
WHO: World Health Organization

Edited by T Sanchez; submitted 14.Oct.2025; peer-reviewed by K Tetteh, PL Aung; revised version received 14.Mar.2026; accepted 01.Apr.2026; published 12.Jun.2026.

Please cite as:
Ngor P, Rekol H, Thyda E, Rith R, Kimleng S, Rattana Y, Vunsokserey O, Vanna H, Pumeoung C, Ringwald P, Maude RJ, White L, Sovannaroth S
Digital Health Center App for Community and National Malaria Surveillance in Cambodia: Implementation Case Study
JMIR Public Health Surveill 2026;12:e85881
URL: https://publichealth.jmir.org/2026/1/e85881
doi:10.2196/85881

© Pengby Ngor, Huy Rekol, Eng Thyda, Ry Rith, Sok Kimleng, Yoem Rattana, Ou Vunsokserey, Hem Vanna, Chang Pumeoung, Pascal Ringwald, Richard J Maude, Lisa White, Siv Sovannaroth. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 12.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Health Care Providers' Perspectives on Early Warning Systems for Acute Respiratory Infections in Canada: Qualitative Study

Jinfan Qiang[1,2], MSc; Kevin Maruthananth[1], HBSc; Maaha Farrukh[1], MSc; Andrew David Pinto[1,2,3,4,5], MD; Banafshe Hosseini[4,5], PhD

[1]Upstream Lab, MAP Centre for Urban Health Solutions, St. Michael's Hospital, 30 Bond Street, Toronto, ON, Canada
[2]Institute of Health Policy, Management, and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada
[3]Department of Family and Community Medicine, St. Michael's Hospital, Unity Health Toronto, Toronto, ON, Canada
[4]Department of Family and Community Medicine, Faculty of Medicine, University of Toronto, Toronto, ON, Canada
[5]Division of Clinical Public Health, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Jinfan Qiang, MSc
Upstream Lab, MAP Centre for Urban Health Solutions, St. Michael's Hospital, 30 Bond Street, Toronto, ON, Canada

Abstract

Background: Acute respiratory infections (ARIs) remain a significant global health challenge and are the second leading cause of disease burden and mortality. Early warning systems (EWS) play a key role in detecting clinical deterioration, alerting health care providers (HCPs), and supporting pandemic surveillance. While existing literature highlights HCPs' positive experiences with EWS in confirming clinical assessments and guiding escalation, perspectives on how these systems can be optimized for ARI management remain underexplored.

Objective: As Canada continues to develop and operationalize EWS for outbreak and pandemic preparedness, this study aims to explore the experiences and insights of primary care providers, emergency department (ED) physicians, and researchers regarding the use of EWS for ARI management in Canada.

Methods: Eleven participants, including primary care providers, ED physicians, and researchers from urban and rural settings across 5 Canadian provinces (Ontario, Newfoundland and Labrador, Quebec, British Columbia, and Manitoba), were recruited in 2024. All participants regularly managed patients with ARIs or played key roles in pandemic response. A codebook thematic analysis was conducted to identify patterns and themes, with subthemes organized under broader thematic categories. Data saturation was assessed during the analysis phase. The study adhered to the COREQ (Consolidated Criteria for Reporting Qualitative Research) guidelines.

Results: Among the 11 participants, there was approximately equal representation across gender and age groups, and more than 90% had over ten years of experience in ARI management. Three overarching themes emerged. First, participants demonstrated general awareness of the use of EWS in ARI management, including outbreak detection, screening and triage support, and informing clinical decision-making. Technologies and surveillance tools used during the COVID-19 pandemic were frequently referenced; however, understanding of specific EWS and their application to ARI management was often limited. Second, participants identified key attributes of an effective EWS as accuracy, timeliness, integration, and equity, emphasizing the need for seamless integration into existing Canadian health care workflows without increasing administrative burden. Third, anticipated challenges were described across 4 stages of EWS development, including initiation (funding and privacy concerns), implementation (outdated data systems and limited legislation), use (staff shortages and capacity constraints), and evaluation (lack of standardized and innovative evaluation approaches).

Conclusions: This study engaged 11 experienced HCPs and researchers who were directly involved in patient care and public health response to ARI outbreaks and qualitatively explored their perspectives on EWS for ARI management and pandemic preparedness. The findings identified 3 overarching themes regarding the general knowledge, desired attributes, and anticipated challenges of EWS in ARI management, highlighting the importance of co-designing EWS with clinicians, researchers, and other key stakeholders to improve their effectiveness and integration into clinical practice and pandemic preparedness across Canada.

(JMIR Public Health Surveill 2026;12:e85244) doi:10.2196/85244

KEYWORDS

early warning system; pandemic preparedness; qualitative method; primary care; health care

XSL•FO
RenderX

Introduction

Acute respiratory infections (ARIs) affect the upper and/or lower respiratory tract, ranging from mild to life-threatening illnesses, and remain a significant public health concern [1], particularly among children [2]. Over the past 2 decades, severe ARIs, including SARS-CoV-2, have caused significant morbidity and mortality worldwide [3-5]. In 2021, SARS-CoV-2 was the second leading cause of global mortality and the leading cause of disability-adjusted life years [6]. Given the highly contagious nature and rapid transmission of ARIs, early detection and timely intervention are essential to mitigating their impact and preventing further transmission.

Early warning systems (EWS) are designed to identify clinical deterioration, alert health care providers (HCPs), and facilitate timely intervention [7,8]. These systems have been widely implemented in acute care settings such as hospitals and emergency departments (EDs) [9-11] and have also been integrated into outpatient and community settings, including general practices and health centers [12,13]. Primary care and ED settings play complementary roles in early outbreak detection; primary care often serves as the first point of contact for patients with early or nonspecific symptoms, while EDs act as sentinel sites during outbreaks by managing surges of patients with acute presentations. Syndromic surveillance using primary care and ED data has been shown to be a valuable tool for early outbreak detection during the COVID-19 pandemic [14]. However, both settings face distinct implementation challenges, including data interoperability and real-time integration in primary care and crowding, time pressure, and workforce constraints in EDs [15,16]. These contextual and workflow-related factors influence how EWS are used in practice, underscoring the importance of understanding frontline clinicians' perspectives when considering EWS implementation and optimization.

In Canada, multiple surveillance and EWS platforms monitor ARIs at various levels of the health care system. FluWatchers, a voluntary, self-reported system operated by the Public Health Agency of Canada, collects data on flu-like symptoms to track respiratory virus activity across the country [17]. The TARRANT Viral Watch initiative links primary care clinics and laboratories to monitor ARI trends in real time [18]. The Sentinel Practitioner Surveillance Network (SPSN) engages health care providers in tracking clinical and laboratory data to assess the epidemiology of influenza and other respiratory pathogens [19]. The Serious Outcomes Surveillance (CIRN SOS) network operates within hospitals to monitor severe ARI cases and evaluate disease burden and vaccine effectiveness [20]. However, challenges related to data interoperability and fragmented governance across jurisdictions remain.

Internationally, EWS-related approaches for ARI include both clinical risk stratification tools and population-based surveillance systems. For example, the National Early Warning Score (NEWS and NEWS2) is used in the United Kingdom, and it has been evaluated during COVID-19 for predicting mortality and Intensive Care Unit admission [21]. In contrast, population-based hospitalization surveillance networks such as the Respiratory Virus Hospitalization Surveillance Network (RESP-NET) in the United States monitor laboratory-confirmed hospitalizations for major respiratory viruses to support situational awareness and public health response [22].

While these systems provide valuable surveillance data, they are not consistently integrated into real-time clinical decision-making, and frontline patient management remains limited. As a result, little research examined how EWS are experienced and used within primary care and ED settings in Canada [23,24] or how system-level limitations affect their practical operations [25]. Understanding these perspectives is crucial for informing optimal EWS implementation and effectiveness in ARI management and pandemic preparedness. Accordingly, this study aims to explore the experiences and insights of primary care providers, ED physicians, and researchers regarding the use of EWS for ARI management in Canada.

Methods

Overview

This study was conducted following the COREQ (Consolidated Criteria for Reporting Qualitative Research (Checklist 1) [26].

Participant Recruitment

Individual online interviews were conducted with health care providers and researchers involved in pandemic response using the Zoom video platform (Zoom Video Communications). Eligible participants included family physicians, ED physicians, and infectious disease specialists whose work primarily involves managing ARIs. Researchers with expertise in pandemic response were also included. Inclusive eligibility criteria ensured representation across a range of clinical roles and practice settings, including providers from both urban and rural contexts. This approach enabled the capture of diverse perspectives from professionals directly impacted by EWS and supported a comprehensive understanding of potential challenges and needs in ARI surveillance. Participants were recruited through professional clinical and research networks focused on primary care, emergency medicine, infectious diseases, and pandemic preparedness, including investigator networks associated with adaptive platform trials, respiratory infection surveillance initiatives, and health services research collaborations in Canada. Recruitment was conducted via email using a standardized recruitment script approved by the Unity Health Toronto Communications Department. Recruitment took place between August and October 2024. The response rate was approximately 52%, with eleven of twenty-one invited individuals participating in the study.

Data Collection

Interviews were conducted by 2 researchers (KM and MF) between August and October 2024. At the beginning of each session, participants completed a brief demographic questionnaire consisting of 9 questions. The demographic questionnaire collected information on participants' professional role, years of relevant experience, practice setting (eg, urban or rural), and involvement in ARI management or pandemic response. Responses were manually recorded to ensure

XSL•FO
RenderX

confidentiality, and this portion was not audio-recorded. A semistructured interview guide was informed by relevant literature on EWS and ARI surveillance, as well as the study objectives and the research team's research experience (see Multimedia Appendix 1). The guide was designed using Bearman's step-wise approach [27] to interview schedule development and followed DeJonckheere and Vaughn's framework [28] for semistructured interviewing in primary care research. The interview guide was not formally pilot tested; however, it was reviewed and refined by the research team prior to data collection and iteratively adapted during the study to reflect participants' experiences and explore perspectives on the design, implementation, and evaluation of EWS. The full semistructured interview guide, including all questions, is provided in Multimedia Appendix 1.

Data Analysis

All interviews were audio-recorded, transcribed, and deidentified by a professional transcription service. Participants were offered the opportunity to review their transcripts for accuracy; however, no participants requested transcript review or provided feedback.

Data were analyzed using a codebook-informed reflexive thematic analysis, following Braun and Clarke's 6-phase framework [29-31]. An inductive approach was used to allow themes to emerge directly from the data. All transcripts were coded and analyzed by a single researcher (JQ), who has training and prior experience in qualitative methodology. Iterative feedback from the principal investigator (BH) was used to refine the coding framework and interpretation. Following familiarization with the data, an initial set of codes was generated and iteratively refined as analysis progressed [32,33]. A finalized codebook was developed and validated by applying it to a new transcript before proceeding with full coding. A hierarchical coding structure outlining themes, subthemes, and example codes is provided in Multimedia Appendix 2.

A codebook was developed to support analytic organization and was refined through iterative application across transcripts. Subthemes were developed when related codes clustered under a broader theme. Reflexivity was maintained throughout the analysis, with process through memoing and regular discussions within the research team, which were used to document reflections, potential biases, interpretive decisions, and support analytic rigor [34].

Interviews were coded using NVivo software (version 14 plus; Lumivero). Data analysis began after all interviews were completed. Data saturation was assessed through ongoing discussions, where it was determined that additional data collection was unlikely to yield new insights [35]. Regular meetings were held with the study team to ensure alignment in interpretation and contextual accuracy.

Ethical Consideration

This study was approved by the Unity Health Toronto Research Ethics Board (REB 24‐101) prior to the initiation of data collection. The informed consent process, including the consent form, was reviewed and approved by the REB. Written informed consent was obtained via email from all participants before participation. Participants were informed that their participation was voluntary and that they could withdraw from the study at any time without consequence.

To protect participant privacy and confidentiality, interviews were transcribed and deidentified by a professional transcription service before being shared with the research team. All study materials and data were securely stored on a restricted-access drive on St. Michael's Hospital's server and were accessible only to research staff approved under the REB application. Interviews were scheduled for one hour, and participants received a CAD $40 (approximately US $ 30, based on an exchange rate of 1 CAD ≈ US $ 0.74 in 2024) honorarium as compensation for their time.

Results

Participant Characteristics

Eleven participants from across Canada (Ontario, Newfoundland and Labrador, Quebec, British Columbia, and Manitoba) were interviewed between August and October 2024, with most interviews lasting 30-45 minutes. Demographic characteristics are summarized in Table 1. Participants were evenly distributed by sex (6/11, 55% male) and age group (6/11, 55% aged 30‐44 years). The most common occupation was family physician (5/11, 45%), followed by ED physicians (2/11, 18%), infectious disease physicians (2/11, 18%), a critical care physician (1/11, 9%), and a primary care scientist (1/11, 9%). Most (10/11, 90%) had over 10 years of experience, and 9 out of 11 (82%) worked in urban settings. Regarding EWS knowledge, most participants reported moderate familiarity, while 27% (3/11) had minimal knowledge, and 9% (1/11) were highly knowledgeable.

XSL•FO
RenderX

Table . Demographic characteristics of participants in a qualitative interview study examining early warning systems (EWS) for acute respiratory infection management among health care providers and researchers in Canada, 2024 (n=11).

Participant characteristics	Values, n (%)
Sex	
Male	6 (55)
Female	5 (45)
Age range (years)	
30 - 44	6 (55)
45 - 59	5 (45)
Occupation	
Family physician	5 (45)
Emergency department physician	2 (18)
Others[a]	4 (36)
Years of experience (years)	
≤10	1 (9)
11 - 15	6 (55)
16 - 19	2 (18)
≥20	2 (18)
Geographic Setting	
Urban	9 (82)
Rural	2 (18)
Experience with EWS[b]	
Minimal	3 (27)
Moderate	7 (64)
High	1 (9)

[a]Others: critical care physician, infectious disease and critical care physician specialist, infectious disease physician, and scientist.
[b]EWS: early warning system.

Identified Themes

Three main themes were identified from the analysis: (1) general knowledge of EWS for ARI, (2) desired attributes of EWS for ARI, and (3) anticipated challenges in EWS implementation. Within each theme, 3-4 subthemes were identified.

Theme 1: General Knowledge of EWS for ARI

Participants reported varying levels of experience with EWS but demonstrated a broad understanding of their role in ARI outbreak monitoring and detection. Most discussions focused on lessons learned from the COVID-19 pandemic, with specific emphasis on the utility of EWS in ARI detection and management.

General Awareness

While all participants were familiar with the general concept of EWS, many struggled to name specific systems used for ARI in Canada. When asked for examples, some participants mentioned “wastewater surveillance” (n=3) or “nasopharyngeal swab testing” (n=2), often describing these as ways to track infection trends at the population level rather than as formal EWS. Others referred to public health outbreak updates or case reporting systems as early warning tools. EWS was often described as “a vague term” (P02, ED physician) or “an abstract concept, something of the future” (P05, primary care scientist). Overall, participants’ understanding of EWS appeared to be shaped by their experiences during the COVID-19 pandemic, with limited distinction made between diagnostic testing, population-level surveillance, and structured EWS. One family physician reflected on their limited familiarity,

> *I might have to downgrade my other response… not as knowledgeable as I thought I was in terms of specific early warning systems.* [P08, Family physician]

There was a perceived lack of well-developed national surveillance systems, with some participants stating that provincial-level wastewater surveillance was the only system they had direct experience with. One participant explained,

> *National surveillance is not really yet fully developed… at the provincial level, we have wastewater surveillance, but we don’t have experience with a better EWS.* [P01, Infectious disease physician]

XSL•FO
RenderX

A participant with public health training described the disconnect between EWS and clinical practice, stating,

> *I have training in public health… worked in low-income settings and infectious disease control … so I have some concept of what they are. If this question is about how it's relevant to my current practice, I don't think I truly understand how the early warning or early detection systems work and maybe that's because they don't work or don't exist in our current clinical environment.* [P07, ED physician]

Participants' self-reported familiarity with early warning systems was collected descriptively and was not a primary analytic focus.

Insights and Tools Learned From the COVID-19 Pandemic

Participants referenced machine learning, artificial intelligence, and electronic communication as tools that could be integrated into EWS for ARI, although discussions on these technologies were brief. COVID-19 was frequently mentioned as a turning point in EWS discussions, highlighting both challenges and innovations. The complexity of SARS-CoV-2, particularly its "asymptomatic transmission" and "atypical symptoms," was noted as an obstacle to outbreak surveillance and management (P02, ED physician). However, the pandemic also led to rapid adoption of new technologies and scaling of public health interventions.

> *The wastewater surveillance is … a novel approach. We did not do this prior to COVID-19, at least in a widespread way. So, the lessons learned were adopt new technology, do it quickly and we can scale it up during the high-pressure pandemic situation.* [P01, Infectious disease physician]

Four participants also indicated that public health agencies played a crucial role in providing outbreak updates and disseminating information during COVID-19. One participant noted,

> *We would rely on public health to tell us … there's an outbreak… then we… institute… more formal screening process at triage so that we are isolating people quickly…It would be public health looking at the collective data… to let us know that they're starting to see something new come up and to be on the lookout.* [P07, ED physician]

Role of EWS in ARI Management

Participants identified 3 key functions of EWS in ARI control and management:

First, most participants highlighted the role of EWS in detecting early signals of outbreaks, particularly by identifying rising case numbers or unusual clusters of symptoms before formal confirmation. As a primary care scientist explained:

> *I think it has tremendous potential to… policymakers… public health authorities aware of what's happened on the ground… (if) we have noticed … cluster of patients who are showing up with … not part of the usual respiratory group of symptoms, and they are all associated with this mutation… this variant, so heads up* [P05, Primary care scientist]

Second, some participants described EWS as supporting the screening and triage process by helping identify individuals at higher risks early in the care pathway, particularly in EDs. This was often discussed in relation to flagging symptoms or recent travel history associated with affected areas, enabling timely isolation and infection control measures. As an ED physician explained,

> *Screen specifically for symptoms and travel exposures at triage in the Emergency Department… usually have a system… to isolate individuals who might be coming from affected areas… where new and emerging diseases have been flagged, it then gets flagged at triage, and then to help tell us whether people have to be isolated and then we very quickly get… the Infection Prevention and Control involved.* [P07, ED physician]

Finally, some participants described EWS as supporting informed clinical decision-making and outbreak preparedness by aggregating data across sources and helping anticipate which infections may be circulating in the community. As one infectious disease physician explained,

> *We're watching the respiratory tract infection epidemiology so we can predict what may be presenting in the clinic based on what is circulating in the community… We are in a pertussis epidemic in [province name]… cases are higher than baseline and when patients present with a cough, we consider that they have pertussis based on the epidemiology. So, this is highly relevant to my clinical practice.* [P01, Infectious Disease physician]

Theme 2: Desired Attributes of EWS for ARI

Participants identified 4 key attributes essential for an effective EWS for ARI, including accuracy, timeliness, integration, and equity. These qualities were viewed as critical for ensuring that the system reliably detects outbreaks, provides timely alerts, seamlessly integrates into clinical workflows, and addresses disparities in health care access.

Accuracy

Most participants emphasized that an effective EWS must balance sensitivity and specificity, often describing this balance in practical terms. Participants highlighted the importance of detecting "clear signals" of emerging outbreaks early (P02, ED physician), while avoiding frequent false alarms that could overwhelm clinicians or reduce confidence in the system. Concerns were raised that over-reliance on an inaccurate or overly sensitive system could erode trust and limit its usefulness in clinical practice.

> *Doesn't shoot out red flags constantly leading to alarm fatigue, but it also needs to be calibrated so that real early warnings that need to be warned about get the appropriate thought and reaction from policymakers and the public.* [P04, Infectious disease and critical care physician specialist]

XSL•FO
RenderX

To enhance accuracy, participants favored a multifactorial approach rather than reliance on a single metric, such as swab testing. A more comprehensive surveillance system was suggested to reduce biases and improve reliability.

> *would have a multi-factorial kind of approach and wouldn't rest on a single metric. Because these things are complex, the health system … and people…and communities…, I don't think there is a silver bullet warning system like only swabs. Because if you pick one thing, it may have blind spots or biases baked into that particular approach. I would advocate a more holistic multi-factorial kind of approach for surveillance to recognize the complexity of the challenge of the problem we're trying to solve.* [P10, Family physician]

Timeliness

A key expectation was that an EWS should detect potential outbreaks in "Real-time" (P02, eED physician) and alert health care providers "within 24 to 72 hours" (P06, family physician). The ability to rapidly disseminate information was seen as critical for early clinical and public health interventions:

> *Timeliness… that it could impact clinical care at the frontline* [P08, Family physician]

> *It would be timely in terms of getting the information out to the providers. And I would think within…short days. Ideally 48, 72 hours… would be helpful especially with respiratory illnesses that have potential to impact morbidity and mortality and impact kind of the need for healthcare providers that are looking after patients. So, timeliness would be the biggest one probably from the surveillance system perspective.* [P08, Family physician]

Integration

Seamless integration into existing clinical systems was highlighted as a fundamental requirement. Participants stressed that the EWS should draw from multiple data sources, including electronic health records, swab testing data, wastewater surveillance, and self-reported symptoms, while minimizing disruptions to workflow.

> *It's ideally integrated into the electronic health records, so they only have to log into one system… I think a few things are ideal… everything from social media, wastewater, slight patterns, prescription medication use, symptoms that are recorded in health records and also people's own self-reporting.* [P06, Family physician]

Equity

Equity emerged as a core and widely emphasized attribute of an effective EWS across participants. Participants emphasized that an effective EWS must prioritize equity, ensuring equal access to testing and surveillance across urban and rural areas and marginalized populations. Key concerns included representation of "Indigenous and newcomer communities" (P03, Family physician), "equitable geographic coverage" (P04, Infectious disease and critical care physician specialist), and strict data privacy safeguards to prevent misuse.

> *It would be really important that particular vulnerable groups that have been marginalised by the healthcare system over the last few decades were incorporated into the decision making around this so that an early warning system doesn't further stigmatise them and that they would be wanting to participate. And… they could provide guidance on how it would need to be designed and implemented. So…our Indigenous communities… the example of having these separate TB clinics that did create a lot of inequities in the past. So, just being thoughtful about that piece. Because then we risk developing something that doesn't actually help support the individuals that may be greatest risk for contracting respiratory illnesses.* [P11, Family Physician]

Across practice contexts, participants consistently identified accuracy, timeliness, integration, and equity as core attributes of an effective EWS.

Theme 3: Anticipated Challenges in EWS Implementation

Participants identified challenges at 4 key stages of EWS development, including initiation, implementation, use, and evaluation. While multiple concerns were raised for each stage, only the most predominant issues are reported below.

Initiation

Three primary challenges were highlighted regarding the initiation of an EWS for ARI: (1) lack of long-term commitment and high costs, (2) complexity of ARI mutations, and (3) public resistance due to data privacy concerns.

The lack of sustained commitment and financial investment was a major concern. Participants noted that enthusiasm and perhaps funding for pandemic-related technologies peaks during crises but diminishes between pandemics.

> *One of the main lessons learned in the pandemic is everyone is interested during the pandemic, but no one is interested in the pandemic between the pandemics…you get all kinds of government funding and interest and global engagement when there is the threat but when there is no threat nobody wants to fund pandemic technology.* [P01, Infectious disease physician]

The complexity of ARI mutations also raised concerns about the system's practical utility. Participants questioned how an EWS could detect novel pathogens that had never been identified before.

> *Now, there is new technology that might help us with PCR, and gene sequencing, and maybe we can hope to detect viruses that are previously unidentified more easily, but that stuff is really difficult… to then think about translating that at a clinical level, I mean, I don't know what to say about that… you can't detect anything that hasn't been detected before.* [P02, ED physician]

XSL•FO
RenderX

Finally, public trust and concerns about data misuse were seen as potential barriers to initiating an EWS.

> *The governance of these types of systems, because… we've seen lots of concerns about corporations or others sucking up this type of data. So, it would need to be governed in a way that actually serves the people and the public and is not just generating profit.* [P06, Family physician]

Implementation

Challenges related to data system incompatibility and lack of legislation on data sharing were emphasized as major barriers to implementation.

Participants described fragmentation across electronic health records and data systems, with different provinces and organizations using incompatible platforms.

> *Different provinces have different electronic medical records. And a system would have to synchronize data from a variety of different sources so if it required a lot of tuning to the local electronic medical record, that would delay implementation.* [P01, Infectious disease physician]

In addition, the absence of clear legislation on data sharing was noted as a challenge. The lack of standardized protocols for tracking and tracing infections during COVID-19 was cited as an example of system limitations.

> *The contact tracing was non-existent in Ontario for the most part. It failed dramatically …restaurant sign up…(were) not be used by anyone. So, when people would test positive, there was no way to then see where they've been or anything… being able to contact trace better, that probably requires some sort of work from legislations at the legislation level.* [P02, ED physician]

Use

Challenges related to workforce shortages, system capacity, and public engagement were the most frequently discussed barriers to EWS functionality. Almost all participants explicitly expressed their concerns related to the "administrative and workload burden" (10) associated with having a new EWS, especially given the chronic "human workforce shortage" (11). In one Canadian province, "2.5 million residents don't have a primary care doctor" (11), and many rely on EDs, which are "typically really busy" (09). A participant stressed that,

> *It's hard for folks (in the ED) to add something else onto their busy list. They see so many people with very vague symptoms, it can be difficult to I think nail that down within say the triage or the first encounter with the patient.* [09, Critical care physician]

Concerns were also raised about Canada's ability to sustain robust testing and surveillance.

> *If you have an atypical presentation and people do not have RSV (respiratory syncytial virus), flu, or COVID, then that might lead you down the road of trying to investigate and discover that we're dealing with something new. So, my worry is that our systems currently are not robust in terms of testing and isolating people with respiratory illness for the things that are circulating.* [P07, ED physician]

Participants also noted that public trust and willingness to engage would influence the success of an EWS.

> *We don't have the greatest track record of using data in that equitable manner… so there is going to be an inherent amount of distress of the system because of our inability to really respond to data with high quality equitable interventions.* [04, Infectious disease and critical care physician specialist]

Additionally, screening for travel history was seen as a critical function of EWS but also a potential source of discrimination.

> *Everyone from Senegal is isolated for Ebola even though they're presenting with they hurt their leg or something. I'm exaggerating, but so it can lead to a sense of discrimination when we're making this based on where people have travelled from.* [07, ED physician]

Evaluation

Discussions on evaluation focused less on challenges and more on suggested methods for assessing EWS effectiveness. Participants proposed that evaluation should consider data synchronization across sources, ease of use, and sensitivity and specificity of screening.

> *A participant envisioned an easy-to-use EWS that "communicates with the EMR" that "flags" warning as they "input data", so they can perform "point-of-care testing".* [P11, Family physician]
>
> *Another participant suggested evaluation through "simulations and screening at airports… the evaluation happens frequently by people trying to get through and seeing if you pick that up through simulated events.*" [P09, Critical care physician]

Discussion

Principal Findings

This qualitative study explored the perspectives of frontline clinicians and researchers involved in pandemic response on the role of EWS in monitoring and managing ARI in Canada. Three key themes emerged, including general knowledge of EWS, desired attributes, and anticipated challenges. While participants recognized the potential benefits of EWS, the absence of a widely recognized and established EWS for ARI in Canada limited their ability to engage in in-depth discussions about specific platforms. Participants identified accuracy, timeliness, integration, and equity as essential attributes of an effective system and highlighted challenges at each stage of development, including initiation, implementation, use, and evaluation. These findings underscore the complexities of designing and implementing a robust, sustainable, and equitable EWS for ARI in Canada. Including participants from diverse specialties and practice settings strengthened the relevance of our findings across varied clinical and health system contexts.

XSL•FO
RenderX

Our findings align with existing literature emphasizing that EWS primarily serves to detect outbreaks, guide clinical decision-making, and inform public health responses [36,37]. EWS has been shown to support real-time tracking of disease trends, early identification of emerging risks, and allocation of health care resources [38]. Similar to our participants' perspectives, previous reviews highlight that effective EWS must be timely, accurate, and seamlessly integrated into existing health care workflows to maximize their utility [39]. Participants' references to a 24 - 72 hour alert window reflect the need for timely situational awareness to support public health and clinical response, rather than a specific evidence-based mortality threshold. Previous studies further suggest that "accuracy" in EWS reflects trade-offs between sensitivity and specificity. For example, increasing sensitivity can reduce missed events but often results in more false-positive alerts and greater alert burden, whereas prioritizing specificity may lower alert frequency but increase the risk of delayed detection of deterioration [39]. Participants' concerns about alert fatigue and missed signals are consistent with these documented trade-offs.

In Canada, multiple surveillance systems monitor respiratory infections, including FluWatchers [17], TARRANET [18], and SPSN [19]. However, these systems primarily function as population-level surveillance tools and are not designed to provide real-time clinical decision support at the point of care. In contrast, several countries, including Australia, China, and parts of Europe, operate more integrated national respiratory infection surveillance systems that combine data from multiple sources, such as laboratory, syndromic, and sentinel surveillance [40,41]. In addition, some international EWS frameworks have begun integrating machine learning and artificial intelligence approaches to enhance early signal detection, identify high-risk populations, and support timely public health alerts [42,43]. Potential reasons for the disparity in EWS for respiratory infection between Canada and other countries may reflect a combination of system-level factors. First, Canada's health system is highly decentralized, with responsibility for health care delivery and public health surveillance largely residing at the provincial and territorial levels. This governance structure can present challenges for the development of a unified, nationally integrated EWS. Second, ongoing variation in data infrastructure, as well as differences in privacy legislation and data-sharing agreements across jurisdictions, may limit the timely integration of surveillance data at a national level. Finally, persistent gaps in access to primary care in Canada, including a substantial proportion of individuals who are unattached to a regular primary care provider, may further constrain surveillance capacity. Individuals experiencing socioeconomic marginalization are disproportionately represented among this population, which may contribute to gaps in population-level monitoring. These findings point to the need for implementation-focused research to design and evaluate integrated EWS that link surveillance, clinical care, and public health decision-making in real time while minimizing alert fatigue and inequitable coverage.

Challenges in EWS Development and Implementation

Participants identified challenges across all 4 stages of EWS development (initiation, implementation, use, and evaluation), with challenges at the initiation stage emerging as the most pressing concern. In particular, participants emphasized the lack of sustained investment and policy commitment to pandemic preparedness outside of crisis periods. While funding and interest peak during outbreaks, maintaining long-term support for EWS was viewed as difficult, a concern that aligns with prior research describing intermittent funding in surveillance infrastructure [44]. Participants noted that challenges during the initiation phase, such as limited sustained investment and unclear long-term priorities, often shaped downstream issues related to implementation, system fit, and long-term use.

Challenges related to implementation and use were also frequently discussed. Participants highlighted the fragmentation of electronic health records across provinces and the absence of standardized data-sharing protocols as major barriers to integrating EWS into clinical workflows. These issues are consistent with broader challenges observed in decentralized health care systems, where data silos and variable policies limit timely information exchange and generate real-time decision support [45]. In addition, workforce shortages, increased administrative burden, and limited system capacity were seen as factors that could further impede the adoption and effective use of EWS, reflecting system pressures that have been widely documented in the Canadian health context [46]. Prior studies of digital surveillance and decision-support implementation have shown that workforce shortages and competing clinical demands can limit engagement with alerting systems and reduce their effective use [39].

Public trust and engagement were identified as critical to the effectiveness of EWS, particularly in relation to data governance and equity. Participants emphasized concern that poorly designed or inadequately governed systems could exacerbate existing health disparities or undermine public confidence. Similar concerns have been noted in the digital health literature, where issues related to data validity, transparency, and legal framework have been shown to influence the utility and acceptability of surveillance technologies [47]. Prior Canadian and international studies have identified equity-related challenges in EWS implementation, including limited applicability to Indigenous, rural, and remote contexts and concerns around data governance and trust [48]. Participants raised similar issues, emphasizing Indigenous leadership, community-specific needs, and contextual relevance. To address these concerns, participants advocated for the importance of co-designing EWS with communities, particularly Indigenous and marginalized populations. Participants described co-design as involving early engagement with community leadership, shared decision-making around data governance and system priorities, and ongoing feedback during pilot testing and implementation. Evidence from disaster management and health equity suggests that co-design approaches can enhance cultural responsiveness and system relevance, supporting more equitable and effective implementation of digital surveillance tools [49,50].

Opportunities and Future Directions

Despite these challenges, there is growing global momentum toward strengthening respiratory surveillance. Many countries

XSL•FO
RenderX

have adopted disease-specific and syndromic surveillance models that rely on real-time data collection and clinical case reporting [51]. For example, the RESP-NET in the United States and Canada's FluWatchers program demonstrates the feasibility of large-scale respiratory surveillance [52,53]. However, our findings suggest that existing surveillance efforts in Canada are not yet configured as a unified system that can support both public health monitoring and real-time frontline clinical decision-making. Moving forward, efforts should focus on designing an EWS that bridges these domains. Key considerations include improving interoperability between electronic health records, laboratory networks, and syndromic surveillance tools; establishing clear and consistent frameworks for data governance, privacy, and equitable access; and involving key stakeholders in system design to support cultural safety and equity in implementation.

In the Canadian context, advancing transparent governance and equitable data use will require clearer coordination across federal, provincial, and territorial jurisdictions. Participants' emphasis on trust also points to the importance of embedding equity considerations early in system design, including meaningful engagement with communities that experience greater social and structural barriers. Together, these considerations may help enable Canada's future EWS to be perceived as legitimate, trustworthy, and responsible to the needs of its diverse population.

Limitations

This study has several limitations. First, we did not include policymakers, public health officials, and data scientists involved in the design and governance of EWS, whose perspectives may have provided additional insights for policy development, data integration strategies, and long-term sustainability. Second, interviews were conducted between August and October 2024, outside peak winter respiratory virus season. Participants' perspectives may therefore reflect experiences during relatively stable surveillance periods rather than peak seasonal demand; however, discussions frequently drew on pandemic-era and outbreak experiences. Finally, patient perspectives were not included, limiting our understanding of public engagement and trust in EWS adoption.

Conclusion

This study explored health care providers' and researchers' perspectives on EWS for ARI in Canada. Participants highlighted the potential of EWS to support outbreak detection, clinical decision-making, and pandemic preparedness, while also identifying persistent system-level challenges related to data integration, governance, and trust. These findings underscore the need for accurate, timely, integrated, and equitable systems that meaningfully bridge public health surveillance and clinical decision-making. Future research should incorporate policymaker and patient perspectives and examine how emerging digital and AI-enabled approaches can be responsibly integrated into EWS. With sustained investment and inclusive system design, Canada has the opportunity to strengthen its EWS capacity and improve preparedness for future respiratory outbreaks.

Funding

This project was supported by funding from EPIC-IfP initative at the University of Toronto. Author BH is supported as a Research Scientist by the Upstream Lab at the MAP center forUrban Health Solutions, at the Li Ka Shing Knowledge Institute. The opinions, results andconclusions reported in this article are those of the authors and are independent from any fundingsources.

Disclaimer

The authors declare the use of generative AI in the research and writing process. According to

the GAIDeT taxonomy (2025), the following tasks were delegated to GAI tools under full human

supervision:

- Proofreading and editing

The GAI tool used was: ChatGPT 5.

Responsibility for the final manuscript lies entirely with the authors.

GAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Declaration submitted by: JQ

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: BH, ADP
Methodology: BH, ADP
Supervision: BH

XSL•FO
RenderX

Funding acquisition: BH, ADP
Investigation: MF, KM
Formal analysis: JQ
Writing – original draft: JQ
Writing – review & editing: BH, JQ

Conflicts of Interest

None declared.

Multimedia Appendix 1
Semistructured interview guide.
[DOCX File, 21 KB - publichealth_v12i1e85244_app1.docx]

Multimedia Appendix 2
Hierarchical coding structure.
[DOCX File, 19 KB - publichealth_v12i1e85244_app2.docx]

Checklist 1
COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist.
[DOCX File, 21 KB - publichealth_v12i1e85244_app3.docx]

References

1. Chen C, You Y, Du Y, et al. Global epidemiological trends in the incidence and deaths of acute respiratory infections from 1990 to 2021. Heliyon 2024 Aug;10(16):e35841. [doi: 10.1016/j.heliyon.2024.e35841]
2. Vos T, Lim SS, Abbafati C, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. The Lancet 2020 Oct;396(10258):1204-1222. [doi: 10.1016/S0140-6736(20)30925-9]
3. Kaye AD, Okeagu CN, Pham AD, et al. Economic impact of COVID-19 pandemic on healthcare facilities and systems: International perspectives. Best Pract Res Clin Anaesthesiol 2021 Oct;35(3):293-306. [doi: 10.1016/j.bpa.2020.11.009]
4. Faramarzi A, Norouzi S, Dehdarirad H, Aghlmand S, Yusefzadeh H, Javan-Noughabi J. The global economic burden of COVID-19 disease: a comprehensive systematic review and meta-analysis. Syst Rev 2024 Feb 16;13(1):68. [doi: 10.1186/s13643-024-02476-6] [Medline: 38365735]
5. Shi T, Denouel A, Tietjen AK, et al. Global disease burden estimates of respiratory syncytial virus-associated acute respiratory infection in older adults in 2015: a systematic review and meta-analysis. J Infect Dis 2020 Oct 7;222(Suppl 7):S577-S583. [doi: 10.1093/infdis/jiz059] [Medline: 30880339]
6. Global health estimates: life expectancy and leading causes of death and disability. World Health Organization. 2021. URL: https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates [accessed 2026-03-24]
7. Wood C, Chaboyer W, Carr P. How do nurses use early warning scoring systems to detect and act on patient deterioration to ensure patient safety? A scoping review. Int J Nurs Stud 2019 Jun;94:166-178. [doi: 10.1016/j.ijnurstu.2019.03.012] [Medline: 31002971]
8. Roland D, Stilwell PA, Fortune PM, Alexander J, Clark SJ, Kenny S. Case for change: a standardised inpatient paediatric early warning system in England. Arch Dis Child 2021 Jul;106(7):648-651. [doi: 10.1136/archdischild-2020-320466] [Medline: 33419727]
9. Flenady T, Dwyer T, Sobolewska A, et al. Developing a sociocultural framework of compliance: an exploration of factors related to the use of early warning systems among acute care clinicians. BMC Health Serv Res 2020 Aug 11;20(1):736. [doi: 10.1186/s12913-020-05615-6] [Medline: 32782002]
10. Graetz D, Kaye EC, Garza M, et al. Qualitative study of pediatric early warning systems' impact on interdisciplinary communication in two pediatric oncology hospitals with varying resources. JCO Glob Oncol 2020 Jul;6:1079-1086. [doi: 10.1200/GO.20.00163] [Medline: 32673079]
11. McGaughey J, Fergusson DA, Van Bogaert P, Rose L. Early warning systems and rapid response systems for the prevention of patient deterioration on acute adult hospital wards. Cochrane Database Syst Rev 2021 Nov 22;11(11):CD005529. [doi: 10.1002/14651858.CD005529.pub3] [Medline: 34808700]
12. Brangan E, Banks J, Brant H, Pullyblank A, Le Roux H, Redwood S. Using the National Early Warning Score (NEWS) outside acute hospital settings: a qualitative study of staff experiences in the West of England. BMJ Open 2018 Oct 27;8(10):e022528. [doi: 10.1136/bmjopen-2018-022528] [Medline: 30368449]
13. Hoffmann M, Stengel S, Szecsenyi J, Peters-Klimm F. Health care professionals' perspectives on the utilisation of a remote surveillance and care tool for patients with COVID-19 in general practice: a qualitative study. BMC Prim Care 2022 Sep 27;23(1):254. [doi: 10.1186/s12875-022-01863-z] [Medline: 36167532]

XSL•FO
RenderX

14. Morgan OW, Aguilera X, Ammon A, et al. Disease surveillance for the COVID-19 era: time for bold changes. The Lancet 2021 Jun;397(10292):2317-2319. [doi: 10.1016/S0140-6736(21)01096-5]
15. Martin D, Razak F, Bayoumi I, et al. Primary care in the COVID-19 pandemic and beyond: lessons from Ontario. Can Fam Physician 2025 Jan;71(1):31-40. [doi: 10.46747/cfp.710131] [Medline: 39843197]
16. Rowe BH, McRae A, Rosychuk RJ. Temporal trends in emergency department volumes and crowding metrics in a western Canadian province: a population-based, administrative data study. BMC Health Serv Res 2020 Apr 26;20(1):356. [doi: 10.1186/s12913-020-05196-4] [Medline: 32336295]
17. Government of Canada Flu (influenza): symptoms and treatment. 2023. URL: https://www.canada.ca/en/public-health/services/diseases/flu-influenza/fluwatcher.html [accessed 2026-03-24]
18. University of Calgary. University of calgary TARRANT viral watch. Welcome to TARRANT Viral Watch. URL: https://cumming.ucalgary.ca/research/alberta-recording-network/home-tarrant [accessed 2026-03-24]
19. Sentinel practitioner surveillance network (SPSN) — influenza vaccine effectiveness program. Public Health Ontario. 2019. URL: http://www.publichealthontario.ca/en/Health-Topics/Immunization/SPSN [accessed 2026-03-24]
20. McNeil S, Andrew M. Serious outcomes surveillance (SOS) network. Canadian Immunization Research Network (CIRN). 2013. URL: https://cirnetwork.ca/network/serious-outcomes/ [accessed 2026-03-24]
21. Kostakis I, Smith GB, Prytherch D, et al. The performance of the National Early Warning Score and National Early Warning Score 2 in hospitalised patients infected by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Resuscitation 2021 Feb;159:150-157. [doi: 10.1016/j.resuscitation.2020.10.039] [Medline: 33176170]
22. Respiratory virus hospitalization surveillance network (RESP-NET). Centers for disease control and prevention. 2025 Jan 19. URL: https://www.cdc.gov/resp-net/dashboard/index.html [accessed 2026-03-24]
23. Lee L, Desroches M, Mukhi S, Bancej C. FluWatchers: evaluation of a crowdsourced influenza-like illness surveillance application for Canadian influenza seasons 2015-2016 to 2018-2019. Can Commun Dis Rep 2021 Sep 10;47(9):357-363. [doi: 10.14745/ccdr.v47i09a02] [Medline: 34650332]
24. McElroy T, Swartz EN, Hassani K, et al. Implementation study of a 5-component pediatric early warning system (PEWS) in an emergency department in British Columbia, Canada, to inform provincial scale up. BMC Emerg Med 2019 Nov 27;19(1):74. [doi: 10.1186/s12873-019-0287-5] [Medline: 31771517]
25. Petersen JA, Rasmussen LS, Rydahl-Hansen S. Barriers and facilitating factors related to use of early warning score among acute care nurses: a qualitative study. BMC Emerg Med 2017 Dec 1;17(1):36. [doi: 10.1186/s12873-017-0147-0] [Medline: 29191159]
26. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
27. Bearman M. Focus on methodology: eliciting rich data: a practical approach to writing semi-structured interview schedules. FoHPE 2019 Nov 28;20(3):1-11. [doi: 10.11157/fohpe.v20i3.387]
28. DeJonckheere M, Vaughn LM. Semistructured interviewing in primary care research: a balance of relationship and rigour. Fam Med Community Health 2019;7(2):e000057. [doi: 10.1136/fmch-2018-000057] [Medline: 32148704]
29. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
30. Braun V, Clarke V. Thematic analysis. In: APA Handbook of Research Methods in Psychology, Vol 2: Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological 2012:57-71 URL: http://content.apa.org/books/13620-000 [accessed 2026-03-24] [doi: 10.1037/13620-004]
31. Braun V, Clarke V. Conceptual and design thinking for thematic analysis. Qualitative Psychology 2022 Feb;9(1):3-26. [doi: 10.1037/qup0000196]
32. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004 Feb;24(2):105-112. [doi: 10.1016/j.nedt.2003.10.001] [Medline: 14769454]
33. Lindgren BM, Lundman B, Graneheim UH. Abstraction and interpretation during the qualitative content analysis process. Int J Nurs Stud 2020 Aug;108:103632. [doi: 10.1016/j.ijnurstu.2020.103632] [Medline: 32505813]
34. Jamieson MK, Govaart GH, Pownall M. Reflexivity in quantitative research: a rationale and beginner's guide. Social and Personality Psych 2023 Apr;17(4):e12735 [FREE Full text] [doi: 10.1111/spc3.12735]
35. Fusch P, Ness L. Are we there yet? Data saturation in qualitative research. Qual Rep 2015;20(9) [FREE Full text] [doi: 10.46743/2160-3715/2015.2281]
36. Soucie JM. Public health surveillance and data collection: general principles and impact on hemophilia care. Hematology 2012 Apr;17 Suppl 1(1):S144-S146. [doi: 10.1179/102453312X13336169156537] [Medline: 22507804]
37. German RR, Lee LM, Horan JM, et al. Updated guidelines for evaluating public health surveillance systems: recommendations from the Guidelines Working Group. MMWR Recomm Rep 2001 Jul 27;50(RR-13):1-35. [Medline: 18634202]
38. Patel D, Ayesha IE, Monson NR, et al. The effectiveness of metformin in diabetes prevention: a systematic review and meta-analysis. Cureus 2023;15(9):e46108. [doi: 10.7759/cureus.46108]

XSL•FO
RenderX

39. Meckawy R, Stuckler D, Mehta A, Al-Ahdal T, Doebbeling BN. Effectiveness of early warning systems in the detection of infectious diseases outbreaks: a systematic review. BMC Public Health 2022 Nov 29;22(1):2216. [doi: 10.1186/s12889-022-14625-4] [Medline: 36447171]
40. de Fougerolles TR, Damm O, Ansaldi F, et al. National influenza surveillance systems in five European countries: a qualitative comparative framework based on WHO guidance. BMC Public Health 2022 Jun 9;22(1):1151. [doi: 10.1186/s12889-022-13433-0] [Medline: 35681199]
41. El Guerche-Séblain C, Rigoine De Fougerolles T, Sampson K, et al. Comparison of influenza surveillance systems in Australia, China, Malaysia and expert recommendations for influenza control. BMC Public Health 2021 Sep 26;21(1):1750. [doi: 10.1186/s12889-021-11765-x] [Medline: 34563151]
42. Gao Y, Cai GY, Fang W, et al. Machine learning based early warning system enables accurate mortality risk prediction for COVID-19. Nat Commun 2020 Oct 6;11(1):5033. [doi: 10.1038/s41467-020-18684-2]
43. El Morr C, Ozdemir D, Asdaah Y, Saab A, El-Lahib Y, Sokhn ES. AI-based epidemic and pandemic early warning systems: a systematic scoping review. Health Informatics J 2024;30(3):14604582241275844. [doi: 10.1177/14604582241275844] [Medline: 39172555]
44. Yang L, Weston C, Cude C, Kincl L. Evaluating Oregon's occupational public health surveillance system based on the CDC updated guidelines. Am J Ind Med 2020 Aug;63(8):713-725. [doi: 10.1002/ajim.23139] [Medline: 32483871]
45. Quigley L, Lacombe-Duncan A, Adams S, Hepburn CM, Cohen E. A qualitative analysis of information sharing for children with medical complexity within and across health care organizations. BMC Health Serv Res 2014 Jun 30;14:283. [doi: 10.1186/1472-6963-14-283] [Medline: 24981205]
46. Baumann A, Crea-Arsenio M. The crisis in the nursing labour market: Canadian policy perspectives. Healthcare (Basel) 2023 Jul 6;11(13):1954. [doi: 10.3390/healthcare11131954] [Medline: 37444788]
47. Donelle L, Comer L, Hiebert B, et al. Use of digital technologies for public health surveillance during the COVID-19 pandemic: a scoping review. Digit Health 2023;9:20552076231173220. [doi: 10.1177/20552076231173220] [Medline: 37214658]
48. Govorchin A, Leduc M, Atleo CG, Hoogeveen D, Borgos I, Patrick L. The right to health: indigenous data sovereignty in Canada during and beyond the COVID-19 pandemic. Lancet Reg Health Am 2026 Feb;54:101335. [doi: 10.1016/j.lana.2025.101335]
49. Tsai CH, Huang C, Chen YC, Zendejas E, Krafka S, Zendejas J. Co-design smart disaster management systems with indigenous communities. Digit Gov: Res Pract 2024 Sep 30;5(3):1-21. [doi: 10.1145/3660643]
50. Gerrard J, Godwin S, Whiteley K, Charles J, Sadler S, Chuter V. Co-design in healthcare with and for First Nations Peoples of the land now known as Australia: a narrative review. Int J Equity Health 2025 Jan 6;24(1):2. [doi: 10.1186/s12939-024-02358-2] [Medline: 39762922]
51. Abat C, Chaudet H, Rolain JM, Colson P, Raoult D. Traditional and syndromic surveillance of infectious diseases and pathogens. Int J Infect Dis 2016 Jul;48:22-28. [doi: 10.1016/j.ijid.2016.04.021] [Medline: 27143522]
52. Respiratory virus hospitalization surveillance network (RESP-NET). Centers for Disease Control and Prevention (CDC). 2024. URL: https://www.cdc.gov/resp-net/dashboard/index.html [accessed 2026-03-24]
53. Ben Moussa M, Rahal A, Lee A, Mukhi S. Syndromic surveillance performance in Canada throughout the COVID-19 pandemic, March 1, 2020 to March 4, 2023. Can Commun Dis Rep 2023 Mar 1;48(11-12):501-509. [doi: 10.14745/ccdr.v49i1112a06]

Abbreviations

ARI: Acute Respiratory Infection
CIRN SOS: Canadian Immunization Research Network Serious Outcomes Surveillance
COREQ: Consolidated Criteria for Reporting Qualitative Research
ED: emergency department
EWS: early warning systems
HCP: health care provider
NEWS: National Early Warning Score
NEWS2: National Early Warning Score 2
RESP-NET: Respiratory Virus Hospitalization Surveillance Network
RSV: Respiratory Syncytial Virus
SPSN: Sentinel Practitioner Surveillance Network

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 03.Oct.2025; peer-reviewed by Q Liu, Y Sarikhani; revised version received 29.Jan.2026; accepted 30.Jan.2026; published 23.Apr.2026.

Please cite as:
Qiang J, Maruthananth K, Farrukh M, Pinto AD, Hosseini B
Health Care Providers' Perspectives on Early Warning Systems for Acute Respiratory Infections in Canada: Qualitative Study
JMIR Public Health Surveill 2026;12:e85244
URL: https://publichealth.jmir.org/2026/1/e85244
doi:10.2196/85244

© Jinfan Qiang, Kevin Maruthananth, Maaha Farrukh, Andrew David Pinto, Banafshe Hosseini. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 23.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Requirements and Implementation Challenges of Automated Outbreak Detection Systems in the European Union: Mixed Methods Study

Paulina vom Felde genannt Imbusch, MScPH; Ann Christin Vietor, MSc; Inessa Markus, MD, MSAE; Michaela Diercke, MA, MSAE, MD; Alexander Ullrich, Dr rer nat

Department of Infectious Diseases Epidemiology, Robert Koch Institute, Seestr. 10, Berlin, State of Berlin, Germany

Corresponding Author:
Paulina vom Felde genannt Imbusch, MScPH
Department of Infectious Diseases Epidemiology, Robert Koch Institute, Seestr. 10, Berlin, State of Berlin, Germany

Abstract

Background: Automated outbreak detection can enhance infectious disease surveillance by enabling early identification of outbreaks and supporting timely public health measures. However, information on its current use by national public health institutes remains limited.

Objective: This study aimed to provide an updated and extended overview of automated outbreak detection use in the European Union and the United Kingdom by (1) assessing current demand, (2) examining the availability of key prerequisites within existing surveillance systems, and (3) identifying challenges and requirements for implementation.

Methods: A mixed methods approach was applied as part of the Joint Action UNITED4Surveillance. Data were collected between April 2023 and January 2024 through an online survey sent to 25 countries, an in-person workshop with experts from 16 countries, and follow-up meetings with participating national public health institutes. Additional information for selected countries was obtained from the literature. Data were analyzed descriptively.

Results: Twenty-one countries completed the survey. In total, 33.3% (n=7) of the countries had established automated outbreak detection systems, 19% (n=4) were planning implementation, and 47.6% (n=10) participated in a pilot project. All surveyed countries reported acting on detected surveillance signals. Most had suitable surveillance infrastructure, including case-based data (n=20, 95.2%), daily reporting (n=21, 100%), and at least 3 years of historical data (n=21, 100%). Main barriers to implementation included limited funding (n=15, 71.4%), insufficient IT capacity, and data quality issues (n=9, 42.9%). Despite heterogeneity in methods and system design, outputs and user requirements were largely similar across countries, with needs for flexible outputs, stratification, and user-friendly interfaces.

Conclusions: While the specific methods in existing automated outbreak detection systems differ, overall demands and outputs are similar, suggesting that a single tool could serve multiple countries. European Union–funded Joint Actions may help address some of these barriers by supporting collaborative tool development and knowledge exchange.

(JMIR Public Health Surveill 2026;12:e96576) doi:10.2196/96576

KEYWORDS

public health surveillance; communicable diseases; disease outbreaks; algorithms; public health informatics

Introduction

Timely detection of infectious disease outbreaks, enabling targeted investigations and control measures, is a core objective of public health surveillance [1,2]. Outbreak detection relies on identifying case numbers that exceed the expected threshold, typically recognized through accumulated case reports of notifiable diseases or outbreak notifications from reporting entities. With the growing availability of electronic case reports, automated outbreak detection has gained significant importance, offering a systematic approach to identify outbreaks [1,3,4].

Outbreak detection refers to the process of identifying unusual increases in case numbers by comparing the current epidemiological situation with a baseline, usually established from historical or geographical data. Statistical methods estimate the expected number of cases and compute a threshold. When observed values deviate significantly from these expectations, surpassing the threshold, signals are generated and often presented in tables, visualizations, or descriptive reports. These signals necessitate epidemiological investigations to determine whether they represent potential outbreaks or false positives [5].

XSL•FO
RenderX

Various methods for outbreak detection exist and have evolved over time [6-8]. Among the most widely used are the method by Farrington et al [9] and adaptations by Noufaily et al [10], which use generalized linear models to predict the expected number of cases accounting for seasonality, time trend, and overdispersion. Building on these methods, we can distinguish a hierarchy of components: statistical methods underlie tools that operationalize them, and some of these tools allow for the automated analysis of complete surveillance datasets. Here, automation means performing outbreak detection across relevant dimensions such as region, age, or subtype rather than sequentially inspecting individual strata and time series. Together with the necessary infrastructure, standard operating procedures, and trained personnel, these tools constitute comprehensive automated outbreak detection systems (AODSs).

Despite methodological advances, information on AODS implementation in the European Union (EU) is limited. The most recent overview from 2010 reported 6 institutes in 5 countries (Denmark, Germany, the Netherlands, Sweden, and the United Kingdom) using automated detection for routine surveillance. On the basis of the existing AODSs and practical experiences, recommendations were drawn, including the regular evaluation of AODSs and aiming for a user-friendly approach directed at epidemiologists [11,12]. Additionally, a 2017 workshop involving stakeholders from 17 institutions across the EU emphasized the exchange of data and experiences across countries, the need for visual outputs for users, and assessment and comparison of detection methods [13].

The EU4Health program is the EU's funding instrument for strengthening public health systems and preparedness. Within this program, Joint Action UNITED4Surveillance (U4S) brings together national public health institutes (NPHIs) across Europe to improve integrated infectious disease surveillance, including strengthening capacities for outbreak detection and response. One aim of U4S was to identify the requirements of EU countries for effective outbreak detection [14]. Understanding these needs and current practices is a necessary first step before developing and implementing new AODSs. While the project also involved developing and piloting an open-source R Shiny app for automated outbreak detection by multiple NPHIs, this study provides an updated and extended overview of automated outbreak detection use in EU countries and the United Kingdom. The analysis encompassed (1) assessing current demand for automated outbreak detection, (2) examining the availability of necessary prerequisites of surveillance systems, and (3) identifying challenges and requirements for running AODSs.

Methods

Overview

A mixed methods approach was used, combining a survey, a workshop, and online meetings as part of the U4S project. The United Kingdom and France did not participate in the survey but were included as additional case studies because the 2010 overview and the workshop described in the Introduction section identified these countries as operating relevant AODSs [11,13]. In the 2010 overview, UK systems were reported separately for England, Wales, and Northern Ireland. Therefore, our analysis distinguishes these jurisdictions where relevant. Reporting of this work followed the SRQR (Standards for Reporting Qualitative Research) [15].

Questionnaire Development

A questionnaire was drafted with the goal of identifying gaps and needs from the perspective of epidemiology and data science. Respondents were asked to select one national surveillance system either in current use or considered most suitable for automated outbreak detection. The questionnaire had two sections: (1) description of public health surveillance systems, including types, legal frameworks, disease groups, data sources, aggregation level, reporting frequency, indicators, geographical resolution, and historical data availability; and (2) automated outbreak detection capacities, covering resource availability; signal follow-up; and details of existing methods, resources, and outputs.

The draft was reviewed by 6 project partners and a contact point at the European Centre for Disease Prevention and Control and revised accordingly. The final version (Multimedia Appendix 1) contained 33 questions, of which 22 were mandatory, primarily single- or multiple-choice rankings and free-text fields.

Survey Administration

The survey was created using the EUSurvey platform and distributed on April 17, 2023, by the U4S coordinator to designated contact points of the 25 countries participating in the project. These contact points were representatives of the participating NPHIs or partner institutions and served as the official coordinators for project activities within their respective countries. They were asked to coordinate questionnaire completion with relevant national experts, including epidemiologists and data scientists, where necessary, and submit one consolidated response describing the national surveillance system considered most suitable for automated outbreak detection.

Workshop

An on-site workshop was held in May 2023 at the Robert Koch Institute in Berlin, Germany, as part of the U4S project. Participants were invited from institutions participating in the U4S project, with active partners involved in the outbreak detection work package encouraged to attend in person. To ensure balanced representation across participating institutions, attendance was limited to a maximum of 2 representatives from active outbreak detection partners and 1 representative from other beneficiary institutions. In total, 30 public health experts, epidemiologists, and data scientists from NPHIs representing 16 EU countries participated. The workshop aimed to discuss survey findings, identify relevant use cases, and plan the pilot phase. Participating countries presented existing AODS and surveillance system overviews. Participants were divided into 2 breakout groups. Data scientists worked on defining outbreak detection tools achievable within the following 12 months, including requirements, vision, data input and output, and handling (operation, users, interfaces, and reporting). Epidemiologists, as future tool users, worked on outlining the ideal tool, specifying intended functionality, data sources, pathogens, use cases, and operational aspects.

XSL•FO
RenderX

One-on-One Meetings and Calls

In January 2024, structured one-on-one meetings were conducted with the 10 piloting countries participating in the tool implementation phase. Between April 2024 and November 2024, these countries piloted the outbreak detection tool by integrating their national surveillance data into the required input format, using the tool in routine surveillance, providing feedback, reporting bugs, suggesting additional requirements, and participating in the tool evaluation. Meetings included the national project representatives and technical experts involved in surveillance and data management. Discussions followed a predefined agenda covering country-specific surveillance structures, administrative divisions, data availability, and functional requirements for the outbreak detection tool. Additional informal online meetings conducted between 2023 and 2024 further supported the collection and clarification of tool requirements.

Data Analysis

Survey responses were exported from EUSurvey and analyzed using R (version 4.4.0; R Foundation for Statistical Computing) [16]. As this study aimed to describe the characteristics of national surveillance systems and automated outbreak detection capacities across participating countries, analyses were descriptive. Categorical variables are presented as counts and percentages where appropriate, whereas free-text responses were reviewed and summarized thematically to capture common approaches and country-specific characteristics.

Data Verification

Survey results were first presented during the workshop in May 2023, allowing participants to review and validate the data. During a project webinar in June 2023, the survey and workshop results were shared with the entire project consortium, giving all survey respondents the opportunity to verify their data.

Ethical Considerations

No formal ethics committee approval was sought because this study did not involve patients, clinical interventions, or the collection of sensitive personal data, and findings were analyzed and reported only in aggregated country-level form. The processing of participants' contact information was reviewed in consultation with the Robert Koch Institute's Data Protection Office and conducted in accordance with the General Data Protection Regulation (Regulation [EU] 2016/679) and the German Federal Data Protection Act (*Bundesdatenschutzgesetz*).

Participants provided electronic informed consent before completing the survey. Names and email addresses were used solely for study administration and were not included in the analysis. All results are presented in aggregated form, ensuring that individual participants cannot be identified.

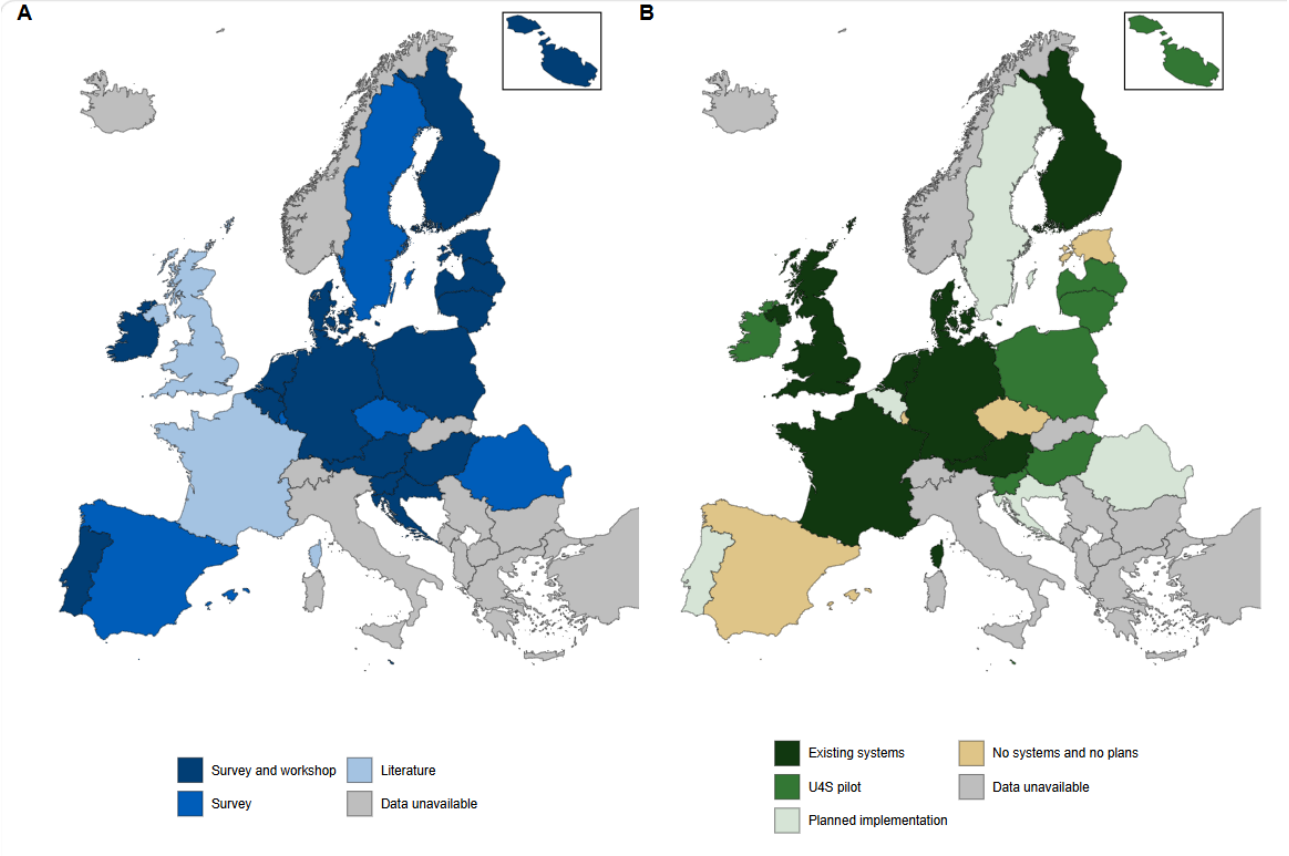
Results

Overview

The analysis included 26 countries, comprising 25 (96.2%) EU member states and the United Kingdom. Of the 25 U4S consortium members invited to participate in the survey, 21 (84%) completed it. France did not respond. The United Kingdom was not invited because it was not part of the U4S consortium. The largest group of countries (16/26, 61.5%) contributed data through the survey and participation in the workshop. A total of 19.2% (5/26) of the countries participated in the survey only; data for 7.7% (2/26) of the countries were obtained solely from the literature; and for 11.5% (3/26) of the countries, no data were available (Figure 1). Most survey respondents (19/21, 90.5%) were associated with NPHIs, and the remainder were associated with ministries of health (2/21, 9.5%).

XSL•FO
RenderX

Figure 1. Map showing the current status of automated outbreak detection systems (AODSs) in the European Union (EU) and the United Kingdom and the corresponding source of information (2024). (A) Source of information used to assess the status of AODSs across the EU and the United Kingdom. All countries that participated in the workshop also completed the survey. The upper right inset shows Malta. (B) Countries are categorized by their current level of implementation of AODSs. In cases in which countries meet the criteria for multiple categories (eg, Finland, Denmark, and the Netherlands, which have existing AODSs and are also UNITED4Surveillance [U4S] pilot partners), the higher level of implementation is displayed. "U4S pilot" refers to those countries that used the automated outbreak detection tool developed within U4S in a pilot phase from April 2024 to December 2024. The upper right inset shows Malta.



Current Use of AODSs in EU Countries

On the basis of the data collected from EU countries and the United Kingdom, 26.9% (7/26) of the countries had existing AODSs in place. A total of 38.5% (10/26) of the countries participated in the U4S pilot project, whereas 15.4% (4/26) were planning to implement such AODSs. Another 15.4% (4/26) had neither AODSs nor plans for implementation, and for 11.5% (3/26) of the countries, no data were available (Figure 1). Country-specific information on data sources and implementation status for each of the 26 countries is also provided in Multimedia Appendix 2.

Identified Need for Outbreak Detection

Beyond the current level of implementation, our survey examined how countries responded to signals generated from surveillance data, whether detected manually or automatically. While a diverse level of implementation across the EU was observed, all countries that completed the survey (21/25, 84%) stated that they would act on a signal identified through their surveillance data, whether from manual inspection or automated analysis. Typically, this response involved verification of the signal, followed by the initialization of an investigation, a risk assessment, further communication, and introduction of measures if needed. In most cases (14/21, 66.7%), the corresponding NPHI received feedback on actions triggered by a signal via outbreak reports (6/14, 42.9%), via direct communication with the investigation authorities (3/14, 21.4%), or by observing the performed public health action (2/14, 14.3%). In one case, investigations were conducted directly by the NPHI. In many countries (11/21, 52.4%), feedback on actions was collected systematically and often also used for evaluation (8/11, 72.7%). In addition, piloting countries grew from initially 7 to 10, reflecting the high interest in the tool.

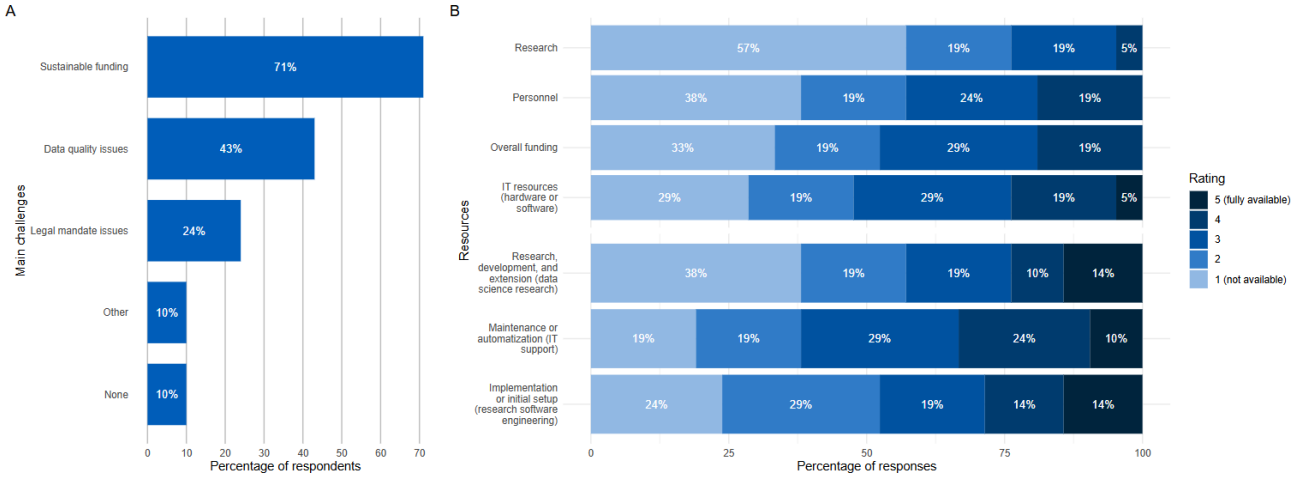
Identified Challenges to Implementing AODSs

When asked about main challenges in implementing AODSs, lack of sustainable funding and resources was indicated as the primary obstacle (15/21, 71.4%). This was followed by data quality (9/21, 42.9%) and legal mandate issues (5/21, 23.8%). In one country, data quality specifically referred to limited data availability and reporting delays. A total of 9.5% (2/21) of the countries reported no barriers, whereas another 9.5% (2/21) reported "other" issues. Of these countries, one respondent (Sweden) noted that an automated detection system had previously been implemented in the surveillance system but had not yet been reimplemented following a system upgrade in 2022. The other country (Lithuania) cited concerns about potential functional incompatibilities between systems, referring to technical issues that can cause delays or inconsistencies when data are synchronized across platforms before integration into the national eHealth system.

XSL•FO
RenderX

Regarding funding, respondents reported particularly limited funding for research capacity, specifically staff knowledge and time for research activities (12/21, 57.1%), personnel (8/21, 38.1%), overall funding (7/21, 33.3%), and IT resources such as hardware and software (6/21, 28.6%). Similarly, no funding was available for IT competences, including research software engineering for implementation (6/21, 28.6%), IT support for maintenance and automatization (6/21, 28.6%), and data science expertise for development and extension (8/21, 38.1%; Figure 2).

Figure 2. Reported challenges and resources for implementing automated outbreak detection systems (AODSs) in the European Union based on a survey of national public health institutes in 2023. (A) Main challenges reported by countries in implementing AODSs. Multiple answers per respondent were possible. (B) Aggregated ratings (1-5) of funding availability and IT competences to run and support implementation. Higher scores indicate greater perceived availability.



Surveillance System Prerequisites

Despite these implementation barriers, many countries already possessed surveillance infrastructures that provided favorable conditions for AODS implementation. The performance of AODSs primarily relies on the information available in the surveillance system, including historical data used to establish baselines, detect trends, and adjust for seasonality [5]. Almost all respondents stated that data were reported in a case-based format (20/21, 95.2%) and on a daily basis (19/21, 90.5%) at the national level. A total of 57.1% (12/21) stated that additional personally identifiable data (eg, addresses, date of birth, and personal identity number) were available. Furthermore, all countries (21/21, 100%) had at least 3 years of historic data available, and 38.1% (8/21) had historic data for 20 years or more.

All countries (21/21, 100%) received information on age, region, pathogen, and sex of infectious disease cases through their surveillance systems at the national level. These are the most common stratification dimensions in AODSs and allow for extensive investigation. Most countries also had additional information on outcomes such as death (19/21, 90.5%), hospitalization (17/21, 81%), and date of symptom onset (17/21, 81%) or specific symptoms (14/21, 66.7%), which would allow for further assessment of signals regarding the severity or even outbreak detection on different indicators (eg, number of hospitalizations). In many countries, information on vaccination status (15/21, 71.4%), risk factors (11/21, 52.4%), or exposures (15/21, 71.4%) was available, which would enable further disease-specific analyses.

Identified Use Cases for AODSs

Given these surveillance capacities, we next explored the disease areas and scenarios in which countries saw the greatest potential for AODSs. Identifying potential use cases for automated outbreak detection is essential for designing an effective and widely used AODS. This includes specifying data input formats, defining user-configurable input options such as filters and parameters, and ensuring that analysis and outputs are actionable and relevant for various disease surveillance systems [17].

The survey collected information on disease groups targeted by the surveillance systems, revealing potential use cases for AODSs and groups that could benefit the most from these systems. Food- and waterborne diseases (19/21, 90.5%), airborne diseases (18/21, 85.7%), and vector-borne and zoonotic diseases (17/21, 81%) were the most common disease groups. At the workshop, gastroenteritis and respiratory diseases emerged as the main use cases, with stronger interest in gastroenteritis. Accordingly, pilot use cases focused mainly on gastrointestinal pathogens, with one other country selecting group A streptococci.

Identified Requirements for AODSs

The identified requirements for AODSs drawn from one-on-one meetings and calls with the piloting countries showed significant overlap despite different organizational and geographical structures. Key needs included flexible output formats, adaptability for pandemic periods, and various filtering and stratification options. The systems should also provide information on cases related to already known outbreaks; enable extraction of detailed case line lists; support different temporal aggregations, such as daily, weekly, and monthly; and offer multiple language options. Several countries also highlighted the need to accommodate whole-genome sequencing data within future outbreak detection tools. The workshop discussion highlighted that signals were managed by different experts across countries with varying levels of IT expertise, underscoring the need for user-friendly, accessible, and

XSL•FO
RenderX

adaptable tools. These requirements highlight the need for a flexible system that can accommodate diverse user needs effectively.

Current AODSs Implemented by Countries

Table 1 presents an overview of current AODSs implemented in 26.9% (7/26) of the countries. The results are presented in 2 sections highlighting the differences in type and depth of information between countries that participated in the U4S project and France and the United Kingdom, for which only literature-based data were available.

XSL•FO
RenderX

Table . Characteristics of automated outbreak detection systems for infectious disease surveillance in 5 countries from the UNITED4Surveillance project (survey data) and France and the United Kingdom (literature; 2024).

Country	Pathogens and surveillance data	Methods	Implementation and automation	Reporting and frequency	Remarks
Austria	*Campylobacter* and *Salmonella*; case data from the national reference laboratory	Farrington Flexible algorithm and CUSUM[a]	R, *surveillance* package; fully automated	Weekly report	COVID-19–era hiring added data scientist capacity; plan to use data from the national electronic reporting system and extend to all notifiable diseases
Denmark	Approximately 20 assorted notifiable diseases, *Campylobacter*, and COVID-19; laboratory data from the national microbiology database	Threshold rules for approximately 20 diseases; Farrington Flexible algorithm; in-house method for COVID-19 prevalence levels	SQL+R+SAS; *surveillance* package; automation varies	Daily automatic email alerts (approximately 20 diseases); weekly ad hoc R reports or web figure	No single framework; evaluation pending; initiative to unify
Finland	Influenza; cases from Register of Primary Health Care Visits	MEM[b]	R, *mem* package; fully automated	Daily website updates on case numbers, thresholds, and epidemic levels	MEM since 2015; focused on influenza; plan to monitor other infectious diseases with data from the National Infectious Diseases Register
Germany	Approximately 30 diseases; national case reports from German surveillance database	Farrington Flexible algorithm	R pipeline, *surveillance* package; fully automated	Comprehensive weekly report (maps, charts, and tables)	Extensive coverage of notifiable pathogens
The Netherlands	Approximately 40 diseases; case-based data in the RIVM[c] database	HMM[d]+spline seasonality [18]	R Shiny dashboard (“survboard”); ad hoc computation during use of dashboard	Interactive dashboard; frequency dependent on use	Real-time geographic, temporal, and demographic views; plan to use WGS[e] data for outbreak detection with the new Shiny app
France (SurSaUD)—syndromic	Influenza+11 other syndromes; ED[f] and SOS Médecins visits+mortality (approximately 700 EDs and 60 GP[g] units); >90% ED coverage	Ensemble of GLMs[h], robust GLMs (weighted), and HMM [19]	Daily R scripts at Santé publique France; alarms fully automated; human validation of alert levels	Daily national dashboard and weekly reports	Performance benchmarked comparing ensemble vs individual methods and application to different data sources
United Kingdom—England (syndromic)	Comprehensive coverage of syndromes; NHS[i] 111 calls, GP consultations, and ED attendances	RAMMIE[j] [20]	Automated Stata script; daily	Daily internal alerts and public weekly bulletins	RAMMIE demonstrated high sensitivity (approximately 92%) and specificity (approximately 99%); adjusts for weekday and holiday effects
United Kingdom—England and Wales (laboratory isolates)	>3300 distinct organisms; HPA[k] surveillance data	Farrington Flexible algorithm (quasi-Poisson GLM)	R scripts (*surveillance* package); fully automated	Weekly reports emailed to national and regional epidemiologists	One of Europe’s oldest automated systems; continuously refined

[a]CUSUM: cumulative sum control chart.

[b]MEM: moving epidemic method.

[c]RIVM: National Institute for Public Health and the Environment.

[d]HMM: hidden Markov model.

[e]WGS: whole-genome sequencing.

[f]ED: emergency department.

[g]GP: general practitioner.

XSL•FO
RenderX

[h]GLM: generalized linear model.

[i]NHS: National Health Service.

[j]RAMMIE: rising-activity multilevel mixed-effects indicator emphasis.

[k]HPA: Health Protection Agency.

The 19.2% (5/26) of countries participating in the U4S project (Austria, Denmark, Finland, Germany, and the Netherlands) all used routinely collected notifiable infectious disease surveillance data and made use of R-based workflows, including the *surveillance* R package [21]. They differed, however, in the number and type of pathogens monitored, the specific outbreak detection method applied, the degree of automation, and the type of reporting. While some countries used their AODSs for individual pathogens, others covered a broad range of notifiable diseases. Despite this heterogeneity, all systems generated near–real-time outputs. Several countries were currently working to expand pathogen coverage, integrate national surveillance data, or harmonize their workflows.

France and the United Kingdom operated long-established R-based AODSs across diverse national data streams and used established R-based analytic frameworks. Both countries had comprehensive coverage of syndromes (France and the United Kingdom) or pathogens (United Kingdom) and automated R scripts run daily or weekly, producing internal alerts, dashboards, and bulletins. The systems in France and the United Kingdom have been established for many years; reflect a high level of methodological maturity, including systematic evaluations; and offer guidance for further development and application of AODSs in other countries.

Discussion

Principal Findings

This study provides an updated overview of the current landscape of AODSs in the EU and the United Kingdom by assessing current demand for automated outbreak detection, evaluating the availability of surveillance system prerequisites, and identifying barriers and requirements for implementation. Our findings reveal a heterogeneous state of implementation, with only 26.9% (7/26) of the countries included in the analysis operating AODSs. However, our results also highlight important opportunities for broader adoption, as surveillance systems in many countries already meet key data-related prerequisites for automated outbreak detection, including the availability of case-based, frequently reported, and sufficiently extensive historical surveillance data.

Compared to the 2010 overview, which outlined AODSs in Denmark, England, Germany, the Netherlands, and Sweden, our findings show that the overall landscape has changed little [11]. Most of these countries continued to operate AODSs, except Sweden, which discontinued its system following the implementation of a new surveillance system. Austria and Finland were the only countries that had introduced new systems: Austria's enhanced capacities during the COVID-19 pandemic may reflect increased prioritization of AODSs enabled by the additional resources made available during the pandemic. In contrast, Finland's adoption using the moving epidemic method focused more on retrospective trend analysis than on real-time signal detection. The findings also suggest that high-resource countries are more likely to establish outbreak detection systems but maintaining them can become challenging when substantial changes in surveillance systems occur. This limited progress appears to reflect persistent structural barriers rather than a lack of demand as respondents consistently identified sustainable funding, IT capacity, data quality, and legal mandates as the main obstacles to implementation.

Deficits in software engineering, IT support, and data science restrict the ability of NPHIs to develop and sustain systems. These constraints help explain why, despite clear interest, only a minority of countries currently operated AODSs. Similarly, the 15.4% (4/26) with neither an operational AODS nor plans for implementation reported barriers including limited funding, data quality, legal mandates, and IT capacity, with 2 also citing low political priority or a limited need due to a small number of notifications. At the same time, surveillance environments are evolving. An increasing number of countries are routinely generating whole-genome sequencing data as part of surveillance activities, which introduces new requirements for outbreak detection tools, particularly with regard to data input formats, interoperability, and the ability to integrate sequencing information alongside epidemiological data.

Our survey and pilot project further indicate growing momentum toward automation. In total, 38.5% (10/26) of the countries actively piloted the developed tool, demonstrating both the technical feasibility of implementation and the willingness of NPHIs to integrate methods into routine practices. In parallel, all survey respondents (21/21, 100%) confirmed that they would act on signals detected in their surveillance data. This finding indicates that the principle of responding to signals is firmly embedded in national public health practice. Structured procedures for response were in place in most countries (18/21, 85.7%), yet the way feedback was collected and integrated varied considerably. Reliance on manual verification may cause delays and inconsistencies, whereas greater automation could enable timelier signal identification, more consistent risk assessment, and systematic feedback.

Moreover, the results show significant potential for implementing new outbreak detection systems as many countries reported having extensive historical data, often exceeding 8 years, which is crucial for establishing baselines and trend analysis. Identifying specific use cases such as gastrointestinal pathogens highlighted the practical value of AODSs in routine surveillance tasks. Another potential application of new methods such as spatial clustering at the address level arises from the availability of certain personally identifiable data. The ability to incorporate such data allows for more precise mapping and detection of patterns.

Experience from pilot and established systems showed substantial alignment in outbreak detection needs across the EU. Despite differences in epidemiological contexts and

XSL•FO
RenderX

organizational structures, system requirements showed considerable overlap. Pilot countries highlighted needs such as flexible output formats, temporal aggregation, filtering, and multilingual support, whereas established systems used different statistical methods and infrastructures yet generated largely comparable outputs. The convergence of outputs despite methodological differences suggests that key outputs, including signals, visualizations, maps, and reports, could be standardized. These shared requirements indicate that a single flexible tool could accommodate diverse national contexts, meeting practical needs without requiring separate systems for each country. The widespread use of the *surveillance* package is a good example of the utility of open-source tools developed by the public health community.

Limitations

Our study has some limitations. First, the survey focused on one surveillance system per country, which might not fully capture the use of AODSs across all national surveillance systems. Second, varying interpretations of automated outbreak detection could have influenced reported use, including differences between signal detection and data visualization. Third, reliance on self-reported data may have introduced bias or overestimation, with response depth varying by how extensively experts were consulted. Fourth, workshop participation was limited to 30 experts from 16 countries, not representing all EU contexts. Fifth, while the systems in the United Kingdom and France were still operational, the data may not reflect all subsequent updates. Sixth, saturation was not formally assessed or used as a criterion for determining the extent of data collection. Countries were predefined by their involvement in their respective U4S project activities. The qualitative components were used to complement and contextualize the country-level survey findings and further explore requirements for AODSs. Despite these limitations, this study provides valuable insights into existing practices and priority areas for improvement.

Future Perspectives and Conclusions

AODSs offer the potential to enhance routine surveillance of infectious diseases across EU countries. However, the results of this study indicate that resource limitations and technical barriers continue to hinder their wider implementation. Broader adoption of flexible, open-source solutions could accelerate cross-country collaboration and promote more harmonized outbreak detection in the EU. The experience across countries with different surveillance capacities further suggests that AODSs can be adapted to diverse settings, although their implementation in more resource-limited contexts will depend on the availability of reliable surveillance data, digital infrastructure, and sustained technical support.

These findings have broader implications for infectious disease surveillance within the EU. Increasing harmonization of outbreak detection approaches could facilitate knowledge exchange between countries, support more comparable surveillance practices, and reduce duplication of effort through the shared development and maintenance of open-source tools. Collaborative initiatives such as U4S may help bridge existing implementation gaps by bringing together expertise, building technical capacity, and fostering collaboration across countries.

Future work should focus on integrating the R Shiny app within the U4S project into local systems; expanding its use beyond the project phase; and refining methods, interfaces, and data integration. Broader adoption of such flexible, open-source solutions could accelerate cross-country collaboration and promote more harmonized outbreak detection in the EU. As automated systems can generate large numbers of signals with varying specificity, systematic evaluation in routine public health practice will be essential to assess their accuracy, usability, workload implications, and overall added value for infectious disease surveillance.

Acknowledgments

The authors sincerely thank the UNITED4Surveillance consortium for their participation in the survey and acknowledge their project partners at the Austrian Agency for Health and Food Safety; Statens Serum Institut, Denmark; the Finnish Institute for Health and Welfare; the National Institute for Public Health and the Environment, the Netherlands; the Infectious Disease Prevention and Control Unit at the Health Promotion and Disease Prevention Directorate, Malta; the Health Promotion and Disease Prevention Directorate, National Public Health Center, Hungary; the National Institute of Public Health, Slovenia; the Latvian Centre for Disease Prevention and Control; the Health Protection Surveillance Centre, Ireland; the National Centre for Communicable Diseases, Lithuania; and the Directorate-General of Health, Portugal, for their collaboration, expertise, and ongoing support throughout the project. The authors used ChatGPT (GPT-5.2) from OpenAI (2026) under full human supervision to assist with linguistic editing of the manuscript, refinement of the code used to generate figures, and support the preparation of a summary table. All AI-generated outputs were reviewed and verified by the authors, who take full responsibility for the final content of the manuscript.

Funding

The research activities reported in this paper were conducted as part of the UNITED4Surveillance project cofunded by the European Union under the EU4Health program (project ID 101102070). The positions of PvFgI and ACV were funded through this project. The funder had no role in the design or conduct of the research, analysis or interpretation of the data, preparation of the manuscript, or decision to submit the manuscript for publication. The views and opinions expressed are, however, those of

XSL•FO
RenderX

the authors only and do not necessarily reflect those of the European Union or granting authority. Neither the European Union nor the granting authority can be held responsible for them.

Data Availability

The survey data generated and analyzed during this study are not publicly available because participants did not provide explicit consent for public data sharing at the time of enrollment. Requests for access to data may be directed to the corresponding author and will be considered in accordance with the study's ethics approval, applicable data protection regulations, and institutional requirements.

Authors' Contributions

Conceptualization: PvFgI, IM, AU
Data curation: PvFgI, ACV, IM, AU
Formal analysis: PvFgI, ACV, AU
Funding acquisition: IM, MD, AU
Methodology: IM, AU
Visualization: PvFgI, ACV, AU
Writing—original draft: PvFgI, ACV, AU
Writing—review and editing: PvFgI, ACV, IM, MD, AU

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey on needs and gaps analysis of automated outbreak detection tools for routine surveillance data.
[PDF File, 104 KB - publichealth_v12i1e96576_app1.pdf]

Multimedia Appendix 2
Overview of country-specific source of information and automated outbreak detection system implementation status across 26 countries, 2024.
[DOCX File, 20 KB - publichealth_v12i1e96576_app2.docx]

Checklist 1
SRQR checklist.
[PDF File, 481 KB - publichealth_v12i1e96576_app3.pdf]

References

1. Yuan M, Boston-Fisher N, Luo Y, Verma A, Buckeridge DL. A systematic review of aberration detection algorithms used in public health surveillance. J Biomed Inform 2019 Jun;94:103181. [doi: 10.1016/j.jbi.2019.103181] [Medline: 31014979]
2. Steele L, Orefuwa E, Dickmann P. Drivers of earlier infectious disease outbreak detection: a systematic literature review. Int J Infect Dis 2016 Dec;53:15-20. [doi: 10.1016/j.ijid.2016.10.005] [Medline: 27777092]
3. Buehler JW, Hopkins RS, Overhage JM, Sosin DM, Tong V, CDC Working Group. Framework for evaluating public health surveillance systems for early detection of outbreaks: recommendations from the CDC Working Group. MMWR Recomm Rep 2004 May 7;53(RR-5):1-11. [Medline: 15129191]
4. Branch-Elliman W, Sundermann AJ, Wiens J, Shenoy ES. The future of automated infection detection: innovation to transform practice (part III/III). Antimicrob Steward Healthc Epidemiol 2023;3(1):e26. [doi: 10.1017/ash.2022.333] [Medline: 36865708]
5. Zareie B, Poorolajal J, Roshani A, Karami M. Outbreak detection algorithms based on generalized linear model: a review with new practical examples. BMC Med Res Methodol 2023 Oct 14;23(1):235. [doi: 10.1186/s12874-023-02050-z] [Medline: 37838735]
6. Page ES. Continuous inspection schemes. Biometrika 1954;41(1-2):100-115. [doi: 10.1093/biomet/41.1-2.100]
7. Hutwagner L, Thompson W, Seeman GM, Treadwell T. The bioterrorism preparedness and response Early Aberration Reporting System (EARS). J Urban Health 2003 Jun;80(2 Suppl 1):i89-i96. [doi: 10.1007/pl00022319] [Medline: 12791783]
8. Unkel S, Farrington CP, Garthwaite PH, Robertson C, Andrews N. Statistical methods for the prospective detection of infectious disease outbreaks: a review. J R Stat Soc Ser A Stat Soc 2012 Jan;175(1):49-82. [doi: 10.1111/j.1467-985X.2011.00714.x]
9. Farrington CP, Andrews NJ, Beale AD, Catchpole MA. A statistical algorithm for the early detection of outbreaks of infectious disease. J R Stat Soc Ser A Stat Soc 1996;159(3):547. [doi: 10.2307/2983331]

XSL•FO
RenderX

10. Noufaily A, Enki DG, Farrington P, Garthwaite P, Andrews N, Charlett A. An improved algorithm for outbreak detection in multiple surveillance systems. Stat Med 2013 Mar 30;32(7):1206-1222. [doi: 10.1002/sim.5595] [Medline: 22941770]
11. Hulth A, Andrews N, Ethelberg S, et al. Practical usage of computer-supported outbreak detection in five European countries. Euro Surveill 2010 Sep 9;15(36):19658. [Medline: 20843470]
12. Salmon M, Schumacher D, Burmann H, Frank C, Claus H, Höhle M. A system for automated outbreak detection of communicable diseases in Germany. Euro Surveill 2016;21(13). [doi: 10.2807/1560-7917.ES.2016.21.13.30180] [Medline: 27063588]
13. Ghozzi S, Ullrich A. Bericht zum workshop “automatic detection of infectious disease outbreaks” [Article in German]. Robert Koch Institute. 2017. URL: https://www.rki.de/DE/Themen/Infektionskrankheiten/Meldewesen/Signale/Aktivitaeten/OutbreakDetectionWorkshop2017_Bericht.pdf?__blob=publicationFile&v=1 [accessed 2026-08-30]
14. UNITED4Surveillance. URL: https://united4surveillance.eu/ [accessed 2025-11-16]
15. O’Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251. [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
16. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2024. URL: https://www.R-project.org/ [accessed 2026-08-30]
17. Buckeridge DL. Outbreak detection through automated surveillance: a review of the determinants of detection. J Biomed Inform 2007 Aug;40(4):370-379. [doi: 10.1016/j.jbi.2006.09.003] [Medline: 17095301]
18. Vega T, Lozano JE, Meerhoff T, et al. Influenza surveillance in Europe: establishing epidemic thresholds by the moving epidemic method. Influenza Other Respir Viruses 2013 Jul;7(4):546-558. [doi: 10.1111/j.1750-2659.2012.00422.x] [Medline: 22897919]
19. Pelat C, Bonmarin I, Ruello M, et al. Improving regional influenza surveillance through a combination of automated outbreak detection methods: the 2015/16 season in France. Euro Surveill 2017 Aug 10;22(32):30593. [doi: 10.2807/1560-7917.ES.2017.22.32.30593] [Medline: 28816649]
20. Morbey RA, Elliot AJ, Charlett A, Verlander NQ, Andrews N, Smith GE. The application of a novel “rising activity, multi-level mixed effects, indicator emphasis” (RAMMIE) method for syndromic surveillance in England. Bioinformatics 2015 Nov 15;31(22):3660-3665. [doi: 10.1093/bioinformatics/btv418] [Medline: 26198105]
21. Salmon M, Schumacher D, Höhle M. Monitoring count time series in R: aberration detection in public health surveillance. J Stat Softw 2016;70(10):1-35. [doi: 10.18637/jss.v070.i10]

Abbreviations

AODS: automated outbreak detection system
EU: European Union
NPHI: national public health institute
SRQR: Standards for Reporting Qualitative Research
U4S: UNITED4Surveillance

Edited by A Mavragani, T Sanchez; submitted 30.Mar.2026; peer-reviewed by A Noufaily, JY Carter; revised version received 18.Aug.2026; accepted 19.Aug.2026; published 15.Sep.2026.

Please cite as:
vom Felde genannt Imbusch P, Vietor AC, Markus I, Diercke M, Ullrich A
Requirements and Implementation Challenges of Automated Outbreak Detection Systems in the European Union: Mixed Methods Study
JMIR Public Health Surveill 2026;12:e96576
URL: https://publichealth.jmir.org/2026/1/e96576
doi:10.2196/96576

© Paulina vom Felde genannt Imbusch, Ann Christin Vietor, Inessa Markus, Michaela Diercke, Alexander Ullrich. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 15.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Barcelona Injury Surveillance System (BISS) for Safer Cities: Observational, Descriptive Study Using Routine Health and Police Information Databases

Katherine Pérez[1,2,3], PhD; Elena Santamariña-Rubio[1], PhD; Mònica Cortés-Albaladejo[1,3], MPH; Lucía Artazcoz[1,2,3,4], PhD; Adnan A Hyder[5], Prof Dr; Carme Borrell[1,2,3], PhD

[1]Agència de Salut Pública de Barcelona (ASPB), Pl Lesseps, 1, Barcelona, Spain
[2]CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, Spain
[3]Institut de Recerca de Sant Pau (IR SANT PAU), Barcelona, Spain
[4]Department of Experimental and Health Sciences, Universitat Pompeu Fabra, Barcelona, Spain
[5]School of Public Health, Boston University, Boston, MA, United States

Corresponding Author:
Katherine Pérez, PhD
Agència de Salut Pública de Barcelona (ASPB), Pl Lesseps, 1, Barcelona, Spain

Abstract

Background: Injuries are a major cause of death and disability, but cities often lack surveillance systems that can monitor injury burden across mechanisms, severity levels, and population groups. In Spain, no comprehensive city-level injury surveillance system routinely captures the full spectrum of injuries. The Barcelona Injury Surveillance System (BISS) was developed to address this gap by integrating routine health and police data.

Objective: This study aims to describe the BISS, including its scope, data sources, and public health rationale, and illustrate its utility through the analysis of recent injury data in Barcelona.

Methods: We conducted a descriptive study using routinely collected emergency department, hospital discharge, mortality, and police data integrated into the BISS. We analyzed nonfatal injuries in 2024, fatal injuries in 2023, and trends from 2018 onward. Injury indicators were examined by sex, age, mechanism, type, and severity. Crude and age-adjusted rates per 100,000 residents were calculated.

Results: In 2024, BISS recorded 123,420 emergency department injury episodes and 18,749 injury-related hospitalizations; among residents, these figures were 99,379 and 14,319, respectively. In 2023, 695 injury-related deaths were recorded among residents. Nonfatal injuries were slightly more frequent in females, especially at older ages, whereas injury-related mortality was higher in males. Falls were the leading specified mechanism and were concentrated among older females, particularly those aged ≥75 years. Self-harm hospitalization rates were highest among females aged 15 years to 24 years, whereas self-harm mortality was higher in males. Road traffic injury and overall mortality rates were also higher in males. From 2018 to 2024, most nonfatal injury indicators increased after the decline observed in 2020, while road traffic injuries declined overall.

Conclusions: BISS demonstrates the value of integrating routine health and police data to generate actionable urban injury intelligence. The findings highlight priorities for prevention, particularly falls in older females, self-harm in young females, and the persistently higher fatal injury burden among males. Integrated city-level surveillance systems such as BISS can support monitoring, equity-oriented prevention, and data-informed public health policy.

(JMIR Public Health Surveill 2026;12:e82079) doi:10.2196/82079

KEYWORDS

injury; surveillance; monitoring; inequalities; evaluation; injury severity

Introduction

Injuries are a leading cause of death and disability worldwide and remain a major public health challenge across the life course [1,2]. Beyond fatalities, millions of nonfatal injuries require care in emergency departments (EDs), hospitals, and other health services, often resulting in long-term physical, psychological, and social consequences [2]. Injury burden is not evenly distributed: Marked differences are observed by age, sex, and socioeconomic status as well as by injury mechanism. Young men are disproportionately affected by road traffic injuries and interpersonal violence, whereas falls account for a substantial and increasing burden among older adults, particularly women [1,2]. These patterns underscore that injuries are not random

XSL•FO
RenderX

events but largely predictable and preventable outcomes shaped by social, environmental, and policy contexts [3].

Effective prevention requires timely, valid, and policy-relevant information on the frequency, characteristics, mechanisms, and severity of injuries as well as on the population groups most affected. Public health surveillance has been defined as the continuous and systematic collection, analysis, and interpretation of health-related data for the planning, implementation, and evaluation of public health practice [4]. Applied to injuries, surveillance systems can support prevention by identifying priorities, monitoring trends, detecting inequalities, and informing the design and evaluation of interventions [5,6]. Because injury risks and consequences differ by sex and intersect with other axes of inequality, injury surveillance should also support an equity-oriented perspective [5,6].

International organizations have long emphasized the importance of standardized injury surveillance. The World Health Organization (WHO) and the US Centers for Disease Control and Prevention (CDC) developed guidance on the establishment, coding, and use of injury surveillance systems [6,7]. In Europe, injury surveillance has also been promoted through harmonized initiatives such as the European Injury Database, which supports the production of comparable indicators across countries [8-10]. Together, these initiatives have highlighted the value of routine, standardized data from health services for understanding injury burden and guiding prevention policies.

In Spain, however, there is no comprehensive injury surveillance system routinely monitoring the full spectrum of injuries at the national, regional, or city level. Existing systems have been limited in scope, with the most established example being the Barcelona Traffic Injury Information System, which has monitored road traffic injuries using multiple sources since the late 1990s [11], or fatal injuries or injury discharges at the national level reported via Eurostat [10]. At the same time, other routine data sources—such as ED records, hospital discharge data, mortality registers, and police records—offer an important but underused opportunity to build broader injury surveillance capacity particularly at the local level, where prevention can be implemented. Integrating these sources is essential to moving beyond a fragmented understanding of injuries and toward a more complete picture of their magnitude, mechanisms, severity, and social patterning.

The Barcelona Injury Surveillance System (BISS) was developed by the Public Health Agency of Barcelona to address this gap. BISS brings together existing routine health and police data to support the epidemiological surveillance of injuries in the city, including road traffic injuries, falls, violence, self-harm, and other external causes treated in health services or recorded in mortality data. By generating standardized indicators and enabling analyses by sex, age, injury mechanism, severity, and other axes of inequality, BISS aims to support priority setting, prevention planning, monitoring, and evaluation from an urban and equity-oriented public health perspective.

This study aimed to describe the BISS, including its scope, data sources, and public health rationale, and illustrate its utility through the analysis of recent injury data in Barcelona.

Methods

Study Design

This observational, descriptive study of injury epidemiology used routinely collected health data in Barcelona integrated into a city-wide injury surveillance system. The BISS compiles and analyzes data on injuries from multiple existing databases (health, administrative, and police records) in a continuous manner. The study design was cross-sectional and retrospective, using pre-existing data. Key elements of the design include the integration of an established road traffic injury registry (the Barcelona Traffic Injury Information System, now a subsystem called BISS-RTI for road traffic injuries) into a broader all-injury surveillance framework. The analysis presented focused on injury data from the year 2024 for emergencies and hospitalizations, 2023 for mortality, and 2018 onward for trends.

Setting

The study was set in the city of Barcelona, Spain, an urban area with approximately 1.6 million residents within a larger metropolitan region. The BISS covers injuries occurring among all persons in Barcelona, including residents, workers, visitors, and commuters within the city. We report on injuries that occurred or were treated in Barcelona during 2024 (with trend analyses from 2018 onward), encompassing incidents across the entire city.

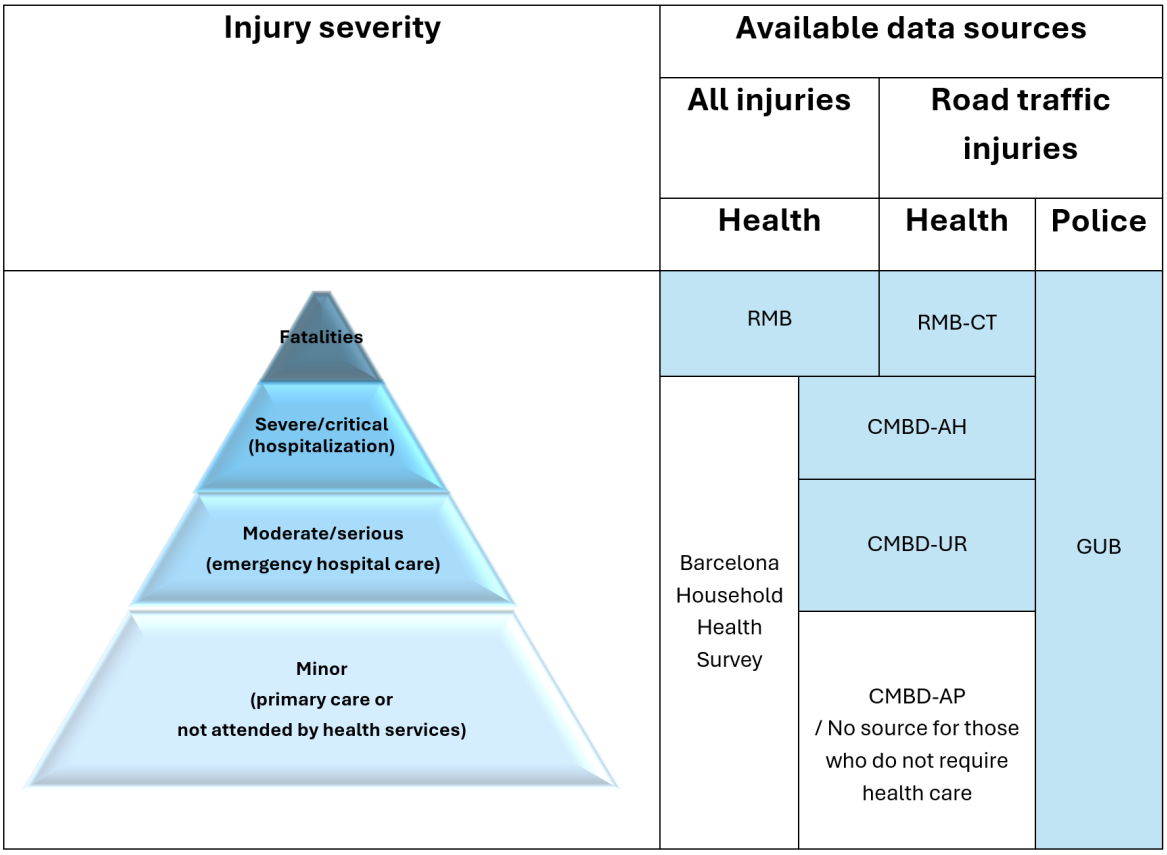
Participants and Study Population

The BISS does not enroll participants directly but instead includes all injury cases meeting specific criteria from each data source. The target population for surveillance is everyone who lives, works, or spends time in Barcelona—effectively the entire population at risk in the city. We included injury episodes (events) rather than unique individuals because one person may suffer multiple injuries over time.

We provide a visual overview of the data and inclusion process: Figure 1 illustrates the “injury pyramid” for Barcelona, showing all levels of injury severity, the available data sources at each level, and the portion covered by BISS multiple (not linked) databases (shaded in the figure). At the base, there are minor injuries, which are the most common and are either treated in primary care or do not require medical attention. The next level consists of moderate to severe injuries, which are typically treated in EDs. Serious or critical injuries, although less common, require emergency medical attention and frequently lead to hospital admission for further treatment. At the top of the pyramid are fatal injuries, some of which are treated in health services (emergency care and hospitalization), while others result in death at the site of the injury before medical help can be provided.

XSL•FO
RenderX

Figure 1. Pyramidal distribution of injuries due to external causes according to severity and available data sources for all injuries and specifically for traffic collision injuries; blue shading shows the coverage of the Barcelona Injury Surveillance System (BISS). CMBD: Register of the Minimum Basic Data Set in Spain; CMBD-AH: CMBD of hospital discharges; CMBD-AP: CMBD of primary care visits; CMBD-UR: CMBD of hospital emergency department visits; GUB: Register of Accidents and Victims by the Barcelona City Police, which does not include information on the nature and severity of injuries; RMB: Barcelona Mortality Register; RMB-CT: Barcelona Traffic Collision Mortality Register.



BISS covers moderate to fatal injuries by leveraging existing routine data sources in Barcelona, primarily health-related records that document injury diagnoses and mechanisms but also health surveys (no data included here). It includes moderate to critical injuries due to external causes using data from the Register of the Minimum Basic Data Set in Spain (CMBD; acronym in Catalan) for hospital emergencies and discharges. Additionally, fatalities resulting from external causes among Barcelona residents are recorded through the Barcelona Mortality Register (RMB; acronym in Catalan). Currently, milder injuries are not included, despite the availability of CMBD data for primary care, as the quality of injury coding from this source has not yet been assessed. However, this may be considered in future phases of the system.

In the specific case of BISS-RTI, in addition to moderate to critical injuries from the CMBD of emergencies and hospital discharges, there are 2 sources for traffic injuries: the Barcelona Traffic Collision Mortality Register (RMB-CT; acronym in Catalan), which provides information on all individuals who have died from traffic-related incidents in Barcelona including injury diagnoses, and the Road Traffic Collisions Register from the Municipal Police (GUB; acronym in Catalan), which is an exhaustive record of all traffic collisions in Barcelona and the individuals injured, ranging from minor to fatal. Although the GUB does not include injury diagnoses, it provides detailed information on the circumstances of each collision.

Eligibility criteria were defined by injury mechanism and severity using diagnostic codes and external cause codes. In general, BISS covers injuries requiring medical attendance up to fatalities, while minor injuries not requiring medical attention (or only primary care) are not yet captured. Specifically, to be included in BISS analyses, a case had to satisfy the data inclusion criteria for one of the BISS databases as outlined in Table 1. Table 1 also details the algorithms and codes used to identify eligible injury records from each source, such as *International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM*) external cause codes indicating injuries due to external causes, and location filters (eg, resident of Barcelona for mortality data). When there was an injury-related fatality at the ED or during hospitalization, records were classified according to the outcome (fatal vs nonfatal) to avoid double-counting deaths. In contrast, ED and hospitalization databases were not linked at the person level due to the lack of a personal identifier; consequently, some cases initially treated in the ED and later hospitalized may be counted in both sources.

XSL•FO
RenderX

Table . Barcelona Injury Surveillance System (BISS) databases and their sources with the inclusion criteria as well as injury subset selection and defined indicators for monitoring each injury subset.

Data source	Database inclusion criteria	Database	Definition criteria for the injury subset	Injury subset	Indicators
CMBD_UR[a]	• Hospital in the city of Barcelona, regardless of the municipality of residence of the injured person • Any of the following *ICD-10-CM*[b] codes in any of the diagnostic fields: S00-S99, T07, T14, T30-T32, T15-T28, T33-T34, T51-T76, T79, T36-T50 with a 6th digit equal to 1, 2, 3, or 4 (from codes T36.9, T37.9, T39.9, T41.4, T42.7, T43.9, T45.9, T47.9, and T49.9, including those in the 5th digit with 1, 2, 3, or 4), ending with the letter A, B, or C ("initial contact") or without a letter OR any of the following *ICD-10-CM* codes in any of the external cause fields: V00-V99, W00-X58, X71-X83, X92-Y09, Y21-Y33, Y35-Y38 • Assistance in hospital emergency departments	ED[c]	• Exclude deaths	A_Episodes of nonfatal injuries treated in the hospital emergency departments of Barcelona, including residents and nonresidents of Barcelona	• 1 Nonfatal Emergency Department Visits for All Injuries • 2 Nonfatal Emergency Department Visits for Severe Injuries • 3 Nonfatal Assault-Related Emergency Department Visits • 4 Nonfatal Intentional Self-Harm Emergency Department Visits • 5 Nonfatal Unintentional Fall-Related Injury Emergency Department Visits • 6 Nonfatal Unintentional Motor Vehicle Traffic-Related Emergency Department Visits • 7 Nonfatal Unintentional Fire-Related Emergency Department Visits • 8 Nonfatal Unintentional Nondrug Poisoning Emergency Department Visits • 9 Nonfatal Unintentional Drug Poisoning Emergency Department Visits • 10 Nonfatal TBI[d] Emergency Department Visits

XSL•FO
RenderX

Data source	Database inclusion criteria	Database	Definition criteria for the injury subset	Injury subset	Indicators
CMBD_UR	• Hospital in the city of Barcelona, regardless of the municipality of residence of the injured person • Any of the following *ICD-10-CM* codes in any of the diagnostic fields: S00-S99, T07, T14, T30-T32, T15-T28, T33-T34, T51-T76, T79, T36-T50 with a 6th digit equal to 1, 2, 3, or 4 (from codes T36.9, T37.9, T39.9, T41.4, T42.7, T43.9, T45.9, T47.9, and T49.9, including those in the 5th digit with 1, 2, 3, or 4), ending with the letter A, B, or C ("initial contact") or without a letter OR any of the following *ICD-10-CM* codes in any of the external cause fields: V00-V99, W00-X58, X71-X83, X92-Y09, Y21-Y33, Y35-Y38 • Assistance in hospital emergency departments	ED	• Any of the following codes in any of the external cause fields: *ICD-10-CM* codes for road traffic injuries: V00.0, V00.1, V00.8, except V00.81, V00.82; V01-V06.1/.9; V09.2/.3/.9, V10-V18/V20-V28.3/.4/.5/.9; V29-V30/V35-V38/V40-V48/V50-V58/V60-V68/V70-V78.4/.5/.6/.7/.9; V81.1, V82.1/.9, V83-V86.0/.1/.2/.3/.4, V87, V89.2/.3/.9 ending with the letter A, B, or C ("initial contact") or without a letter OR an economic financing regime that appears in the CMBD as mutual society for traffic accidents • Exclude deaths	B_Episodes of nonfatal injuries due to traffic collisions treated in the hospital emergency departments of Barcelona, including residents and nonresidents of Barcelona	• 23 Nonfatal Emergency Department Visits for Road Traffic–Related Injuries • 24 Nonfatal Emergency Department Visits for Road Traffic–Related Severe Injuries • 25 Nonfatal Emergency Department Visits for Road Traffic–Related Pedestrian Injuries • 26 Nonfatal Emergency Department Visits for Road Traffic–Related Pedestrian Transport User Injuries • 27 Nonfatal Emergency Department Visits for Road Traffic–Related Cyclist Injuries • 28 Nonfatal Emergency Department Visits for Road Traffic–Related Motorcyclist Injuries • 29 Nonfatal Emergency Department Visits for Road Traffic–Related Car User Injuries • 30 Nonfatal Emergency Department Visits for Road Traffic–Related TBI
CMBD_AH[e]		HOSP[f]	• Exclude deaths	C_Episodes of nonfatal injury discharged by admission to a hospital in Barcelona, including residents and nonresidents of Barcelona	

XSL•FO
RenderX

Data source	Database inclusion criteria	Database	Definition criteria for the injury subset	Injury subset	Indicators
	• Hospital in the city of Barcelona, regardless of the municipality of residence of the injured person • Any of the following *ICD-10-CM* codes in any of the diagnostic fields: S00-S99, T07, T14, T30-T32, T15-T28, T33-T34, T51-T76, T79, T36-T50 with 6th digit equal to 1, 2, 3, or 4 (of the codes T36.9, T37.9, T39.9, T41.4, T42.7, T43.9, T45.9, T47.9, and T49.9, those in which the 5th digit is 1, 2, 3, or 4), ending with the letter A, B, or C (“initial contact”) or without a letter • Urgent health care (unscheduled)				• 11 Nonfatal Hospitalizations for All Injuries • 12 Nonfatal Hospitalizations for Severe Injuries • 13 Nonfatal Assault-Related Hospitalizations • 14 Nonfatal Intentional Self-Harm Hospitalizations • 15 Nonfatal Unintentional Fall-Related Injury Hospitalization • 16 Nonfatal Unintentional Motor Vehicle Traffic-Related Hospitalizations • 17 Nonfatal Unintentional Fire-Related Injury Hospitalizations • 18 Nonfatal Unintentional Nondrug Poisoning Hospitalizations • 19 Nonfatal Unintentional Drug Poisoning Hospitalizations • 20 Nonfatal TBI Hospitalizations • 21 Nonfatal Hospitalizations for Long Bone Fracture • 22 Nonfatal Fall-Related Hip Fracture Hospitalizations in Persons ≥65 Years Old
CMBD_AH		HOSP		D_Episodes of nonfatal injuries due to traffic collisions discharged due to admission to a hospital in Barcelona, including residents and nonresidents of Barcelona	

XSL•FO
RenderX

Data source	Database inclusion criteria	Database	Definition criteria for the injury subset	Injury subset	Indicators
	• Hospital in the city of Barcelona, regardless of the municipality of residence of the injured person • Any of the following *ICD-10-CM* codes in any of the diagnostic fields: S00-S99, T07, T14, T30-T32, T15-T28, T33-T34, T51-T76, T79, T36-T50 with 6th digit equal to 1, 2, 3, or 4 (of the codes T36.9, T37.9, T39.9, T41.4, T42.7, T43.9, T45.9, T47.9, and T49.9, those in which the 5th digit is 1, 2, 3, or 4), ending with the letter A, B, or C (“initial contact”) or without a letter • Urgent health care (unscheduled)		• Any of the following codes in any of the external cause fields: *ICD-10-CM* codes for road traffic injuries: V00.0, V00.1, V00.8, except V00.81, V00.82; V01-V06.1/.9; V09.2/.3/.9, V10-V18/V20-V28.3/.4/.5/.9; V29-V38/V40-V48/V50-V58/V60-V68/V70-V78.4/.5/.6/.7/.9; V81.1, V82.1/.9, V83-V86.0/.1/.2/.3/.4, V87, V89.2/.3/.9 ending with the letter A, B, or C (“initial contact”) or without a letter OR an economic financing regime that appears in the CMBD as mutual society for traffic accidents • Exclude deaths		• 31 Nonfatal Hospitalizations for Road Traffic–Related Injuries • 32 Nonfatal Hospitalizations for Road Traffic–Related Severe Injuries • 33 Nonfatal Hospitalizations for Road Traffic–Related Pedestrian Injuries • 34 Nonfatal Hospitalizations for Road Traffic–Related Pedestrian Transport User Injuries • 35 Nonfatal Hospitalizations for Road Traffic–Related Cyclist Injuries • 36 Nonfatal Hospitalizations for Road Traffic–Related Motorcyclist Injuries • 37 Nonfatal Hospitalizations for Road Traffic–Related Car User Injuries • 38 Nonfatal Hospitalizations for Road Traffic–Related TBI
RMB[g]	• Any of the following *ICD-10* codes in the underlying cause of death field: V01–Y36, Y85–Y87, Y89, U01–U03 to the CBD[h]	MORTALITY[i]	• Municipality of death: Barcelona	E_People residing in Barcelona who died from injuries due to external causes in Barcelona	• 39 Injury Fatalities • 40 Unintentional Drowning Fatalities • 41 Unintentional Fall-Related Fatalities • 42 Unintentional Fire-Related Fatalities • 43 Unintentional Motor Vehicle Traffic Fatalities • 44 Unintentional Nondrug Poisoning Fatalities • 45 Unintentional Drug Overdose Fatalities • 46 Unintentional Suffocation Fatalities • 47 Suicides • 48 Homicides

XSL•FO
RenderX

Data source	Database inclusion criteria	Database	Definition criteria for the injury subset	Injury subset	Indicators
RMB_CT[j]	• Death in Barcelona due to traffic collision	MORTALITY_RTI[k]	• Place of collision: Barcelona	F_Residents and nonresidents of Barcelona who died in Barcelona due to road traffic injuries that occurred in Barcelona	• 49 Road Traffic–Related Injury Fatalities • 50 Pedestrian Road Traffic–Related Injury Fatalities • 51 Cyclist Road Traffic–Related Injury Fatalities • 52 Motorcyclist Road Traffic–Related Injury Fatalities • 53 TBI Road Traffic–Related Injury Fatalities
GUB[l]	• People involved in a road traffic collision who have been injured according to police criteria (who have required some type of medical assistance)	INJURED_RTI[m]	• For people injured in the same collision, information about the collision summarized in a single record	G_Road traffic collisions in Barcelona in which a person, resident or nonresident of Barcelona, has been injured (required some type of medical assistance)	• 54 Road Traffic–Related Injury Collisions
GUB	• People involved in a road traffic collision who have been injured according to police criteria (who have required some type of medical assistance)	INJURED_RTI	• People injured	H_Residents and nonresidents of Barcelona injured (required some type of medical assistance) due to a traffic collision in Barcelona	• 55 Road Traffic–Injured People • 56 Road Traffic–Injured Pedestrian • 57 Road Traffic–Injured Pedestrian Transport Users • 58 Road Traffic–Injured Motorcyclist • 59 Road Traffic–Injured Cyclist • 60 Road Traffic–Injured Car User

[a]CMBD_UR: register of the Minimum Basic Data Set of hospital emergency department visits.

[b]*ICD-10-CM*: *International Classification of Diseases, Tenth Revision, Clinical Modification*.

[c]Emergency department visit data.

[d]TBI: traumatic brain injury.

[e]CMBD_AH: register of the Minimum Basic Data Set of hospital discharges

[f]Hospital discharge data.

[g]RMB: Barcelona Mortality Register.

[h]CBD: basic cause of death.

[i]Injury-related mortality data.

[j]RMB_CT: Barcelona Traffic Collision Mortality Register.

[k]Road traffic injury–related mortality data.

[l]GUB: Register of Accidents and Victims by the Barcelona City Police.

[m]Data for people with a road traffic–related injury.

XSL•FO
RenderX

Data Sources

Multiple data sources were used to construct the BISS databases, each providing information on different segments of the injury pyramid. All sources consist of routinely collected administrative or medical records (and police records for road traffic injuries). The sources and their use in BISS are described in the following sections.

Hospital Emergency Department Data (CMBD-UR)

The Urgencies Register (CMBD-UR) is a component of the regional CMBD and captures all ED visits in hospitals and emergency centers. For BISS, we extracted all ED records with an injury-related diagnosis or external cause code representing moderate or serious nonfatal injuries that received medical treatment. Each ED record includes up to 8 *ICD-10-CM* diagnosis codes and corresponding external cause codes for the injury. These data were considered comprehensive for hospital-attended injuries in the city, as the CMBD covers all public and most private hospitals in Catalonia.

Hospital Discharge Data (CMBD-AH)

The Hospital Discharge Register, another part of the CMBD, records all inpatient admissions in Catalonia's hospitals. We included records of patients hospitalized for injury (again identified via *ICD-10-CM* external cause and diagnosis codes). These represent serious-to-critical injuries requiring admission. Each record can contain up to 15 diagnosis codes. By using both ED and admission data, BISS distinguishes injuries treated and released from the hospital versus those severe enough to require hospitalization.

Mortality Register (RMB)

The RMB collects data on all deaths of Barcelona residents, with cause of death coded in *ICD-10*. For BISS, we extracted injury-related deaths (external cause of mortality codes V01-Y89). The RMB data ensure coverage of fatal injuries among city residents, whether the death occurred in Barcelona or elsewhere. However, the mortality database does not provide detailed injury diagnoses (beyond the underlying cause of death), which limits certain indicators (eg, we cannot identify trauma subtype for all fatalities).

Traffic Fatalities Register (RMB-CT)

The RMB-CT is a specialized subset of the mortality data focusing on road traffic injury–related deaths in Barcelona. It is compiled by deterministically linking multiple sources: the general mortality data, records from the Institute of Legal Medicine and Forensic Sciences of Catalonia (ie, autopsy and forensic reports), civil death certificates, and police reports. This person-level data linkage allows the collection of detailed information on each traffic fatality, including injury diagnoses from an autopsy, circumstances of the crash, and victim details. The RMB-CT thus provides rich verification of road traffic death cases beyond the death certificate alone. Data from RMB-CT feed into the BISS MORTAL-RTI database (fatal road traffic injuries). The RMB includes all deaths of individuals residing in Barcelona, regardless of whether the death occurred within the city or elsewhere (manual available at [12]).

Police Traffic Injury Database (GUB)

The GUB records all traffic collisions in Barcelona that result in personal injury (from minor to fatal). This source does not include medical information like injury diagnoses, but it captures the context of each collision (eg, date, location, types of vehicles, road conditions) and basic information on persons injured (age, sex, road user role). The police criteria consider a [13] person injured if they require any form of medical attention at the scene or afterward. The GUB data are used for the BISS INJURED-RTI database (nonfatal traffic injuries, including minor injuries not seen in hospital records) and for calculating total collisions and total injured persons in traffic (indicators 54 - 60). Police data were not directly linked to hospital records at the individual level for this analysis (ie, we did not attempt to deduplicate police-reported injuries against hospital-treated injuries) but were rather analyzed as a complementary dataset to provide information on minor injuries and collision circumstances [14].

Therefore, the BISS-RTI is fed by the 3 data sources from the Barcelona Traffic Injuries and Accidents Information System [14-17]:

- The CMBD-AH and CMBD-UR, using an injury definition for traffic-related incidents established by the "Working Group of the Spanish Society of Epidemiology on the measurement of the health impact of injuries in Spain" [18]
- The GUB, which is based on traffic collision reports from the Barcelona City Police, providing comprehensive information on the circumstances of collisions that occur within the city
- The RMB-CT

BISS Databases and Injury Subsets and Surveillance Indicators

Based on the data sources from the CMBD-UR, CMBD-AH, RMB, RMB-RTI, and GUB and applying the data inclusion criteria outlined in Table 1, 5 databases were generated for the BISS: ED: emergency department visit data; HOSP: hospital discharge data; MORTALITY: injury-related mortality data; MORTAL-RTI: road traffic injury–related mortality data; and INJURED-RTI: data for people with a road traffic–related injury.

From these 5 databases (ED, HOSP, MORTALITY, MORTAL-RTI, and INJURED-RTI), 8 subsets of injuries to be monitored were generated (one for general injuries and another for traffic collision injuries) and labeled as A, B, C, D, E, F, G, and H.

Table 1 presents the data sources feeding into the BISS, the 5 derived databases, and the 8 resulting injury subsets. They further list the selection criteria applied to each data source and subset, along with the corresponding surveillance indicators used for monitoring. An episode of injury refers to an incident that, due to an external cause, results in harm to an individual. For nonfatal injuries, we referred to episodes rather than individuals, as a single person may experience multiple injury episodes throughout their lifetime.

XSL•FO
RenderX

Variables

The primary outcomes of interest were counts and rates of injury cases across various categories (nonfatal injuries treated in EDs, nonfatal injuries requiring hospitalization, and fatal injuries). Rather than traditional "exposures" and "confounders" as in an etiologic study, this surveillance system tracked descriptive variables such as demographics, injury mechanisms, and severities for each case. Key variables included in BISS databases are listed in Table 2. These include basic demographics (sex, age, and country and municipality of residence), injury descriptors (date of injury or admission, diagnosis codes, and external cause codes specifying mechanism and intent), clinical outcomes (hospital discharge disposition, such as survival or death, and injury severity), and traffic-specific variables for road traffic injuries (eg, road user type, collision circumstances from police records). Table 2 displays which variables are available in each of the 5 BISS databases: ED (ED visits), HOSP (hospital admissions/discharges), MORTALITY (fatal injuries among Barcelona residents), MORTAL-RTI (fatal road traffic injuries in Barcelona), and INJURED-RTI (police-reported traffic collision injuries). A "✓" indicates the variable is present in a given database. For example, injury diagnosis codes (*ICD-10-CM*) are recorded in the health databases (ED, HOSP) with multiple diagnosis fields available and in the forensic mortality data (MORTAL-RTI) but are not recorded in the police injury database. External cause codes (*ICD-10-CM*) for the injury mechanism are captured in ED and HOSP records but not in mortality data. The severity of injuries is determined differently in police records versus medical records: Medical records allow derivation of severity via the Maximum Abbreviated Injury Scale (MAIS; from *ICD-10-CM* diagnoses), whereas the police database (INJURED-RTI) uses its own criterion of whether an injury required medical attention (not anatomic severity score). All coding frameworks (*ICD-10-CM* for diagnoses and external causes and *ICD-10* for causes of death) and classification schemes (eg, CDC's injury matrices for intent and mechanism) used to define variables are consistent with international standards.

Table 3 provides a description of the variables considered in the BISS, including the possible values assigned to each. A more detailed description of each variable is provided in Multimedia Appendix 1.

XSL•FO
RenderX

Table . Variables included in each of the Barcelona Injury Surveillance System (BISS) databases.

Variables	Databases				
	ED[a]	HOSP[b]	MORTALITY[c]	MORTAL-RTI[d]	INJURED-RTI[e]
Admission date	✓	✓	N/A[f]	N/A	N/A
Discharge date	✓	✓	N/A	N/A	N/A
Date of death	N/A	N/A	✓	✓	N/A
Collision date	N/A	N/A	N/A	✓	✓
Sex	✓	✓	✓	✓	✓
Age	✓	✓	✓	✓	✓
Country of origin	N/A	N/A	✓	✓	✓
Municipality of residence	✓	✓	Only includes residents of Barcelona	✓	✓
Basic primary health care area (ABS)	✓	✓	✓	✓	✓
Neighborhood of residence for Barcelona residents	N/A	N/A	✓	N/A	✓
Economic financing regime	✓	✓	N/A	N/A	N/A
Situation of the person injured at discharge	✓	✓	N/A	N/A	N/A
Diagnosis of the lesion (*ICD-10-CM*[g] codes)	✓ (8 diagnostic fields)	✓ (15 diagnostic fields)	N/A	✓ (15 diagnostic fields)	N/A
External cause (*ICD-10-CM* codes)	✓ (8 fields of external cause)	✓ (5 fields of external cause)	N/A	N/A	N/A
Basic cause of death (*ICD-10* codes)	N/A	N/A	✓	✓	N/A
Type of lesion and anatomical region affected	✓	✓	N/A	✓	N/A
Mechanism and intentionality of the injury	✓	✓	✓	N/A	N/A
Type of user	✓	✓	N/A	✓	✓
Severity	✓	✓	N/A	✓	✓[h]
Collision variables (eg, type of collision, type of road, time, use of safety accessories, type of lighting)	N/A	N/A	N/A	N/A	✓

[a]Emergency department visit data.

[b]Hospital discharge data.

[c]Injury-related mortality data.

[d]Road traffic injury–related mortality data.

[e]Road traffic injury data from the Register of Accidents and Victims by the Barcelona City Police (GUB).

[f]N/A: not applicable.

[g]*ICD-10-CM*: *International Classification of Diseases, Tenth Revision, Clinical Modification*.

[h]Severity according to police criteria, which considers that a person involved in the collision is injured if they require some type of medical attention

XSL•FO
RenderX

Table . Description and values of variables considered in the Barcelona Injury Surveillance System (BISS).

Variables	Description	Values
Sex	N/A[a]	Female, male
Age (years)	N/A	0 - 14, 15 - 24, 25 - 44, 45 - 64, 65 - 74, ≥75
Country of origin	Country according to National Statistic Institute codes	N/A
Municipality of residence	N/A	Barcelona, outside Barcelona
Basic Health Area (ABS; acronym in Catalan)	Elementary territorial unit through which primary health care services are organized	67 areas
Neighborhood residence	For residents of Barcelona	73 neighborhoods
Economic financing regime	N/A	Work accident insurance, traffic collision insurance, CatSalut, and other
Situation of the person injured at discharge	N/A	Death, hospital admission, discharge at home, others
Injury diagnoses	*ICD-10-CM*[b] codes	8 fields for the ED[c] database; 15 fields for the HOSP[d] and MORTAL-RTI[e] databases
External cause	*ICD-10-CM* codes	8 fields for the ED database; 5 fields for the HOSP database
Basic cause of death	*ICD-10* codes	N/A
Type of injury	Derived from the injury diagnostic *ICD-10-CM* codes and classified using the matrix proposed by the CDC[f] [19]	N/A
Anatomical region affected	Derived from the injury diagnostic *ICD-10-CM* codes and classified using the matrix proposed by the CDC [19]	N/A
Mechanism and intentionality of fatal injuries	Derived from the Basic Cause of Death *ICD-10* codes following the CDC's classification system [20]	N/A
User type for RTI[g]	Derived from the external cause V-codes in nonfatal RTI from the ED and HOSP databases; in contrast, a dedicated variable already included in the MORTAL-RTI and INJURED-RTI[h] databases	Pedestrian, passenger car users, motorcyclist, cyclist, van or truck user, bus user, and other users
Severity	Derived from the MAIS[i] (ranging from 1 to 6), calculated from the AIS[j] (ranging from 1 to 6) of each of the valid injury codes [21]	MAIS<3; MAIS≥3
Collision variables	The INJURED-RTI database, managed by the Barcelona Urban Police, includes a wide range of variables related to traffic collisions and their circumstances, such as the type of collision, type of road, time of the incident, type of lighting, weather conditions, and the use of safety accessories by the individuals involved.	N/A

[a]N/A: not applicable.

[b]*ICD-10-CM*: *International Classification of Diseases, Tenth Revision, Clinical Modification*.

[c]Emergency department visits.

[d]Hospital discharges.

[e]Fatal road traffic–related injuries in Barcelona.

[f]CDC: Centers for Disease Control and Prevention.

[g]RTI: road traffic injury.

[h]Police-reported traffic collision injuries in Barcelona.

[i]MAIS: Maximum Abbreviated Injury Score.

[j]AIS: Abbreviated Injury Score.

XSL•FO
RenderX

Selection and Definition of Indicators

In addition to these variables, the BISS defines a set of surveillance indicators. Each indicator corresponds to a specific subset of injuries (eg, by severity, mechanism) and a measure (number of cases, rates). For example, indicators include counts of nonfatal hospital-treated injuries, counts of traffic-related injuries, counts of severe injuries, mortality counts by cause, and population-based rates for these events. These indicators were carefully selected to ensure comparability at the international level, aligning with global standards. The system relies on standard classifications, as outlined in the previous section, to provide a consistent and reliable framework for injury surveillance and analysis.

The selection of the indicator package for the BISS was conducted through an extensive review of scientific literature and official sources. The process involved searches in PubMed, as well as portals of the Spanish Ministry of Health, the European Commission, and the US CDC. Key reference publications included the report by the Spanish Ministry of Health, *Indicators of Morbidity and Mortality of Injuries Due to Traffic Accidents* [22], and the European Core Health Indicators proposal [23].

Additionally, publications and tools from the European Injury Database of the European Commission [24] were considered, as well as the CDC's *State Injury Indicators Report: Instructions for Preparing 2020 Data* [7] and associated tools [25]. These references provided a robust basis for selecting internationally aligned and evidence-based indicators.

The BISS compiles a comprehensive set of 60 indicators organized into 5 categories to monitor the morbidity and mortality associated with injuries from external causes and to track their temporal evolution (Table 4):

1. Nonfatal injury indicators (indicators 1‑22): These indicators track nonfatal injury episodes based on the type of health care required (eg, hospital emergency or hospitalization departments), the mechanism and intentionality of the injury, and the severity.
2. Traffic collision nonfatal injury indicators (indicators 23‑38): This subset includes 16 indicators specific to nonfatal injuries resulting from traffic collisions and classified by health care requirements (eg, hospital emergency or hospitalization), severity, and user type (eg, pedestrian, cyclist, driver).
3. Mortality indicators for external causes (indicators 39‑48): These 10 indicators focus on mortality caused by external factors and are categorized by the mechanism of the basic cause of death.
4. Traffic collision mortality indicators (indicators 49‑53): This set includes 5 indicators specific to traffic collision–related mortality distinguished by the type of user involved.
5. Traffic collision and injury indicators (indicators 54‑60): These 7 indicators measure the total number of traffic collisions where medical assistance was required and the total number of injured persons without differentiating severity or type of assistance but classified by the type of user.

XSL•FO
RenderX

Table . Injury indicators, subset, definitions, and limitations of each indicator from the Barcelona Injury Surveillance System (BISS).

Indicator	Injury subset	Indicator-specific definition[a]	Limitations
1_Nonfatal_Emerg_All	A[b]	All episodes	• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. There may be assists from the same episode that are separated by >31 days and are considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH[c] is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
11_Nonfatal_Hosp_All	C[d]	All episodes	• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. There may be assists from the same episode that are separated by >31 days and are considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.

XSL•FO
RenderX

Indicator	Injury subset	Indicator-specific definition[a]	Limitations
2_Nonfatal_Emerg_Severe	A	Episodes with MAIS[e] ≥3	• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. There may be assists from the same episode that are separated by >31 days and are considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
12_Nonfatal_Hosp_Severe	C	Episodes with MAIS ≥3	• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. There may be assists from the same episode that are separated by >31 days and are considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
3_Nonfatal_Emerg_Assault	A		

XSL•FO
RenderX

Indicator	Injury subset	Indicator-specific definition[a]	Limitations
		Episodes with any of the following codes in any diagnosis or external cause field: *ICD-10-CM*[f] codes for assault-related injuries: X92-Y09; T36-T50 with 6th character=3 (note: include T36.9, T37.9, T39.9, T41.4, T42.7, T43.9, T45.9, T47.9, and T49.9 with 5th character=3 [intent information for these codes included in the 5th character and not the 6th]); T51-T65 with 6th character=3 (note: include T51.9, T52.9, T53.9, T54.9, T56.9, T57.9, T58.0, T58.1, T58.9, T59.9, T60.9, T61.0, T61.1, T61.9, T62.9, T63.9, T64, and T65.9 with a 5th character =3 [intent information for these codes included in the 5th character and not the 6th]); T71 with 6th character=3; T74, T76, Y38	• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. There may be assists from the same episode that are separated by >31 days and are considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
13_Nonfatal_Hosp_Assault	C	Episodes with any of the following codes in any diagnosis or external cause field: *ICD-10-CM* codes for assault-related injuries: X92-Y09; T36-T50 with 6th characte=3 (note: include T36.9, T37.9, T39.9, T41.4, T42.7, T43.9, T45.9, T47.9, and T49.9 with 5th character=3 [intent information for these codes included in the 5th character and not the 6th]); T51-T65 with 6th character=3 (note: include T51.9, T52.9, T53.9, T54.9, T56.9, T57.9, T58.0, T58.1, T58.9, T59.9, T60.9, T61.0, T61.1, T61.9, T62.9, T63.9, T64, and T65.9 with a 5th character=3 [intent information for these codes included in the 5th character and not the 6th]); T71 with 6th character=3; T74, T76, Y38	• isodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. There may be assists from the same episode that are separated by >31 days and are considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
4_Nonfatal_Emerg_Self-Harm	A		

XSL•FO
RenderX

Indicator	Injury subset	Indicator-specific definition[a]	Limitations
		Episodes with any of the following codes in any diagnosis or external cause field: *ICD-10-CM* codes for intentional self-harm injuries: X71-X83; T36-50 with 6th character=2 (note: include T36.9, T37.9, T39.9, T41.4, T42.7, T43.9, T45.9, T47.9, and T49.9 with 5th character=2 [intent information for these codes included in the 5th character and not the 6th]); T51-T65 with 6th character=2 (note: include T51.9, T52.9, T53.9, T54.9, T56.9, T57.9, T58.0, T58.1, T58.9, T59.9, T60.9, T61.0, T61.1, T61.9, T62.9, T63.9, T64, and T65.9 with a 5th character=2 [intent information for these codes included in the 5th character and not the 6th]); T71 with 6th character=2; T14.91	• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. There may be assists from the same episode that are separated by >31 days and are considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
14_Nonfatal_Hosp_Self-Harm	C	Episodes with any of the following codes in any diagnosis or external cause field: *ICD-10-CM* codes for intentional self-harm injuries: X71-X83; T36-50 with 6th character=2 (note: include T36.9, T37.9, T39.9, T41.4, T42.7, T43.9, T45.9, T47.9, and T49.9 with 5th character =2 [intent information for these codes included in the 5th character and not the 6th]); T51-T65 with 6th character =2 (note: include T51.9, T52.9, T53.9, T54.9, T56.9, T57.9, T58.0, T58.1, T58.9, T59.9, T60.9, T61.0, T61.1, T61.9, T62.9, T63.9, T64, and T65.9 with a 5th character=2 [intent information for these codes included in the 5th character and not the 6th]); T71 with 6th character=2; T14.91	• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. There may be assists from the same episode that are separated by >31 days and are considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
5_Nonfatal_Emerg_Unint_Fall	A	Episodes with any of the following codes in external cause field: *ICD-10-CM* codes for unintentional fall-related injuries: V00.1-V00.8 if 6th character=1; W00-W15, W17, W19; W16 if 6th character=2, W16.42, W16.92; W18.1-W18.3	

XSL•FO
RenderX

Indicator	Injury subset	Indicator-specific definition[a]	Limitations
			• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances f the same person <31 days apart are considered part of the same episode. There may be assists from the same episode that are separated by >31 days and are considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
15_Nonfatal_Hosp_Unint_Fall	C	Episodes with any of the following codes in external cause field: *ICD-10-CM* codes for unintentional fall-related injuries: V00.1-V00.8 if 6th character=1; W00-W15, W17, W19; W16 if 6th character=2, W16.42, W16.92; W18.1-W18.3	• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. There may be assists from the same episode that are separated by >31 days and are considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
6_Nonfatal_Emerg_Unint_MVT[g]	A		

XSL•FO
RenderX

Indicator	Injury subset	Indicator-specific definition[a]	Limitations
		Episodes with any of the following codes in external cause field: *ICD-10-CM* codes for unintentional motor vehicle traffic-related injuries: V02-V04 (.1-.9), V09.2; V12-V14 (.3-.9), V19.4-V19.6; V20-V28 (.3-.9), V29.4-V29.9; V30-V39 (.4-.9); V40-V49 (.4-.9); V50-V59 (.4-.9); V60-V69 (.4-.9); V70-V79 (.4-.9); V80.3-V80.5, V81.1, V82.1, V83-V86 (.0-.3), V87.0-V87.8, V89.2	• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. There may be assists from the same episode that are separated by >31 days and are considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
16_Nonfatal_Hosp_Unint_MVT	C	Episodes with any of the following codes in external cause field. *ICD-10-CM* codes for unintentional motor vehicle traffic-related injuries: V02-V04 (.1-.9), V09.2; V12-V14 (.3-.9), V19.4-V19.6; V20-V28 (.3-.9), V29.4-V29.9; V30-V39 (.4-.9); V40-V49 (.4-.9); V50-V59 (.4-.9); V60-V69 (.4-.9); V70-V79 (.4-.9); V80.3-V80.5, V81.1, V82.1, V83-V86 (.0-.3), V87.0-V87.8, V89.2	• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. There may be assists from the same episode that are separated by >31 days and are considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
7_Nonfatal_Emerg_Unint_Fire	A	Episodes with any of the following codes in external cause field: *ICD-10-CM* codes for unintentional fire-related injuries: X00-X08	

XSL•FO
RenderX

Indicator	Injury subset	Indicator-specific definition[a]	Limitations
			• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. There may be assists from the same episode that are separated by >31 days and are considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
17_Nonfatal_Hosp_Unint_Fire	C	Episodes with any of the following codes in external cause field: *ICD-10-CM* codes for unintentional fire-related injuries: X00-X08	• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. There may be assists from the same episode that are separated by >31 days and are considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
8_Nonfatal_Emerg_Unint_Nondrug	A		

XSL•FO
RenderX

Indicator	Injury subset	Indicator-specific definition[a]	Limitations
		Episodes with any of the following codes in any diagnosis or external cause field: *ICD-10-CM* codes for unintentional nondrug poisonings: T51-T52 (.0-.3 and .8 with 6th character=1 and .91); T53 (.0-.7 with 6th character=1 and .91); T55 (.0, .1 with 6th character=1); T56 (.0-.7, .81, and .89 with 6th character =1 and .91); T57 (.0-.3 and .8 with 6th character=1 and .91); T58 (.01, .11, .2, and .8 with 6th character=1 and .91); T59 (.0-.7, .81, and .89 with 6th character=1 and .91); T60 (.0-.8 with 6th character=1 and .91); T61 (.01, .11, .77, .78, and .8 with 6th character=1 and .91); T62 (.0-.8 with 6th character=1 and .91); T64 (.01 and .81); T54.0 with 6th character=1; T65 (.0, .1, .21, .22, .29, .3-.6, .81, .83, and .89 with 6th character=1 and .91)	• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. There may be assists from the same episode that are separated by >31 days and are considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
18_Nonfatal_Hosp_Unint_Nondrug	C	Episodes with any of the following codes in any diagnosis or external cause field: *ICD-10-CM* codes for unintentional nondrug poisonings: T51-T52 (.0-.3 and .8 with 6th character=1 and .91); T53 (.0-.7 with 6th character=1 and .91); T55 (.0, .1 with 6th character=1); T56 (.0-.7, .81, and .89 with 6th character=1 and .91); T57 (.0-.3 and .8 with 6th character=1 and .91); T58 (.01, .11, .2, and .8 with 6th character=1 and .91); T59 (.0-.7, .81, and .89 with 6th character=1 and .91); T60 (.0-.8 with 6th character=1 and .91); T61 (.01, .11, .77, .78, and .8 with 6th character=1 and .91); T62 (.0-.8 with 6th character=1 and .91); T64 (.01 and .81); T54.0 with 6th character=1; T65 (.0, .1, .21, .22, .29, .3-.6, .81, .83, and .89 with 6th character=1 and .91)	• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. There may be assists from the same episode that are separated by >31 days and are considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
9_Nonfatal_Emerg_Unint_Drug	A		

XSL•FO
RenderX

Indicator	Injury subset	Indicator-specific definition[a]	Limitations
		Episodes with any of the following codes in any diagnosis or external cause field: *ICD-10-CM* codes for unintentional drug poisonings: T36, T37, T47 (.0-.8 with 6th character=1 and .91); T38 (.0-.7, .80, .81, .89, .90, and .99 with 6th character=1); T39 (.09, .1, .2, .31, .39, .4, and .8 with 6th character=1 and .91); T40 (.0-.4, .41, .42, .49, .5, .60, .69, .7, .71, .72, .8, .90, and .99 with 6th character=1); T41 (.0, .1, .20, .29, .3, .41, and .5 with 6th character=1); T42 (.0-.6, .71, and .8 with 6th character=1); T43 (.01, .02, .1, .20-.22, .29, .3, .4, .50, .59, .60-.65, .69, and .8, with 6th character=1 and .91); T44 and T46 (.0-.8, .90, and .99 with 6th character=1); T45 .0-.4, .51, .52, .60-.62, .69, .7, and .8 with 6th character=1 and .91)	• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. There may be assists from the same episode that are separated by >31 days and are considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
19_Nonfatal_Hosp_Unint_Drug	C	Episodes with any of the following codes in any diagnosis or external cause field: *ICD-10-CM* codes for unintentional drug poisonings: T36, T37, T47 (.0-.8 with 6th character=1 and .91); T38 (.0-.7, .80, .81, .89, .90, and .99 with 6th character=1); T39 (.09, .1, .2, .31, .39, .4, and .8 with 6th character=1 and .91); T40 (.0-.4, .41, .42, .49, .5, .60, .69, .7, .71, .72, .8, .90, and .99 with 6th character=1); T41 (.0, .1, .20, .29, .3, .41, and .5 with 6th character=1); T42 (.0-.6, .71, and .8 with 6th character=1); T43 (.01, .02, .1, .20-.22, .29, .3, .4, .50, .59, .60-.65, .69, and .8, with 6th character=1 and .91); T44 and T46 (.0-.8, .90, and .99 with 6th character=1); T45 .0-.4, .51, .52, .60-.62, .69, .7, and .8 with 6th character=1 and .91)	• • Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. There may be assists from the same episode that are separated by >31 days and are considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
10_Nonfatal_Emerg_TBI[h]	A	Episodes with any of the following codes in any diagnosis field: *ICD-10-CM* diagnosis codes for the CDC[i] proposed TBI case definition: S02.0, S02.1; S02.80, S02.81, S02.82, S02.91; S04.02, S04.03, S04.04; S06; S07.1; T74.4	

XSL•FO
RenderX

Indicator	Injury subset	Indicator-specific definition[a]	Limitations
			• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. There may be assists from the same episode that are separated by >31 days and are considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
20_Nonfatal_Hosp_TBI	C	Episodes with any of the following codes in any diagnosis field: *ICD-10-CM* diagnosis codes for the CDC proposed TBI case definition: S02.0, S02.1; S02.80, S02.81, S02.82, S02.91; S04.02, S04.03, S04.04; S06; S07.1; T74.4	• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. There may be assists from the same episode that are separated by >31 days and are considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
21_Nonfatal_Hosp_Long-Bone	C	Episodes with any of the following codes in any diagnosis field: *ICD-10-CM* diagnosis codes proposed for long bone fracture: S72.0- S72.9; S42.2-S42.9; S82.1-S82.9	

XSL•FO
RenderX

Indicator	Injury subset	Indicator-specific definition[a]	Limitations
			• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. There may be assists from the same episode that are separated by >31 days and are considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
22_Nonfatal_Hosp_Fall-Hip-65	C	Episodes in people >65 years old with any of the following hip fracture codes in any diagnosis field and any of the following unintentional fall codes in any external cause field: *ICD-10-CM* codes for hip fractures: S72.0, S72.1, S72.2, M97.0; ICD-10-CM codes for unintentional fall-related injuries: V00.1-V00.8 with the 6th character=1; W00-W15, W17, W19; W16 with the 6th character=2, W16.42, W16.92; W18.1-W18.3	• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. There may be assists from the same episode that are separated by >31 days and are considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
23_Nonfatal_Emerg_RTI[j]	B[k]	All episodes	

XSL•FO
RenderX

Indicator	Injury subset	Indicator-specific definition[a]	Limitations
			• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. However, there could be assists derived from the same episode that are separated in time by >31 days and are being considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
31_Nonfatal_Hosp_RTI	D[l]	All episodes	• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. However, there could be assists derived from the same episode that are separated in time by >31 days and are being considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
24_Nonfatal_Emerg_Severe-RTI	B	Episodes with MAIS ≥3	

XSL•FO
RenderX

Indicator	Injury subset	Indicator-specific definition[a]	Limitations
			• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. However, there could be assists derived from the same episode that are separated in time by >31 days and are being considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
32_Nonfatal_Hosp_Severe-RTI	D	Episodes with MAIS ≥3	• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. However, there could be assists derived from the same episode that are separated in time by >31 days and are being considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
25_Nonfatal_Emerg_Pedestrian	B	Episodes with any of the following codes in any external cause field: *ICD-10-CM* codes for pedestrian RTI-related injuries: V00.0; V01-V06.1/.9; V09.2/.3/.9	

XSL•FO
RenderX

Indicator	Injury subset	Indicator-specific definition[a]	Limitations
			• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. However, there could be assists derived from the same episode that are separated in time by >31 days and are being considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
33_Nonfatal_Hosp_Pedestrian	D	Episodes with any of the following codes in any external cause field: *ICD-10-CM* codes for pedestrian RTI-related injuries: V00.0; V01-V06.1/.9; V09.2/.3/.9	• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances of the same person <31 days apart are considered part of the same episode. However, there could be assists derived from the same episode that are separated in time by >31 days and are being considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
26_Nonfatal_Emerg_PedesTransp	B	Episodes with any of the following codes in external cause field: *ICD-10-CM* codes for pedestrian transport RTI-related injuries: V00.1; V00.8, except V00.81, V00.82	

XSL•FO
RenderX

Indicator	Injury subset	Indicator-specific definition[a]	Limitations
			• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances of the same person <31 days apart are considered part of the same episode. However, there could be assists derived from the same episode that are separated in time by >31 days and are being considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
34_Nonfatal_Hosp_PedesTransp	D	Episodes with any of the following codes in external cause field: *ICD-10-CM* codes for pedestrian transport RTI-related injuries: V00.1; V00.8, except V00.81, V00.82	• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. However, there could be assists derived from the same episode that are separated in time by >31 days and are being considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
27_Nonfatal_Emerg_Cyclist	B	Episodes with any of the following codes in external cause field: *ICD-10-CM* codes for cyclist RTI-related injuries: V10-V18.3/.4/.5/.9; V19.4/.5/.6/.9	

XSL•FO
RenderX

Indicator	Injury subset	Indicator-specific definition[a]	Limitations
			• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. However, there could be assists derived from the same episode that are separated in time by >31 days and are being considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
35_Nonfatal_Hosp_Cyclist	D	Episodes with any of the following codes in external cause field: *ICD-10-CM* codes for cyclist RTI-related injuries: V10-V18.3/.4/.5/.9; V19.4/.5/.6/.9	• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. However, there could be assists derived from the same episode that are separated in time by >31 days and are being considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
28_Nonfatal_Emerg_Motorcyclist	B	Episodes with any of the following codes in any external cause field: *ICD-10-CM* codes for motorcyclist RTI-related injuries: V20-V28.3/.4/.5/.9; V30-V38.4/.5/.6/.7/.9; V29-V39.4/.5/.6/.9; V39	

XSL•FO
RenderX

Indicator	Injury subset	Indicator-specific definition[a]	Limitations
			• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances of the same person <31 days apart are considered part of the same episode. However, there could be assists derived from the same episode that are separated in time by >31 days and are being considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
36_Nonfatal_Hosp_Motorcyclist	D	Episodes with any of the following codes in any external cause field: *ICD-10-CM* codes for motorcyclist RTI-related injuries: V20-V28.3/.4/.5/.9; V30-V38.4/.5/.6/.7/.9; V29-V39.4/.5/.6/.9; V39	• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. However, there could be assists derived from the same episode that are separated in time by >31 days and are being considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
29_Nonfatal_Emerg_CarUser	B	Episodes with any of the following codes in any external cause field: *ICD-10-CM* codes for car user RTI-related injuries: V40-48.4/.5/.6/.7/.9; V49.4/.5/.6/.9	

XSL•FO
RenderX

Indicator	Injury subset	Indicator-specific definition[a]	Limitations
			• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. However, there could be assists derived from the same episode that are separated in time by >31 days and are being considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
37_Nonfatal_Hosp_CarUser	D	Episodes with any of the following codes in any external cause field: *ICD-10-CM* codes for car user RTI-related injuries: V40-48.4/.5/.6/.7/.9; V49.4/.5/.6/.9	• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. However, there could be assists derived from the same episode that are separated in time by >31 days and are being considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
30_Nonfatal_Emerg_TBI-RTI	B	Episodes with any of the following codes in any diagnosis field: *ICD-10-CM* codes proposed for TBIs: S02.0, S02.1; S02.80, S02.81, S02.82, S02.91; S04.02, S04.03, S04.04; S06; S07.1; T74.4	

XSL•FO
RenderX

Indicator	Injury subset	Indicator-specific definition[a]	Limitations
			• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. However, there could be assists derived from the same episode that are separated in time by >31 days and are being considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
38_Nonfatal_Hosp_TBI-RTI	D	Episodes with any of the following codes in any diagnosis field: *ICD-10-CM* codes proposed for TBIs: S02.0, S02.1; S02.80, S02.81, S02.82, S02.91; S04.02, S04.03, S04.04; S06; S07.1; T74.4	• Episodes of injury treated in hospitals in Barcelona are included, but it is unknown whether the incident took place in the city or outside. • As an international consensus criterion, different attendances by the same person <31 days apart are considered part of the same episode. However, there could be assists derived from the same episode that are separated in time by >31 days and are being considered different episodes. Conversely, the same episode could be considered care separated by <31 days that belong to different episodes, in case you are not well informed whether they are successive visits or due to sequelae. • The CMBD-AH is a register of hospital discharges; therefore, there may be cases of people with the injury and were admitted in the year of study but who do not appear because they were registered the following year.
39_Fatal_All	E[m]	All deaths	

XSL•FO
RenderX

Indicator	Injury subset	Indicator-specific definition[a]	Limitations
			• Only the deaths of people residing in Barcelona are available. • People who died in Barcelona are selected regardless of where the episode of injury took place, since this information is not available. • The RMB[n] does not have information on injuries, so it is not possible to define mortality indicators based on diagnostic codes, such as TBI fatalities.
40_Fatal_Unintent-Drowning	E	Deaths with any of the following codes in the underlying cause of death field: *ICD-10* codes for unintentional drowning–related deaths: W65-W74	• Only the deaths of people residing in Barcelona are available. • People who died in Barcelona are selected regardless of where the episode of injury took place, since this information is not available. • The RMB does not have information on injuries, so it is not possible to define mortality indicators based on diagnostic codes, such as TBI fatalities.
41_Fatal_Unintent-Fall	E	Deaths with any of the following codes in the underlying cause of death field: *ICD-10* codes for unintentional fall–related deaths: W00-W19	• Only the deaths of people residing in Barcelona are available. • People who died in Barcelona are selected regardless of where the episode of injury took place, since this information is not available. • The RMB does not have information on injuries, so it is not possible to define mortality indicators based on diagnostic codes, such as TBI fatalities.
42_Fatal_Unintent-Fire	E	Deaths with any of the following codes in the underlying cause of death field: *ICD-10* codes for unintentional fire–related deaths: X00-X19	• Only the deaths of people residing in Barcelona are available. • People who died in Barcelona are selected regardless of where the episode of injury took place, since this information is not available. • The RMB does not have information on injuries, so it is not possible to define mortality indicators based on diagnostic codes, such as TBI fatalities.
43_Fatal_Unintent-MVT	E	Deaths with any of the following codes in the underlying cause of death field: *ICD-10* codes for unintentional motor vehicle traffic-related deaths: [V02-V04] (.1,.9), V09.2, [V12-V14] (.3-.9), V19 (.4-.6), [V20-V28] (.3-.9), [V29-V79] (.4-.9), V80 (.3-.5), V81.1, V82.1, [V83-V86] (.0-.3), V87 (.0-.8), V89.2	• Only the deaths of people residing in Barcelona are available. • People who died in Barcelona are selected regardless of where the episode of injury took place, since this information is not available. • The RMB does not have information on injuries, so it is not possible to define mortality indicators based on diagnostic codes, such as TBI fatalities.
44_Fatal_Unintent-Nondrug	E		

XSL•FO
RenderX

Indicator	Injury subset	Indicator-specific definition[a]	Limitations
		Deaths with any of the following codes in the underlying cause of death field: *ICD-10* codes for unintentional nondrug-related deaths: X45-X49	• Only the deaths of people residing in Barcelona are available. • People who died in Barcelona are selected regardless of where the episode of injury took place, since this information is not available. • The RMB does not have information on injuries, so it is not possible to define mortality indicators based on diagnostic codes, such as TBI fatalities.
45_Fatal_Unintent-Drug	E	Deaths with any of the following codes in the underlying cause of death field: *ICD-10* codes for unintentional drug-related deaths: X40-X44	• Only the deaths of people residing in Barcelona are available. • People who died in Barcelona are selected regardless of where the episode of injury took place, since this information is not available. • The RMB does not have information on injuries, so it is not possible to define mortality indicators based on diagnostic codes, such as TBI fatalities.
46_Fatal_Unintent-Suffocation	E	Deaths with any of the following codes in the underlying cause of death field: *ICD-10* codes for unintentional suffocation-related deaths: W75-W84	• Only the deaths of people residing in Barcelona are available. • People who died in Barcelona are selected regardless of where the episode of injury took place, since this information is not available. • The RMB does not have information on injuries, so it is not possible to define mortality indicators based on diagnostic codes, such as TBI fatalities.
47_Suicide	E	Deaths with any of the following codes in the underlying cause of death field: *ICD-10* codes for suicide: X60-X84, Y87.0, U03	• Only the deaths of people residing in Barcelona are available. • People who died in Barcelona are selected regardless of where the episode of injury took place, since this information is not available. • The RMB does not have information on injuries, so it is not possible to define mortality indicators based on diagnostic codes, such as TBI fatalities.
48_Homicide	E	Deaths with any of the following codes in the underlying cause of death field: *ICD-10* codes for homicide: X85-Y09, Y87.1, U01-U02	• Only the deaths of people residing in Barcelona are available. • People who died in Barcelona are selected regardless of where the episode of injury took place, since this information is not available. • The RMB does not have information on injuries, so it is not possible to define mortality indicators based on diagnostic codes, such as TBI fatalities.
49_Fatal_RTI	F[o]	All deaths	

XSL•FO
RenderX

Indicator	Injury subset	Indicator-specific definition[a]	Limitations
			The IMLCFC[p] selects the cases and includes some cases in which the CBD[q] code according to the judicial mortality register of Barcelona is a code that does not correspond to traffic, such as V43.0, V29.3, and V79.3.

XSL•FO
RenderX

Indicator	Injury subset	Indicator-specific definition[a]	Limitations
50_Fatal_Pedestrian	F	Deaths selected according to information about the type of user from the 2 sources from which it is fed: GUB and/or IMLCFC; based on the information on the vehicle in which the person was traveling at the time of the collision or by which they were hit and the position they occupied in relation to the vehicle, the type of user variable is generated using the following categories: pedestrian, passenger car user, motorcyclist, cyclist, van or truck user, bus user, and other user	• The IMLCFC selects the cases and includes some cases in which the CBD code according to the judicial mortality register of Barcelona is a code that does not correspond to traffic, such as V43.0, V29.3, and V79.3.
51_Fatal_Cyclist	F	Deaths selected according to information about the type of user from the 2 sources from which it is fed: GUB and/or IMLCFC; based on the information on the vehicle in which the person was traveling at the time of the collision or by which they were hit and the position they occupied in relation to the vehicle, the type of user variable is generated using the following categories: pedestrian, passenger car user, motorcyclist, cyclist, van or truck user, bus user, and other user	• The IMLCFC selects the cases and includes some cases in which the CBD code according to the judicial mortality register of Barcelona is a code that does not correspond to traffic, such as V43.0, V29.3, and V79.3.
52_Fatal_Motorcyclist	F	Deaths selected according to information about the type of user from the 2 sources from which it is fed: GUB and/or IMLCFC; based on the information on the vehicle in which the person was traveling at the time of the collision or by which they were hit and the position they occupied in relation to the vehicle, the type of user variable is generated using the following categories: pedestrian, passenger car user, motorcyclist, cyclist, van or truck user, bus user, and other user	• The IMLCFC selects the cases and includes some cases in which the CBD code according to the judicial mortality register of Barcelona is a code that does not correspond to traffic, such as V43.0, V29.3, and V79.3.
53_Fatal_TBI-RTI	F	Deaths with any of the following codes in any diagnosis field: *ICD-10-CM* diagnosis codes proposed for TBIs: S02.0, S02.1; S02.80, S02.81, S02.82, S02.91; S04.02, S04.03, S04.04; S06; S07.1; T74.4	• The IMLCFC selects the cases and includes some cases in which the CBD code according to the judicial mortality register of Barcelona is a code that does not correspond to traffic, such as V43.0, V29.3, and V79.3.
54_RTI_Colisions	G[r]	For people who were injured in the same collision, information about the collision summarized in a single record	• N/A[s]
55_RTI_Injured_People	H[t]	All injured people	• N/A
56_RTI_Injured_Pedestrian	H		• N/A

XSL•FO
RenderX

Indicator	Injury subset	Indicator-specific definition[a]	Limitations
		People selected according to information about the type of user from the GUB; based on the information on the vehicle in which the person was traveling at the time of the collision or by which they were hit and the position they occupied in relation to the vehicle, the type of user variable is generated using the following categories: pedestrian, passenger car user, motorcyclist, cyclist, van or truck user, bus user, and other user	
57_RTI_Injured_PedestrianTransport	H	People selected according to information about the type of user from the GUB; based on the information on the vehicle in which the person was traveling at the time of the collision or by which they were hit and the position they occupied in relation to the vehicle, the type of user variable is generated using the following categories: pedestrian, passenger car user, motorcyclist, cyclist, van or truck user, bus user, and other user	• N/A
58_RTI_Injured_Motorcyclist	H	People selected according to information about the type of user from the GUB; based on the information on the vehicle in which the person was traveling at the time of the collision or by which they were hit and the position they occupied in relation to the vehicle, the type of user variable is generated using the following categories: pedestrian, passenger car user, motorcyclist, cyclist, van or truck user, bus user, and other user	• N/A
59_RTI_Injured_Cyclists	H	People selected according to information about the type of user from the GUB; based on the information on the vehicle in which the person was traveling at the time of the collision or by which they were hit and the position they occupied in relation to the vehicle, the type of user variable is generated using the following categories: pedestrian, passenger car user, motorcyclist, cyclist, van or truck user, bus user, and other user	• N/A
60_RTI_Injured_Car	H	People selected according to information about the type of user from the GUB; based on the information on the vehicle in which the person was traveling at the time of the collision or by which they were hit and the position they occupied in relation to the vehicle, the type of user variable is generated using the following categories: pedestrian, passenger car user, motorcyclist, cyclist, van or truck user, bus user, and other user	• N/A

XSL•FO
RenderX

[a]Only codes (both diagnostic and external cause) ending in A, B, C, or missing (reflects initial encounter, active treatment) were included.

[b]Episodes of nonfatal injuries treated in the hospital emergency departments of Barcelona, for both residents and nonresidents of the city.

[c]CMBD-AH: Register of the Minimum Basic Data Set of hospital discharges.

[d]Episodes of nonfatal injuries discharged after admission to a hospital in Barcelona for both residents and nonresidents of the city.

[e]MAIS: Maximum Abbreviated Injury Score.

[f] *ICD-10-CM: International Classification of Diseases, Tenth Revision, Clinical Modification.*

[g]MVT: motor vehicle traffic.

[h]TBI: traumatic brain injury.

[i]CDC: Centers for Disease Control and Prevention.

[j]RTI: road traffic injury.

[k]Episodes of nonfatal injuries due to traffic collisions treated in the hospital emergency departments of Barcelona, for both residents and nonresidentsof the city.

[l]Episodes of nonfatal injuries due to traffic collisions discharged after admission to a hospital in Barcelona, for both residents and nonresidents of the city.

[m]Individuals residing in Barcelona who died from injuries due to external causes within the city.

[n]RMB: Barcelona Mortality Register.

[o]Individuals, both residents and nonresidents of Barcelona, who died in Barcelona due to traffic collision–related injuries.

[p]IMLCFC: L'Institut de Medicina Legal i Ciències Forenses de Catalunya.

[q]CBD: basic cause of death.

[r]Traffic collisions in Barcelona where a person, either a resident or nonresident of the city, has been injured (requiring some form of medicalassistance).

[s]Not applicable.

[t]Individuals, both residents and nonresidents of Barcelona, injured in traffic collisions in Barcelona (requiring some form of medical assistance).

Table 4 provides a comprehensive overview of the 60 indicators used by the BISS, detailing their corresponding data sources, databases, and injury subsets from which they are derived. It also specifies the inclusion and definition criteria applied to each database, ensuring clarity and consistency in categorization and monitoring. Complementing this, Table 4 outlines the specific definition criteria for each indicator, offering a detailed account of how they are constructed and classified, including any known limitations. This table serves as key references for understanding the methodological framework and scope of the BISS injury surveillance system.

Quantitative Variables

Several continuous or quantitative variables in the dataset were categorized or summarized for analysis.

Age was analyzed in age groups rather than as a continuous variable. We used the following standard age bands: 0 - 14 years, 15 - 24 years, 25 - 44 years, 45 - 64 years, 65 - 74 years, and ≥75 years. These groupings align with public health reporting standards and facilitate age-specific rate calculations and comparisons across demographic subpopulations.

Injury severity (as derived from *ICD-10-CM* diagnoses using Abbreviated Injury Score [AIS] or MAIS scores) was dichotomized into severe/critical versus nonsevere. Specifically, injuries with MAIS≥3 were considered “severe” (including critical injuries), while those with MAIS<3 were considered moderate (or minor) injuries. This threshold is commonly used in injury epidemiology to identify serious injuries and was chosen based on international definitions. Some indicators (eg, indicators 2, 12, 24, 32) specifically track the subset of cases meeting this MAIS≥3 criterion.

Time variables such as dates (admission date, discharge date, and collision date) were mainly used for record criteria and calculating durations (eg, length of hospital stay, time trends by year). For annual surveillance reporting, we aggregated data by calendar year. No seasonal nor monthly analyses are presented in this paper, so dates were not explicitly categorized by month or season in this analysis.

Population denominators for rate calculations (number of residents of Barcelona) were treated as continuous counts but were effectively fixed values for each year and each demographic stratum. For standardization, we used the average population of Barcelona across the analysis period as a standard population.

All other variables (eg, sex , mechanism categories, intent) were analyzed as categorical variables, with categories as defined in Table 3. For instance, mechanism of injury has categories like traffic, falls, and burns; intentionality has categories such as unintentional, assault, and self-harm. These categorizations were predetermined based on public health definitions and the coding structure. Grouping of continuous data (eg, age, injury severity scores) was done to enhance interpretability and align with existing frameworks for injury surveillance.

Bias

Because this study was based on routinely collected data, we considered several potential sources of bias and took steps to address them where possible.

Selection Bias

The BISS was designed to be population-based and exhaustive for the defined inclusion criteria, minimizing selection bias. All eligible injury events recorded in the source databases were included (no sampling). However, by design, we excluded minor injuries not captured in hospital or police records, which could bias the overall injury counts toward more severe cases. We acknowledge that the current system underrepresents minor injuries (those only seen in primary care or not treated), an issue

XSL•FO
RenderX

we plan to address in future expansions of BISS. In terms of geographic inclusion, hospital records included injuries treated in Barcelona hospitals even if the injury occurred outside the city; thus, there was potential overinclusion of cases not actually stemming from Barcelona incidents. To mitigate this, for mortality, we restricted to city residents, and for nonfatal indicators, we interpreted results with caution, noting that some hospital-treated injuries might be from outside incidents (this is listed as a limitation in relevant indicators, and population rate denominators use city residents who could slightly overestimate true incidence). Conversely, some injuries that occurred in Barcelona may have been treated outside the city (eg, in neighboring municipalities), leading to underascertainment of certain cases. This was likely minimal for serious injuries given the concentration of tertiary hospitals in Barcelona.

Information Bias (Misclassification)

We relied on coding in administrative data, which could have led to misclassification of injuries or causes (eg, errors in *ICD-10* coding). To address this, we used previously validated coding algorithms and standard definitions for identifying injury mechanisms and outcomes, and we referenced validation studies of *ICD* coding in Spain when available. For example, the definition of traffic injury cases was based on a published algorithm [26], and all cause-of-death coding followed official mortality coding standards, reducing misclassification of injury deaths. Still, coding inconsistencies were possible (eg, variation in external cause coding practices between hospitals). The inclusion of multiple sources and the focus on broad categories of injuries (rather than single *ICD* codes) helped dilute some random misclassification. We did not have access to independent verification of each case's details, so some misclassification bias (in injury mechanism or severity) may remain.

Confounding

Since this was primarily a descriptive study, confounding (as in risk factor/outcome relationships) was not a major concern. We were not examining an exposure-outcome causal hypothesis but rather summarizing injury occurrence. Therefore, we did not need to control for confounders via stratification or modeling. However, when comparing subgroups (eg, injury rates by sex), we used age-standardized rates to account for confounding by age structure in sex comparisons.

Summary of Bias

In summary, the design inherently covered the population of interest and used robust data sources, which helped mitigate selection bias. We made efforts to ensure consistency and accuracy in case definitions to minimize information bias.

Study Size

No a priori sample size calculation was needed because this study included the entire target population of injury events meeting the inclusion criteria in 2018 through 2024. The effective sample size was determined by the number of eligible injury cases recorded in the data sources. By leveraging whole-population data, the study aimed for comprehensive coverage rather than a targeted enrollment number. The study size was thus fixed by data availability: All relevant records for the year were included. This yielded sufficient numbers to produce stable estimate rates for most indicators at the city level. Where certain subcategory counts were small (for rare causes of injury or specific subpopulations), the results were still reported but interpreted with caution.

Statistical Methods

All statistical methods applied in this study were descriptive and were planned with the goal of comprehensive surveillance reporting. We did not use inferential statistical tests or complex models; instead, we produced frequency distributions and rate calculations as described in the following paragraphs.

We generated annual summary tables of injury counts and proportions stratified by key variables. For example, we tabulated the number of injury cases by sex, age group, injury mechanism, and injury severity. In these tables, each cell shows the absolute number of cases and its percentage of the column or row total (eg, the proportion of all injuries that falls represent or the percentage of injuries occurring in a given age group). Separate tabulations were done for males and females to facilitate sex-specific insights.

For each of the defined indicators, we calculated the annual incidence rate per 100,000 population. The population denominators were drawn from the Barcelona Municipal Register of Inhabitants (Padron) for the corresponding year. We produced two main types of rate tables for each indicator: (1) overall and sex-specific rates (including crude and age-standardized rates for males and females) and (2) age-specific rates by sex. Age standardization was done using the direct method and the combined 2018 - 2024 Barcelona population as the standard, to adjust for differences in age distribution when comparing sexes.

There were very few missing data for core variables due to the nature of the sources (eg, age and sex are almost always recorded in administrative data). If any key variable was missing, those records were still included in overall counts but excluded from stratified analyses requiring that variable. We did not need to apply any imputation, as missingness was minimal and not central to the indicators.

We did not perform analyses of subgroups or interactions beyond the stratifications by sex and age already described. The design of the indicators inherently stratified certain subgroups. We did not explicitly test interactions or perform subgroup hypothesis tests, since our focus was on reporting descriptive metrics rather than estimating effect sizes.

Given that our study was descriptive, we did not conduct formal sensitivity analyses. However, we did examine whether results were consistent over the time series (2018 - 2024) as an informal check.

All analyses were performed using Stata 16 statistical software. Data cleaning and rate calculations were done using standard Stata procedures and validated by cross-checking selected results with an independent calculation. The statistical outputs (tables and figures) were then compiled for reporting, with no adjustments for multiple comparisons and no derivation of *P* values.

XSL•FO
RenderX

Ethical Considerations

This study was based exclusively on anonymized secondary data collected within the routine mandate of public health authorities, which did not require review by a research ethics committee. All data were processed in compliance with applicable data protection regulations.

Data Access and Cleaning Methods

Access to the underlying data was governed by institutional agreements, but within the research team, the investigators had full access to all relevant de-identified records used to create the study datasets. Specifically, the Public Health Agency of Barcelona has authorized access to the CMBD and mortality databases for surveillance purposes as well as to the police collision database through a collaboration with the city police. All data were handled in compliance with privacy regulations. Therefore, although raw data cannot be made public, the authors (analysts) could use the complete datasets for the purposes of this study. We also note that details on how to access supplemental materials (eg, coding lists or aggregate data tables) are provided in the manuscript or Multimedia Appendix 1 as needed.

Data cleaning procedures were carried out to ensure a high-quality analytic dataset. Key steps included filtering and case selection, merging and deduplication, coding harmonization, missing or inconsistent values, and linkage quality (if applicable).

For filtering and case selection, we applied the inclusion criteria algorithms (per Table 1) to each source to extract injury cases. This involved querying diagnosis and external cause code fields for relevant *ICD-10-CM* codes, selecting only Barcelona residents for the mortality dataset, and selecting only traffic-related cases for the BISS-RTI subsets using cause codes. We verified that these filters were correctly implemented by cross-checking data.

For merging and deduplication, records are uniquely identified within each data source. In the health datasets (ED and HOSP), it is possible that the same patient's care could generate multiple records (eg, an ED visit followed by an admission). Currently, due to the unavailability of a common patient ID, it was not possible to avoid double counting. We treated ED and HOSP data as separate strata and did not merge them at the patient level; an ED visit that became an admission would appear in both ED and HOSP databases. It was expected that, in the future, patient ID will be available for BISS and consequently, no double-counting would occur across ED versus HOSP indicators, making them mutually exclusive. In contrast, we ensured that deaths were only counted in the mortality datasets and not simultaneously in ED or HOSP (for example, if a person died in hospital, that case is recorded in the mortality register and in the HOSP data; for BISS reporting, we counted it as a death, not a nonfatal hospitalization). In the police data, each injured person in a collision is a separate record; if one person had multiple injury-causing collisions, they would appear multiple times, but each corresponds to a distinct event (this is consistent with the definition of counting episodes, not unique individuals).

Regarding coding harmonization, we standardized variable coding across data sources. For instance, we created common categorical variables for mechanism and intent of injury using *ICD* codes from health data and analogous categories from police data. We also mapped the police-defined road user categories to match the categories used in health data where possible. All *ICD-10* and *ICD-10-CM* codes were checked against their descriptions to confirm they were correctly grouped under each mechanism or intent category.

We also performed checks for impossible, inconsistent, or missing values (eg, negative ages, admission dates after discharge dates). No systematic issues were found; any obvious data errors were minimal, and those records were either corrected if possible or excluded if not critical. For example, if an ED record was missing age or sex, it would not contribute to age- or sex-specific analyses but would still be counted in total injuries.

Regarding linkage quality, in the RMB-CT (traffic fatality), data linkage between sources (forensic, police, and death certificate) is managed by a deterministic match (using identifiers like national ID, name, date of birth, and date of incident). The quality of these linkages is maintained by the Public Health Agency's longstanding procedures; for this study, we did not perform independent linkage, but we trusted the quality control in the existing RMB-CT process. All traffic deaths included have been vetted through that system. We did not link the police nonfatal records with hospital records due to lack of identifiers; thus, no probabilistic linkage procedure was required beyond what is done within the RMB-CT for fatalities.

Throughout, the data cleaning process was thoroughly documented. The code used for data processing and analysis is available upon request, and the study protocol and variable code lists are provided in Multimedia Appendix 1. These steps ensured that the final analytical datasets were accurate representations of the source data, with a harmonized structure and ready for calculating the injury indicators.

Stakeholders Involved in the Injury Surveillance System

A wide range of stakeholders were involved in the development, implementation, and use of the injury surveillance system. These included public health officials, hospital and emergency care providers, epidemiologists, injury prevention specialists, data analysts, policy-makers, researchers, and representatives from affected communities. Each stakeholder group played a critical role in ensuring that the system generates accurate, relevant, and timely data to inform prevention strategies and health policies. Their active engagement enhanced the system's relevance, facilitated data interpretation, and support the translation of findings into effective public health action.

Evaluation of the Surveillance System

The surveillance system will be evaluated following CDC guidelines to assess its performance and alignment with public health objectives. Key attributes—such as usefulness, simplicity, flexibility, data quality, acceptability, sensitivity, predictive value, representativeness, timeliness, and stability—will be

XSL•FO
RenderX

analyzed to identify strengths and areas for improvement. Evaluation findings will guide targeted recommendations, developed in consensus with stakeholders, and communicated in formats adapted to relevant audiences to support system enhancement.

Dissemination of Surveillance System Results

The dissemination of results from the BISS is a critical step to ensure that collected data translate into effective public health action. Results will be shared systematically with key stakeholders. The planned dissemination includes a web-based dashboard featuring indicators displayed in tables, graphs, and maps across multiple years; reports, infographics, and policy briefs; peer-reviewed publications; and oral and poster presentations, including those at individual, community, and professional meetings. Efforts will be made to ensure that the results are communicated in a timely, accessible, and actionable manner, facilitating data-driven decision-making and strengthening injury prevention strategies across sectors.

Results

Age and Sex Distribution

Overall, 123,420 nonfatal injury episodes were recorded in 2024 in hospital EDs, of which 99,379 (80.5%) occurred among Barcelona residents. ED episodes were slightly more frequent in females than in males, both in total episodes (63,108 vs 60,312) and among residents (52,323 vs 47,056). A total of 18,749 nonfatal hospitalized injury episodes were recorded in 2024, of which 14,319 (76.4%) were among Barcelona residents. Hospitalized episodes were also somewhat more frequent in females than in males, both in total episodes (9645 vs 9104) and among residents (7916 vs 6403). In contrast, mortality data (year 2023) included 695 fatal injuries, all among Barcelona residents, of whom 388 were males and 307 were females (Table 5).

XSL•FO
RenderX

Table . Number, age-specific and age-adjusted rates by sex for nonfatal emergency department (ED) and hospitalized injury episodes as well as injury-related mortality from the Barcelona Injury Surveillance System (BISS).

Age group (years) by sex[a]	Nonfatal ED episodes (injury subset A: ED[b] database) in 2024			Nonfatal hospitalized episodes (injury subset C: HOSP[c] database) in 2024			Fatally injured people (injury subset E: MORTALITY[d] database) in 2023	
	Total injuries, n	Residents in Barcelona		Total injuries, n	Residents in Barcelona		Residents in Barcelona	
		Injuries, n	Age-specific rate (per 100,000)[e]		Injuries, n	Age-specific rate (per 100,000)[e]	Injuries, n	Age-specific rate (per 100,000)[e]
Male								
0 - 14	6874	5868	5929.1	540	246	249	1	1
15 - 24	8258	5325	6072.3	653	332	379	13	15.2
25 - 44	19,711	14,305	5368.1	1754	1048	393	43	16.4
45 - 64	12,650	9955	4498.2	2059	1270	574	78	35.5
65 - 74	4281	3641	5074.3	979	735	1024	36	50.6
≥75	8468	7897	10,983.5	3119	2772	3855	217	306.6
Unknown	70	65	N/A[f]	0	0	N/A	0	N/A
Total	60,312	47,056	5743.9 (AR[g]=5873.1)	9104	6403	782.7 (AR=875.0)	388	47.9 (AR=55.6)
Female								
0 - 14	4806	4067	4356.7	356	176	189	1	1.1
15 - 24	6359	4014	4774.3	398	260	309	4	4.9
25 - 44	13,869	10,449	3957.3	775	514	195	11	4.2
45 - 64	13,197	10,768	4522.8	1320	908	381	28	11.8
65 - 74	7080	6145	6595.5	1058	838	899	17	18.4
≥75	17,787	16,873	13,953.0	5738	5220	4317	246	205.6
Unknown	10	7	N/A	0	0	N/A	0	N/A
Total	63,108	52,323	5854.2 (AR=5619.6)	9645	7916	885.8 (AR=788.5)	307	34.6 (AR=30.0)

[a]For residents of Barcelona, the sex of 14 injured people was unknown, and the ages of 7 females and 65 males were also unknown. For nonresidents of Barcelona, the sex of 5 injured people was unknown, and the ages of 3 femals and 5 males were also unknown.

[b]Emergency department visit data.

[c]Hospital discharge data.

[d]Injury-related mortality data.

[e]Rates were calculated using as the numerator only men and women for whom age information was available.

[f]N/A: not applicable.

[g]AR: age-adjusted rate.

By sex, crude resident rates for nonfatal injuries were slightly higher for females than for males for both ED episodes (5854.2 vs 5743.9 per 100,000) and hospitalized episodes (885.8 vs 782.7 per 100,000). However, age-adjusted rates were higher for ED episodes in males younger than 45 years and hospitalized episodes in males younger than 75 years. For mortality, both crude and age-adjusted rates were clearly higher in males than in females (47.9 vs 34.6 per 100,000 and 55.6 vs 30.0 per 100,000, respectively), indicating a greater fatal injury burden among males.

Injury rates varied markedly by age. ED rates were higher in males at younger ages but higher in females at older ages, especially from the age of 65 years onward. Hospitalized injuries showed a similar pattern, with the highest rates in the oldest age groups, particularly among females. Fatal injury rates increased sharply with age in both sexes and were highest among adults aged 75 years or older.

Injury Mechanism

Table 6 shows the distribution of injury episodes and deaths by injury mechanism in the ED, hospitalized, and mortality data sources. Missing or unspecified information was substantial, particularly in the ED database, where episodes without an external cause code accounted for 63.6% (38,337/60,312) of records for males and 62.8% (39,662/63,108) for females; the

XSL•FO
RenderX

corresponding proportions in hospitalized episodes were 20.4% (1854/9104) and 16.5% (1591/9645), respectively, while nonspecific mechanisms also represented a relevant share across all sources, especially mortality records. Among specified mechanisms, unintentional falls were the most frequent in both ED and hospitalized data and were consistently more common in females than in males (ED: 7052/63,108, 11.2% vs 4761/60,312, 7.9%; hospitalized: 4583/9645, 47.5% vs 2890/9104, 31.7%). Intentional self-harm was also frequent and occurred at higher proportions for females in the nonfatal data (ED: 2260/63,108, 3.6% vs 1515/60,312, 2.5%; hospitalized: 508/9645, 5.3% vs 311/9104, 3.4%), whereas males more often presented with aggression, motor vehicle traffic injuries, and several other unintentional mechanisms. In the mortality data, unintentional asphyxia was the leading specified mechanism and was more frequent for females (113/307, 36.8% vs 114/388, 29.4%), while intentional self-harm had the largest sex difference in the opposite direction, with markedly higher proportions among males than females (75/388, 19.3% vs 31/307, 10.1%).

XSL•FO
RenderX

Table . Number and percentage of nonfatal episodes treated in hospital emergency services, nonfatal episodes that were discharged after hospitalization, and fatally injured people by principal mechanism and sex from the Barcelona Injury Surveillance System (BISS).

Principal mechanism	Nonfatal ED[a] episodes (injury subset A: ED[b] database) in 2024		Nonfatal hospitalized episodes (injury subset C: HOSP[c] database) in 2024		Fatally injured people (injury subset E: MORTALITY[d] database) in 2023	
	Males (n=60,312), n (%)	Females (n=63,108), n (%)	Males (n=9104), n (%)	Females (n=9645), n (%)	Males (n=388), n (%)	Females (n=307), n (%)
Intentional self-harm	1515 (2.5)	2260 (3.6)	311 (3.4)	508 (5.3)	75 (19.3)	31 (10.1)
Aggression	882 (1.5)	381 (0.6)	289 (3.2)	58 (0.6)	9 (2.3)	2 (0.7)
Unintentional fall	4761 (7.9)	7052 (11.2)	2890 (31.7)	4583 (47.50)	39 (10.1)	35 (11.4)
Traffic with a motor vehicle, unintentional	311 (0.5)	287 (0.5)	251 (2.8)	155 (1.6)	17 (4.4)	3 (1)
Cyclist nontraffic/non-motor vehicle, unintentional	265 (0.4)	102 (0.2)	121 (1.3)	34 (0.4)	—[e]	—
Pedestrian not traffic/no vehicle motor, unintentional	13 (0)	33 (0.1)	19 (0.2)	30 (0.3)	2 (0.5)	0 (0)
Nontraffic with a motor vehicle, unintentional	15 (0)	14 (0)	12 (0.1)	7 (0.1)	1 (0.3)	0 (0)
Other transport, unintentional	20 (0)	22 (0)	18 (0.2)	12 (0.1)	—	—
Exposure to fire/smoke, unintentional	215 (0.4)	270 (0.4)	349 (3.8)	278 (2.9)	1 (0.3)	5 (1.6)
Unintentional cutting or piercing	518 (0.9)	304 (0.5)	107 (1.2)	23 (0.2)	0 (0)	1 (0.3)
Unintentional overexertion	633 (1)	735 (1.2)	91 (1)	102 (1.1)	—	—

XSL•FO
RenderX

Principal mechanism	Nonfatal ED[a] episodes (injury subset A: ED[b] database) in 2024		Nonfatal hospitalized episodes (injury subset C: HOSP[c] database) in 2024		Fatally injured people (injury subset E: MORTALITY[d] database) in 2023	
	Males (n=60,312), n (%)	Females (n=63,108), n (%)	Males (n=9104), n (%)	Females (n=9645), n (%)	Males (n=388), n (%)	Females (n=307), n (%)
Hit by or against object, unintentional	801 (1.3)	611 (1)	155 (1.7)	73 (0.8)	—	—
Drowning/submersion, unintentional	0 (0)	0 (0)	4 (0)	0 (0)	4 (1)	3 (1)
Machinery, unintentional	51 (0.1)	9 (0)	105 (1.2)	9 (0.1)	—	—
Natural/environmental, unintentional	217 (0.4)	312 (0.5)	55 (0.6)	43 (0.4)	4 (1)	3 (1)
Asphyxia, unintentional	178 (0.3)	171 (0.3)	172 (1.9)	102 (1.1)	114 (29.4)	113 (36.8)
Drug poisoning, unintentional	754 (1.3)	485 (0.8)	241 (2.6)	235 (2.4)	23 (5.9)	8 (2.6)
Nondrug poisoning, unintentional	1640 (2.7)	1190 (1.9)	90 (1)	61 (0.6)	5 (1.3)	2 (0.7)
Other specified, unintentional	3400 (5.6)	2981 (4.7)	1259 (13.8)	946 (9.8)	0 (0)	2 (0.7)
Nonspecific	5786 (9.6)	6227 (9.9)	711 (7.8)	795 (8.3)	94 (24.2)	99 (32.2)
No external cause code	38,337 (63.6)	39,662 (62.8)	1854 (20.4)	1591 (16.5)	—	—

[a]ED: emergency department.

[b]Emergency department visit data.

[c]Hospital discharge data.

[d]Injury-related mortality data.

[e]Data not available in the mortality register.

Nature of Injury, Anatomical Region, and Injury Severity

Table 7 shows the distribution of nonfatal injury episodes in 2024 by type of injury, anatomical region, and injury severity, separately for ED and hospitalized episodes and for males and females. In both care settings, substantial proportions of cases were coded as unspecified injury, particularly in ED episodes, where this category accounted for 37.3% (22,417/60,312) of episodes in males and 40.4% (25,470/63,108) of episodes in females, compared with 4% (362/9104) and 4.3% (411/9645), respectively, among hospitalized episodes. Among specified injury types, fractures were more frequent in females than in males and were especially prominent among hospitalized episodes (4855/9645, 50.3% in females vs 3812/9104, 41.9% in males), whereas open wounds were more common in men in both the ED and hospitalized data. Extremities and head and neck were the most frequently affected anatomical regions in both sexes, although trunk injuries represented a larger proportion of hospitalized episodes in males than in females.

XSL•FO
RenderX

Regarding injury severity, most episodes were classified as nonsevere, but severe injuries accounted for a considerable proportion of hospitalized episodes and were slightly more frequent in females than in males (4540/9645, 47.1% vs 4025/9104, 44.2%), while the proportion of episodes with unknown severity remained notable in both data sources.

Table . Nature of injury, anatomical region, and injury severity for episodes treated in hospital emergency services and discharged after hospitalization, by sex in the Barcelona Injury Surveillance System (BISS).

Variables	Nonfatal ED[a] episodes (injury subset A: ED[b] database) in 2024[c]		Nonfatal hospitalized episodes (injury subset C: HOSP[d] database) in 2024[c]	
	Males (n=60,312), n (%)	Females (n=63,108), n (%)	Males (n=9104), n (%)	Females (n=9645), n (%)
Type of injury				
Burns and corrosions	1093 (1.8)	1061 (1.7)	425 (4.7)	321 (3.3)
Dislocation	1763 (2.9)	1300 (2.1)	217 (2.4)	117 (1.2)
Effect of foreign body entering orifice	3529 (5.9)	3100 (4.9)	1468 (16.1)	1111 (11.5)
Fracture	9071 (15)	11,081 (17.6)	3812 (41.9)	4855 (50.3)
Internal organ injury	5594 (9.3)	6399 (10.1)	2111 (23.2)	1927 (20)
Open wound	7520 (12.5)	4641 (7.4)	1093 (12)	679 (7)
Other effects of external causes	2046 (3.4)	1872 (3)	49 (0.5)	55 (0.6)
Poisoning	1617 (2.7)	1508 (2.4)	542 (6)	731 (7.6)
Superficial and contusion	9097 (15.1)	9147 (14.5)	861 (9.5)	1012 (10.5)
Toxic effects	2238 (3.7)	1843 (2.9)	233 (2.6)	147 (1.5)
Other specified injury[e]	14,122 (23.3)	15,127 (23.9)	607 (6.7)	298 (3.1)
Unspecified injury	22,471 (37.3)	25,470 (40.4)	362 (4)	411 (4.3)
Anatomical regions				
Extremities	25,630 (42.5)	26,494 (42)	3997 (43.9)	4831 (50.1)
Head and neck	20,832 (34.5)	22,467 (35.6)	2625 (28.8)	2471 (25.6)
Spine and back	725 (1.2)	1012 (1.6)	475 (5.2)	488 (5.1)
Trunk	4713 (7.8)	5342 (8.5)	2800 (30.8)	2297 (23.8)
Unclassified region	5511 (9.1)	4840 (7.7)	923 (10.1)	977 (10.1)
Nonspecific region	12,267 (20.3)	11,736 (18.6)	551 (6.1)	459 (4.8)
Injury severity				
Severe (MAIS[f]≥3)	19,915 (33)	19,206 (30.4)	4025 (44.2)	4540 (47.1)
Not severe (MAIS<3)	33,490 (55.5)	37,294 (59.1)	3059 (33.6)	3195 (33.1)
Unknown[g]	6907 (11.5)	6608 (10.5)	2020 (22.2)	1910 (19.8)

[a]ED: emergency department.

[b]Emergency department visit data.

[c]Each category of type of injury and anatomical region shows the number of people with an injury of that type or in that anatomical region; therefore, percentages sum to more than 100%, as one person may have multiple injuries affecting multiple regions.

[d]Hospital discharge data.

[e]Amputation, blood vessel, crushing, sprains and strains, injury to nerves, injury to muscles and tendons, other injury.

[f]MAIS: Maximum Abbreviated Injury Score.

[g]Includes episodes with diagnosis codes considered traumatic injuries according to the study definition and Centers for Disease Control and Prevention criteria but not mapped by the ICD Programs for Injury Categorization in R (ICDPIC-R) [27] because they do not correspond to anatomical injuries.

Indicators

Table 8 shows a selection of indicators included in the BISS. Age-adjusted rates of all injuries, severe injuries, and traumatic brain injury were higher in males than in females in both the ED and hospitalization data. In the ED data, males had higher rates of all injuries (5873.1 vs 5619.6), severe injuries (1896.6 vs 1691.3), and traumatic brain injury (1120.8 vs 917.0).

XSL•FO
RenderX

However, mechanism-specific ED indicators (indicators 3 - 9) should be interpreted with caution because of the high proportion of missing values. In hospitalized episodes, males also had higher age-adjusted rates of all injuries (875.0 vs 788.5), severe injuries (283.0 vs 258.6), and traumatic brain injury (230.9 vs 187.3), whereas females showed higher rates of fall-related injuries (401.8 vs 341.4), long-bone fractures (253.5 vs 181.0), fall-related hip fractures among people aged ≥65 years (115.6 vs 69.5), and self-harm (50.4 vs 30.5). In contrast, assault-related hospitalization rates were higher in males (19.7 vs 4.8). Mortality rates were consistently higher in males, both overall (55.6 vs 30.0) and for specific causes, particularly unintentional suffocation (17.9 vs 10.8), self-harm/suicide (9.4 vs 3.3), falls (5.7 vs 3.4), drug poisoning/overdose (2.8 vs 0.9), and traffic-related injuries (2.1 vs 0.3).

XSL•FO
RenderX

Table . Barcelona Injury Surveillance System (BISS) indicator results by sex for people residing in Barcelona.

Indicator	Nonfatal ED[a] episodes (injury subset A: ED[b] database) in 2024						Nonfatal hospitalized episodes (injury subset C: HOSP[c] database) in 2024						Fatally injured people (injury subset E: MORTALITY[d] database) in 2023					
	Males			Females			Males			Females			Males			Females		
	Injuries, n	Crude rate (per 100,000)	Age-adjusted rate (per 100,000)	Injuries, n	Crude rate (per 100,000)	Age-adjusted rate (per 100,000)	Injuries, n	Crude rate (per 100,000)	Age-adjusted rate (per 100,000)	Injuries, n	Crude rate (per 100,000)	Age-adjusted rate (per 100,000)	Injuries, n	Crude rate (per 100,000)	Age-adjusted rate (per 100,000)	Injuries, n	Crude rate (per 100,000)	Age-adjusted rate (per 100,000)
1/11/39: All injuries	46,991	5743.90	5873.10	52,316	5854.20	5619.60	6403	782.7	875	7916	885.8	788.5	388	47.9	55.6	307	34.6	30
2/12: Severe injuries	15,228	1861.40	1896.60	15,867	1775.50	1691.30	2053	250.9	283	2644	295.9	258.6	—[e]	—	—	—	—	—
3/13/48: Assault/homicide	515	63	60	270	30.2	30.8	167	20.4	19.7	43	4.8	4.8	9	1.1	1.1	2	0.2	0.2
4/14/47: Self-harm/suicide	1222	149.4	143.2	1919	214.7	218.8	254	31	30.5	444	49.7	50.4	75	9.3	9.4	31	3.5	3.3
5/15/41: Unintentional fall-related injuries	3911	478.1	521.9	6132	686.2	623.8	2382	291.2	341.4	4149	464.3	401.8	39	4.8	5.7	35	3.9	3.4
6/16/43: Unintentional motor vehicle traffic-related injuries	200	24.4	24	181	20.3	19.8	148	18.1	18.4	95	10.6	10.2	17	2.1	2.1	3	0.3	0.3
7/17/42: Fire-related injuries	113	13.8	13.6	159	17.8	17.8	113	13.8	13.6	105	11.7	11.9	1	0.1	0.1	5	0.6	0.5

XSL•FO
RenderX

Indicator	Nonfatal ED[a] episodes (injury subset A: ED[b] database) in 2024						Nonfatal hospitalized episodes (injury subset C: HOSP[c] database) in 2024						Fatally injured people (injury subset E: MORTALITY[d] database) in 2023					
	Males			Females			Males			Females			Males			Females		
	Injuries, n	Crude rate (per 100,000)	Age-adjusted rate (per 100,000)	Injuries, n	Crude rate (per 100,000)	Age-adjusted rate (per 100,000)	Injuries, n	Crude rate (per 100,000)	Age-adjusted rate (per 100,000)	Injuries, n	Crude rate (per 100,000)	Age-adjusted rate (per 100,000)	Injuries, n	Crude rate (per 100,000)	Age-adjusted rate (per 100,000)	Injuries, n	Crude rate (per 100,000)	Age-adjusted rate (per 100,000)
8/18/44: Unintentional non-drug poisoning injuries	1230	150.3	146.2	975	109.1	110.1	98	12	12.2	60	6.7	6.6	5	0.6	0.6	2	0.2	0.2
9/19/45: Unintentional drug poisoning injuries/drug overdose	820	100.2	96.5	698	78.1	78.2	188	23	24.2	206	23.1	21.3	23	2.8	2.8	8	0.9	0.9
46: Unintentional suffocation fatalities	—	—	—	—	—	—	—	—	—	—	—	—	114	14.1	17.9	113	12.7	10.8
40: Unintentional drowning fatalities	—	—	—	—	—	—	—	—	—	—	—	—	4	0.5	0.5	3	0.3	0.3
10/20: TBI[f]	8436	1031.2	1120.8	8966	1003.3	917	1624	198.5	230.9	1934	216.4	187.3	—	—	—	—	—	—
21: Long bone fracture	—	—	—	—	—	—	1329	162.4	181	2571	287.7	253.5	—	—	—	—	—	—

XSL•FO
RenderX

Indicator	Nonfatal ED[a] episodes (injury subset A: ED[b] database) in 2024						Nonfatal hospitalized episodes (injury subset C: HOSP[c] database) in 2024						Fatally injured people (injury subset E: MORTALITY[d] database) in 2023					
	Males			Females			Males			Females			Males			Females		
	Injuries, n	Crude rate (per 100,000)	Age-adjusted rate (per 100,000)	Injuries, n	Crude rate (per 100,000)	Age-adjusted rate (per 100,000)	Injuries, n	Crude rate (per 100,000)	Age-adjusted rate (per 100,000)	Injuries, n	Crude rate (per 100,000)	Age-adjusted rate (per 100,000)	Injuries, n	Crude rate (per 100,000)	Age-adjusted rate (per 100,000)	Injuries, n	Crude rate (per 100,000)	Age-adjusted rate (per 100,000)
22: Fall-related hip fracture in persons ≥65 years old	—	—	—	—	—	—	452	55.2	69.5	1226	137.2	115.6	—	—	—	—	—	—

[a]ED: emergency department.

[b]Emergency department visit data.

[c]Hospital discharge data.

[d]Injury-related mortality data.

[e]Data not available in the mortality register.

[f]TBI: traumatic brain injury.

Self-harm rates were highest among females aged 15 years through 24 years, who showed the greatest burden of all age-sex groups. In contrast, unintentional falls increased steeply with age and were particularly concentrated among older females, with the highest rates observed in those aged ≥65 years, especially among females aged ≥75 years (Table 9).

Table . Number of nonfatal episodes discharged after hospitalization and age-specific rates by sex for unintentional falls and self-harm related injuries in the resident Barcelona population, as recorded in the Barcelona Injury Surveillance System (BISS) in 2024.

Mechanism by age group (years)	Males		Females	
	Episodes, n	Episodes, age-specific rate × 100,000 population (residents)	Episodes, n	Episodes, age-specific rate × 100,000 population (residents)
Self-harm				
0 - 14	3	3	36	38.6
15 - 24	53	60.4	130	154.6
25 - 44	88	33	102	38.6
45 - 64	76	34.3	127	53.3
65 - 74	15	20.9	26	27.9
≥75	19	26.4	23	19
Unintentional falls				
0 - 14	79	79.8	42	45
15 - 24	52	59.3	32	38.1
25 - 44	191	71.7	108	40.9
45 - 64	354	160	281	118
65 - 74	299	416.7	425	456.2
≥75	1407	1956.9	3261	2696.7

XSL•FO
RenderX

Trends

From 2018 through 2024, nonfatal injury rates showed a dip in 2020 followed by a sustained increase in both sexes. Overall, ED-treated injuries rose slightly for males (+3.4%) and females (+3.2%), whereas severe ED-treated injuries increased markedly (+54.5% and +49.1%, respectively). Discharged injuries also increased in both sexes, with females consistently showing higher rates than males. Fall-related injuries rose sharply and remained substantially higher in females, while self-harm injuries increased in both sexes but increased more among females. In contrast, road traffic injuries declined overall, although rates remained consistently higher in males. Fatal injury rates (2018 - 2023) increased in both sexes, especially in males, who also had persistently higher mortality than females throughout the period (Table 10).

XSL•FO
RenderX

Table . Barcelona Injury Surveillance System (BISS) indicator trends by sex for people residing in Barcelona from 2018 through 2024.

Indicators by sex	2018		2019		2020		2021		2022		2023		2024	
	Episodes, n	Episodes, crude rate (per 100,000)	Episodes, n	Episodes, crude rate (per 100,000)	Episodes, n	Episodes, crude rate (per 100,000)	Episodes, n	Episodes, crude rate (per 100,000)	Episodes, n	Episodes, crude rate (per 100,000)	Episodes, n	Episodes, crude rate (per 100,000)	Episodes, n	Episodes, crude rate (per 100,000)
Males														
Nonfatal ED[a] injuries														
1_All injuries	43,221	5556.4	45,682	5738.2	32,483	4124.7	38,977	4945.1	42,413	5413.8	45,21	5579.2	46,991	5743.9
2_Severe (MAIS[b]≥3) injuries	9373	1205.0	9974	1252.8	8256	1048.3	11,150	1414.6	13,284	1695.6	14,821	1828.8	15,228	1861.4
Nonfatal discharged injuries														
11_All injuries	4671	600.5	4963	623.4	4536	576	5325	675.6	5761	735.4	5956	734.9	6403	782.7
12_Severe (MAIS≥3) injuries	1372	176.4	1505	189	1458	185.1	1732	219.7	1899	242.4	2118	261.4	2053	250.9
15_Unintentional fall-related injuries	1568	201.6	1442	181.1	1414	179.5	1702	215.9	2160	275.7	2238	276.2	2382	291.2
14_Intentional self-harm injuries	171	22	168	21.1	189	24	174	22.1	218	27.8	190	23.4	254	31
31_Road traffic injuries	472	60.7	484	60.8	352	44.7	462	58.6	404	51.6	377	46.5	450	55
Fatal injuries														
39_All injuries	303	39	296	37.2	366	46.5	336	42.6	355	45.3	388	47.9	N/A[c]	N/A
41_Unintentional fall-related injuries	36	4.6	18	2.3	45	5.7	33	4.2	28	3.6	39	4.8	N/A	N/A
47_Intentional self-harm injuries	64	8.2	59	7.4	76	9.7	59	7.5	69	8.8	75	9.3	N/A	N/A
Females														
Nonfatal ED injuries														
1_All injuries	49,011	5,674.6	51,131	5818.3	36,103	4157.3	41,838	4837.1	46,848	5426.2	49,989	5634.8	52,316	5854.2

XSL•FO
RenderX

Indicators by sex	2018		2019		2020		2021		2022		2023		2024	
	Episodes, n	Episodes, crude rate (per 100,000)	Episodes, n	Episodes, crude rate (per 100,000)	Episodes, n	Episodes, crude rate (per 100,000)	Episodes, n	Episodes, crude rate (per 100,000)	Episodes, n	Episodes, crude rate (per 100,000)	Episodes, n	Episodes, crude rate (per 100,000)	Episodes, n	Episodes, crude rate (per 100,000)
2_Severe (MAIS≥3) injuries	10,287	1191.0	10,635	1210.2	8835	1017.4	12,159	1405.8	14,781	1712.0	15,836	1785.0	15,867	1775.5
Nonfatal discharged injuries														
11_All injuries	6330	732.9	6652	756.9	6283	723.5	6866	793.8	6948	804.8	7639	861.1	7916	885.8
12_Severe (MAIS≥3) injuries	1813	209.9	1946	221.4	2031	233.9	2153	248.9	2296	265.9	2717	306.3	2644	295.9
15_Unintentional fall-related injuries	2758	319.3	2678	304.7	2620	301.7	3024	349.6	3528	408.6	4068	458.5	4149	464.3
14_Intentional self-harm injuries	239	27.7	270	30.7	303	34.9	366	42.3	394	45.6	399	45	444	49.7
31_Road traffic injuries	235	27.2	203	23.1	164	18.9	202	23.4	208	24.1	184	20.7	214	23.9
Fatal injuries														
39_All injuries	258	29.9	260	29.6	315	36.3	303	35	318	36.8	307	34.6	N/A	N/A
41_Unintentional fall-related injuries	35	4.1	33	3.8	71	8.2	52	6	34	3.9	35	3.9	N/A	N/A
47_Intentional self-harm injuries	32	3.7	23	2.6	32	3.7	30	3.5	29	3.4	31	3.5	N/A	N/A

[a]ED: emergency department.

[b]MAIS: Maximum Abbreviated Injury Score.

[c]N/A: not available.

Road Traffic Injuries

Table 11, which includes data from the road traffic injury subsystem and police records (GUB), shows a consistently higher burden of traffic-related injuries among males than among females across most indicators and data sources. All rates were calculated using the resident population as the denominator; therefore, the analysis was restricted to residents. For all injuries combined, males had higher numbers and crude rates than females for nonfatal ED episodes, nonfatal hospitalized episodes, fatalities, and injured people recorded in police data. Severe injuries (MAIS≥3) were also more frequent among males in both ED and hospitalized episodes. Motorcyclist injuries accounted for a substantial share of the burden, particularly among males, with markedly higher counts and rates across all available sources. Cyclist injuries and injuries

XSL•FO
RenderX

among users of personal transport devices were also more common in males. In contrast, pedestrian injuries showed a more balanced sex pattern: Although hospitalized pedestrian injuries were more frequent in males, police data recorded more injured female pedestrians than male pedestrians, and pedestrian fatalities were slightly higher among females. Car user injuries had smaller sex differences in the ED and hospitalized data, although police records still indicated higher numbers among males. Finally, traumatic brain injury was more frequent among males than females across ED, hospitalized, and fatal cases. Importantly, these resident-based estimates do not capture the full number of traffic-injured people occurring in the city, as many additional injured individuals were also recorded by the Guàrdia Urbana but were not included in the analytical rates because the denominator was restricted to residents only.

XSL•FO
RenderX

Table . Barcelona Injury Surveillance System for road traffic injury (BISS-RTI) indicator results by sex for people residing in Barcelona in 2024 RTI fatalities among Barcelona residents due to collisions occurring in Barcelona

Indicator	RTI[a] nonfatal ED[b] episodes				RTI nonfatal hospitalized episodes				RTI fatalities[c]				Injured people police data			
	Males		Females		Males		Females		Males		Females		Males		Females	
	n	Crude rate (per 100,000)	n	Crude rate (per 100,000)	n	Crude rate (per 100,000)	n	Crude rate (per 100,000)	n	Crude rate (per 100,000)	n	Crude rate (per 100,000)	n	Crude rate (per 100,000)	n	Crude rate (per 100,000)
2349_All injuries	1839	224.8	1180	132	450	55	214	23.9	5	0.6	3	0.3	3346	409	2354	263.4
2432_Severe (MAIS≥2[d]) injuries	1155	141.2	668	74.7	221	27	109	12.2	N/A[e]	N/A	N/A	N/A	N/A	N/A	N/A	N/A
2530_Pedestrian injuries	86	10.5	88	9.8	65	7.9	44	4.9	1	0.1	2	0.2	337	41.2	409	45.8
2634_Pedestrian transport users injuries[f]	121	14.8	66	7.4	41	5	21	2.3	N/A	N/A	N/A	N/A	250	30.6	133	14.9
2751_Cyclist injuries	180	22	67	7.5	42	5.1	12	1.3	1	0.1	1	0.1	335	40.9	149	16.7
2852_Motorcyclist injuries	458	56	181	20.3	222	27.1	55	6.2	2	0.2	0	—[g]	1915	234.1	1006	112.6
2957_Car user injuries	37	4.5	40	4.5	14	1.7	17	1.9	N/A	N/A	N/A	N/A	425	51.9	314	35.1
3063_TBI[h]	202	24.7	177	19.8	95	11.6	59	6.6	5	0.6	3	0.3	N/A	N/A	N/A	N/A

[a]RTI: road traffic injury.

[b]ED: emergency department.

[c]RTI fatalities among Barcelona residents due to collisions occurring in Barcelona.

[d]MAIS: Maximum Abbreviated Injury Score.

[e]N/A: not available.

[f]Transport examples include skates, skateboards, nonmotorized scooters, and shoes with wheels.

[g]Not applicable.

[h]TBI: traumatic brain injury.

Discussion

This study provides a comprehensive assessment of the burden of nonfatal and fatal injuries in Barcelona using an integrated injury information system based on ED, hospitalization, mortality, and, for road traffic injuries, police data. Several findings deserve emphasis. First, nonfatal injuries were slightly more frequent among females, particularly at older ages, whereas fatal injuries were clearly more frequent among males. Second, falls were the leading specified mechanism in nonfatal care and were concentrated among older females, while self-harm represented a marked burden among young females in hospitalization data but a higher mortality burden among males. Third, from 2018 through 2024, most nonfatal injury

XSL•FO
RenderX

indicators increased after the decline observed in 2020, whereas road traffic injuries showed an overall downward trend. Taken together, these findings highlight the heterogeneous distribution of injury burden by sex, age, mechanism, and severity and show the added value of combining multiple routine data sources to identify prevention priorities at the city level.

Our findings also underline the public health relevance of integrated injury surveillance. Surveillance has been defined as the ongoing, systematic collection, analysis, interpretation, and dissemination of health data needed for prevention and control, and WHO injury surveillance guidance similarly emphasizes that the ultimate purpose of injury data collection is to support prevention strategies that reduce death and disability. In this context, this study showed that a multisource system can do more than quantify service use: It can identify high-risk population groups, characterize the severity and mechanisms of injury, and monitor trends over time. This is particularly important in urban settings, where injury risks are shaped by mobility, aging, social context, and health care use and where prevention requires coordinated action across sectors [6,28].

The predominance of falls, especially among older females, is one of the most relevant findings of this study. Falls were the leading specified mechanism in both ED and hospitalized injury data, and rates increased sharply with age, with the highest burden observed in females aged ≥65 years, particularly those aged ≥75 years. This pattern is consistent with previous evidence showing that falls are a major cause of injury-related morbidity and mortality in older adults worldwide. WHO identifies falls as a major public health problem in older populations, and recent studies have shown that frailty substantially increases the long-term risk of fall- and fracture-related hospitalization, particularly among older women. Our findings therefore support the need to frame injury prevention in cities not only around transport safety but also around healthy aging, fall prevention, and the prevention of severe consequences such as fractures, loss of autonomy, and institutionalization [29,30].

Self-harm emerged as another key priority. In our data, hospitalization rates for self-harm were highest among females aged 15 years to 24 years, whereas mortality from self-harm was substantially higher among males. This distinction between nonfatal and fatal burden is important because it suggests that mortality data alone provide an incomplete picture of self-harm epidemiology. The findings are consistent with recent European evidence. In France, hospitalizations for self-harm have increased sharply among adolescent girls and young women since late 2020, and a nationwide registry-based study in Spain reported rising hospitalization rates for nonlethal intentional self-harm between 2018 and 2023, particularly among young women. This pattern is also supported by meta-analytic evidence showing a higher prevalence of self-injurious behaviors among adolescent females. Together, these findings reinforce the need for sex- and age-sensitive prevention strategies that connect injury surveillance with youth mental health services and suicide prevention policies [31,32].

The temporal trends observed in this study are also informative. The decline in nonfatal injuries in 2020 followed by a sustained increase is consistent with evidence from other settings showing reduced injury-related emergency visits during the first phase of the COVID-19 pandemic, likely reflecting both changes in exposure and changes in care-seeking behavior. After 2020, however, most nonfatal injury indicators, including severe injuries, falls, and self-harm, increased in Barcelona. In contrast, road traffic injuries declined overall. Although our study does not permit causal attribution, this decline is consistent with the long-standing urban mobility and road safety policies implemented in Barcelona and with previous evaluations showing that sustained interventions were associated with substantial reductions in traffic-related injuries over time. Recent work from Barcelona estimated that more than 34,800 road traffic injuries may have been prevented between 2008 and 2023, including approximately 1000 severe injuries. These contrasting trends suggest that, although progress may have been made in road safety, other injury mechanisms are emerging as equally important or growing priorities for urban public health action [13,33].

The Barcelona system provides actionable local evidence across multiple sectors, including health, mobility, mental health, and social care. At the same time, the results make clear that improvements in data quality remain essential. In particular, the high proportion of missing external cause coding in ED data, together with the share of unspecified injury diagnoses and unknown severity, limits the precision of mechanism-specific analyses in nonadmitted cases. In addition, because the system is primarily episode-based, repeated contacts related to the same injury event may not always be distinguishable in the absence of full linkage across all sources. Regarding harmonization, linkage and the potential for double counting are essential in multisource systems. Although previous work in Barcelona demonstrated the feasibility of probabilistic linkage between police and hospital ED data for road traffic injuries, this has not yet been adopted routinely within the surveillance system, partly because of data protection constraints [34]. Future development should therefore prioritize improved external cause coding, harmonized data standards, and secure linkage capacity across health and nonhealth data sources. Recent work on digital injury surveillance platforms and data science strategies for injury prevention suggests that these steps are critical to move from fragmented routine data to robust prevention intelligence [28,35]. Several limitations should be considered when interpreting these findings. First, missing or unspecified external cause information in ED data limits mechanism-specific interpretation for nonfatal injuries treated without admission. Second, differences in coding practices, completeness, and purpose across emergency, hospitalization, mortality, and police data sources may affect comparability. Third, because the system is largely episode-based, some repeated contacts may not be fully distinguishable without individual-level linkage across all sources. Finally, although the study provides a detailed and policy-relevant picture of injury burden in Barcelona, the findings are most directly applicable to this urban context and should be generalized cautiously to settings with different health systems, coding practices, or mobility patterns.

In conclusion, this study showed that an integrated city-level injury information system can provide a substantially richer picture of injury burden than isolated routine data sources. In

XSL•FO
RenderX

Barcelona, the current burden of injury is characterized by a high and increasing burden of falls among older females, an important burden of self-harm among young females, and persistently higher injury mortality among males alongside an overall decline in road traffic injuries. These findings suggest that urban injury prevention should now be framed more broadly than road safety alone and should incorporate healthy aging, fall prevention, and youth mental health as central priorities. More broadly, they support the role of integrated injury intelligence systems as essential public health tools for identifying priorities, monitoring change, and informing equitable prevention policy in cities.

Acknowledgments

The authors declare the use of generative artificial intelligence (GAI) in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to GAI tools under full human supervision: proofreading, editing, and translation. The GAI tool used was Microsoft Copilot with OpenAI GPT 5.4. Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes. Declaration submitted by collective responsibility.

Funding

The authors declare that financial support was received for the research, authorship, and/or publication of this article. Support was received from the Department of Research and Universities of the Generalitat de Catalunya (AGAUR), N. 2021 SGR 01462 and Partnership for Healthy Cities, Bloomberg Philanthropies, Vital Strategies.

Data Availability

The injury datasets analyzed during this study were obtained from Agència de Salut Pública de Barcelona (ASPB; health data) and the Barcelona City Council (police data) under data use agreements and are not publicly available. Aggregated data tables are available from the corresponding author on reasonable request. The authors are committed to open science and will consider sharing de-identified individual data in accordance with privacy regulations and with permission of the source agencies.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Description of the variables considered in the Barcelona Injury Surveillance System (BISS).
[DOCX File, 19 KB - publichealth_v12i1e82079_app1.docx]

References

1. Murray CJL, GBD 2021 Collaborators. Findings from the Global Burden of Disease Study 2021. Lancet 2024 May 18;403(10440):2259-2262. [doi: 10.1016/S0140-6736(24)00769-4] [Medline: 38762327]
2. Global Burden of Disease 2021. : Institute for Health Metrics and Evaluation; 2024 URL: https://www.healthdata.org/sites/default/files/2024-05/GBD_2021_Booklet_FINAL_2024.05.16.pdf [accessed 2026-06-27]
3. Norton R, Kobusingye O. Injuries. N Engl J Med 2013 May 2;368(18):1723-1730. [doi: 10.1056/NEJMra1109343] [Medline: 23635052]
4. Preventing injuries and violence: an overview. World Health Organization. 2022. URL: https://www.who.int/publications/i/item/9789240047136 [accessed 2026-06-27]
5. Sorenson SB. Gender disparities in injury mortality: consistent, persistent, and larger than you'd think. Am J Public Health 2011 Dec;101 Suppl 1(Suppl 1):S353-S358. [doi: 10.2105/AJPH.2010.300029] [Medline: 21778511]
6. Holder Y, Peden M, Krug E, Lund J, Gururaj G, Kobusingye O, editors. Injury surveillance guidelines. World Health Organization. 2001. URL: https://www.who.int/publications/i/item/9241591331 [accessed 2026-06-27]
7. Walensky RP, Jones CM, Qualters JR. State injury indicators report: instructions for preparing 2020 data. In: Thomas KE, Johnson RL, editors. : National Center for Injury Prevention and Control, Centers for Disease Control and Prevention; 2020 URL: https://stacks.cdc.gov/view/cdc/117725 [accessed 2026-06-27]
8. Council recommendation of 31 May 2007 on the prevention of injury and the promotion of safety. EUR-Lex.europa.eu.: Official Journal of the European Union URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32007H0718%2801%29 [accessed 2026-06-27]
9. Lyons RA, Turner S, Lyons J, et al. All Wales Injury Surveillance System revised: development of a population-based system to evaluate single-level and multilevel interventions. Inj Prev 2016 Apr;22 Suppl 1(Suppl 1):i50-i55. [doi: 10.1136/injuryprev-2015-041814] [Medline: 26658339]

XSL•FO
RenderX

10. Verschuuren M, Gissler M, Kilpeläinen K, et al. Public health indicators for the EU: the joint action for ECHIM (European Community Health Indicators & Monitoring). Arch Public Health 2013;71(1):12. [doi: 10.1186/0778-7367-71-12] [Medline: 23721296]
11. Pérez K, Ricart I. Lesionats per accident de trànsit atesos als serveis d'urgències hospitalàries de Barcelona. Evolució 1997-2002 Sistema d'informació de Dades d'Urgències Hospitalàries Dels Lesionats per Accident de Trànsit (DUHAT). 2003 URL: https://www.aspb.cat/wp-content/uploads/2016/03/Inf-DUHAT-BARNA-2002.pdf
12. Manual mortalitat i annexos. Agència de Salut Pública. 2026. URL: https://www.aspb.cat/documents/manual-mortalitat/ [accessed 2026-06-25]
13. Law RK, Wolkin AF, Patel N, et al. Injury-related emergency department visits during the COVID-19 pandemic. Am J Prev Med 2022 Jul;63(1):43-50. [doi: 10.1016/j.amepre.2022.01.018] [Medline: 35292198]
14. Bartoll X, Pérez C. Col·lisions de trànsit i persones lesionades a Barcelona. : Agència de Salut Pública de Barcelona; 2021 URL: https://www.aspb.cat/wp-content/uploads/2022/02/ASPB-colisions-transit-2020.pdf [accessed 2026-06-27]
15. Cortés Albaladejo M, Pérez C. Persones mortes a Barcelona per lesions per col·lisions de trànsit 1997-2020. : Agència de Salut Pública de Barcelona; 2021 URL: https://www.aspb.cat/wp-content/uploads/2020/08/ASPB-persones-mortes-barcelona-lesions-collisio-transit-2020.pdf [accessed 2026-06-27]
16. Novoa A, Pérez C. Personas lesionadas por tráfico atendidas en los servicios de urgencias en Barcelona, 2020. : Agència de Salut Pública de Barcelona; 2022 URL: https://www.aspb.cat/es/documents/personas-lesionades-transito-urgencias-barcelona/ [accessed 2026-06-27]
17. Lesions de trànsit. Agència de Salut Pública. URL: https://www.aspb.cat/arees/la-salut-en-xifres/lesions-transit/ [accessed 2026-06-25]
18. Pérez C, Cirera E. Indicadores de morbilidad y mortalidad de lesión por accidente de tráfico: grupo de trabajo de la sociedad española de epidemiología sobre la medida del impacto en la salud de las lesiones por accidentes de tráfico en españa. : Ministerio de Sanidad y Consumo; 2007 URL: https://seepidemiologia.es/wp-content/uploads/2021/03/INDICADORES-SEE.pdf [accessed 2026-07-06]
19. Hedegaard H, Johnson RL, Garnett MF, Thomas KE. The 2020 International Classification of Diseases, 10th Revision, Clinical Modification Injury Diagnosis Framework for Categorizing Injuries by Body Region and Nature of Injury. Natl Health Stat Report 2020 Dec(150):1-27. [Medline: 33395385]
20. National Center for Health Statistics. Instruction manual: part 9: ICD-10 cause-of-death lists for tabulating mortality statistics. : Centers for Disease Control and Prevention (CDC); 2002. URL: https://www.cdc.gov/nchs/data/dvs/im9_2002.pdf.pdf [accessed 2026-07-06]
21. European Commission. On the implementation of objective 6 of the european commission's policy orientations on road safety 2011-2020 – first milestone towards an injury strategy. 2013 URL: https://road-safety.transport.ec.europa.eu/system/files/2021-09/ser_inj_swd.pdf [accessed 2026-07-06]
22. Ministerio de Sanidad Y Consumo. Indicadores de Morbilidad y Mortalidad de Lesión por accidentes de tráfico. Ministerio de Sanidad. 2007. URL: https://www.sanidad.gob.es/en/areas/promocionPrevencion/lesiones/seguridadVial/estudiosEncuestas/indicadores.htm [accessed 2026-06-26]
23. European core health indicators (ECHI). European Commission. URL: https://health.ec.europa.eu/indicators-and-data/european-core-health-indicators-echi_en [accessed 2026-06-26]
24. Toolbox. European Association for Injury Prevention and Safety Promotion (EuroSafe). URL: https://www.eurosafe.eu.com/key-actions/injury-data/toolbox [accessed 2026-06-26]
25. Ftp.cdc.gov - /pub/health_statistics/NCHS/injury/tools. Centers for Disease Control and Prevention. URL: https://ftp.cdc.gov/pub/Health_Statistics/NCHS/injury/tools/ [accessed 2026-06-26]
26. Borrell C, Domínguez-Berjón F, Galán I, Fernández E. Adiós al papel en Gaceta Sanitaria. Gac Sanit 2013 Nov;27(6):479. [doi: 10.1016/j.gaceta.2013.10.003]
27. Black A, Clark D. Icdpicr: "ICD" programs for injury categorization in r. The Comprehensive R Archive Network. URL: https://CRAN.R-project.org/package=icdpicr [accessed 2026-07-06]
28. Zheng L, Li X, Liu Z, et al. Application of a digital injury-surveillance platform. JAMA Netw Open 2025 Apr 1;8(4):e254799. [doi: 10.1001/jamanetworkopen.2025.4799] [Medline: 40227688]
29. Xu Q, Ou X, Li J. The risk of falls among the aging population: a systematic review and meta-analysis. Front Public Health 2022;10:902599. [doi: 10.3389/fpubh.2022.902599] [Medline: 36324472]
30. Dent E, Dalla Via J, Bozanich T, et al. Frailty increases the long-term risk for fall and fracture-related hospitalizations and all-cause mortality in community-dwelling older women. J Bone Miner Res 2024 Apr 19;39(3):222-230. [doi: 10.1093/jbmr/zjad019] [Medline: 38477757]
31. López-Cuadrado T, Mortier P, Alonso J, Martínez-Alés G. Trends of non-lethal intentional self-harm and suicide in Spain 2018-2023: a nationwide registry-based study. Lancet Reg Health Eur 2025 Nov;58:101467. [doi: 10.1016/j.lanepe.2025.101467] [Medline: 41080066]
32. Farkas BF, Takacs ZK, Kollárovics N, Balázs J. The prevalence of self-injury in adolescence: a systematic review and meta-analysis. Eur Child Adolesc Psychiatry 2024 Oct;33(10):3439-3458. [doi: 10.1007/s00787-023-02264-y] [Medline: 37486387]

XSL•FO
RenderX

33. Pérez K, Santamariña-Rubio E, López MJ, et al. 25 years of providing evidence on road safety interventions at the city level. Front Public Health 2024;12:1463878. [doi: 10.3389/fpubh.2024.1463878] [Medline: 39989864]
34. Cirera E, Plasència A, Ferrando J, Arribas P. Probabilistic linkage of police and emergency department sources of information on motor-vehicle injury cases: a proposal for improvement. Journal of Crash Prevention and Injury Control 2001;43(3):229-237. [doi: 10.1080/10286580108902567]
35. Allegrante JP, Mitchell RJ, Taylor JA, Mack KA. Injury surveillance: the next generation. Inj Prev 2016 Apr;22 Suppl 1(Suppl 1):i63-i65. [doi: 10.1136/injuryprev-2015-041943] [Medline: 27044497]

Abbreviations

AIS: Abbreviated Injury Score
BISS: Barcelona Injury Surveillance System
BISS-RTI: Barcelona Injury Surveillance System, specific for Road Traffic Injuries
CDC: Centers for Disease Control and Prevention
CMBD: Register of the Minimum Basic Data Set in Spain
CMBD-AH: Register of the Minimum Basic Data Set of hospital discharges
CMBD-UR: Register of the Minimum Basic Data Set of hospital emergency department visits
ED: emergency department
GUB: Register of Accidents and Victims by the Barcelona City Police
***ICD-10-CM*:** *International Classification of Diseases, Tenth Revision, Clinical Modification*
MAIS: Maximum Abbreviated Injury Score
RMB: Barcelona Mortality Register
RMB-CT: Barcelona Traffic Collision Mortality Register
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 08.Aug.2025; peer-reviewed by L Post, RSD Calder; revised version received 21.Apr.2026; accepted 27.Apr.2026; published 24.Jul.2026.

Please cite as:
Pérez K, Santamariña-Rubio E, Cortés-Albaladejo M, Artazcoz L, Hyder AA, Borrell C
The Barcelona Injury Surveillance System (BISS) for Safer Cities: Observational, Descriptive Study Using Routine Health and Police Information Databases
JMIR Public Health Surveill 2026;12:e82079
URL: https://publichealth.jmir.org/2026/1/e82079
doi:10.2196/82079

© Katherine Pérez, Elena Santamariña-Rubio, Mònica Cortés-Albaladejo, Lucía Artazcoz, Adnan A Hyder, Carme Borrell. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 24.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Establishing One Health Surveillance Platform for Electronic Integrated Disease Surveillance and Response in Malawi: Action Design Research Study

Tsung-Shu Joseph Wu[1,2,3,4], BSc, MSc; Matthew Vundu Mvula[3], BSc; Edward Kada Koma Chado[5], BSc, MPH; Blessings Nthezemu Kamanga[3], BSc, MSc; Daniel Denis Mapemba[5], BSc, MSc; Rajab Enoch Billy[3], BSc, MSc; Louis Nyirongo[3], BSc; Annie Chauma Mwale[5], MBBSMD, MPH; Evelyn Chitsa Banda[6], MSCN, PhD; Matthew Kagoli[5], MD, PhD; Tiwonge Davis Manda[7], PhD; Gunnar Aksel Bjune[8], MD, PhD; Jens Johan Kaasbøll[1], PhD

[1]Department of Informatics, Faculty of Mathematics and Natural Sciences, University of Oslo, Oslo, Norway
[2]Research Department, Luke International, Oslo, Norway
[3]Health Information Systems Department, Luke International, Lilongwe, Malawi
[4]Overseas Mission Department, Pingtung Christian Hospital, Pingtung City, Taiwan
[5]Department of Epidemiology and Surveillance, Public Health Institute of Malawi, Lilongwe, Malawi
[6]National Public Health Research Division, Public Health Institute of Malawi, Lilongwe, Malawi
[7]Department of Computing, University of Malawi, Zomba, Malawi
[8]Institute for Health and Society, Faculty of Medicine, University of Oslo, Oslo, Norway

Corresponding Author:
Tsung-Shu Joseph Wu, BSc, MSc
Department of Informatics
Faculty of Mathematics and Natural Sciences
University of Oslo
Gaustadalléen 23B
Oslo, 0373
Norway
Phone: 47 98603117
Email: wcsg@lukeinternational.no

Abstract

Background: Facing the threats of emerging and reemerging health issues requires One Health surveillance systems to provide information for integrated responses. Malawi started enhancing the electronic integrated disease surveillance and response (eIDSR) system in 2015, progressing with the aim of developing a One Health Surveillance Platform (OHSP) using District Health Information Software 2 (DHIS2) as its technical backbone, thereby supporting the COVID-19 pandemic response more resiliently and impacting the integrated disease surveillance and response (IDSR) performance. Digital solutions are critical components of One Health surveillance; however, evidence of the successful establishment and implementation of adaptive digital One Health surveillance systems is scarce.

Objective: This study aims to report on the establishment of the OHSP in Malawi and how an adaptive digital health solution contributed to strengthening and impacting the country's eIDSR during the COVID-19 pandemic and beyond the pandemic.

Methods: The establishment of Malawi's OHSP was based on the action design research methodology with a transdisciplinary approach. The core team reflected the multiple iterative processes of building the OHSP and formalized its impact on IDSR reporting quality.

Results: The OHSP core team conducted multiple iterative cycles to build the platform, leveraging lessons from previous eIDSR pilots, reusing digital health infrastructure, and developing DHIS2 digital solutions in 2019, right before the COVID-19 pandemic. The initial establishment was to cover 48.3% (14/29) of the country's health districts. Pivoting from the initial plan as the COVID-19 pandemic emerged, the core team swiftly adapted the OHSP to scale up nationwide and assisted the health system in responding to the pandemic. The pandemic shock resulted in a national scale-up of the OHSP and impacted the national weekly IDSR reporting quality from nonexistence in 2015 to 97.8% and 74.5% for completeness and timeliness, respectively, in 2024.

XSL•FO
RenderX

Conclusions: The establishment of the OHSP significantly bolstered the surveillance function for weekly IDSR reporting. Government leadership and good coordination were key to success. Continuous capacity building, enhancement of community-level surveillance with digital innovations, adaptable technical infrastructure, and a reuse strategy can provide long-term sustainability for One Health surveillance. Malawi's experience may apply to other countries with demonstrated value of resilient, government-led digital health interventions. Future efforts should focus on improving interoperability with other One Health domains and investing in infrastructure upgrades with local leadership and domestic funding to prepare for future emergencies.

(JMIR Public Health Surveill 2026;12:e72029) doi:10.2196/72029

KEYWORDS

adaptive digital health system; DHIS2; District Health Information Software; eIDSR; IDSR; integrated disease surveillance and response; One Health surveillance; One Health; pandemic preparedness

Introduction

Emerging and reemerging health threats, coupled with the complexity of modern health systems, underscore the need for an integrated approach to health surveillance and response, known as the One Health approach [1-3]. Leveraging the One Health approach, the One Health surveillance system emphasizes intersectoral and multidisciplinary collaboration and integrated surveillance of human, animal, and environmental health, which offers promising avenues for strengthening public health management and pandemic preparedness [2,3]. The recent COVID-19 pandemic highlights the critical need to adopt and realize the One Health approach–aided surveillance systems to foster the health system's resilience and reduce disease risks [1,3]. However, given the broad scope of One Health, a proper implementation model for examining the One Health surveillance system and resilient health systems has yet to be fully developed [1,4].

In strengthening health system resilience, surveillance is essential for enhancing the health system's governance function during the shock onset, alert, and impact management stage [5]. Considering health system resilience and recent COVID-19 pandemic experience, effective information and communication systems and digital technologies–empowered surveillance can aid health authorities in responding more resiliently [5-7]. The "One Health surveillance system" is considered a collaborative surveillance system that provides the systematic collection, validation, analysis, and interpretation of data, and dissemination of information collected from humans, animals, and the environment to inform decisions for more effective, evidence-based health interventions [2,8]. By linking surveillance data across different sectors and applying a transdisciplinary approach, One Health surveillance can help different health-related actors address the full spectrum of disease and outbreak detection, emergency preparedness, response, management, and global health security and health system resilience [2,4].

However, establishing, realizing, and implementing a functional One Health surveillance system has proven challenging, with findings showing that the absence of a clear vision for the future of One Health is a barrier to interdisciplinary collaboration and that siloed approaches by different sectors restrict the ability of professionals to work collaboratively across disciplines [9,10]. At the policy level, political will and legal basis are essential for integrating surveillance systems yet remain difficult to attain the collaborative goal [2,8,10]. At the institutional level, establishing a formal governance body with representatives from each sector could assist in overcoming long-standing barriers of privacy and distrust, where sociopolitical factors were considered as organizational attributes contributing to establishing a One Health surveillance system [2,8-10]. At the operational level, technical barriers highlight that the requirements for a successful One Health surveillance system should include understanding disease transmission and the importance of domain experts' knowledge as scientific support, as well as technical mechanisms to support collaboration [2,10].

Malawi adopted the integrated disease surveillance and response (IDSR) system as its national surveillance system for human health and recognized the need for improvements through advanced information technologies [11]. The latest IDSR technical guidelines highlight the importance of using electronic and digital tools to strengthen its function, known as electronic integrated disease surveillance and response (eIDSR) [12]. Countries are urged to implement electronic tools with an interoperable approach to fortify the eIDSR system and information-sharing platforms between various health units to expedite data transmission and enable a swift response to public health threats [12].

In 2014, the Ministry of Health (MoH) sought to enhance the nation's health emergency preparedness plan and piloted an in-house–developed eIDSR system for the first time [13]. After the World Bank–funded eIDSR pilot project, the Public Health Institute of Malawi (PHIM) of the MoH carried out another pilot by using SMS text message technology—Argus—to aid periodic eIDSR reporting [14]. However, these 2 pilots could not be sustained and scaled up. Meanwhile, Malawi recognized the importance of adopting a One Health approach and leveraging digital tools for surveillance by strengthening the eIDSR system as a starting point for the human health domain in the country [11,15,16].

After Cyclone Idai in 2019, the postdisaster needs assessment revealed the urgent need for strengthening and sustaining epidemiological surveillance and emergency preparedness for overall health system resilience [17]. As Malawi has responded continuously to numerous health disasters, including floods, droughts, and disease outbreaks, the idea of the One Health Surveillance Platform (OHSP) originated amidst the pilots of eIDSR, a successful response to an animal-originated anthrax outbreak and the postdisaster period learning [18]. As a result, the MoH chose the District Health Information Software 2

XSL•FO
RenderX

(DHIS2), a technology that has been used for Malawi's health management information system (HMIS) with a successful nationwide scale-up [16,19], as the technical backbone to establish the OHSP.

Health system resilience is the ability of the system to prepare for and respond to sudden shocks and everyday challenges, and its capacity to absorb deteriorations, adapt, and transform to cope with them [5,20]. It requires strong leadership, political will, a collaborative and coordinated governance structure, and robust surveillance systems to detect shocks and their impact in a timely manner [5,6]. In response to the emerging challenges posed by the COVID-19 pandemic, the MoH decided to accelerate the speed of the establishment of the OHSP to accommodate emerging needs. Malawi was able to use the OHSP as one of the digital health innovations, thereby bouncing forward resiliently during the COVID-19 pandemic [21]. This study aims to report on the establishment of OHSP in Malawi and how an adaptive digital health solution contributed to strengthening and impacting the country's eIDSR during the COVID-19 pandemic and beyond the pandemic.

Methods

Overview

The establishment of Malawi's OHSP for eIDSR started in 2015 based on the action design research (ADR) methodology [22] with the initial intention to strengthen the country's IDSR system. The ADR method has four stages: (1) problem formulation; (2) building, intervention, and evaluation (BIE); (3) reflection and learning; and (4) formalization of learning [22]. We chose this method to ensure that practical insights, stakeholder feedback, and emerging needs were continually incorporated into the platform's development.

Context of the Public Health Surveillance System in Malawi

Malawi is a landlocked country situated in southeastern Africa with a population of almost 16 million. It borders Tanzania, Zambia, and Mozambique [23]. Administratively, the country is divided into 3 regions—northern, central, and southern—with further demarcation into 5 zones, 29 health districts, and 4 major cities. The epidemiology department of the PHIM within the MoH is the main custodian of the IDSR system.

Malawi adopted the IDSR strategy in 2002, and the third edition technical guidelines were published in May 2014 with incremental notifiable diseases and health conditions to fulfill the International Health Regulations (IHR) 2005 and public health needs [24]. Malawi conducted a nationwide assessment of the IDSR system in 2017 [11] and found that the IDSR system showed relatively good completeness but poor timeliness of monthly reports nationwide and zero weekly reports. The use of IT to overcome challenges and improve the surveillance system to have better timeliness for IDSR reporting was proposed [11]. Ever since the assessment, PHIM has been working on strengthening the nation's IDSR system activities with electronic and digital solutions, and later on added the One Health surveillance vision after successfully responding to a hippopotamus' original anthrax outbreak in 2019 [18].

Research Materials and Analysis Methods

Our analysis centered around various documents, including policies, meeting minutes from focus group meetings and key informant interviews, field notes, technical documentation, and digital health products used during the establishment of the OHSP. The OHSP is a digital health innovation accelerated during the COVID-19 pandemic for resilient responses [21] and was sustained for eIDSR functions beyond the pandemic. We summarize the findings from the current establishment to fill the existing knowledge gaps in the implementation of the digitalized OHSP. The impact of the OHSP digital health innovations on the weekly IDSR reporting quality was analyzed using 2024 weekly IDSR reporting quality data with the built-in function of the Reporting Rate Summary in the DHIS2, using the same approach as applied in the previous study [11].

Reflection and Learning Analytic Tool for OSHP and Impact on IDSR System

In order to rigorously reflect on how the OHSP establishment, and the use in responding to the COVID-19 pandemic resiliently, an analytic tool against the resilience framework was synthesized from the available literature, covering the One Health surveillance systems establishment, attributes of engineering resilient information systems for emergency response, strengthening health systems resilience, and known challenges of implementing DHIS2 [2,5,6,9,10,25,26]. Based on the synthesized findings, we first defined the following five analysis domains: (1) One Health surveillance characteristics, (2) addressing known DHIS2 challenges, (3) health system resilience phases, (4) health system resilience attributes, and (5) health system resilience tools [2,5,6,25]. We then identified thematic areas [2,5,6,25,26] and their corresponding establishment level of a One Health surveillance system, namely policy, institutional, and operational levels, based on Bordier et al [2]. Key attributes and challenges were identified from the reviewed literature [1,2,5,6,9,10,25-28]. We selected the relevant strategies and principles applied to the thematic areas and finally selected attributes for our analysis. The high-level analytic tool is shown in Table 1, and the indicators are detailed in Multimedia Appendix 1 [1,2,5,6,9,10,25,26].

In the One Health surveillance characteristics domains, we focused on analyzing the One Health surveillance–related policies, legal instruments, governance, and leadership in Malawi, reflecting on the known characteristics of setting up and sustaining a One Health surveillance at the policy and institutional levels [2,9,10]. We further analyzed the stakeholders' engagement processes and strategy for data integration and the technical infrastructure of the OHSP to see if it fulfills the One Health surveillance needs at the operational level [2,9].

XSL•FO
RenderX

Table 1. High-level analytic tool for health system resilience reflection and learning of the One Health Surveillance Platform establishment.

Domain	Thematic area	Establishment levels	Attributes and challenges
One Health surveillance characteristic	• Policy and governance • Sustainability • Data integration • Technical infrastructure	• Policy level • Institutional level • Operational level	• Policies, legal and operational framework • Funding and scalability • Data exchange across domains • Hardware, software, and communication infrastructures
Addressing known DHIS2[a] implementation challenges	• Political, cultural, social, and structural infrastructure • Appropriate data • Workforce • Education and training	• Institutional level • Operational level	• Political challenges • Stability • Integrity of the health system or fragmentation • Structural infrastructure • Sufficient and qualified data • Data quality • Adequate technical support • Comprehensive training programs
Health system resilience phases	• Anticipation • Preparation	• Policy level • Institutional level • Operational level	• Early detection and risk analysis • Scenario planning • Resource mobilization
Health system resilience attributes	• Stakeholder Ecosystem • Awareness • Adaptive resilience • Robustness	• Policy level • Institutional level • Operational level	• Stakeholder engagement • Sectoral interaction • Situation awareness • Operational environment • Adaptive structures • Interoperability • Coping strategies • Withstanding challenges • Capacity to handle increased demand • Mobilization of essential resources
Health system resilience tools	• Institutionalization	• Policy level • Institutional level • Operational level	• Capacity building • System and infrastructure enhancement

[a]DHIS2: District Health Information Software 2.

Since the DHIS2 was chosen as the backbone technology for the OHSP, we applied the Addressing Known DHIS2 Implementation Challenges domain for analysis [25]. Among the 11 themes, we selected 4 relevant themes, focusing on analyzing the infrastructure issues, the data pipeline, and the workforce to provide technical support, as well as education and training to implement the OHSP, and synthesized the rest in other domains.

For the health system resilience phases and attributes, adaptive resilience was selected to assess the ability of the OHSP and development team to swiftly adjust to immediate challenges with flexibility while continuously evolving and improving based on past experiences and new information. Resilient digital health systems should be designed to adapt and maintain functionality, as well as enable smooth recovery and reintegration amid changing circumstances to accommodate new requirements, emerging threats, or evolving user needs. The practice of adaptability includes integrating new business requirements and new technologies, scaling up or down as necessary, and supporting interoperability with other systems [1,9,27,28].

Finally, since the OHSP is now institutionalized in Malawi and clearly stated in Malawi's Health Sector Strategic Plan III 2023-2030 (HSSP III) [29], we applied the same reporting quality indicators, completeness, and timeliness, to assess the OHSP's preliminary impact on weekly IDSR reporting at the national level [11]. Completeness was measured using the number of actual reports received against the number of expected reports from health facilities according to the health system level. At the same time, the timeliness of reporting was calculated based on the proportion of health facilities submitting surveillance reports on time to the district, with a national target of 80% for both indicators [11].

Ethical Considerations

This study was approved by the Malawi National Health Science Research Committee with approval number 16/4/1563 and the Regional Ethics Committee, Regionale Komiteer For Medisinsk

XSL•FO
RenderX

Of Helsefaglig Forskningsetikk sør-øst, with approval number 2015/2441/REK sør-øst A for its ethical conduct as part of the first author's PhD study. The establishment of the OHSP was conducted under the supervision of the MoH to strengthen the country's surveillance system in accordance with the relevant guidelines and regulations. All key informants were informed and obtained their verbal consent to provide information. The policies for analysis are publicly available. The meeting minutes from focus group meetings and key informant interviews, and field notes are kept safely by the first author and restricted access unless approval is obtained from the MoH. PHIM and the Digital Health Division of the MoH keep the technical documentation and digital health products. No compensation or reimbursement was provided to the meeting participants and key informants.

Results

The observable implementation activities and outputs, including the establishment, implementation, and experience of the OHSP with the COVID-19 pandemic shock, and the post–COVID-19 pandemic impact on the IDSR reporting, are presented in this section based on the ADR method.

Constitution of the OHSP Development Team

At the beginning of the study, the first author, TSJW, fulfilled his roles as an ADR researcher and practitioner in his capacity and was deeply immersed in the country's digital health development context. He facilitated the constitution of the OHSP core team with the MoH leadership. The OHSP core team (n=12) comprises 2 government officials responsible for the nation's IDSR system, 3 epidemiology unit and institution leaders at the PHIM, 5 digital health and IT staff from the Digital Health Division and Information Communication Technology Division of the MoH, a coordination officer, and a technical advisor. TSJW, the first author and technical advisor who was involved in the 2017 IDSR performance assessment, executed eIDSR pilots and provided technical inputs and guidance on developing the OHSP throughout the BIE stages. The extended OHSP development team consists of government technical officers from different sectors, in their role as the IHR focal person in the respective departments. The OHSP core team inherited technical officers supported by partners and sustained the institutional memories within the government from the previous eIDSR pilot projects by integrating them into the OHSP core team, and started working on establishing the OHSP technical solutions in 2019.

Problem Formulation Stage

The establishment of the OHSP to support the eIDSR function was inspired by surveillance practices in the country and international recommendations, aligning with the third edition of the IHR 2005. The initial problems of the IDSR system were identified, highlighting the lack of technological solutions for weekly IDSR reporting. In order to tackle this problem, the first author, TSJW, continuously engaged the MoH and development partners to strengthen the IDSR system with 2 eIDSR pilots conducted from 2015 to 2018.

The World Bank supported the first pilot for the post-Ebola era with the aim of developing and implementing an eIDSR system at all levels of the health system in 2015. The following three main activities were conducted: (1) development of a comprehensive eIDSR system architecture and system logical design, (2) software development, and (3) system deployment and field support. During the first eIDSR pilot, we managed to develop the architecture design, establish a border health fever screening system, and develop the IDSR immediate notifiable disease reporting and syndromic reporting automated functions in the outpatient module of a partner-funded electronic medical record system (EMRS). The eIDSR architecture design used the organizational architecture level to guide respective information product developments and information flow; at the same time, the dashboard was designed to enable monitoring of the diseases and syndromic patterns of the outpatient visits together with a geographic information system and global health news Really Simple Syndication feeds function, as shown in Figure 1.

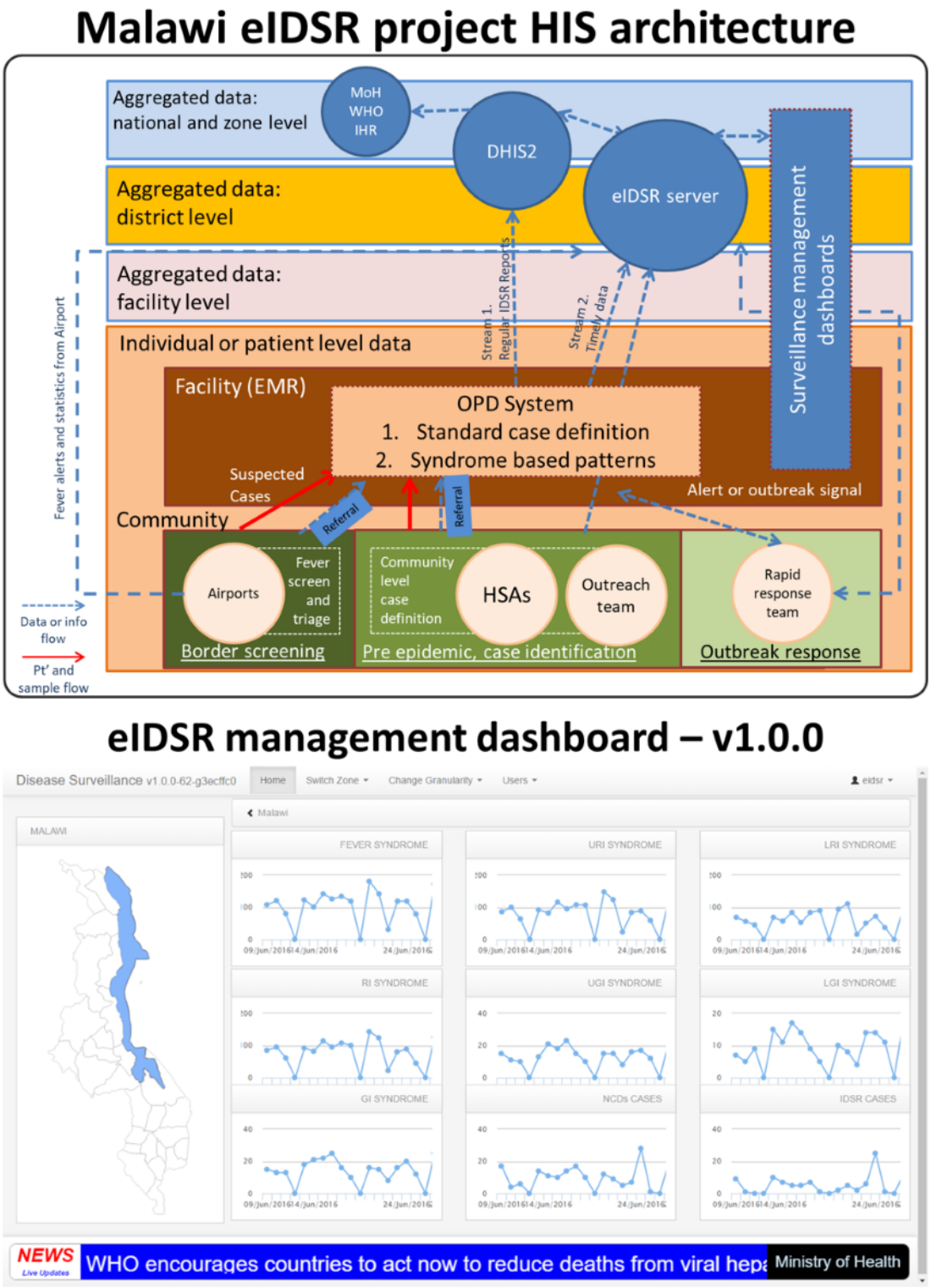
The following two main problems were identified based on the eIDSR pilot implementation results:

1. Sustainability: most of the project team members were dismissed after the project, leaving a minimal capacity within the MoH to maintain the developed systems. The management dashboard and eIDSR central database were designed in-house without the full transition of the source codes to the ministry, and the eIDSR functions in the EMRS were removed after the pilot due to a lack of funding for continuous support.
2. Technical infrastructure: the eIDSR pilot at the facility level was implemented in 59 health facilities with EMRS; however, it covered only 3.2% (59/1849) of the total public health facilities in the country at that time. By the time the pilot ended, the eIDSR pilot also did not cover the whole country for the needed periodic IDSR reporting (weekly and monthly) and community-level case detection.

In order to continue the eIDSR enhancement, the MoH obtained support from the United Nations International Children's Emergency Fund (UNICEF) to pilot the World Health Organization (WHO) Argus solution and facilitate system integration for the country in mid-2017. The UNICEF support was to address the need for a timely weekly IDSR report using the technology. The solution was built and deployed to 220 health care workers for weekly IDSR reporting. However, sustainability and technical infrastructure challenges remained, with no subsequent financial resources and inadequate coverage of the facilities. During the pilot, an additional need was identified to have interoperability between the WHO-Argus solution and the nation's HMIS, which uses DHIS2 technology. The 2 pilots provided lessons for the subsequent establishment of the OHSP. The prevention of the hippopotamus' original anthrax outbreak in late 2018 inspired the MoH to decide to build a suitable digital system to accommodate the One Health surveillance needs and serve the eIDSR functions with the DHIS2 as the technological backbone support.

XSL•FO
RenderX

Figure 1. Preliminary electronic integrated disease surveillance and response (eIDSR) architecture and dashboard concept of the pilot in 2015. DHIS2: District Health Information Software 2; EMR: electronic medical record; HSA: health surveillance assistant; IHR: International Health Regulations; MoH: Ministry of Health; OPD: outpatient department, Pt': patient; WHO: World Health Organization.



BIE Cycles

Pre–COVID-19 Pandemic BIE Cycle

The OHSP establishment formally started in September 2019. Based on the pilot experience, the OHSP core team identified sustainability and technical infrastructure as the 2 key problems to be addressed in the forthcoming OHSP. In addition, previous core team members studied IDSR implementation gaps and community-level surveillance barriers in relation to technology for timely reporting; inadequate training, limited resources, and health-seeking behaviors affecting surveillance capability were also taken into consideration when the team started the OHSP establishment, as summarized in Textbox 1.

XSL•FO
RenderX

Textbox 1. Problems identified from the Malawi electronic integrated disease surveillance and response (IDSR) pilots and previous studies during 2015-2019.

Sustainability

- Organizational sustainability
  - Lack of coordination leading to duplication of efforts
  - Lack of engagement from key personnel
  - Unavailability of key Ministry of Health (MoH) staff
  - Coordination issues for systems deployment
  - Conflicting project schedules with some implementing partners
- Human resource sustainability
  - User non-compliance with system usage
  - Technical staff were dismissed without a full transition
  - Inadequate technical personnel in the MoH to continue the development and maintenance
- Financial sustainability
  - Unforeseen communication expenses
  - Cash flow issues are delaying hardware procurement
  - Lack of financial resources for scale-up and sustainability

Technical Infrastructure

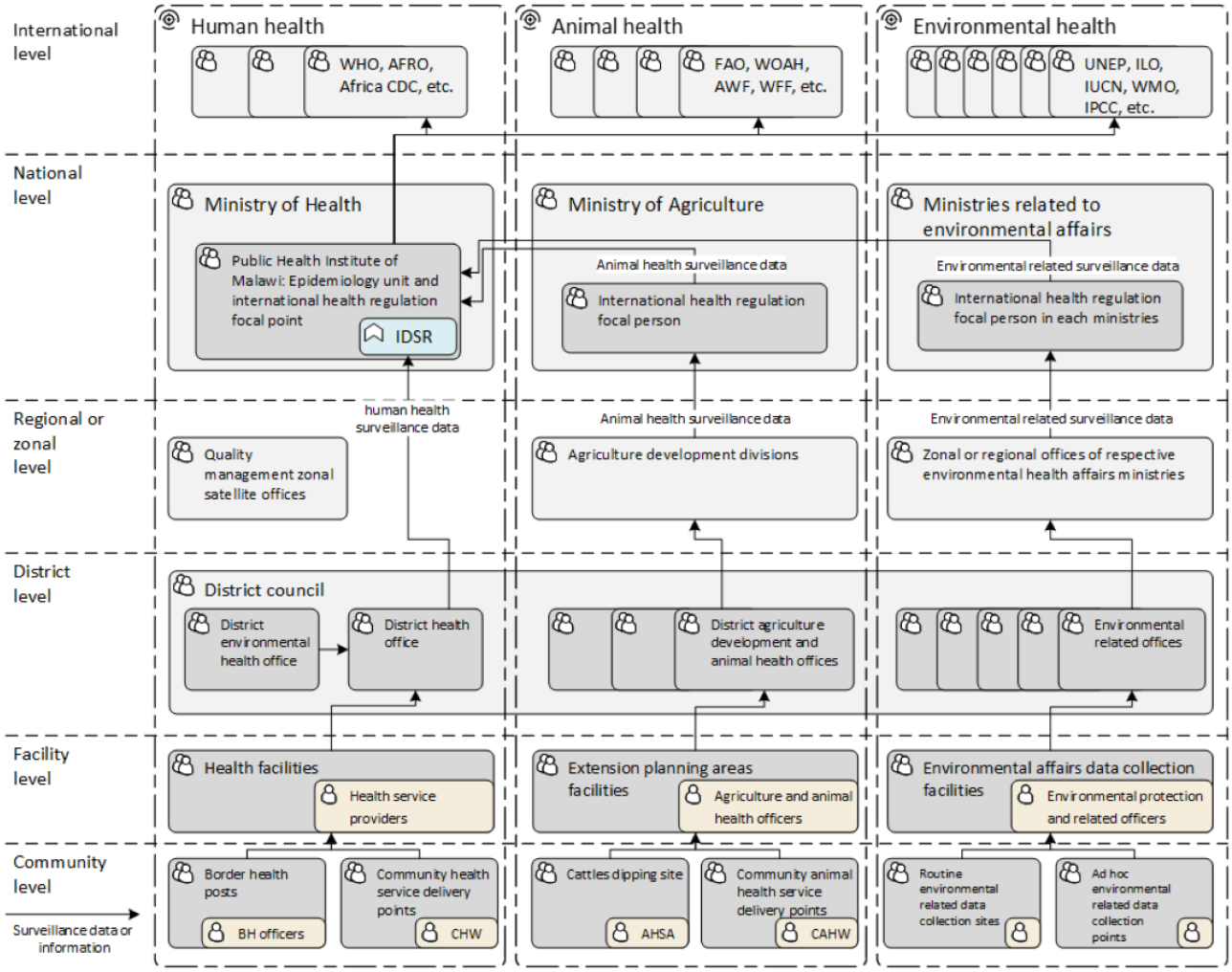
- Incomplete functionality for the full IDSR system needs, including the community-level surveillance
- Low coverage of the solution
- Extended power outages are affecting deployment
- Connectivity issues affecting system functionality
- Labor-intensive setup process for the terminal device (smartphone)
- Congestion and delivery failures of the SMS service

The OHSP core team initiated the BIE stage to address organizational sustainability issues by reviewing the organizational structure across One Health domains, government documents, and literature to inform the design of the OHSP organizational architecture to accommodate One Health surveillance's needs and the IDSR weekly electronic reporting function. Addressing the coordination and leadership issue was the first step. The core team conducted technical meetings with MoH leadership and learned that the One Health task force was formed during the anthrax response. Through these iterative meetings, we transparently communicated the governance concerns and affirmed that the ownership and leadership of the OHSP belong to PHIM within the MoH. PHIM's leadership then led the core team to further engage with stakeholders to contribute to the architecture design of the OHSP. The OHSP core team began engaging with the animal and environmental sectors, with the animal health domain, represented by the Ministry of Agriculture, being fully consulted. While for the environmental domain, due to its vast fields, the team decided to engage further in the next phase of the OHSP implementation after the consultation with relevant stakeholders.

The core team developed a high-level business process architecture of the OHSP (Figure 2), and aligned information flows based on the identified organizational units, surveillance personnel, and studied data collection tools. The hierarchy of the surveillance units is similar across domains but still shows differences at the subnational level across the 3 One Health domains, as listed in Table 2.

XSL•FO
RenderX

Figure 2. High-level business process architecture of the One Health Surveillance Platform in Malawi, 2019. AFRO: World Health Organization Regional Office for Africa; AHSA: animal health surveillance assistant; AWF: African Wildlife Foundation; BH: border health; CAHW: community animal health worker; CDC: Centers for Disease Control and Prevention; CHW: community health worker; FAO: Food and Agriculture Organization; IDSR: integrated disease surveillance and response; ILO: International Labour Organization; IPCC: Intergovernmental Panel on Climate Change; IUCN: International Union for Conservation of Nature; UNEP: United Nations Environment Programme; WFF: World Wildlife Fund; WHO: World Health Organization; WMO: World Meteorological Organization; WOAH: World Organisation for Animal Health.



XSL•FO
RenderX

Table 2. Organizational units mapping between the 3 One Health domains in the context of Malawi.

Administrative level	Human health	Animal health	Environmental health
National level	• Ministry of Health	• Ministry of Agriculture	• Ministries of Water and Sanitation, Natural Resources and Climate Change, Mining, Labour, and Land
Subnational level: region, zone, or division	• 3 regions, 5 health zones	• 8 agricultural development divisions	• Different from each ministry
District level	• 29 health districts with district health offices	• 28 administrative districts with district agricultural development offices	• 28 administrative districts
Facility level	• Central hospitals • District hospitals • Rural and community hospitals • Health centers	• Extension planning areas and sections where agriculture extension development officers work	• Different from each ministry
Community level	• Health post • Village clinic • Outreach clinic • Village health committees (operated by community health workers)	• Area development committees • Village development committees (operated by animal health surveillance assistants	• Different from each ministry

Upon stakeholder engagement, we noticed that only the human health surveillance system had standardized reporting and data collection tools using IDSR technical guidelines. Animal health surveillance has a reporting structure, but the standardized reporting forms and tools have yet to be updated and made available. Environmental health surveillance encompasses an even larger degree of complexity, including climate change, weather forecasts, air and water quality monitoring, waste management, pollution incident surveillance, occupational hazards, and risk surveillance.

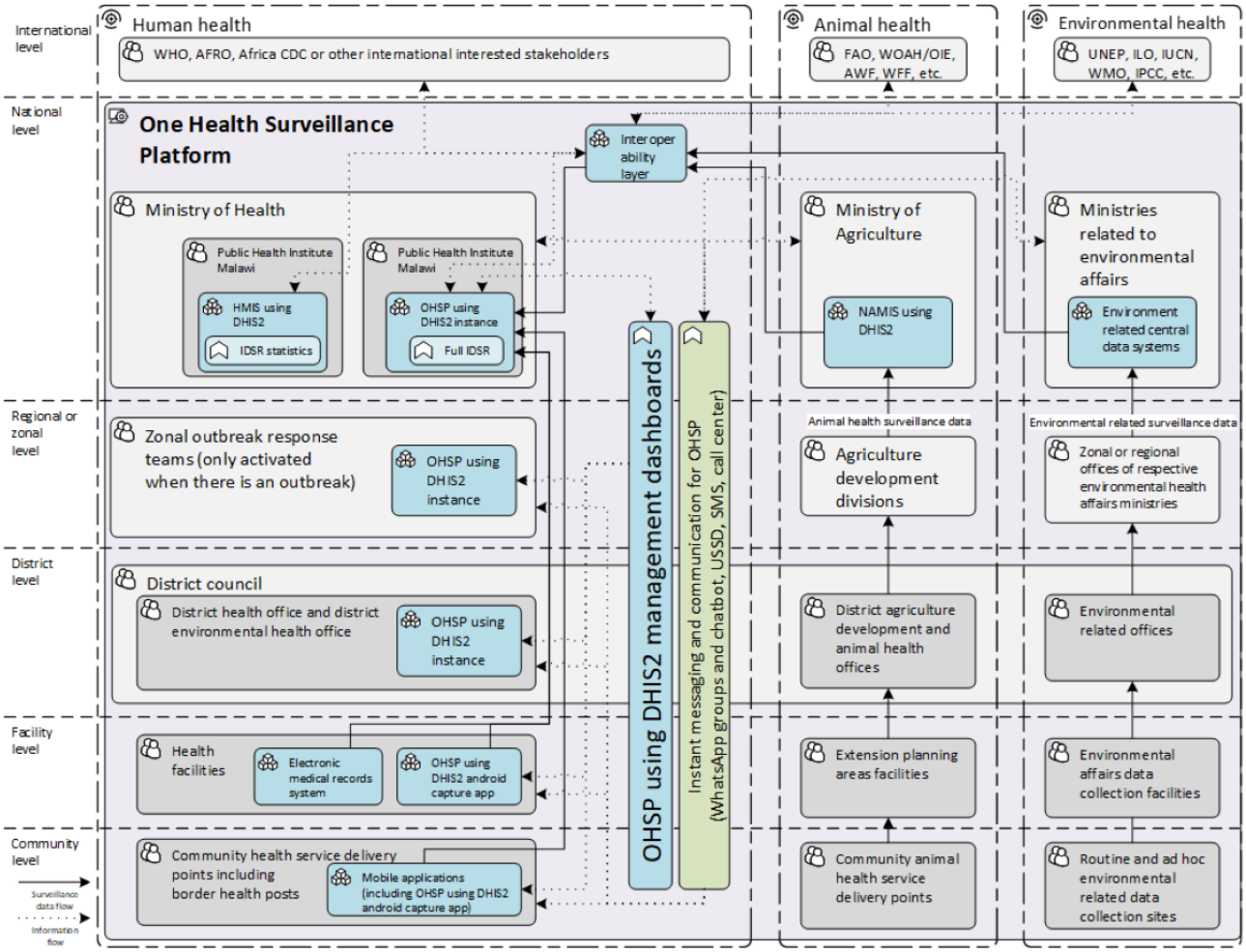
After the high-level business process architecture was developed, the core team focused on establishing the main data repository to enable facility-level periodic IDSR reporting and to prepare for integrating animal and environmental health surveillance data pipelines in the next phase. The core team developed a strategy to mobilize future surveillance-related investment for human resource retention and to continue the enhancement of the OHSP to anchor and align the One Health surveillance architecture and existing solutions. At the beginning of the OHSP establishment, UNICEF supported the initial OHSP development and committed to initial implementation covering 48.3% (14/29) of the country's health districts.

Once the high-level OHSP architectures were designed, the core team started working on the technical infrastructure for the establishment of the OHSP. Technical factors were considered to favor collaboration and integration of data from different domains, including DHIS2, as an existing national DHIS2-enabled digital health infrastructure used in HMIS, and an interoperability layer (IL), as an existing national information exchange infrastructure, to address these needs.

The core team started the development of the OHSP in November 2019. In terms of roles, the technical staff were responsible for software development and deployment. The first author worked as a technical advisor serving leading and mediating functions in the communication between different stakeholders, including the requirements and limitations from different disciplines to the technical staff. Using a transdisciplinary approach, he translated the business process, information flow, and available tools into programmable technical requirements for developers. After the technical execution, the core team designed the technical architecture for the OHSP, as shown in Figure 3.

XSL•FO
RenderX

Figure 3. Technical architecture of the One Health Surveillance Platform (OHSP) after the transdisciplinary translation from public health needs to programmable digital health solutions in Malawi, 2019. AFRO: World Health Organization Regional Office for Africa; AWF: African Wildlife Foundation; CDC: Centers for Disease Control and Prevention; DHIS2: District Health Information Software 2; FAO: Food and Agriculture Organization; HMIS: health management information system; IDSR: integrated disease surveillance and response; ILO: International Labour Organization; IPCC: Intergovernmental Panel on Climate Change; IUCN: International Union for Conservation of Nature; NAMIS: National Agriculture Management Information System; UNEP: United Nations Environment Programme; USSD: unstructured supplementary service data; WFF: World Wildlife Fund; WHO: World Health Organization; WMO: World Meteorological Organization; WOAH: World Organisation for Animal Health.



Based on the architectural design, the core team technical staff applied the Agile method, executing Scrum sprints to build the OHSP digital products. An instant messaging and communication collaboration mechanism was established and included in the OHSP. The One Health forum WhatsApp group was created for all IHR focal persons from different ministries to collaborate on health emergency preparedness and responses for real-time communications during the 2018-2019 anthrax outbreak response and was managed by the MoH. The functional One Health forum WhatsApp group served as one of the digital solutions in the OHSP, providing alerts to the group members. Similar WhatsApp groups have existed to facilitate IDSR coordination and collaboration since 2018. The experts in the One Health forum WhatsApp group anticipated that COVID-19 would become a great threat to the nation's public health in mid-January 2020. The signal was provided to the core team for the team to focus on leveraging available infrastructure resources and reusing technologies from previous eIDSR pilots to develop the following three DHIS2-related digital products: (1) a national DHIS2 instance for the OHSP, (2) a DHIS2 Android Capture app for OHSP, and (3) an interoperability mediator between HMIS and OHSP.

The primary focus during January 2020 was setting up the national DHIS2 instance, especially the metadata design for weekly, monthly, and quarterly IDSR reporting. The core team paid attention to building the OHSP with integration and interoperability infrastructure, especially in relation to the HMIS. The IDSR periodic report and case-based surveillance form were standardized to provide appropriate data and replicated on the OHSP to match those in the HMIS to maintain data integrity. The integration was done by first matching the data element IDs, including the data field IDs and the dataset IDs. The organizational unit IDs were mapped and aligned with national standards using a master health facility registry. The team cloned existing facility mappings and data element mapping identification from the existing HMIS instance and registered the OHSP as a client in the nation's IL so that it could recognize the OHSP instance as a source for data exchange.

Technical solutions were applied to ensure graceful degradation and seamless reintegration, including the DHIS2 features of PostgreSQL (PostgreSQL Global Development Group) performance tuning to improve query performance and scalability. The MoH-owned Proxmox (Proxmox Server

XSL•FO
RenderX

Solutions GmbH) cluster was created to tolerate server downtime without affecting user access through multiple web servers. Load balancing within Nginx (F5 Inc) was configured to distribute load evenly across server instances and to detect any unavailable instances and reroute traffic as needed. The core team applied the "sticky sessions" solution of the DHIS2 to help maintain user session continuity by routing requests from the same client to the same server. The core team configured daily data backups of the OHSP to ensure data restoration capabilities. Fire and water safety controls and a continuous power backup supply at the national digital health server room were applied for the OHSP to operate continuously with minimal downtime. For security, the OHSP DHIS2-related products used password protection and controlled access.

In February 2020, the team concentrated on ensuring HMIS-OHSP interoperability. When building the interoperability, an initial script was created to directly transfer the IDSR data from the HMIS instance into the OHSP. However, when running the script, there was a time-out due to the huge payload that was being sent from HMIS to OHSP; hence, the team decided to use the IL for chunking functionality. Process-wide, the IDSR periodic reporting process was not standardized; some facilities reported in HMIS, while some reported in the OHSP. The MoH leadership addressed the issue by disseminating an internal memo to regulate all surveillance data to be entered in OHSP. The core upgraded the HMIS and OHSP instance, so the payload was channeled through the IL to use the chunking of the payload when sending it from the OHSP to the HMIS.

Simultaneously, the team worked on the integration and user interface design of the DHIS2 Android Capture app to interact with the OHSP instance for data collection and reporting. Also, the team worked on obtaining additional public IP addresses with the government domain for OHSP to facilitate Secure Sockets Layer acquisition and deployment of the solutions through the internet. By March 2020, OHSP was tested by PHIM and was fully ready for implementation to support Malawi's surveillance and response needs.

During the COVID-19 Pandemic BIE Cycle

Right before the core team was ready to deploy the OHSP products, new requests emerged as the country faced the impending threats from the COVID-19 pandemic. According to the original plan, the first release of the OHSP products was to facilitate timely IDSR periodic aggregated data reports. However, with the emergency, the MoH and the national COVID-19 response team at the emergency operation center raised requests and digital health solutions needs, which required additional functions, integration, and interoperability between the OHSP products and other digital health solutions.

The COVID-19 response emergency operation center was formed according to the Africa Centres for Disease Control and Prevention recommendations in April 2020. The first author, TSJW, was appointed acting manager at the COVID-19 emergency operation center that same month, supporting response activities while continuing to serve on the core team for OHSP development. The core team responded to the requests raised and prioritized them according to the consensus and guidance from the government by leveraging available digital solutions and the IL. The case-based surveillance form of the IDSR system in OHSP was harmonized and standardized with various stakeholders' expectations and updated with COVID-19 conditions in May 2020. Additional forms for border health declaration and contact tracing were developed and added to the OHSP during May-June 2020. The validation rules for mandatory data entry were programmed in the OHSP to ensure data quality at the same time. After the new features of OHSP products were developed, they were included in the new national emergency preparedness and response plans to support the response from May 2020.

The core team applied the agreed strategy to mobilize resources for the national scale-up of OHSP instead of only implementing it in 14 districts when funders and partners showed interest in investment resources for pandemic response. Through the government leadership and collaboration mechanism, multiple funders accepted the advocacy at the beginning of the COVID-19 pandemic, and the core team secured access to additional technical infrastructure resources, especially the server, terminal devices (tablets), and reverse billing approach of connectivity, for deploying the OHSP products countrywide since June 2020. However, the team was unable to mobilize adequate resources for comprehensive redundancy and backup servers of OHSP DHIS2-related products till the end of the pandemic in May 2024.

During the national rollout, all partners followed the MoH leadership to provide standardized training, follow-up supervision, and refresher training. A training program for OHSP products users was designed and implemented comprehensively with user manuals, standardized presentations, a deployment checklist, and detailed schedules, covering computer literacy from the beginner level. A centralized digital health helpdesk at the Digital Health Division was established to provide support to OHSP users. The core team members received further education and training for their DHIS2-related capacity building through online and in-person participation in DHIS2 academies from 2021 to 2023.

A national OHSP enhancement plan was developed to institutionalize OHSP, and a dedicated technical working group was established to facilitate One Health surveillance and regular updates and the One Health bulletin at PHIM in 2023. A comprehensive capacity-building plan and a system infrastructure enhancement plan of OHSP were developed to integrate animal health and environmental health data pipelines in 2024. After the deployment, the OHSP core team has constantly collected the needs and feedback from relevant stakeholders through the established collaboration mechanism. Since then, the OHSP digital products have been sustained and continue to be used for IDSR reporting to date.

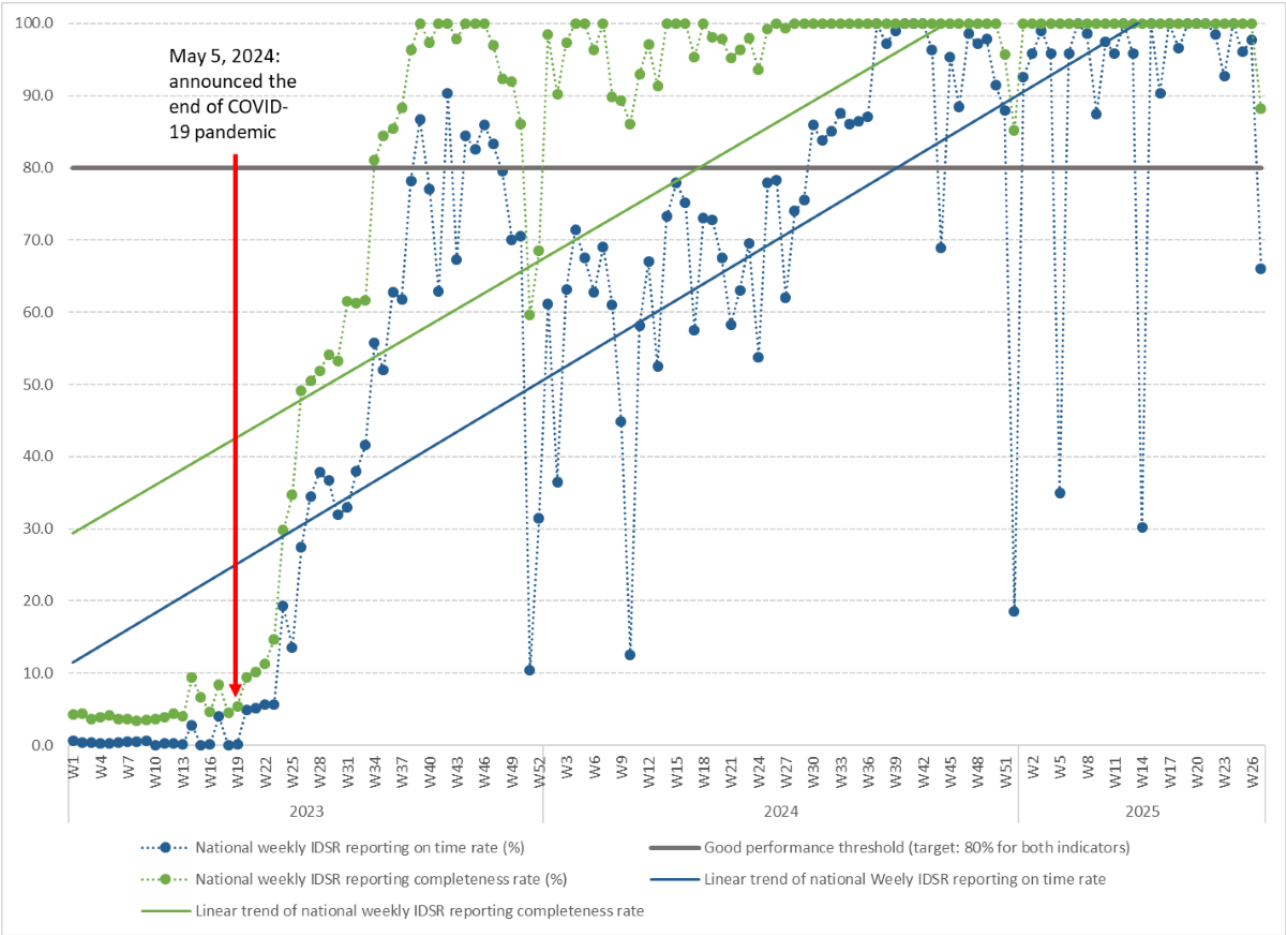
Post–COVID-19 Pandemic OHSP's Impact on IDSR Reporting

The OHSP became the main digital data repository for all IDSR surveillance periodic reports after the pandemic. The interoperability mediator is functioning with the MoH as the custodian, ready for reuse and tapping into other One Health surveillance domains' data. The One Health forum WhatsApp

XSL•FO
RenderX

group is still functioning to facilitate the ongoing One Health surveillance activities. The completeness and timeliness of IDSR periodic reports' data quality indicators were defined according to the national guideline and configured in the OHSP's national DHIS2 instance. The MoH continuously provides appropriate feedback on data quality issues through the dashboards, bulletins, and WhatsApp groups to reporting facilities to ensure data quality, with the OHSP core team maintained within the MoH governance structure. The OHSP dashboards developed provide functionality to monitor the epidemiological patterns of notifiable diseases routinely for outbreak detection. As a result, the weekly IDSR reporting transformed from nonexistence in 2015 to a state where the nation reached 97.8% and 74.5% reporting quality for completeness and timeliness, respectively, in 2024. The secular trends and weekly IDSR reporting rate quality are illustrated in Figure 4.

Figure 4. The integrated disease surveillance and response (IDSR) weekly reports completeness and timeliness indicators from January 2, 2023, (epidemiological week 1) to June 29, 2025, (epidemiological week 26) in Malawi.



Discussion

Summary of Main Findings

This study aims to report on the establishment of the OHSP in Malawi and how an adaptive digital health solution contributed to strengthening and impacting the country's IDSR system during the COVID-19 pandemic and beyond the pandemic. Our findings show that Malawi successfully established a government-led digital tools–aided OHSP right before the COVID-19 pandemic, building on the ongoing effort and further strengthening the country's IDSR system. The OHSP was built corresponding to the characteristics of One Health surveillance systems across policy, institutional, and operational levels [1,2]. Amid the COVID-19 pandemic, the MoH evidently used the OHSP to mitigate the spread and impact of the pandemic [21], and the weekly IDSR reporting improved significantly after the pandemic, from zero reporting in 2015 [11] to 97.8% completeness and 74.5% timeliness by 2024, while PHIM is currently undertaking a full impact assessment of the OHSP.

Findings from this study indicate that the alignment of the national health policies and the establishment of a formal coordination body with government leadership were essential to the One Health surveillance governance. The government-owned and government-led strategy was useful for communicating with all potential funders to invest in the OHSP and align with the government's direction.

The developed holistic OHSP architecture guided the data and information integration design, especially on organizational units and reporting tools standardization. The mapping of the corresponding organizational units with the human health system hierarchy at each administrative level for animal and environmental domains, and having standardized data elements, was a complex task. Considering the complexity, the OHSP core team sensitized all stakeholders engaged in helping with the proper mapping of the units and understanding of the

XSL•FO
RenderX

importance of the data standards for future One Health surveillance data exchange and interoperability.

The reuse of the eIDSR pilots' hardware and available digital solutions provided adequate hardware, software, and communication infrastructure for the initial establishment of the OHSP and accelerated in responding to the COVID-19 pandemic needs. Several known DHIS2 implementation challenges were addressed during the establishment of the OHSP. The OHSP anticipation and adaptive resilience abilities were observed in the study. A detailed discussion, comparison, and contrast of findings with existing literature is followed.

Interpretations and Implications of Lessons Learnt

One Health Surveillance Characteristics

Our findings were reflected based on four thematic areas using the analytic tool developed (Table 1) under the One Health surveillance characteristics domain, corresponding to the policy, institutional, and operations levels according to the One Health surveillance framework developed by Bordier et al [2].

Policy and Governance at the Policy and Institutional Levels

As known challenges, robust policy frameworks and governance structures are essential factors toward a successfully integrated One Health surveillance system [2,8-10]. The third edition of the IHR 2005 [30] was the main international legal instrument used to justify the One Health surveillance initiative in Malawi. The Joint External Evaluation conducted in Malawi in 2019 highlights the importance of operationalizing and strengthening the One Health platform, improving surveillance systems, and enhancing the data-sharing platform [15], providing an operational framework for Malawi to take action.

The Public Health Act (1948) [31] and the Control and Diseases of Animals Act (1967) [32] provide legal frameworks for disease surveillance, reporting, and control in both human and animal health sectors in Malawi. The Environment Management Act (2017) [33] complements the One Health approach by promoting environmental protection, monitoring, reporting, and public participation in decision-making processes. Malawi's health sector policy, especially the HSSP III [34], highlights the adoption of the One Health approach and OHSP, particularly in areas of antimicrobial resistance and climate change, echoing the One Health surveillance focuses found by Bordier et al [2] and Malawi's emerging needs [17]. At present, the OHSP is institutionalized at PHIM in Malawi to strengthen the IDSR reporting quality, avoiding potential abandonment of the digital health initiative after the COVID-19 temporary use [29]. The developed OHSP enhancement plan and HSSP III [34] guide the gradual integration of different One Health domain data pipelines using the OHSP digital tools. These are useful legal instruments that solidify the government-led coordination mechanism at the policy and institutional level, as advocated in relevant literature [2,5,6], and may be applied in other settings.

Concerning the governance and operation of collaborative surveillance across different sectors, the political will is essential in establishing a One Health surveillance system and may foster its success [2,9,10,35]. Our findings clearly indicated that Malawi showed strong political will, solidified soon after the successful anthrax outbreak response, and gained momentum from the resilient response to the COVID-19 pandemic. The 2 instances affirmed the leadership role of PHIM and strengthened the commitment of the government to take the same One Health approach for future health emergencies. As advocated, government-owned and led digital health interventions offer an opportunity to strengthen health systems and attain health-related goals on the continent [36]. The reinforced government commitment to OHSP's development had effectively reduced the creation of "silos" in the health surveillance system. Political will and leadership are the two key enabling factors contributing to the sustainable operation of the One Health surveillance system, echoing the findings from Vietnam and Australia's experience [10,35]. Coordinated implementation with government leadership maintained continuity and OHSP ownership, further ensuring financial and organizational sustainability in Malawi at the institutional level [2,26].

Sustainability for All Levels

The affirmation of the One Health leadership, established coordination mechanisms, and policy alignment secured the organizational sustainability of Malawi's One Health surveillance initiatives, reflecting the advocacies from others [2,9,10]. The government officials at the forefront led the implementation and delivery of the OHSP digital products to avoid district and facility health care workers conceiving the OHSP as an external system. This approach further reduced operational costs and silo creation, ensuring long-term sustainability as argued by academics in the same field [2,8,10,35]. Meanwhile, the reuse of technology strategy with open-source digital solutions, such as the global DHIS2 and Open Health Information Exchange communities [25,37], provides technical advancement and long-term software sustainability, coupled with the ongoing maintenance by the core team, making the OHSP a resilient information system [26].

Concerning financial sustainability, the strategy of advocating and lobbying for surveillance-related investment to anchor and align the One Health surveillance architecture was a useful approach for the OHSP to scale up nationwide, with the COVID-19 pandemic providing the MoH with an opportunity to align and direct partners' investment to support it. Government ownership and leadership of the OHSP avoid the potential interruption, silos, or hiccup of funding availability, responding to the challenges observed and concepts reported before [2,5,9,10,35]. The OHSP, as a government system, with or without external financial support, requires all civil servants to commit to making it work.

Data Integration at the Operational Level

Most of the One Health surveillance systems in action are focused on zoonotic diseases [38,39] for data to be integrated with human surveillance. The established OHSP can swiftly connect with animal and environmental health surveillance data with its adaptability. In fact, the recent development of the National Agriculture Management Information System (NAMIS) in Malawi also uses the same DHIS2 technology to support the

XSL•FO
RenderX

Ministry of Agriculture in improving livestock production, enhancing herd health, and addressing transboundary animal diseases [40]. The OHSP can provide its affordance to interoperate with NAMIS for surveillance needs. Environmental health, by the nature of its vast domain, has more work to do. Therefore, the OHSP enhancement plan focuses on climate change in alignment with HSSP III can be supported by the same DHIS2 solutions [41]. However, as per our findings, the standardization of organizational units to cover all One Health surveillance domains will be a key issue that needs to be addressed, ideally using a similar master facility registry for animal and environmental surveillance sites.

Technical Infrastructure Reuse at the Operational Level

The functional OHSP, as an effective information system, addressed Malawi's Joint External Evaluation report issues and lack of weekly IDSR reporting by operationalizing and strengthening the One Health platform [11,15]. The high-level architecture, inspired by the concept of the integrated health information system architecture [42] and its alignment with IDSR processes, supported the design of OHSP and the strategy to reuse the existing infrastructure. Although the collaboration with stakeholders during the COVID-19 pandemic secured adequate technical infrastructure for the nationwide OHSP scale-up; however, the current collaborative modalities and technical support mechanisms are still limited in the human health domain, similar to other settings' One Health surveillance progress [38,43,44], where investment will be needed in other domains.

Digital health tools have proven instrumental in improving resilience and maintaining essential health services during crises [21]. However, the challenges noted in the infrastructure, particularly the lack of comprehensive redundancy and the limitations of the server room, point to the need for continuous investment, as highlighted by engineering resilience [26], to have a robust technical infrastructure to ensure OHSP's scalability and reliability.

Corresponded to Known DHIS2 Implementation Challenges

Findings were mainly reflected and compared with Dehnavieh et al [25].

Political, Cultural, Social, and Structural Infrastructure

The legal instruments, MoH's leadership, and institutionalization of the OHSP at PHIM addressed known politically related challenges [25]. However, as discussed earlier, sustaining stable financing for technical infrastructure remains a challenge.

Appropriate Data—Quality and Standards

The standardized IDSR data collected through the OHSP—including updates and additions introduced during the COVID-19 pandemic—along with validation rules for mandatory data entry, defined data quality indicators, data quality monitoring dashboards, and a weekly epidemiological bulletin ensured appropriate surveillance data supplies. These results directly addressed the challenges identified in Dehnavieh et al [2,25].

Workforce and Capacity Building

Malawi has a relatively adequate workforce for the development and implementation of OHSP DHIS2-related products, as evidenced by the constitution of the core team consisting of government officials and technical support officers from partner organizations who programmed the DHIS2 solutions locally. The OHSP enhancement plan, capacity-building plan, and online DHIS2 academies [45] addressed the human resource challenges [25].

Health System Resilience—Anticipation and Adaptive Resilience

Concerning the health system resilience, Malawi's health emergency awareness capacity was enhanced through real-time communication using WhatsApp groups for timely information sharing and decision-making. The use of such a digital solution with the One Health approach corresponds to the Riyadh Declaration to improve preparedness and response to future pandemics [46]. However, although the MoH manages these WhatsApp groups, risks remain for intentional information leakage by group members.

The adaptability of the OHSP was evident in the rapid response to COVID-19, with adjustments made to the system based on emerging needs. This finding echoes the recent organizational resilience studies during COVID-19 about adaptation, innovation, capabilities, and the use of new technology [47]. The Agile software development method, a skilled and committed core team, and affordable digital tools made the success possible.

The result demonstrated that Malawi's OHSP is a resilient system with high adaptability that supports the surveillance goals during the COVID-19 pandemic [21]. The OHSP core team consistently managed to build effective governance with strong government leadership to provide clear policy directions, which are essential for building resilient health systems [1,4]. The government's strong leadership in the establishment of the OHSP paved the way for OHSP digital health solutions to be developed with resilience and rapidly scaled nationwide, echoing conclusions on Africa's digital transformation [36].

Remaining Gaps and Future Directions

The OHSP addressed some identified barriers to support the IDSR system [11]; however, the current functionality has yet to address the community-level surveillance gaps fully [48]. Event-based surveillance (EBS) is one of the recommended surveillance methods to support community-level surveillance, as stated in the recent IDSR technical guidelines [12]. Other countries' EBS implementation experiences and practices encountered some challenges, including the trade-off between sensitivity and specificity, low awareness of EBS, and inadequate infrastructure [49,50]. The role of middle-layer actors in participating in community-level surveillance, community engagement, and additional digital solutions should be considered as the OHSP is enhanced to support the EBS in the country, not only for human health but also animal, including wild and domestic, and the surrounding environment as a whole [1,48,50]. Engagement with stakeholders using existing coordination mechanisms and applying the participatory

XSL•FO
RenderX

approach should be encouraged for future OHSP enhancement [35]. Despite Malawi managing to address the OHSP establishment challenges through strategic planning and implementation, continuous capacity building, training, supervision, and retaining skilled personnel for the OHSP to be resilient still exist in Malawi with constrained resources, similar to the challenges that Tanzania encountered trying to establish a similar platform [51]. Mobilizing and lobbying resources to fully operationalize the OHSP enhancement plan, with highlighting multisectoral benefits as argued by Fasina et al [44], is the recommended direction.

Limitations

Our study focused on the establishment experience of the OHSP in Malawi; despite its impact on the weekly IDSR reporting quality, its further use, linkage to other One Health domains, and its impact on the overall One Health surveillance are limited. Regarding the health system resilience, although the frameworks were developed [5,6], they are limited to emergency management and health care systems, with none specifically for One Health surveillance systems [26,52,53]. The synthesis analytic tool we developed for the OHSP establishment analysis can be used in similar settings for assessing the One Health surveillance system, despite the fact that our analysis was limited to the One Health surveillance system establishment stage; further use of the tool is encouraged.

Conclusions

The success of Malawi's OHSP establishment highlights the importance of government leadership and coordination in ensuring financial and organizational sustainability. Key findings informed the need for continuous capacity building, adaptable technical infrastructure, and enhanced community-level surveillance in the long run.

As global health systems continue to face emerging threats to respective contexts, Malawi's experience demonstrates that a well-integrated, government-led digital OHSP can enhance resilience and preparedness to these health challenges. Amid the dynamic of the global development policy changes and complex human-animal-environmental interactions, findings from this study can be applied to broader audiences to enhance their surveillance and health systems in combating the pandemics to come.

Future investments, with emphasis on local leadership and domestic financial resources, in interoperability, infrastructure, and community engagement, will be essential for maximizing the impact of OHSP and similar initiatives worldwide.

Acknowledgments

We acknowledge support from multiple funding and implementing partners whose contributions supported the establishment of the One Health Surveillance Platform in Malawi before and during the COVID-19 pandemic; including United Nations International Children's Emergency Fund, the Gates Foundation through Kuunika Data for Action Project via Cooper/Smith, United States Agency for International Development/United States President's Emergency Plan for AIDS Relief through Elizabeth Glaser Pediatric AIDS Foundation, Management Sciences for Health, International Training and Education Center for Health, World Bank, and the Global Fund. We also acknowledge the useful and resourceful technical support and contribution from the District Health Information Software 2 global community and the Health Information Systems Programme Centre at the University of Oslo. We extend our sincere gratitude to the senior management team of the Ministry of Health for their strong leadership and unwavering commitment to realizing the establishment of the One Health Surveillance Platform.

The authors declare that Grammarly was used as a machine tool to suggest language improvement within the language. No other generative AI tools were used in this manuscript.

Funding

The funding support to this study of the One Health Surveillance Platform establishment and learning was supported by the Pingtung Christian Hospital (grants PS-IR-109001, PS-IR-111001, and PS-IR-111001), Taiwan, through Luke International, Norway, as part of the first author's PhD work.

Data Availability

All meeting minutes and internal notes are available upon request. The One Health Surveillance Platform–related source codes and documentation, for public access according to the Ministry of Health regulations, are available in the following GitHub repositories: (1) Ministry of Health of Malawi in production services GitHub repository [54] and (2) Ministry of Health of Malawi Project-Based GitHub repository [55].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Detailed analytic tool for health system resilience reflection and learning of the One Health Surveillance Platform establishment. [XLSX File (Microsoft Excel File), 26 KB - publichealth_v12i1e72029_app1.xlsx]

XSL•FO
RenderX

References

1. Okello A. One Health: a cross-sectoral, holistic approach to reduce disease risk and build resilience. Development Co-operation Report 2020: Learning from Crises, Building Resilience. Paris: OECD Publishing; 2020. URL: https://www.oecd.org/en/publications/development-co-operation-report-2020_f6d42aa5-en.html [accessed 2026-01-08]
2. Bordier M, Uea-Anuwong T, Binot A, Hendrikx P, Goutard FL. Characteristics of one health surveillance systems: a systematic literature review. Prev Vet Med 2020;181:104560. [doi: 10.1016/j.prevetmed.2018.10.005] [Medline: 30528937]
3. Strengthening WHO preparedness for and response to health emergencies. World Health Organization. 2021. URL: https://apps.who.int/gb/ebwha/pdf_files/WHA74/A74_R7-en.pdf [accessed 2025-12-16]
4. Strengthening collaboration on One Health. World Health Organization. 2022. URL: https://apps.who.int/gb/ebwha/pdf_files/WHA75/A75_19-en [accessed 2025-12-16]
5. Thomas S, Sagan A, Larkin J, Cylus J, Figueras J, Karanikolos M. Strengthening health systems resilience: Key concepts and strategies (Policy Brief 36). European Observatory on Health Systems and Policies. 2020. URL: https://iris.who.int/handle/10665/332441 [accessed 2026-01-08]
6. Foroughi Z, Ebrahimi P, Aryankhesal A, Maleki M, Yazdani S. Toward a theory-led meta-framework for implementing health system resilience analysis studies: a systematic review and critical interpretive synthesis. BMC Public Health 2022;22(1):287 [FREE Full text] [doi: 10.1186/s12889-022-12496-3] [Medline: 35151309]
7. Budd J, Miller BS, Manning EM, Lampos V, Zhuang M, Edelstein M, et al. Digital technologies in the public-health response to COVID-19. Nat Med 2020;26(8):1183-1192. [doi: 10.1038/s41591-020-1011-4] [Medline: 32770165]
8. Stärk KDC, Arroyo Kuribreña M, Dauphin G, Vokaty S, Ward MP, Wieland B, et al. One Health surveillance--more than a buzz word? Prev Vet Med 2015;120(1):124-130. [doi: 10.1016/j.prevetmed.2015.01.019] [Medline: 25722055]
9. Bordier M, Delavenne C, Nguyen DTT, Goutard FL, Hendrikx P. One Health surveillance: a matrix to evaluate multisectoral collaboration. Front Vet Sci 2019;6:109 [FREE Full text] [doi: 10.3389/fvets.2019.00109] [Medline: 31106210]
10. Johnson I, Hansen A, Bi P. The challenges of implementing an integrated One Health surveillance system in Australia. Zoonoses Public Health 2018;65(1):e229-e236 [FREE Full text] [doi: 10.1111/zph.12433] [Medline: 29226606]
11. Joseph Wu T, Kagoli M, Kaasbøll JJ, Bjune GA. Integrated disease surveillance and response (IDSR) in Malawi: implementation gaps and challenges for timely alert. PLoS One 2018;13(11):e0200858 [FREE Full text] [doi: 10.1371/journal.pone.0200858] [Medline: 30496177]
12. Technical guidelines for integrated disease surveillance and response in the African region: Third edition. World Health Organization. Regional Office for Africa. 2019. URL: https://iris.who.int/items/9868f3f2-6971-425a-b2b4-7e56cd9dff3a [accessed 2026-01-08]
13. Implementation completion and results report for Malawi Nutrition and HIV/AIDS Project (P125237). World Bank. 2019. URL: https://documents1.worldbank.org/curated/en/099616005202218266/pdf/IDU00fbcc4350092f04a70098b10c7c802c215b0.pdf [accessed 2025-12-16]
14. Guerra J, Davi KM, Rafael FC, Assane H, Imboua L, Diallo FBT, et al. Case study of Argus in Togo: an SMS and web-based application to support public health surveillance, results from 2016 to 2019. PLoS One 2020;15(12):e0243131 [FREE Full text] [doi: 10.1371/journal.pone.0243131] [Medline: 33259550]
15. Joint external evaluation of IHR core capacities of the Republic of Malawi. World Health Organization. 2019. URL: https://www.who.int/publications/i/item/WHO-WHE-CPI-2019.58 [accessed 2025-12-16]
16. Health Sector Strategic Plan II. World Health Organization. 2017. URL: https://extranet.who.int/countryplanningcycles/sites/default/files/planning_cycle_repository/malawi/health_sector_strategic_plan_ii_030417_smt_dps.pdf [accessed 2025-12-16]
17. Government of Malawi. Malawi 2019 floods post disaster needs assessment (PDNA). Government of Malawi. 2019. URL: https://www.unicef.org/malawi/reports/malawi-2019-floods-post-disaster-needs-assessment-report [accessed 2026-01-08]
18. IANPHI in the new decade - the 2019 IANPHI annual meeting and our priorities for the year ahead. IANPHI. 2020. URL: https://ianphi.org/_includes/documents/sections/news/insider-february-2020.pdf [accessed 2025-12-16]
19. Moyo C, Nkhonjera T, Kaasbøll J. Assessing the use of mobile technology to improve timeliness of HMIS reports in Malawi. : IEEE; 2015 Presented at: IST-Africa Conference; 2015 May 6-8; Lilongwe, Malawi. [doi: 10.1109/istafrica.2015.7190573]
20. Barasa E, Mbau R, Gilson L. What is resilience and how can it be nurtured? A systematic review of empirical literature on organizational resilience. Int J Health Policy Manag 2018;7(6):491-503 [FREE Full text] [doi: 10.15171/ijhpm.2018.06] [Medline: 29935126]
21. Mghamba J, Gilmour E, Robinson L, Simba A, Tuyishime A, Persaud A, et al. The use of innovative approaches to strengthen health system resilience during the COVID-19 pandemic: case studies from selected Commonwealth countries. Front Public Health 2023;11:1115415 [FREE Full text] [doi: 10.3389/fpubh.2023.1115415] [Medline: 37181718]
22. Sein MK, Henfridsson, Purao S, Rossi M, Lindgren R. Action design research. MIS quarterly 2011;35(1):37-56. [doi: 10.2307/23043488]
23. National Statistical Office. Malawi Statistical Yearbook. Zomba: National Statistical Office; 2015.
24. Ministry of Health. Technical Guidelines for Integrated Disease Surveillance and Response in Malawi. 3rd ed. Geneva: World Health Organization; 2014.

XSL•FO
RenderX

25. Dehnavieh R, Haghdoost A, Khosravi A, Hoseinabadi F, Rahimi H, Poursheikhali A, et al. The district health information system (DHIS2): a literature review and meta-synthesis of its strengths and operational challenges based on the experiences of 11 countries. Health Inf Manag 2019;48(2):62-75. [doi: 10.1177/1833358318777713] [Medline: 29898604]
26. Vecchiola C, Anjomshoa H, Bernstein Y, Dumitrescu I, Garnavi R, von Känel J, et al. Engineering resilient information systems for emergency management. IBM J Res Dev 2013;57(5):1-2. [doi: 10.1147/jrd.2013.2259432]
27. Folke C, Carpenter SR, Walker B, Scheffer M, Chapin T, Rockström J. Resilience thinking: integrating resilience, adaptability and transformability. Ecol Soc 2010;15(4):1-2. [doi: 10.5751/es-03610-150420]
28. Asah F, Chikumba P, Kaasbøll J, Manya A, Moyo C, Msiska B. From sustainable to continuously adaptive information systems. 2018 Presented at: IST-Africa Week Conference; 2018 May 9-11; Gaborone, Botswana.
29. Firouzi F, Farahani B, Daneshmand M, Grise K, Song J, Saracco R, et al. Harnessing the power of smart and connected health to tackle COVID-19: IoT, AI, robotics, and blockchain for a better world. IEEE Internet Things J 2021;8(16):12826-12846 [FREE Full text] [doi: 10.1109/JIOT.2021.3073904] [Medline: 35782886]
30. WHO. Doom: Doom 3, Id Tech 1, Doom II: Hell on Earth, Bestiaire de Doom, Id Tech 4, Final Doom, Doom 3: Resurrection of Evil, Whe. Geneva: World Health Organization; 2014:84.
31. Public Health Act (1948). MalawiLLI. 1948. URL: https://malawilii.org/akn/mw/act/1948/12/eng@2014-12-31 [accessed 2025-12-16]
32. Control and Diseases of Animals Act, Malawi, (1967). MalawiLLI. URL: https://malawilii.org/legislation/ [accessed 2026-01-08]
33. Environment Management Act, (2017). MalawiLLI. 2017. URL: https://malawilii.org/legislation/ [accessed 2025-12-16]
34. Ministry of Health. Health Sector Strategic Plan III (2023–2030): accelerating the move towards Universal Health Coverage. In: Malawi Health Sector Strategic Plan III, 2023-2030. Lilongwe: Government of Malawi; 2023.
35. Bordier M, Goutard FL, Antoine-Moussiaux N, Pham-Duc P, Lailler R, Binot A. Engaging stakeholders in the design of one health surveillance systems: a participatory approach. Front Vet Sci 2021;8:646458 [FREE Full text] [doi: 10.3389/fvets.2021.646458] [Medline: 34109232]
36. Tran Ngoc C, Bigirimana N, Muneene D, Bataringaya J, Barango P, Eskandar H. Conclusions of the digital health hub of the Transform Africa Summit (2018): strong government leadership and public-private-partnerships are key prerequisites for sustainable scale up of digital health in Africa. BMC Proc 2019;13(Suppl 9):13. [doi: 10.1186/s12919-019-0173-6]
37. A community that builds together. OpenHIE Community. URL: https://ohie.org/community/ [accessed 2025-12-16]
38. Latinne A, Nga NTT, Long NV, Ngoc PTB, Thuy HB, Predict Consortium, et al. One health surveillance highlights circulation of viruses with zoonotic potential in bats, pigs, and humans in Viet Nam. Viruses 2023;15(3) [FREE Full text] [doi: 10.3390/v15030790] [Medline: 36992498]
39. Innes GK, Lambrou AS, Thumrin P, Thukngamdee Y, Tangwangvivat R, Doungngern P, et al. Enhancing global health security in Thailand: strengths and challenges of initiating a One Health approach to avian influenza surveillance. One Health 2022;14:100397 [FREE Full text] [doi: 10.1016/j.onehlt.2022.100397] [Medline: 35686140]
40. Mwakabira I, Byson L, Manda T, Phiri Y, Simwela A. Design gaps in configurable systems: adaptability of DHIS2 to other domains. Design Gaps in Configurable Systems: Adaptability of DHIS2 to other Domains. 2023 IST-Africa Conference (IST-Africa); 2023 Presented at: IST-Africa Conference; 2023 May 31; Tshwane, South Africa. [doi: 10.23919/ist-africa60249.2023.10187872]
41. Katwesige J, Nabende P, Kasasa S, Ssenyonga R, Nakafeero M, Kissa J. Integration of DHIS2 and climatic data for disease prediction: a case of malaria occurrence in Gulu district, Uganda. 2020 Presented at: IST-Africa Conference (IST-Africa); 2023 May 20; Kampala, Uganda.
42. Braa J, Sahay S. Integrated Health Information Architecture Power to the Users. Mumbai: Matrix publishers; 2012.
43. Oberin M, Badger S, Faverjon C, Cameron A, Bannister-Tyrrell M. Electronic information systems for One Health surveillance of antimicrobial resistance: a systematic scoping review. BMJ Glob Health 2022;7(1) [FREE Full text] [doi: 10.1136/bmjgh-2021-007388] [Medline: 34983786]
44. Fasina FO, Bett B, Dione M, Mutua F, Roesel K, Thomas L, et al. One Health gains momentum in Africa but room exists for improvement. One Health 2022;15 [FREE Full text] [doi: 10.1016/j.onehlt.2022.100428] [Medline: 36277101]
45. Academy-DHIS2. HISP UiO. URL: https://dhis2.org/academy/ [accessed 2025-12-16]
46. Al Knawy B, Adil M, Crooks G, Rhee K, Bates D, Jokhdar H, et al. The riyadh declaration: the role of digital health in fighting pandemics. Lancet 2020;396(10262):1537-1539. [doi: 10.1016/s0140-6736(20)31978-4]
47. Paeffgen T, Lehmann T, Feseker M. Comeback or evolution? Examining organizational resilience literature in pre and during COVID-19. CRR 2023;6(1):1-27. [doi: 10.1108/crr-07-2023-0012]
48. Wu T, Kumwanje M, Nyasulu Y, Kaasbøll J, Bjune G. Health seeking behaviour for acute respiratory illness among adults in rural Malawi: implication to emerging respiratory infections surveillance. Research Square 2023. [doi: 10.21203/rs.3.rs-2672808/v1]
49. Kuehne A, Keating P, Polonsky J, Haskew C, Schenkel K, Le Polain de Waroux O, et al. Event-based surveillance at health facility and community level in low-income and middle-income countries: a systematic review. BMJ Glob Health 2019;4(6):e001878 [FREE Full text] [doi: 10.1136/bmjgh-2019-001878] [Medline: 31908863]

XSL•FO
RenderX

50. Usman R, Bakare L, Akpan UO, Dalhat M, Dada AO, Okudo I, et al. Establishing event-based surveillance system in Nigeria: a complementary information generating platform for improved public health performance, 2016. Pan Afr Med J 2022;42:63 [FREE Full text] [doi: 10.11604/pamj.2022.42.63.29621] [Medline: 35949466]
51. Mremi IR, Rumisha SF, Sindato C, Kimera SI, Mboera LE. Comparative assessment of the human and animal health surveillance systems in Tanzania: opportunities for an integrated one health surveillance platform. Glob Public Health 2023;18(1):1-17 [FREE Full text] [doi: 10.1080/17441692.2022.2110921] [Medline: 35951768]
52. Garcia-Perez A, Cegarra-Navarro J, Sallos M, Martinez-Caro E, Chinnaswamy A. Resilience in healthcare systems: cyber security and digital transformation. Technovation 2023;121:102583 [FREE Full text] [doi: 10.1016/j.technovation.2022.102583]
53. Epizitone A. Framework to develop a resilient and sustainable integrated information system for health care applications: a review. Int J Adv Comp Sci Appl 2022;13(7). [doi: 10.14569/ijacsa.2022.0130758]
54. Malawi Ministry of Health. GitHub. URL: https://github.com/Malawi-Ministry-of-Health [accessed 2025-12-18]
55. Kuunika. GitHub. URL: https://github.com/Kuunika [accessed 2025-12-18]

Abbreviations

ADR: action design research
BIE: building, intervention, and evaluation
DHIS2: District Health Information Software 2
EBS: event-based surveillance
eIDSR: electronic integrated disease surveillance and response
EMRS: electronic medical record system
HMIS: health management information system
HSSP III: Health Sector Strategic Plan 2023-2030
IDSR: integrated disease surveillance and response
IHR: International Health Regulations
IL: interoperability layer
MoH: Ministry of Health
NAMIS: National Agriculture Management Information System
OHSP: One Health Surveillance Platform
PHIM: Public Health Institute of Malawi
UNICEF: United Nations International Children's Emergency Fund
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 05.Mar.2025; peer-reviewed by S Thawillarp, K Gooding; comments to author 22.Apr.2025; accepted 02.Dec.2025; published 26.Jan.2026.

Please cite as:
Wu TSJ, Mvula MV, Chado EKK, Kamanga BN, Mapemba DD, Billy RE, Nyirongo L, Mwale AC, Banda EC, Kagoli M, Manda TD, Bjune GA, Kaasbøll JJ
Establishing One Health Surveillance Platform for Electronic Integrated Disease Surveillance and Response in Malawi: Action Design Research Study
JMIR Public Health Surveill 2026;12:e72029
URL: https://publichealth.jmir.org/2026/1/e72029
doi:10.2196/72029
PMID:

©Tsung-Shu Joseph Wu, Matthew Vundu Mvula, Edward Kada Koma Chado, Blessings Nthezemu Kamanga, Daniel Denis Mapemba, Rajab Enoch Billy, Louis Nyirongo, Annie Chauma Mwale, Evelyn Chitsa Banda, Matthew Kagoli, Tiwonge Davis Manda, Gunnar Aksel Bjune, Jens Johan Kaasbøll. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 26.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Interactive Digital Visualization Counseling for Lifestyle Change in Patients at Risk of Cardiovascular Diseases: Randomized Controlled Trial

Adrijana Svenšek[1], PhD; Mateja Lorber[1], PhD; Zalika Klemenc-Ketiš[2,3,4], PhD; Lucija Gosak[1], PhD; Dominika Muršec[1], MSc; Gregor Štiglic[1,5,6], PhD

[1]Faculty of Health Sciences, University of Maribor, Žitna ulica 15, Maribor, Slovenia
[2]Primary Healthcare Research and Development Institute, Community Health Centre Ljubljana, Ljubljana, Slovenia
[3]Department of Family Medicine, Faculty of Medicine, University of Maribor, Maribor, Slovenia
[4]Department of Family Medicine, Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia
[5]Faculty of Electrical Engineering and Computer Science, University of Maribor, Maribor, Slovenia
[6]Usher Institute, University of Edinburgh, Edinburgh, United Kingdom

Corresponding Author:
Adrijana Svenšek, PhD
Faculty of Health Sciences, University of Maribor, Žitna ulica 15, Maribor, Slovenia

Abstract

Background: Cardiovascular disease (CVD) remains the leading cause of death. Primary prevention relies heavily on health risk assessments and lifestyle changes, which can reduce long-term risk and mortality. Digital health offers an accessible and cost-effective approach to support prevention, enabling data sharing and visualization of key indicators such as blood pressure and glucose fluctuations. These visual insights may help patients better understand the effects of lifestyle changes and enhance communication with health care providers.

Objective: This research aims to evaluate whether the use of CVD risk visualization (Petal-X) and continuous glucose monitoring (CGM), alone or in combination, is associated with lifestyle changes and the perception of person-centered care (PCC) among patients at increased risk of CVD.

Methods: We conducted a 4-arm, single-blind, 2×2 factorial randomized controlled feasibility trial in primary care. A total of 119 participants were enrolled, of whom 101 completed the 6-month follow-up. Participants were randomized to 1 of 4 arms: (1) Petal-X CVD risk visualization+CGM; (2) CGM only; (3) Petal-X only; or (4) standard care with routine lifestyle counseling and no digital tools. CGM was used for 10 days in the CGM arms. Since this was a feasibility trial, no formal sample size calculation was performed. Primary outcomes are healthy lifestyle (Health Lifestyle and Personal Control Questionnaire [HLPCQ]) and perception of PCC (Person-Centered Practice Inventory—Service User [PCPI–SU]), and secondary outcomes (Systematic Coronary Risk Evaluation 2 [SCORE2], anthropometrics, and biological age) were assessed at baseline and 6 months. Descriptive statistics and Kruskal-Wallis tests (*K* independent samples) were used for analyses.

Results: At baseline, mean SCORE2 values ranged from 3.84 (SD 2.08) in intervention group 3 to 4.87 (SD 2.61) in intervention group 1, with the control group having a mean value of 4.53 (SD 3.63). Regarding the assessment of a healthy lifestyle, the domain of daily routine had the highest baseline scores across all groups (eg, mean 19.24, SD 5.87 in intervention group 1), and these scores improved by the final evaluation, although there were no statistically significant differences (P=.42) in changes between the groups. The perception of PCC was rated highest across all groups in the domain of shared decision-making, with no statistically significant differences (P=.26) between the groups. Results indicated improvements in healthy lifestyle habits, but the impact of interventions on perceived changes remained insignificant.

Conclusions: Healthy lifestyle and perceived PCC scores improved, although no statistically significant between-group differences were found. Risk visualization appears to be a key tool for increasing CVD awareness and strengthening patient involvement in care planning. Longer interventions with larger samples are needed to clarify these effects and optimize digital tools for lifestyle change.

Trial Registration: ClinicalTrials.gov NCT06226948; https://clinicaltrials.gov/study/NCT06226948

(JMIR Public Health Surveill 2026;12:e83488) doi:10.2196/83488

XSL•FO
RenderX

KEYWORDS

biological age; cardiovascular disease; CGM; healthy lifestyle; Petal-X; RCT; visualization; continuous glucose monitoring; randomized controlled trial

Introduction

Chronic diseases are long-lasting and usually progressive in nature and not transmissible between people. They represent a major global health burden, both in terms of mortality and health care costs. There are 4 main types of chronic disease: cardiovascular disease (CVD), cancer, chronic respiratory diseases (such as chronic obstructive pulmonary disease and asthma), and diabetes. These chronic diseases are responsible for most deaths worldwide [1,2]. Importantly, these diseases share common modifiable lifestyle risk factors such as poor diet, physical inactivity, tobacco use, and harmful alcohol consumption [2-4]. In 2021, they were estimated to cause 20.5 million deaths, approximately a third of deaths worldwide, which affected more than half a billion people [5-8]. The prevalence of CVD is rising due to increases in obesity, hypertension, type 2 diabetes mellitus, and physical inactivity [5,9-12]. Given that unhealthy lifestyles are estimated to contribute to around 60% of chronic disease risk [3,5,13,14], managing behavioral and environmental risk factors remains one of the most effective approaches to alleviating the burden of CVDs. Targeted lifestyle interventions have been shown to reduce CVD incidence and related mortality over the long term [15,16]. According to Visseren et al [17], lifestyle improvement should be the first line of intervention, including dietary changes (eg, reducing salt, saturated fats, and alcohol) and physical activity. The guidelines divide patients into risk categories so that prevention can be targeted. The groups are: I-very high risk of CVD (>40% in the next 10 y), II-high risk of CVD (20% - 40% in the next 10 y), III-moderate risk of CVD (10% - 20% in the next 10 y), IV-low risk of CVD (<10% in the next 10 y) [18]. However, many individuals at risk for CVD do not understand their personal risk or the potential benefits of lifestyle change. Moreover, they may underestimate the consequences of inaction, which underscores the importance of interventions that enhance awareness and promote sustained behavioral change [19-21].

In this context, digital health has emerged as a promising tool. Technological innovations are increasingly being used to deliver scalable interventions aimed at improving health behaviors [22]. Smartphones, which are now owned by approximately 6.9 billion people worldwide (86% of the global population), have become a platform for delivering digital health interventions [23]. Evidence suggests that digital health applications can support lifestyle change, improve glycemic control [24,25], lower blood pressure (BP) [26], and even reduce hospitalizations when integrated into cardiac rehabilitation programs [27,28]. Digital technologies also offer new ways of presenting health information in an engaging and understandable way. Visual representations of risk (graphs, dendrograms, and interactive charts) can improve patients' understanding of their condition and motivate behavior change [8,29-31]. As noted by Visseren et al [17], absolute risk and risk reduction are more easily understood when presented visually, rather than numerically. Advanced tools such as decision support systems and continuous glucose monitoring (CGM) devices provide real-time, personalized insights into patient health. These technologies support prevention and treatment by integrating clinical data with evidence-based knowledge to enable personalized interventions [32]. For example, CGM systems not only help reduce the risk of hypoglycemia and hyperglycemia but also encourage behavior change by providing instant feedback on glucose levels [33-35]. However, the successful implementation of digital health tools depends on both patient engagement and health literacy [36]. Without the ability to understand and act on digital information, even the most advanced tools may have a limited effect. This is where person-centered care (PCC) becomes crucial. By involving patients as active partners in their own care and decisions, PCC has been shown to enhance outcomes across a range of conditions [37,38]. Research focuses on how risk for CVD can be effectively visualized using digital health tools, with a particular focus on CGM and its role in supporting lifestyle change. In line with the principles of PCC, the aim is to understand how these technologies can promote greater patient engagement and ultimately contribute to improved CVD outcomes. Simultaneously, the research recognizes the need for high-quality evaluations of the tools and visualization strategies used in clinical settings.

Methods

The research used a quantitative, multiarm design with data collected through validated questionnaires and direct participant measurements to comprehensively assess behavioral changes as well as subjective and objective health outcomes.

Study Design and Research Protocol

A quantitative research design was used to obtain measurable and generalizable data, which were collected using a combination of validated questionnaires and direct participant measurements, allowing a comprehensive assessment of both subjective experiences and objective health indicators. The trial followed a 4-arm 2×2 factorial design, and the research protocol was registered at ClinicalTrials.gov (NCT06226948). The primary objective was to compare changes in lifestyle behaviors across these 4 arms and to explore whether the combined use of CVD risk visualization and real-time CGM (intervention group 1) leads to greater improvements than standard care. Standard care consisted of routine clinical management at the primary care level, including regular check-ups with a family physician or nurse, brief lifestyle counseling (on diet, physical activity, smoking, and alcohol use) according to national guidelines, and access to general educational materials on CVD prevention. No additional digital tools or enhanced monitoring were provided to the control group.

The secondary objective was to compare changes in estimated CVD risk (SCORE2 [Systematic Coronary Risk Evaluation 2]) across the 4 groups. Additionally, we aimed to determine which lifestyle domains changed in patients at risk of CVD following

XSL•FO
RenderX

exposure to CVD risk visualization, real-time CGM, or their combination.

Type of Study

This research was designed as a single-blind, 4-arm randomized controlled trial (RCT) and was conducted from May to December 2024. Participants were informed only about the intervention they received and were not informed about the full set of study arms or the interventions delivered in other groups (limited disclosure). This RCT was designed as a pragmatic feasibility study embedded in routine primary care. A formal sample size or power calculation was not performed, as this study was designed as a pragmatic feasibility trial. The primary aim was to evaluate feasibility, usability, and behavioral trends rather than to test definitive effectiveness. Therefore, the sample size was determined based on practical feasibility considerations, including the number of eligible patients available in primary care and the logistical capacity for conducting clinical and laboratory assessments. As a result, the study may have been underpowered to detect statistically significant between-group differences.

Participants were randomly assigned to 4 study arms:

- Intervention group 1: Petal-X CVD risk visualization+CGM
- Intervention group 2: CGM only
- Intervention group 3: Petal-X visualization only
- Control group: standard care

Participants were randomized in a 1:1:1:1 ratio into 4 arms using a computer-generated allocation sequence (RAND function in Microsoft Excel). The randomization sequence was generated by the study researcher. The same researcher enrolled participants and assigned them to the trial arms according to the generated sequence. Given the pragmatic feasibility design, formal allocation concealment procedures were not implemented; the enrolling researcher was aware of the upcoming assignment. Participants without a compatible smartphone were randomized only between the 2 non-CGM arms (Petal-X only vs standard care).

The study followed the intention-to-treat principle, whereby all participants were analyzed within the groups to which they were originally assigned, regardless of whether they completed the intervention [39]. This approach preserved the integrity of randomization and reduced bias.

The trial and its reporting were in accordance with the CONSORT (Consolidated Standards of Reporting Trials) 2010 flowchart and the CONSORT extension for pilot and feasibility RCTs (see *Results*), and the corresponding checklist has been completed and submitted (Checklist 1).

Assessment Instrument

The primary outcomes (healthy lifestyle and perception of PCC) were assessed using 2 validated instruments: the Health Lifestyle and Personal Control Questionnaire (HLPCQ) and the Person-Centered Practice Inventory—Service User (PCPI-SU).

The following validated instruments were used in the research.

The HLPCQ [40] was used to assess lifestyle and perceived personal control. The questionnaire included demographic data (sex, age, marital status, number of children, smoking status, and education and employment status) and anthropometric measures (body weight and height). The HLPCQ consists of 26 items grouped into 5 domains. We used a validated Slovenian version of the questionnaire [41]. In the Slovenian validation study conducted among 666 adults at risk of CVD, the total scale showed good internal consistency (Cronbach α=0.85), with subscale α coefficients ranging from 0.59 to 0.88, indicating acceptable to good reliability [41]. In the present feasibility trial, we therefore relied on these previously established psychometric properties and did not repeat internal consistency testing (eg, Cronbach α) in our sample. The main aim of this questionnaire was to identify and quantify lifestyle patterns that reflect lifestyle change as reflected in stress levels and internal control over health. All domains were significantly positively related to each other, suggesting that people who adopt healthy eating habits and avoid harmful diets also follow a daily routine of activities, engage in organized exercise, seek social support, and take care of their mental health. Higher scores indicate a better lifestyle.

The PCPI-SU [42] was used to assess perceptions of PCC. It consists of 20 items and is designed to capture individuals' views on their perceptions of PCC. It covers 5 domains: working with the patient's beliefs and values (4 items), shared decision making (5 items), genuine collaboration (4 items), compassionate presence (3 items), and holistic action (4 items). The questionnaire was a 5-point Likert scale. Higher scores indicated better perceptions of PCC. In a Slovenian study among 426 adults with noncommunicable diseases, the PCPI-SU demonstrated good internal consistency, with Cronbach α=0.82 for the total scale [43]. Consistent with our feasibility design and to avoid redundancy, we did not recalculate Cronbach α in the current sample and relied on these existing reliability estimates.

Participants were shown the Petal-X visualization [44] using a graphical tool that illustrated their current level of CVD risk in a visual representation. This tool was based on the SCORE2 prognostic model [45], which calculates an individual's 10-year risk of developing CVD. We then modified the parameters to reflect target values and used the visualization again to show how an individual's risk could potentially be reduced through lifestyle changes and adherence to the guidelines provided. The code of the Petal-X dashboard is publicly available on GitHub [46], and the English version of the dashboard as a web application is available on the internet [47]. The dashboard was developed using the Observable framework, with the Petal-X visualization [44] implemented using the Observable Plot system [48]. Petal-X was developed using a human-centered design process with iterative prototyping and expert review to ensure clarity and interpretability of the risk visualization [44]. In addition, the underlying surrogate model of SCORE2 was quantitatively validated and demonstrated high fidelity to the original SCORE2 equations, and a controlled experiment with 88 health care students showed that Petal-X supported more accurate interpretation of modifiable risk factor contributions than standard SCORE2 graphical score charts, without reducing perceived transparency, trust, or intention to use the tool [44]. Based on this prior evidence, we considered Petal-X to be of

XSL•FO
RenderX

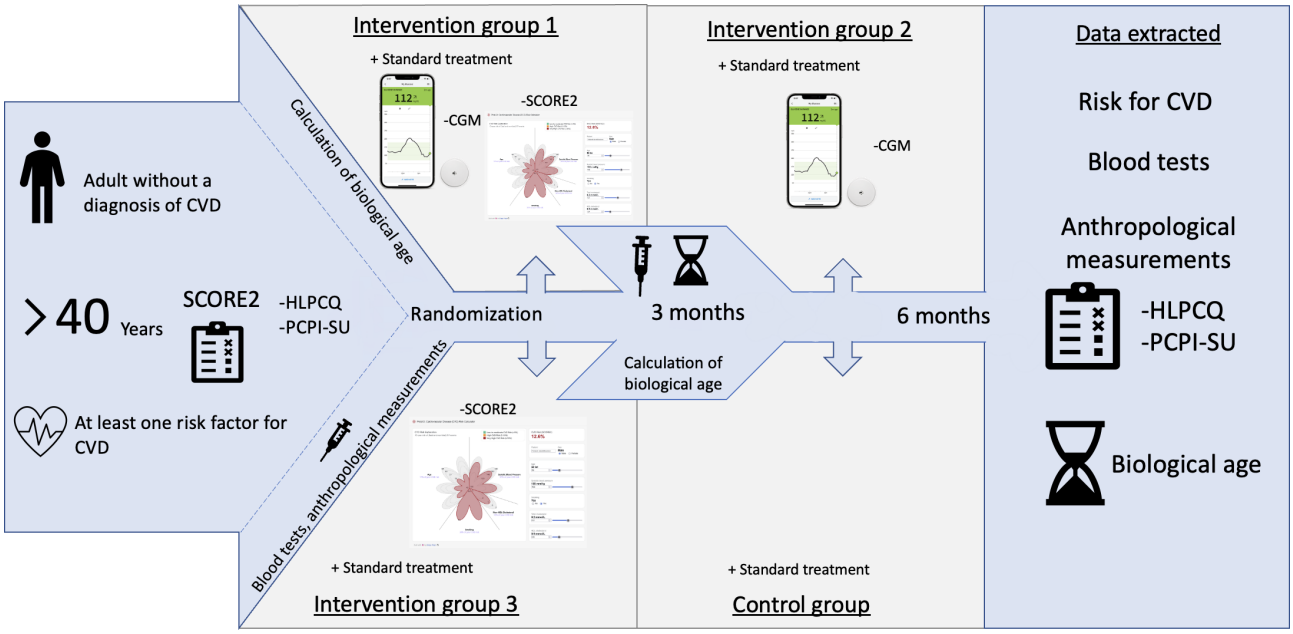
sufficient quality and validity for use as an intervention tool in this feasibility trial. It is based on the validated SCORE2 prognostic model.

Biological age is calculated from plasma GlycanAge analysis [49] and is based on the algorithm developed by Levine et al [50], which incorporates 9 blood biomarkers most strongly associated with mortality: mean corpuscular volume, C-reactive protein, albumin, blood glucose, alkaline phosphatase, red cell distribution width, creatinine, lymphocyte percentage, and chronological age; CGM data is collected using the Dexcom system G7 [51], allowing for real-time monitoring of blood glucose levels.

Description of the Research Process

All participants who met the inclusion criteria completed the HLPCQ and the PCPI-SU at baseline and at the 6-month follow-up, and were also assessed for CVD risk using the SCORE2 prognostic model, incorporating the following variables: age, sex, systolic BP, total cholesterol, high-density lipoprotein (HDL) cholesterol, and smoking status. Venous blood samples were collected at baseline and at 6 months to determine standard biochemical markers (including lipid profile and fasting glucose) and to calculate phenotypic and glycated biological age, with an additional blood sample taken at 3 months for biological age assessment only. The risk assessment using SCORE2 was repeated after 6 months to evaluate potential changes. Additionally, biological age was calculated at 3 and 6 months as a more precise indicator of lifestyle changes compared to chronological age. Two models were used: the model by Levine et al [50], based on blood biomarkers, and GlycanAge, based on glycan analysis, as defined by Lauc and Primorac [49]. Participants were allocated to the following 3 different intervention groups and 1 control group according to the quantitative research design, with the groups differing according to the type of intervention they received during treatment (Figure 1).

Figure 1. Intervention groups and research process. CGM: continuous glucose monitoring; CVD: cardiovascular disease; HLPCQ: Healthy Lifestyle and Personal Control; PCPI-SU: Person-Centered Practice Inventory—Service User; SCORE2: Systematic Coronary Risk Evaluation 2.



Intervention group 1 received both CGM for 10 days and the Petal-X CVD risk visualization tool [44]. Petal-X allowed for a visual presentation of the participant's current risk and an interactive simulation of how lifestyle changes could improve the risk profile.

Intervention group 2 used CGM for 10 days, without using the Petal-X CVD risk visualization tool.

Intervention group 3 received only the Petal-X risk visualization.

The control group received standard care without additional intervention and did not receive any of the digital interventions (Petal-X or CGM) after study completion, as the trial did not include a delayed-intervention phase.

At baseline and after 6 months, all participants underwent anthropometric and blood measurements, completed the HLPCQ and PCPI-SU questionnaires, and received updated CVD risk (SCORE2) and biological age assessments. These data allowed for comparisons between groups and evaluation of changes in lifestyle and CVD. The 6-month duration was selected based on previous studies demonstrating meaningful changes in health behaviors within this period [52,53]. Adverse events and device-related issues (eg, CGM-related problems) were monitored throughout the study. No serious adverse events were reported.

Study Sample

Recruitment was open to individuals across Slovenia, but the promotional campaign and study implementation were concentrated in a Community Health Center in central Slovenia. Purposive sampling was used, and inclusion criteria included adults being 40 years or older with no previous diagnosis of CVD but with at least one of the following risk factors: elevated BP (systolic >140 mmHg), elevated fasting glucose (>6.1 mmol/L), abnormal lipid levels (total cholesterol >5 mmol/L, HDL <1.4 mmol/L for men and <1.6 mmol/L for women,

XSL•FO
RenderX

low-density lipoprotein >3.5 mmol/L, triglycerides >1.7 mmol/L), or BMI >30.

Preexisting CVD, being younger than 40 years, and the absence of risk factors were the exclusion criteria. The sample was selected using purposive sampling. In total, 119 participants were included at baseline: 30 in intervention group 1 (CGM and Petal-X), 24 in intervention group 2 (CGM only), 33 in intervention group 3 (Petal-X only), and 32 in the control group. At follow-up, 101 participants remained: 27 in intervention group 1, 22 in intervention group 2, 26 in intervention group 3, and 26 in the control group. Dropout reasons included newly diagnosed health conditions, inability to attend follow-up appointments, or personal commitments.

Statistical Analysis

Descriptive statistics (means and SDs for continuous variables and absolute and relative frequencies for categorical variables) were used to characterize the sample and summarize outcome measures. Group differences in categorical baseline characteristics (eg, gender, smoking status, and education) were analyzed using chi-square tests. All statistical analyses adhered to the intention-to-treat principle. IBM SPSS Statistics 29.0 was used for all analyses, and statistical significance was set at $P \leq .05$.

For Hypothesis 1, differences in lifestyle change across the 4 randomized groups were analyzed using the Kruskal–Wallis test, with the HLPCQ total score as the dependent variable. For Hypothesis 2, changes in cardiovascular risk across the 4 groups were analyzed with SCORE2 as the dependent variable, also using the Kruskal–Wallis test, due to nonnormal distributions and modest cell sizes. HLPCQ and PCPI-SU scores were treated as continuous variables using their total scale scores.

Spearman's rank correlation coefficients were calculated to explore associations between healthy lifestyle (HLPCQ) and perceived PCC (PCPI-SU) [54].

To reduce detection and analysis bias, outcome data were handled using coded group labels during data cleaning and analysis, and laboratory assessments were performed according to standard procedures without reference to group allocation.

Ethical Considerations

All data collection and intervention procedures were conducted at a community health center in Central Slovenia. The research was carried out by a researcher under the supervision of an associate professor and a physician.

Participants underwent venous blood sampling at baseline, 3 months, and 6 months. Appointment invitations included a brief personalized summary of test results, which served to further motivate participants. The study was single-blind at the participant level: participants were informed only about procedures in their allocated arm and were not informed about the full set of study arms. Due to the pragmatic implementation, study personnel were not blinded.

The study was approved by the Ethics Committee for Nursing Research at the Faculty of Health Sciences, University of Maribor (Approval No. 02/7K-2023), and the National Medical Ethics Committee of the Republic of Slovenia (Approval No. 0120-217/2023/4). We have also registered the research protocol in clinical trials (ClinicalTrials.gov NCT06226948) [55]. All participants were informed that their data would be used exclusively for research purposes.

Participants received written and verbal information about the study and provided written informed consent before any study procedures. Participation was voluntary, and participants could withdraw at any time without consequences for their usual care. To protect privacy and confidentiality, each participant was assigned a unique study identification code; identifiable information was stored separately from research data and was accessible only to authorized study personnel. Data were handled in accordance with applicable data protection regulations (eg, General Data Protection Regulation) and institutional policies, and were stored on password-protected systems with restricted access; only anonymized or aggregated data were used for analysis and reporting. No financial compensation was provided for participation.

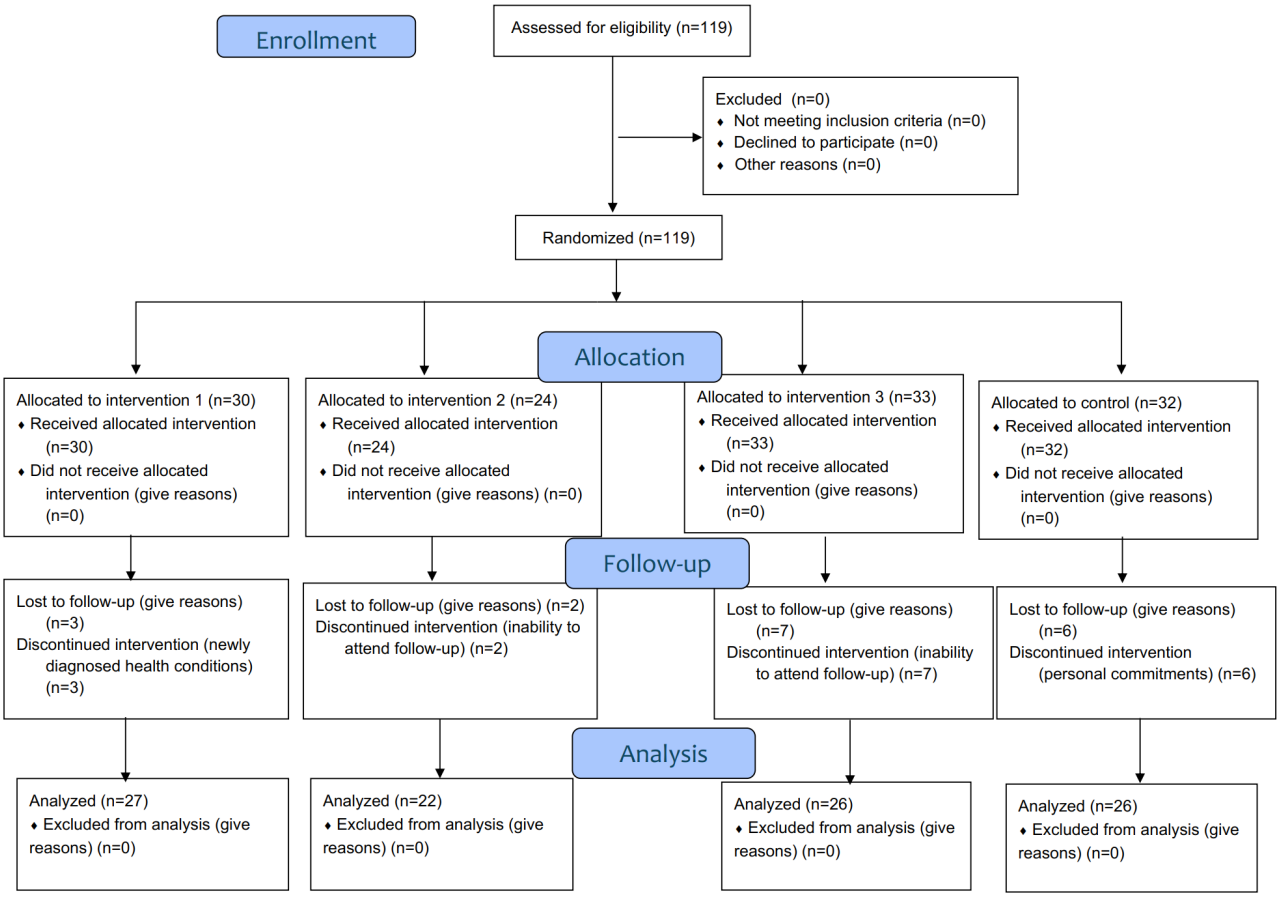
Results

Participant Enrollment

Figure 2 presents the CONSORT participant flow diagram. In total, 119 participants were enrolled and randomized into 4 arms, and 101 participants completed the 6-month follow-up; dropouts were mainly due to newly diagnosed health conditions, inability to attend follow-up appointments, or personal commitments.

XSL•FO
RenderX

Figure 2. The CONSORT (Consolidated Standards of Reporting Trials) 2010 flowchart.



Demographic Characteristics

Of the 119 participants, 71% (n=85) were women and 29% (n=34) were men. At the start of the study, most participants were married (n=56, 47%) and had children (n=107, 90%). Most participants had specialization in a professional higher education program, completed a university program or a master's degree (second degree; n=43, 36%), and were employed (n=110, 92%). Most participants did not smoke (n=82, 69%).

Table 1 shows the basic demographic characteristics of all intervention groups and the control group at baseline. The only demographic variable that differed significantly between the groups was educational level (χ^2_{15}=37.15; *P*=.001). Participants in Intervention group 1 (13/30, 43.3%) and Intervention group 3 (15/33, 45.5%) more frequently reported holding a master's degree compared with the control group (6/32, 18.8%). No other demographic characteristics differed significantly across groups, including gender (χ^2_3=3.15; *P*=.37), age (H_3=3.13; *P*=.37), marital status (χ^2_{12}=9.22; *P*=.68), having children (χ^2_3=0.71; *P*=.87), number of children (H_3=0.94; *P*=.82), employment (χ^2_6=7.62, *P*=.27), smoking status (χ^2_6=2.08; *P*=.91), and number of cigarettes smoked (H_3=3.18; *P*=.36). Any differences in the study were due to the interventions in the intervention groups and the level of education in the context of demographic characteristics.

XSL•FO
RenderX

Table . Descriptive characteristics of participants.

	Intervention group 1	Intervention group 2	Intervention group 3	Control group	*P* value
Gender, n (%)					.37
Man	9 (30)	9 (37.50)	6 (18.18)	11 (34.38)	
Woman	21 (70)	15 (62.50)	27 (81.82)	21 (65.63)	
Nonbinary	0 (0)	0 (0)	0 (0)	0 (0)	
Age (years), mean (SD)	51.87 (7.22)	48.50 (7.02)	50.33 (7.14)	49.84 (8.44)	.37
Marital status, n (%)					.68
Married	15 (50)	9 (37.50)	16 (48.48)	16 (50)	
Extramarital community	11 (36.67)	10 (41.67)	8 (24.24)	12 (37.50)	
Single	4 (13.33)	5 (20.83)	6 (18.18)	3 (9.38)	
Widower	0 (0)	0 (0)	2 (6.06)	1 (3.13)	
Other	0 (0)	0 (0)	1 (3.03)	0 (0)	
Having children, n (%)					.87
Yes	28 (93.33)	21 (87.50)	29 (87.88)	29 (90.63)	
No	2 (6.67)	3 (12.50)	4 (12.12)	3 (9.38)	
Number of children, mean (SD)	1.70 (0.84)	2.00 (1.22)	1.76 (1.03)	1.78 (1.01)	.82
Educational level, n (%)					.001
No education	0 (0)	0 (0)	0 (0)	0 (0)	
Primary education	1 (3.33)	0 (0)	0 (0)	0 (0)	
Secondary school education	8 (26.67)	7 (29.17)	2 (6.06)	13 (40.63)	
Higher university education	4 (13.33)	7 (29.17)	16 (48.48)	13 (40.63)	
Master's degree	13 (43.33)	9 (37.50)	15 (45.45)	6 (18.75)	
PhD	4 (13.33)	0 (0)	0 (0)	0 (0)	
Other	0 (0)	1 (4.17)	0 (0)	0 (0)	
Employment, n (%)					.27
Yes	28 (93.33)	22 (91.67)	32 (96.97)	28 (87.50)	
No	2 (6.67)	0 (0)	1 (3.03)	1 (3.13)	
Retired	0 (0)	2 (8.33)	0 (0)	3 (9.38)	
Other	0 (0)	0 (0)	0 (0)	0 (0)	
Smoking status, n (%)					.91
Yes	6 (20)	5 (20.83)	8 (24.24)	8 (25)	
No	20 (66.67)	18 (75)	23 (69.70)	21 (65.63)	
Stop smoking	4 (13.33)	1 (4.17)	2 (6.06)	3 (9.38)	
Number of cigarettes, mean (SD)	15 (6.32)	9.80 (5.26)	11.50 (5.63)	9.75 (5.63)	.37

Evaluation of the Measurements of the Participants in the Main Research

Intervention group 1 had the highest mean SCORE2 at baseline (4.87, SD 2.61), followed by the control group (4.53, SD 3.63) and intervention group 2 (4.08, SD 2.42), while intervention group 3 had the lowest score (3.84, SD 2.08). This indicates that intervention group 3 had the most favorable CVD risk profile at baseline, as lower scores reflect lower risk. Intervention group 1 had the highest mean total cholesterol value at baseline (5.93, SD 1.10), while intervention group 2 had the lowest mean value (4.82, SD 1.51). Intervention group

XSL•FO
RenderX

3 had the highest mean HDL value, with a final value of 1.62 (SD 0.34), while intervention group 2 had the lowest mean value, with a baseline value of 1.28 (SD 0.31).

The highest mean systolic BP was measured in intervention group 1, with a baseline value of 129.23 mmHg (SD 16.76), while the lowest mean was in intervention group 3, with a final value of 119.69 (SD 11.58).

There were no statistically significant differences between groups in SCORE2 (H_3=3.269; *P*=.35), total cholesterol (H_3=2.402; *P*=.49), HDL (H_3=2.93; *P*=.40), systolic BP (H_3=3.346; *P*=.34), diastolic BP (H_3=2.588; *P*=.46), heart rate (H_3=2.170; *P*=.54), BMI (H_3=4.066; *P*=.25), phenotypic biological age (H_3=2.708; *P*=.44) and glycated biological age (H_3=0.261, *P*=.97). Statistical analysis showed no significant differences between groups. This means that the groups compared with respect to SCORE2, total cholesterol, HDL, BP, heart rate, BMI, phenotypic biological age, and glycated biological age were like each other. The baseline and final values for all measurements are presented in Table 2.

Table . Participant measurements by randomized groups.

	Intervention group 1, mean (SD)	Intervention group 2, mean (SD)	Intervention group 3, mean (SD)	Control group, mean (SD)	Change in *P* value (SD)
SCORE2[a]					.35
Baseline value	4.87 (2.61)	4.08 (2.42)	3.84 (2.08)	4.53 (3.63)	
Final value	4.06 (2.71)	3.53 (2.02)	3.17 (1.99)	4.36 (3.27)	
Total cholesterol					.49
Baseline value	5.93 (1.10)	5.09 (0.66)	5.68 (0.93)	5.28 (1.12)	
Final value	5.10 (2.26)	4.82 (1.51)	5.43 (0.99)	5.44 (1.07)	
HDL[b]					.02
Baseline value	1.43 (0.28)	1.28 (0.31)	1.61 (0.33)	1.33 (0.32)	
Final value	1.42 (0.31)	1.34 (0.31)	1.62 (0.34)	1.38 (0.32)	
Systolic blood pressure					.34
Baseline value	129.23 (16.76)	124.17 (14.21)	122.36 (13.32)	123.38 (17.09)	
Final value	128.26 (10.88)	121.36 (14.98)	119.69 (11.58)	124.77 (12.58)	
Diastolic blood pressure					.46
Baseline value	84.73 (8.73)	82.54 (11.92)	78.48 (9.86)	82.38 (10.93)	
Final value	78.93 (9.14)	77.18 (9.51)	75.35 (7.70)	78.19 (7.99)	
Heart rate					.54
Baseline value	76.07 (12.53)	73.79 (13.08)	72.15 (9.90)	71.66 (11.62)	
Final value	72.11 (8.38)	71.36 (12.69)	71.81 (9.33)	71.31 (9.70)	
BMI					.25
Baseline value	26.35 (4.19)	26.83 (4.50)	25.16 (4.03)	28.01 (5.64)	
Final value	26.48 (4.21)	26,55 (4.22)	24.48 (3.97)	27.65 (5.11)	
Phenotypic biological age					.44
Baseline value	46.82 (7.89)	43.30 (7.53)	43.90 (7.01)	43.34 (9.38)	
Final value	43.07 (14.02)	40.64 (11.36)	36.35 (16.98)	40.44 (16.43)	
Glycan biological age					.68
Baseline value	56.73 (16.17)	48.71 (21.73)	54.33 (16.08)	51.72 (18.31)	
Value after 3 mo	57.39 (16.53)	53.23 (20.79)	54.74 (16.00)	49.37 (19.8)	
Final value	56.04 (18.73)	51.86 (20.95)	54.96 (17.09)	53.50 (15.76)	

[a]SCORE2: Systematic Coronary Risk Evaluation 2.

[b]HDL: high-density lipoprotein cholesterol.

Based on the analyses of the data, we found that there were no significant differences in lifestyle changes between participants in the intervention groups and those in the control group.

XSL•FO
RenderX

Healthy Lifestyle Assessment: HLPCQ

In all 3 intervention groups and the control group, the domain with the highest baseline score was daily routine. The mean scores for this domain were 19.24 (SD 5.87), 19.17 (SD 3.59), 19.93 (SD 5.18), and 17.61 (SD 5.82) for the first, second, and third intervention groups and the control group, respectively. However, the lowest-rated domain in all groups was organized exercise. The mean scores were 4.27 (SD 1.87), 4.00 (SD 1.78), 4.55 (SD 2.08), and 4.09 (SD 1.69).

At the end point, the highest and lowest rated domains remained unchanged but increased, indicating an improvement in HLPCQ scores. The mean scores for the highest-rated domain, Daily Routine, were: 21.27 (SD 4.94) for the intervention group 1; 21.50 (SD 4.30) for the intervention group 2; 21.25 (SD 3.90) for the intervention group 3; and 19.80 (SD 4.85) for the control group. However, organized exercise was the lowest rated domain in all groups. The mean scores were 5.31 (SD 1.67) in the first intervention group, 4.36 (SD 1.94) in the second intervention group, 4.54 (SD 2.00) in the third intervention group, and 4.88 (SD 2.01) in the control group.

Missing data were minimal across variables, with no missing values for healthy dietary choices at baseline (n=0), and only isolated cases for dietary harm avoidance (n=1), daily routine (n=3), organized physical activity (n=1), and social and mental balance (n=2); the highest number of missing values occurred in the total HLPCQ scores, with up to 5 in 1 group at final measurement.

There were no statistically significant differences in change in HLPCQ scores between groups (H_3=2.801; *P*=.42). Thus, the Kruskal–Wallis test showed that group had no significant effect on the change in HLPCQ (Table 3).

Table . Baseline and final HLPCQ[a] scores by randomized group.

HLPCQ	Intervention group 1, mean (SD)	Intervention group 2, mean (SD)	Intervention group 3, mean (SD)	Control group, mean (SD)
Healthy dietary choices				
Baseline value	15.93 (3.50)	14.79 (2.30)	16.94 (2.57)	15.25 (3.44)
Final value	17.72 (4.08)	16.64 (3.55)	18.28 (3.39)	16.42 (2.80)
Dietary harm avoidance				
Baseline value	10.40 (2.39)	10.52 (2.25)	11.31 (2.15)	10.28 (2.73)
Final value	11.08 (2.04)	10.32 (2.23)	11.88 (2.01)	10.46 (1.94)
Daily routine				
Baseline value	19.24 (5.87)	19.17 (3.59)	19.93 (5.18)	17.61 (5.82)
Final value	21.27 (4.94)	21.50 (4.30)	21.25 (3.90)	19.80 (4.85)
Organized physical activity				
Baseline value	4.27 (1.87)	4.00 (1.78)	4.55 (2.08)	4.09 (1.69)
Final value	5.31 (1.67)	4.36 (1.94)	4.54 (2.00)	4.88 (2.01)
Social and mental balance				
Baseline value	14.10 (2.51)	13.92 (2.54)	14.36 (2.36)	13.53 (3.29)
Final value	14.36 (2.04)	14.45 (2.82)	13.88 (3.05)	13.96 (2.73)
HLPCQ total				
Baseline value	64.17 (12.54)	63 (7.53)	66.76 (8.91)	60.58 (12.96)
Final value	69.50 (10.34)	67.27 (8.56)	68.90 (7.78)	65.72 (10.18)

[a]HLPCQ: Healthy Lifestyle and Personal Control Questionnaire.

The Kruskal–Wallis test (H_3=2.801; *P*=.42) revealed no statistically significant differences between the groups regarding changes in HLPCQ. However, the control group showed the most significant improvement in terms of a healthy lifestyle.

Perception of Person-Centered Care: PCPI-SU

At baseline, the highest-rated domain in all 3 intervention groups and the control group was sharing decision-making. The mean scores were 19.57 (SD 4.68) in the first intervention group, 18.00 (SD 4.64) in the second intervention group, 19.39 (SD 4.96) in the third intervention group, and 18.78 (SD 5.08) in the control group. However, engaging authentically was the lowest scoring domain in all groups, with average scores of 11.00 (SD 2.85) in the first intervention group, 10.29 (SD 2.48) in the second intervention group, 11.00 (SD 3.01) in the third intervention group, and 10.31 (SD 2.79) in the control group.

The domains that were highest and lowest rated at the final perception remained unchanged. The average domain scores for sharing decision-making in the first intervention group were 19.77 (SD 4.61), in the second intervention group 20.86 (SD 4.16), in the third intervention group 19.27 (SD 5.80), and in the control group 19.35 (SD 4.91). The lowest-rated domain

XSL•FO
RenderX

was being sympathetically present in all groups. The average scores were 10.96 (SD 2.57) in the first intervention group, 11.55 (SD 2.99) in the second intervention group, 11.42 (SD 3.24) in the third intervention group, and 11.20 (SD 3.16) in the control group.

Missing data were low across variables, with no missing values for several baseline and final measurements. Occasional missing data were observed for working with the person's beliefs and values (n=1), sharing decision-making (n=1), engaging authentically (n=3), being sympathetically present (n=1), and working holistically (n=2). The highest number of missing values occurred in the total PCPI-SU scores, with 3 at baseline and 2 at the final measurement (Table 4).

Table . The perception of PCPI-SU[a] and components at the start and end of the study by the groups that were randomly selected.

	Intervention group 1, mean (SD)	Intervention group 2, mean (SD)	Intervention group 3, mean (SD)	Control group, mean (SD)
Working with the person's beliefs and values				
Baseline value	15.43 (3.99)	14.00 (3.84)	15.15 (3.83)	14.44 (4.26)
Final value	15.23 (3.93)	15.27 (4.24)	14.88 (4.08)	15.35 (3.98)
Sharing decision-making				
Baseline value	19.57 (4.68)	18.00 (4.64)	19.39 (4.6)	18.78 (5.08)
Final value	19.77 (4.61)	20.86 (4.16)	19.27 (5.80)	19.35 (4.91)
Engaging authentically				
Baseline value	14.69 (3.41)	14.17 (3.24)	14.33 (3.37)	13.66 (3.64)
Final value	14.84 (3.14)	15.59 (3.26)	14.65 (3.50)	14.40 (3.69)
Being sympathetically present				
Baseline value	11.00 (2.85)	10.29 (2.48)	11.00 (3.01)	10.31 (2.79)
Final value	10.96 (2.57)	11.55 (2.99)	11.42 (3.24)	11.20 (3.16)
Working holistically				
Baseline value	14.55 (3.93)	13.61 (4.24)	14.42 (3.87)	13.87 (4.40)
Final value	14.23 (3.86)	15.00 (4.08)	15.50 (4.35)	14.77 (4.22)
PCPI-SU total				
Baseline value	75.64 (17.82)	70.13 (15.98)	76.70 (14.98)	71.17 (18.53)
Final value	74.92 (16.16)	79.19 (17.12)	75.73 (19.02)	73.54 (18.23)

[a]PCPI-SU: Person-Centered Practice Inventory—Service User.

This indicates that there are no significant differences between the groups in terms of changes in PCPI-SU. This is consistent with the results of the Kruskal–Wallis test, which showed no statistically significant differences (H_3=4.056; P=.26). However, it shows that there was a difference in the change in the second intervention group, in which the CGM was worn. Based on this, it can be concluded that the intervention in the second group had an impact on improving the perception of PCC.

Discussion

Effectiveness of Risk Visualization and Glucose Monitoring in Reducing CVD Risk

The overall trend for the SCORE2 risk showed a decrease in CVD for all participants over the 6-month period. This suggests that participants showed some improvement in their risk profile regardless of the type of intervention. The group that was exposed to risk visualization through the Petal-X tool as part of the intervention showed the most improvement. When comparing all 3 intervention groups with the control group more explicitly, a clear pattern emerges. The group receiving combined CVD risk visualization and CGM showed the largest numerical reduction in SCORE2, followed by the group receiving visualization only, while the CGM-only group demonstrated a smaller improvement. The control group also exhibited a reduction in SCORE2, but the magnitude of change was consistently lower than in the visualization-based intervention groups. Although these differences did not reach statistical significance, the direction and relative size of the effects suggest that visual risk communication may have stronger motivational value than glucose monitoring alone, and that combining visualization with CGM may provide additional reinforcement. This highlights the potential value of visual feedback in supporting risk awareness and self-regulation. The findings are consistent with previous research indicating that traditional CVD factors, such as elevated BP, high cholesterol levels, and smoking, remain important in predicting future CVD events [56,57]. Despite the use of different methodologies and assessment tools in different studies, the consistency in core risk factors supports the reliability of tools such as SCORE2 for estimating individual risk. In the present study, the use of a

XSL•FO
RenderX

validated prognostic model (SCORE2), combined with robust data collection from the Slovenian population, provided a solid foundation for evaluating risk trends.

Unlike other research designs, such as the focus group methods used by Vivek Dhukaram et al [58] or the large-scale registry analyses conducted in the Taiwanese study using the National Health Insurance Research Database [59], our approach provides clinically relevant insights into how risk evolves in real-world conditions. Furthermore, physician-led evaluations based on European guidelines, as demonstrated in the study by Dudas et al [60], continue to be a fundamental aspect of CVD risk management. Our findings support these approaches by integrating a PCC component through visualization, which may enhance patient engagement and understanding of personal risk. In line with international evidence [57,61], this study confirms that identifying at-risk individuals early on through structured risk assessments enables more precise preventive interventions. Visualizing risk, particularly when paired with ongoing monitoring or tailored counseling, may further encourage behavioral changes and reduce the long-term burden of CVD. The results of the CVD risk assessment allow for early identification of at-risk individuals, which may contribute to more targeted preventive interventions and a reduction of the burden of CVD in the population, and these findings are also consistent with the reviewed literature [56,57,61]. It is important to interpret the results within the context of this feasibility trial. The absence of statistical significance may be related to the modest sample size, baseline variability, and the short duration of exposure to the interventions. These factors limit the ability to detect smaller but potentially meaningful behavioral changes. Therefore, the results should be interpreted as preliminary indicators of trends rather than definitive effect estimates.

Evaluating and Modifying Lifestyle Behaviors in CVD Risk Patients Through Visualization and Glucose Monitoring

Assessment of lifestyle using the HLPCQ revealed improvements in multiple domains, particularly among participants who received CVD risk visualization and CGM. These tools were most effective in promoting healthy dietary choices and enhancing daily routines, emphasizing the importance of combining risk awareness with real-time feedback [62]. Participants who received only visualization or only CGM also showed progress, though less consistently. Organized physical activity remained the weakest area across all groups, suggesting that general interventions may be insufficient to influence this behavior. Further research confirms the importance of routine, nutrition, and physical activity in improving quality of life and reducing the risk of disease [63-68]. Social and mental balance showed minimal change, which suggests that these aspects may be more resistant to digital or informational interventions and more dependent on the broader psychosocial context. The avoidance of dietary harm also decreased, particularly in groups without combined support. Older participants tended to report better organization of their lifestyle and healthier eating habits. This aligns with findings that age can positively influence adherence to preventive behaviors. Although no statistically significant differences were found, there were meaningful behavioral shifts in groups with combined interventions. The HLPCQ proved to be a useful tool for detecting these changes, particularly given its ability to incorporate self-regulation and health-related behaviors. Its future use in primary care and epidemiological research appears promising. It is important to note that lifestyle change depends not only on information but also on health literacy, motivation, and contextual support [69]. Research from Sweden has emphasized the importance of valid methods for engaging communities in sustainable prevention efforts, particularly among populations with low health literacy or limited access to care [70]. Therefore, enhancing health literacy could be central to national strategies for preventing lifestyle-related conditions [71]. In conclusion, visualization and CGM can support behavioral changes in individuals at risk of CVD, particularly regarding dietary and routine habits. However, more targeted approaches are needed for physical activity and psychosocial domains, which remain less responsive.

Lifestyle Change Across Intervention and Control Groups

We tested the hypothesis that there would be significant differences in lifestyle change between groups, but this was not confirmed. The Kruskal–Wallis analysis revealed no statistically significant differences in HLPCQ score changes between the 4 study groups (3 intervention groups and the control group). The control group reported the greatest overall change in healthy lifestyle behavior, followed by smaller positive changes in all intervention groups. This suggests that the inclusion of new technologies, such as CVD risk visualization and CGM, did not result in statistically superior lifestyle improvements over 6 months. Nevertheless, the presence of positive trends in all intervention groups indicates a potential motivational effect of the interventions. A sustainable lifestyle change is a long-term process requiring continuity, reinforcement, and individual adaptation to have a lasting impact on quality of life and health-related well-being [17]. Based on other studies with similar intervention durations [17,53], we expected to detect measurable changes after 6 months. However, the short-term use of CGM and limited intervention exposure may have reduced the observable impact within the timeframe of our research.

Additionally, the Petal-X visualization of CVD risk was described as a concrete and motivating element that helped individuals recognize the necessary behavioral changes in their daily lives. These observations indicate that while measurable outcomes were modest, experiential factors may still play an important role in behavior change. Further research should explore longer intervention periods and enhanced support structures to evaluate the full potential of visual and self-monitoring tools in promoting healthy lifestyles. It is important to note that educational level differed significantly between the randomized groups at baseline. This imbalance is a common limitation in small feasibility samples, where randomization may not fully equalize sociodemographic characteristics. Since higher educational attainment is associated with better health literacy, greater engagement in preventive behaviors, and improved ability to interpret risk information, it is possible that educational level influenced participants'

XSL•FO
RenderX

responses to the interventions. Although we cannot determine the extent of this influence, it may have contributed to behavioral differences observed across groups. Future studies should include stratified randomization or larger sample sizes to minimize such baseline imbalances.

Determining Whether Person-Centered Care Is Associated With Lifestyle Change

A link between PCC and lifestyle change was demonstrated. The results of the correlation analysis revealed a weak, yet statistically significant, positive relationship between perceived PCC and healthy lifestyle behaviors, including self-regulation. The Spearman correlation coefficient showed that higher HLPCQ scores, reflecting healthier lifestyle patterns and greater personal control, were associated with higher perceptions of PCC. Although the correlation was weak, its statistical significance suggests that the association is unlikely to be due to chance. These findings show that improving lifestyle and self-management behaviors could lead to a more positive perception of care quality from the patient's perspective. Therefore, even minor improvements in perceived PCC could be meaningful, particularly when promoting long-term behavioral change and better health outcomes. Researchers emphasize the importance of health care professionals assessing individual risk, designing care plans that are tailored to patients' lifestyles and contexts, and implementing PCC approaches to positively influence health outcomes [72-76]. Furthermore, interventions targeting lifestyle changes have been shown to improve BP and lipid profile control, highlighting the effectiveness of PCC in managing risk factors and encouraging behavioral changes [77]. This evidence highlights the importance of tailoring health care delivery to patients' individual needs and preferences. PCC has the potential to improve patient satisfaction and may reduce CVD risk factors such as hypertension [78-80]. Therefore, integrating lifestyle-oriented PCC strategies into clinical practice should be considered a priority in both prevention and chronic care management.

These findings should be interpreted with caution, as the study may have been underpowered and the intervention period relatively short, both of which could have limited the ability to detect statistically significant effects. It is also important to consider potential clinical significance. Both the HLPCQ and PCPI-SU scores showed improvements across groups, including the control group. Although these changes were not statistically significant, small improvements in lifestyle behaviors or person-centeredness may still be clinically meaningful in a primary care setting. Behavior change is often gradual and incremental, and even minor positive shifts can contribute to long-term cardiovascular risk reduction when sustained over time.

Limitations

This study has several limitations. The lack of a double-blind design may have introduced bias, and participants may have perceived differences between the intervention groups, which could have influenced the outcomes. The sample included only individuals with an increased risk of CVD based on clinical markers, excluding those with behavioral risk factors. Variability in digital literacy, particularly among older participants, may have affected uptake of the technology-supported interventions. Motivation fluctuated due to personal life events, and some technical issues with CGM devices occurred, but these were resolved. Although the lifestyle questionnaire had not been validated in Slovenia, the necessary cultural adaptations were made in line with WHO guidelines.

Additionally, the study was not sufficiently powered to detect small or moderate effects, as the sample size was limited and no formal power calculation was conducted prior to recruitment. This may have contributed to the absence of statistically significant findings despite observable clinical trends. The 6-month intervention period may have been too short to capture measurable changes in long-term lifestyle behaviors or CVD risk, which typically require sustained support and longer follow-up. This likely reduced the detectable impact of both visualization and CGM-based interventions. Consequently, the trial was not adequately powered to detect small or moderate between-group differences, and the absence of statistically significant findings should therefore be interpreted with caution.

Key Findings and Recommendations

Although the quantitative results did not confirm statistically significant improvements, participants in the intervention groups reported enhanced awareness, motivation, and engagement in their health management. Digital tools such as Petal-X, CGM, and biological age assessment supported PCC and encouraged meaningful lifestyle adjustments. These findings demonstrate the potential for integrating digital tools into preventive health care to improve understanding of CVD risk and promote behavior change. Future studies should include larger and more diverse samples, extend the follow-up period, and examine effects in specific subgroups. Strengthening digital literacy among health care professionals and patients will be essential for successful implementation. With better training and user-friendly tools, technologies such as Petal-X and CGM could provide substantial support for lifestyle change, ongoing risk monitoring, and more effective PCC.

Conclusions

This study makes a valuable contribution to the growing field of digital health, demonstrating the potential of combining risk visualization (Petal-X), CGM, and biological age assessment for the lifestyle-based prevention of CVD. While no clinically significant improvements were observed, the findings provide valuable insights into the feasibility, usability, and motivational impact of digital tools in PCC. Integrating such technologies could facilitate the early identification of risk, encourage patient engagement, and promote behavior change. These results emphasize the necessity of further longitudinal research and the development of digital infrastructure to facilitate lifestyle interventions within preventive health care in Slovenia. We also provided a detailed description of the standard care received by the control group to ensure transparency and comparability between intervention arms, in accordance with CONSORT recommendations.

This study highlights the potential of integrating digital tools such as Petal-X visualization, CGM, and biological age

XSL•FO
RenderX

assessment into preventive health care. While statistical significance was limited, participants reported increased awareness, motivation, and involvement in managing their health. These findings support the value of PCC approaches and the need for further research on long-term, technology-supported lifestyle interventions. Expanding digital literacy and developing national guidelines for digital integration will be key to enhancing care quality and reducing CVD risk in the future.

Acknowledgments

The authors would like to thank all participants who took part in this research.

The authors acknowledge partial support from the Slovenian Research Agency (grant numbers N3-0307 and P2-0057).

The authors declare the use of generative artificial intelligence (GAI) in the research and writing process. According to GAIDeT (Generative AI Delegation Taxonomy; 2025), the following tasks were delegated to GAI tools under full human supervision: translation and reformatting. The GAI tool used was ChatGPT 5.1.

Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Funding

This work was supported by the University of Maribor under the Development Pillar of Financing 2025 - 2028 (RSF 3.0) and the Development Goals of the University of Maribor 2025 - 2028. The funder had no role in the study design, data collection, analysis, interpretation of results, or writing of the manuscript. (SI: Razvojni steber financiranja 2025 - 2028 (RSF 3.0), Razvojni cilji Univerze v Mariboru 2025 - 2028).

Authors' Contributions

The research presented in this paper was a collaborative effort involving multiple authors who contributed to various aspects of the review. The review was conceived by GŠ, AS, ML, DM, and ZK-K. AS and GŠ structured the review, synthesized the evidence, drafted the manuscript, reviewed the manuscript, and completed the final version. GŠ, AS, ML, ZK-K, LG, and DM provided evidence and edits for the review and reviewed the full manuscript when completed. AS collected the data, while LG, GŠ, and AS performed the statistical analysis.

Conflicts of Interest

None declared.

Checklist 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File, 314 KB - publichealth_v12i1e83488_app1.pdf]

References

1. Bernell S, Howard SW. Use your words carefully: what is a chronic disease? Front Public Health 2016;4:159. [doi: 10.3389/fpubh.2016.00159] [Medline: 27532034]
2. Ageing and health. World Health Organization. 2024. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2026-01-06]
3. Alabdallah MSH, Almalki ASS, Alsaedi AS, Abdullah AMH, Aljuhine FHA, Attar AK. Managing chronic disease in primary care: a review. Int J Med Developing Countries 2020;4:504-508. [doi: 10.24911/IJMDC.51-1576048326]
4. Kvedar JC, Fogel AL, Elenko E, Zohar D. Digital medicine's march on chronic disease. Nat Biotechnol 2016 Mar;34(3):239-246. [doi: 10.1038/nbt.3495] [Medline: 26963544]
5. Kasturiarachchi S, Rakovac I, Hetz K, Wickramasinghe K, Kluge HHP, Galea G. Commercial determinants of noncommunicable diseases in the WHO European region. : World Health Organization; 2024 URL: https://eprints.whiterose.ac.uk/id/eprint/213530/1/WHO_report_NCDs.pdf [accessed 2026-01-06]
6. Di Cesare M, Perel P, Taylor S, et al. The heart of the world. Glob Heart 2024;19(1):11. [doi: 10.5334/gh.1288] [Medline: 38273998]
7. Di Cesare M, Thomas HB, Gaziano LH, et al. World Heart Report 2023: confronting the world's number one killer. : World Heart Federation (WHF); 2023 URL: https://world-heart-federation.org/wp-content/uploads/World-Heart-Report-2023.pdf?utm_source=chatgpt.com [accessed 2026-01-19]
8. Svenšek A, Lorber M, Gosak L, Verbert K, Klemenc-Ketis Z, Stiglic G. The role of visualization in estimating cardiovascular disease risk: scoping review. JMIR Public Health Surveill 2024 Oct 14;10:e60128. [doi: 10.2196/60128] [Medline: 39401079]

XSL•FO
RenderX

9. Kobat H, Elkonaissi I, Foreman E, et al. Smoking, diabetes mellitus, and previous cardiovascular disease as predictors of anticancer treatment-induced cardiotoxicity in non-small-cell lung cancer: a real-world study. Clin Lung Cancer 2024 Jan;25(1):e35-e42. [doi: 10.1016/j.cllc.2023.09.007] [Medline: 37880075]
10. Reynolds R, Dennis S, Hasan I, et al. A systematic review of chronic disease management interventions in primary care. BMC Fam Pract 2018 Jan 9;19(1):11. [doi: 10.1186/s12875-017-0692-3] [Medline: 29316889]
11. Rostami A, Sabet HR, Azarm E, et al. Self-care management, patient education, and nursing support in patients with diabetes and hypertension. Galen Med J 2024;13:e3166. [doi: 10.31661/gmj.v13i.3166]
12. Timmis A, Vardas P, Townsend N, et al. European Society of Cardiology: cardiovascular disease statistics 2021. Eur Heart J 2022 Feb 22;43(8):716-799. [doi: 10.1093/eurheartj/ehab892] [Medline: 35016208]
13. Bui TT, Nguyen THT, Lee J, Kim SY, Oh JK. Determinants of unhealthy living by gender, age group, and chronic health conditions across districts in Korea using the 2010-2017 community health surveys. Epidemiol Health 2024;46:e2024014. [doi: 10.4178/epih.e2024014] [Medline: 38228089]
14. Cardiovascular diseases (CVDs). World Health Organization. 2021. URL: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) [accessed 2026-01-06]
15. Gong Q, Zhang P, Wang J, et al. Morbidity and mortality after lifestyle intervention for people with impaired glucose tolerance: 30-year results of the Da Qing diabetes prevention outcome study. Lancet Diabetes Endocrinol 2019 Jun;7(6):452-461. [doi: 10.1016/S2213-8587(19)30093-2] [Medline: 31036503]
16. Kaur IP, Arora K, Dhalani M, et al. The synergy of percutaneous coronary intervention and lifestyle modification in reducing mortality and blockage prevention. Cardiol Rev 2026;34(1):40-46. [doi: 10.1097/CRD.0000000000000658] [Medline: 41359572]
17. Visseren FLJ, Mach F, Smulders YM, et al. ESC guidelines on cardiovascular disease prevention in clinical practice: developed by the Task Force for cardiovascular disease prevention in clinical practice with representatives of the European Society of Cardiology and 12 medical societies with the special contribution of the European Association of Preventive Cardiology (EAPC). Eur Heart J 2021;42:3227-3337. [doi: 10.1093/eurheartj/ehab484]
18. Petek Šter M, Barbič Žagar B. Obravnava bolnikov z arterijsko hipertenzijo v ambulantah družinske medicine v Sloveniji – ocena stanja po uvedbi referenčnih ambulant [Article in Slovene]. ZdravVestn 2021;90(9-10):478-487. [doi: 10.6016/ZdravVestn.3069]
19. Gómez-Pardo E, Fernández-Alvira JM, Vilanova M, et al. A comprehensive lifestyle peer group-based intervention on cardiovascular risk factors: the randomized controlled Fifty-Fifty Program. J Am Coll Cardiol 2016 Feb 9;67(5):476-485. [doi: 10.1016/j.jacc.2015.10.033] [Medline: 26562047]
20. Milani RV, Bober RM, Lavie CJ. The role of technology in chronic disease care. Prog Cardiovasc Dis 2016;58(6):579-583. [doi: 10.1016/j.pcad.2016.01.001] [Medline: 26772623]
21. Sharma N, Phadnis DRM. Role of dietary patterns and lifestyle factors in prevention of cardiovascular disease. Int J Home Sci 2024 Jan 1;10(2):453-461. [doi: 10.22271/23957476.2024.v10.i2g.1658]
22. Baumeister H, Montag C. Digital Phenotyping and Mobile Sensing New Developments in Psychoinformatics: Springer Cham; 2022. [doi: 10.1007/978-3-030-98546-2]
23. How many smartphones are in the world? BankMyCell. 2023. URL: https://www.bankmycell.com/blog/how-many-phones-are-in-the-world [accessed 2026-01-06]
24. Zimmermann G, Venkatesan A, Rawlings K, Scahill MD. Improved glycemic control with a digital health intervention in adults with type 2 diabetes: retrospective study. JMIR Diabetes 2021 Jun 2;6(2):e28033. [doi: 10.2196/28033] [Medline: 34075880]
25. Asmat K, Dhamani K, Froelicher ES, Gul R. A patient-centered self-management intervention to improve glycemic control, self-efficacy and self-care behaviors in adults with type 2 diabetes mellitus: a SPIRIT compliant study protocol for randomized controlled trial. Diabetes Metab Syndr Obes 2023;16:225-236. [doi: 10.2147/DMSO.S385715] [Medline: 36760576]
26. Yilmaz E, Uzuner A, Bajgora M, et al. Effect of eTansiyon smartphone application on hypertension control. Prim Health Care Res Dev 2022 Oct 19;23:e64. [doi: 10.1017/S146342362100058X] [Medline: 36259143]
27. Widmer RJ, Allison TG, Lerman LO, Lerman A. Digital health intervention as an adjunct to cardiac rehabilitation reduces cardiovascular risk factors and rehospitalizations. J Cardiovasc Transl Res 2015 Jul;8(5):283-292. [doi: 10.1007/s12265-015-9629-1] [Medline: 25946990]
28. Krzowski B, Peller M, Boszko M, et al. Mobile app and digital system for patients after myocardial infarction (afterAMI): study protocol for a randomized controlled trial. Trials 2022 Jun 21;23(1):522. [doi: 10.1186/s13063-022-06463-x] [Medline: 35729626]
29. Svenšek A, Gosak L, Lorber M, Štiglic G, Fijačko N. Review and comparative evaluation of mobile apps for cardiovascular risk estimation: usability evaluation using mHealth app usability questionnaire. JMIR Mhealth Uhealth 2025 May 8;13:e56466. [doi: 10.2196/56466] [Medline: 40341099]
30. Ooge J, Stiglic G, Verbert K. Explaining artificial intelligence with visual analytics in healthcare. WIREs Data Min Knowl 2022 Jan;12(1):e1427. [doi: 10.1002/widm.1427]
31. Stiglic G, Kocbek P, Fijacko N, Zitnik M, Verbert K, Cilar L. Interpretability of machine learning - based prediction models in healthcare. WIREs Data Min Knowl 2020 Sep;10(5). [doi: 10.1002/widm.1379]

XSL•FO
RenderX

32. Chatterjee P, Armentano RL. Internet of Things for a smart and ubiquitous eHealth system. Presented at: 2015 International Conference on Computational Intelligence and Communication Networks (CICN); Dec 12-14, 2015; Jabalpur, India p. 903-907. [doi: 10.1109/CICN.2015.178]
33. Rodbard D. Continuous glucose monitoring: a review of recent studies demonstrating improved glycemic outcomes. Diabetes Technol Ther 2017 Jun;19(S3):S25-S37. [doi: 10.1089/dia.2017.0035] [Medline: 28585879]
34. Ruedy KJ, Parkin CG, Riddlesworth TD, Graham C, DIAMOND Study Group. Continuous glucose monitoring in older adults with type 1 and type 2 diabetes using multiple daily injections of insulin: results from the DIAMOND trial. J Diabetes Sci Technol 2017 Nov;11(6):1138-1146. [doi: 10.1177/1932296817704445] [Medline: 28449590]
35. Takahashi H, Iwahashi N, Kirigaya J, et al. Glycemic variability determined with a continuous glucose monitoring system can predict prognosis after acute coronary syndrome. Cardiovasc Diabetol 2018 Aug 18;17(1):116. [doi: 10.1186/s12933-018-0761-5] [Medline: 30121076]
36. Vitolo M, Ziveri V, Gozzi G, et al. DIGItal health literacy after COVID-19 outbreak among frail and non-frail cardiology patients: the DIGI-COVID study. J Pers Med 2022 Dec 31;13(1):99. [doi: 10.3390/jpm13010099] [Medline: 36675760]
37. Cederberg M, Alsén S, Ali L, et al. Effects of a person-centered eHealth intervention for patients on sick leave due to common mental disorders (PROMISE study): open randomized controlled trial. JMIR Ment Health 2022 Mar 15;9(3):e30966. [doi: 10.2196/30966] [Medline: 35289756]
38. Busse TS, Nitsche J, Kernebeck S, et al. Approaches to improvement of digital health literacy (eHL) in the context of person-centered care. Int J Environ Res Public Health 2022 Jul 7;19(14):8309. [doi: 10.3390/ijerph19148309] [Medline: 35886158]
39. McCoy CE. Understanding the intention-to-treat principle in randomized controlled trials. West J Emerg Med 2017 Oct;18(6):1075-1078. [doi: 10.5811/westjem.2017.8.35985] [Medline: 29085540]
40. Darviri C, Alexopoulos EC, Artemiadis AK, et al. The Healthy Lifestyle and Personal Control Questionnaire (HLPCQ): a novel tool for assessing self-empowerment through a constellation of daily activities. BMC Public Health 2014 Sep 24;14:995. [doi: 10.1186/1471-2458-14-995] [Medline: 25253039]
41. Svenšek A, Štiglic G, Lorber M. Validation of a Slovenian version of the Healthy Lifestyle and Personal Control Questionnaire (HLPCQ) for use with patients in family medicine. BMC Public Health 2025 May 14;25(1):1778. [doi: 10.1186/s12889-025-22945-4] [Medline: 40369501]
42. Centre for Person-centred Practice (CPcPR). Queen Margaret University. 2019. URL: https://www.qmu.ac.uk/research-and-knowledge-exchange/expertise/centre-for-person-centred-practice-research [accessed 2026-01-06]
43. Lorber M, Reljić NM, Kegl B, et al. Person-centred care: a support strategy for managing non-communicable diseases. Healthcare (Basel) 2024 Feb 23;12(5):526. [doi: 10.3390/healthcare12050526] [Medline: 38470637]
44. Rojo D, Lamqaddam H, Gosak L, Verbert K. Petal-X: human-centered visual explanations to improve cardiovascular risk communication. arXiv. Preprint posted online on Jun 26, 2024. [doi: 10.48550/arXiv.2406.18690]
45. Hageman S, Pennells L, Ojeda F, et al. SCORE2 risk prediction algorithms: new models to estimate 10-year risk of cardiovascular disease in Europe. Eur Heart J 2021 Jul 1;42(25):2439-2454. [doi: 10.1093/eurheartj/ehab309]
46. Petal-X: Cardiovascular Disease (CVD) Risk Calculator - MedInfo. GitHub. 2025. URL: https://github.com/93degree/petal-x-medinfo2025 [accessed 2026-01-19]
47. Petal-X: Cardiovascular Disease (CVD) Risk Calculator. GitHub. URL: https://93degree.github.io/petal-x-medinfo2025/ [accessed 2026-01-19]
48. Petal-X: Cardiovascular Disease (CVD) Risk Calculator. Observable. 2025. URL: https://93degree.observablehq.cloud/petal-x/ [accessed 2026-01-06]
49. Lauc G, Primorac D. The role of glycans in personalization of preventive health care. Croat Med J 2024 Jun 13;65(3):293-296. [doi: 10.3325/cmj.2024.65.293] [Medline: 38868975]
50. Levine ME, Lu AT, Quach A, et al. An epigenetic biomarker of aging for lifespan and healthspan. Aging (Albany NY) 2018 Apr 30;10(4):573-591. [doi: 10.18632/aging.101414] [Medline: 29676998]
51. Dexcom G7. Dexcom. 2025. URL: https://www.dexcom.com/g7-cgm-system [accessed 2026-01-06]
52. Márquez-Celedonio FG, Téxon-Fernández O, Chávez-Negrete A, Hernández-López S, Marín-Rendón S, Berlín-Lascurain S. Clinical effect of lifestyle modification on cardiovascular risk in prehypertensives: PREHIPER I study. Rev Esp Cardiol 2009 Jan;62(1):86-90. [Medline: 19150019]
53. Ge L, Sadeghirad B, Ball GDC, et al. Comparison of dietary macronutrient patterns of 14 popular named dietary programmes for weight and cardiovascular risk factor reduction in adults: systematic review and network meta-analysis of randomised trials. BMJ 2020 Apr 1;369:m696. [doi: 10.1136/bmj.m696] [Medline: 32238384]
54. Field A. Discovering Statistics Using IBM SPSS Statistics: SAGE Publications; 2017.
55. Svenšek A. Using the visualization for lifestyle change in patients at risk of cardiovascular diseases. ClinicalTrials.gov. 2024. URL: https://classic.clinicaltrials.gov/ct2/show/study/NCT06226948 [accessed 2026-01-06]
56. Hernesniemi JA, Tynkkynen J, Havulinna AS, et al. Significant interactions between traditional risk factors affect cardiovascular risk prediction in healthy general population. Ann Med 2015 Feb;47(1):53-60. [doi: 10.3109/07853890.2014.970570] [Medline: 25405541]

XSL•FO
RenderX

57. Reinikainen J, Kuulasmaa K, Oskarsson V, et al. Regional and temporal differences in the associations between cardiovascular disease and its classic risk factors: an analysis of 49 cohorts from 11 European countries. Eur J Prev Cardiol 2024 Mar 27;31(5):569-577. [doi: 10.1093/eurjpc/zwad359] [Medline: 37976098]
58. Vivek Dhukaram A, Baber C, De Stefanis P. Patient - centred cardio-vascular disease management: end - user perceptions. J Assist Technol 2012 Jun 15;6(2):105-122. [doi: 10.1108/17549451211234966]
59. Kornelius E, Chiou JY, Yang YS, Lu YL, Peng CH, Huang CN. The diabetes shared care program and risks of cardiovascular events in type 2 diabetes. Am J Med 2015 Sep;128(9):977-985. [doi: 10.1016/j.amjmed.2015.03.025] [Medline: 25908394]
60. Dudas K, Olsson LE, Wolf A, et al. Uncertainty in illness among patients with chronic heart failure is less in person-centred care than in usual care. Eur J Cardiovasc Nurs 2013 Dec;12(6):521-528. [doi: 10.1177/1474515112472270] [Medline: 23303766]
61. Maggioni AP, Anker SD, Dahlström U, et al. Are hospitalized or ambulatory patients with heart failure treated in accordance with European Society of Cardiology Guidelines? Evidence from 12,440 patients of the ESC Heart Failure Long-Term Registry. Eur J Heart Fail 2013 Oct;15(10):1173-1184. [doi: 10.1093/eurjhf/hft134] [Medline: 23978433]
62. Rocha J, Barroso I, Caramelo A, Rainho C. Promotion of healthy lifestyles in the elderly integrated into an active aging project in the region of Tâmega and Sousa. Presented at: International Congress of Research in Nursing ESEP; Jul 14-15, 2022. [doi: 10.3390/msf2022017009]
63. Gwinnutt JM, Wieczorek M, Balanescu A, et al. 2021 EULAR recommendations regarding lifestyle behaviours and work participation to prevent progression of rheumatic and musculoskeletal diseases. Ann Rheum Dis 2023 Jan;82(1):48-56. [doi: 10.1136/annrheumdis-2021-222020] [Medline: 35260387]
64. Hahn J, Malspeis S, Choi MY, et al. Association of healthy lifestyle behaviors and the risk of developing rheumatoid arthritis among women. Arthritis Care Res (Hoboken) 2023 Feb;75(2):272-276. [doi: 10.1002/acr.24862] [Medline: 35040282]
65. Ibsen DB, Søgaard K, Sørensen LH, et al. Modifiable lifestyle recommendations and mortality in Denmark: a cohort study. Am J Prev Med 2021 Jun;60(6):792-801. [doi: 10.1016/j.amepre.2021.01.003] [Medline: 33775511]
66. Li Y, Fan X, Wei L, Yang K, Jiao M. The impact of high-risk lifestyle factors on all-cause mortality in the US non-communicable disease population. BMC Public Health 2023;23(1):422. [doi: 10.1186/s12889-023-15319-1]
67. Lopez-Fontana I, Perrot A, Krueger KR, Le Scanff C, Bherer L, Castanier C. A global lifestyle assessment: psychometric properties of the general lifestyle questionnaire. Psychologie Française 2020 Dec;65(4):311-323. [doi: 10.1016/j.psfr.2019.06.004]
68. Murray CJL, Abbafati C, Abbas KM, et al. Five insights from the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1135-1159. [doi: 10.1016/S0140-6736(20)31404-5] [Medline: 33069324]
69. van Dijk ML, Te Loo LM, Vrijsen J, et al. LOFIT (Lifestyle front Office For Integrating lifestyle medicine in the Treatment of patients): a novel care model towards community-based options for lifestyle change-study protocol. Trials 2023 Feb 17;24(1):114. [doi: 10.1186/s13063-022-06960-z] [Medline: 36803271]
70. Lundin Gurné F, Svensson PA, Björkman I, Lidén E, Jakobsson S. Seeking lifestyle counselling at primary health care centres: a cross-sectional study in the Swedish population. BMC Prim Care 2023 Mar 20;24(1):74. [doi: 10.1186/s12875-023-02035-3] [Medline: 36941550]
71. Hirooka N, Kusano T, Kinoshita S, Aoyagi R, Hidetomo N. Association between healthy lifestyle practices and life purpose among a highly health-literate cohort: a cross-sectional study. BMC Public Health 2021 Apr 29;21(1):820. [doi: 10.1186/s12889-021-10905-7] [Medline: 33926422]
72. Marques MDC, Pires R, Perdigão M, et al. Patient-centered care for patients with cardiometabolic diseases: an integrative review. J Pers Med 2021 Dec 3;11(12):1289. [doi: 10.3390/jpm11121289] [Medline: 34945763]
73. Deraz O, Van Sloten T, Climie R, et al. Person-centered and measured life's simple 7 cardiovascular health concordance and association with incident cardiovascular disease. Sci Rep 2023 Mar 31;13(1):5247. [doi: 10.1038/s41598-023-32219-x] [Medline: 37002422]
74. Fors A, Swedberg K, Ulin K, Wolf A, Ekman I. Effects of person-centred care after an event of acute coronary syndrome: two-year follow-up of a randomised controlled trial. Int J Cardiol 2017 Dec 15;249:42-47. [doi: 10.1016/j.ijcard.2017.08.069] [Medline: 28893432]
75. Pardo Y, Garin O, Oriol C, Zamora V, Ribera A, Ferrer M. Patient-centered care in coronary heart disease: what do you want to measure? A systematic review of reviews on patient-reported outcome measures. Qual Life Res 2023 May;32(5):1405-1425. [doi: 10.1007/s11136-022-03260-6] [Medline: 36350473]
76. Rossi LP, Granger BB, Bruckel JT, et al. Person-centered models for cardiovascular care: a review of the evidence: a scientific statement from the American Heart Association. Circulation 2023 Aug 8;148(6):512-542. [doi: 10.1161/CIR.0000000000001141] [Medline: 37427418]
77. Aifah A, Okeke NL, Rentrope CR, et al. Use of a human-centered design approach to adapt a nurse-led cardiovascular disease prevention intervention in HIV clinics. Prog Cardiovasc Dis 2020;63(2):92-100. [doi: 10.1016/j.pcad.2020.02.013] [Medline: 32092444]

XSL•FO
RenderX

78. Hwang M, Chang AK. The effect of nurse-led digital health interventions on blood pressure control for people with hypertension: a systematic review and meta-analysis. J Nurs Scholarsh 2023 Sep;55(5):1020-1035. [doi: 10.1111/jnu.12882] [Medline: 36929538]
79. Stephen C, Halcomb E, Batterham M, McInnes S, Zwar NA. Impact of a general practice nurse intervention to improve blood pressure control: the ImPress study. Aust J Gen Pract 2023 Dec;52(12):875-881. [doi: 10.31128/AJGP-09-22-6573] [Medline: 38049138]
80. Zhang J, Zheng X, Ma D, Liu C, Ding Y. Nurse-led care versus usual care on cardiovascular risk factors for patients with type 2 diabetes: a systematic review and meta-analysis. BMJ Open 2022 Mar;12(3):e058533. [doi: 10.1136/bmjopen-2021-058533]

Abbreviations

BP: blood pressure
CGM: continuous glucose monitoring
CONSORT: Consolidated Standards of Reporting Trials
CVD: cardiovascular disease
HDL: high-density lipoprotein
HLPCQ: Healthy Lifestyle and Personal Control Questionnaire
PCC: person-centered care
PCPI-SU: Person-Centered Practice Inventory—Service User
RCT: randomized controlled trial
SCORE2: Systematic Coronary Risk Evaluation 2

Edited by A Mavragani, T Sanchez; submitted 03.Sep.2025; peer-reviewed by A Rakhab, P Soukcharoen; revised version received 22.Dec.2025; accepted 23.Dec.2025; published 22.Jan.2026.

Please cite as:
Svenšek A, Lorber M, Klemenc-Ketiš Z, Gosak L, Muršec D, Štiglic G
Interactive Digital Visualization Counseling for Lifestyle Change in Patients at Risk of Cardiovascular Diseases: Randomized Controlled Trial
JMIR Public Health Surveill 2026;12:e83488
URL: https://publichealth.jmir.org/2026/1/e83488
doi:10.2196/83488

© Adrijana Svenšek, Mateja Lorber, Zalika Klemenc-Ketis, Lucija Gosak, Dominika Muršec, Gregor Stiglic. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 22.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Using Natural Language Processing to Examine State Child Maltreatment Policies and Associations With Outcomes: Multistate Cross-Sectional Study

Zhidi Luo[1,2], PhD; Richard A Epstein[1,2], MPH, PhD; Nethra Sambamoorthi[3], PhD; Lutfiyya N Muhammad[4], PhD, MPH; Neil Jordan[1,4,5], PhD

[1]Department of Psychiatry and Behavioral Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL, United States
[2]Health Sciences Integrated Program, Northwestern University Feinberg School of Medicine, Chicago, IL, United States
[3]School of Professional Studies, Northwestern University, Chicago, IL, United States
[4]Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, United States
[5]Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL, United States

Corresponding Author:
Zhidi Luo, PhD
Department of Psychiatry and Behavioral Sciences
Northwestern University Feinberg School of Medicine
Abbott Hall, 12th Floor, 710 N. Lake Shore Dr
Chicago, IL, 60611
United States
Phone: 1 312 926 2323
Email: zhidi.luo@northwestern.edu

Abstract

Background: Child maltreatment is a major public health issue in the United States, with substantial variation in how states define, report, and respond to abuse and neglect. While prior research has examined individual policy components, less is known about how multiple policies jointly shape broader policy environments and relate to maltreatment outcomes.

Objective: This study used natural language processing (NLP) to characterize state child maltreatment policy environments and examine their associations with maltreatment incidence, recurrence, and fatalities across US jurisdictions.

Methods: A cross-sectional study was conducted using 2021 data from 50 US states, the District of Columbia, and Puerto Rico (N=52). A total of 411 state maltreatment policy items were derived from the State Child Abuse & Neglect Policies Database. A total of 6 NLP models (bidirectional and auto-regressive transformer [BART], bidirectional encoder representations from transformers [BERT], robustly optimized BERT approach [RoBERTa], decoding-enhanced BERT with disentangled attention [DeBERTa], Copilot, and LLaMA 3.1) were applied using a zero-shot classification framework to quantify policy characteristics. Models were evaluated using intrinsic (category consistency and semantic alignment) and extrinsic (factor analysis and clustering performance) metrics. Exploratory factor analysis was used to identify latent policy domains, and k-means clustering was applied to group jurisdictions with similar policy profiles. Maltreatment outcomes, including incidence, recurrence, and fatalities, were obtained from the National Child Abuse and Neglect Data System. Outcome differences across clusters were assessed using ANOVA with post hoc pairwise comparisons.

Results: The study population included 72,838,819 children and 3,774,528 maltreatment reports, of which 751,283 were substantiated or indicated cases. Across 6 NLP models, DeBERTa demonstrated the best overall performance. Exploratory factor analysis identified 3 primary policy domains: maltreatment definition, mandated reporting, and alternative response. A total of 4 policy clusters were identified. Jurisdictions with weaker reporting requirements and fewer penalties (cluster 2) had the highest incidence of maltreatment (mean 17.51, SD 6.13 vs mean 9.65, SD 4.10; mean 10.28, SD 6.47; and mean 11.74, SD 6.47 per 1000 children; P=.07) and recurrence (mean 1.74, SD 0.87 vs mean 0.52, SD 0.44; mean 0.66, SD 0.52; and mean 0.68, SD 0.43 per 1000 children; P<.001). Jurisdictions with stronger reporting requirements, broader definitions, and greater use of alternative response systems (cluster 1) had the lowest incidence and recurrence. No statistically significant differences in maltreatment fatalities were observed across clusters (P=.89). However, in a subanalysis of 18 fatality-related policy items, clusters differed significantly in fatality rates, with jurisdictions characterized by less clearly defined fatality policies exhibiting higher fatality

XSL•FO
RenderX

rates (mean 47.91, SD 8.35 vs mean 22.64, SD 11.82; mean 25.63, SD 18.62; and mean 19.03, SD 11.64 per 1,000,000 children; *P*=.03).

Conclusions: State child maltreatment policies form distinct, multidimensional patterns associated with maltreatment outcomes. The findings highlight the importance of evaluating integrated policy frameworks rather than isolated components and demonstrate the use of NLP for large-scale, data-driven policy analysis.

(JMIR Public Health Surveill 2026;12:e99643) doi:10.2196/99643

KEYWORDS

child abuse; child neglect; cluster analysis; clustering analysis; data mining; factor analysis; health policy; natural language processing; public policy

Introduction

Child maltreatment, which includes both child abuse and neglect, is a critical public health concern in the United States. In fiscal year 2023, the most recent year for which information is available, there were 546,159 reported children who experienced child maltreatment, with an estimated 2000 children dying from maltreatment [1]. Child maltreatment is associated with an increased risk of adverse health outcomes and risk behaviors, including substance use, mental health disorders, chronic physical conditions, and premature mortality [2-5]. The economic burden is substantial, with the total lifetime cost estimated at US $2.94 trillion for investigated cases and US $563 billion for substantiated cases in the federal fiscal year 2018 [6].

Child maltreatment policy refers to states' definitions of child abuse and neglect, along with related regulations that shape child welfare practices. These policies determine how allegations are referred, reported, investigated, and managed [7]. For example, mandated reporting expansion refers to policies that broaden the categories of individuals required to report suspected maltreatment (eg, extending beyond professionals such as teachers and health care providers to include additional occupational groups or, in some states, all adults) [8]. Similarly, policies such as centralized intake systems standardize how reports are received and screened, while differential response systems allow for alternative, noninvestigative pathways for lower-risk cases [9]. Together, these policy components influence both the volume and characteristics of cases that enter and progress through the child welfare system.

At the federal level, the Child Abuse Prevention and Treatment Act (CAPTA), first enacted in 1974 and reauthorized multiple times since, establishes minimum standards for state definitions of abuse and neglect and conditions federal funding on states' compliance with these requirements [10]. However, states retain substantial discretion in translating these federal requirements into policy and practice, resulting in significant cross-state variation [7]. This variation leads to substantial differences in reported maltreatment rates and intervention effectiveness, with substantiation rates ranging from 5% to 45% across states [11].

Research suggests that states spending an additional US $1000 per person in poverty on benefits such as cash aid, child care, or Medicaid assistance saw 4.3% fewer maltreatment reports and 4.0% fewer substantiations [12]. Additionally, states that implemented mandated reporting expansion, centralized intake, and increased state staffing saw a 32% increase in maltreatment reports, although substantiated reports declined by 5% to 6% [7]. In contrast, states adopting both differential response and higher standards of proof experienced a 24% decrease in substantiated reports, with differential response accounting for 11% and higher proof standards contributing between 12% and 13% [7]. A separate national quasi-experimental study using data from 2004 to 2017 found that states with differential response programs had approximately 19% fewer substantiated reports and a 17% reduction in foster care use relative to states without differential response [9]. These literatures reflect the complex and sometimes countervailing ways in which policy components interact with one another and with local implementation contexts.

However, existing studies have largely examined individual policies or small sets of policy features in isolation. Less is known about how multiple policy components co-occur within states to form broader policy environments, or how states differ systematically across key policy domains. This limitation is important because child welfare systems operate as integrated structures, where policies governing reporting, screening, investigation, and service response interact with one another. As a result, the relationship between any single policy component and observed outcomes may depend on the broader policy context in which it is embedded.

Recent advancements in natural language processing (NLP), particularly the development of large language models (LLMs), have begun to transform policy research and child welfare research. In policy text mining, zero-shot and few-shot classification have been applied to large corpora of legislative text, demonstrating scalability without extensive hand-labeled training data [13]. Other work has used NLP to construct interpretable, theory-driven measures of legislative quality from statutory text [14] and to track policy status across large sets of climate and environmental documents [15-17]. In child welfare research, NLP and related methods have focused on case-level rather than policy-level data. Predictive risk modeling using structured administrative records has estimated individual maltreatment risk, in some cases outperforming existing screening tools [18,19], while NLP applied to case notes and health records has identified instances of maltreatment [20-23]. These efforts enhanced data-driven decision-making within child welfare agencies by allowing for more nuanced insights into family dynamics and case histories [24]. These approaches demonstrate NLP's value for surfacing information from unstructured text, but case note narratives have also been shown

XSL•FO
RenderX

to embed caseworker bias and uncertainty, raising caution about their use in downstream predictive tasks [25]. However, none of this literature has applied NLP to the formal statutory and regulatory policy documents that define state child welfare systems, as distinct from case-level clinical or casework text.

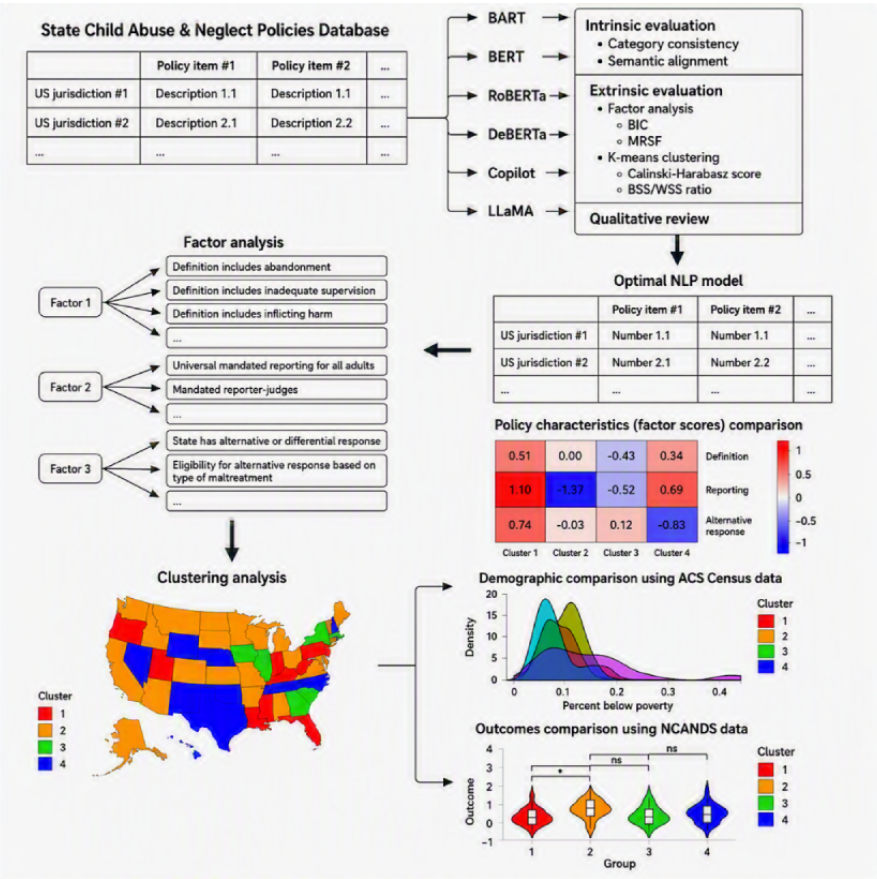
This study uses NLP to identify key components of state child maltreatment policies, examine how states differ across these components, and assess how these policy configurations are associated with maltreatment outcomes in the United States. By characterizing the multidimensional structure of state policy environments, this study aims to provide a more comprehensive understanding of how policy variation corresponds to differences in observed child welfare outcomes.

Methods

Overview

This study was reported in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines. The overall analysis proceeded through 3 stages, as illustrated in Figure 1: comparing and evaluating 6 NLP models and quantifying the 411 state policy items ("Quantification of State Maltreatment Policy Using NLP" section), identifying latent policy dimensions via factor analysis and jurisdiction grouping via k-means clustering ("Exploratory Factor Analysis" and "Clustering Analysis" sections), and comparison of demographic, economic, and maltreatment outcomes across the resulting policy clusters ("Demographic, Economic, and Outcome Comparisons" section).

Figure 1. Analytical workflow for evaluating state child maltreatment policy configurations using natural language processing (NLP), exploratory factor analysis, and clustering among 52 US jurisdictions in 2021. ACS: American Community Survey; BART: bidirectional and auto-regressive transformer; BERT: bidirectional encoder representations from transformers; BIC: Bayesian information criterion; BSS/WSS: between-cluster sum of squares to within-cluster sum of squares ratio; DeBERTa: decoding-enhanced BERT with disentangled attention; MRSF: multiple R-squared of scores with factors; NCANDS: National Child Abuse and Neglect Data System; RoBERTa: robustly optimized BERT approach.



Data Sources

This study used data from 3 sources. State-level child abuse and neglect policy data from 2021 were obtained from the State Child Abuse & Neglect Policies Database (SCAN), which documents state child welfare laws and policies across all 50 US states, the District of Columbia, and Puerto Rico. The database contains 411 standardized policy variables covering state-specific definitions of child abuse and neglect, reporting policies, screening policies, investigation policies, child welfare responses, child welfare system context, and child fatality–related policies in calendar year 2021 [26]. SCAN captures written state laws and policies from publicly available or state-provided sources, including state statutes, regulations, and agency policy documents. Policy variables are coded using a standardized protocol and are verified by state agency contacts when feasible.

Additionally, child welfare agency– and child-level maltreatment data from 2021 were sourced from the National Child Abuse and Neglect Data System (NCANDS), a federally sponsored

XSL•FO
RenderX

program administered by the US Department of Health and Human Services [27,28]. NCANDS collects and analyzes state-reported data on child abuse and neglect known to child protective services, as mandated by the CAPTA [27,28]. Child-level records were used to derive state-level measures of maltreatment incidence and recurrence by aggregating individual cases. Maltreatment fatality rates were obtained from the NCANDS agency-level data because some fatality information is not available in the child-level files for confidentiality and data security reasons.

Finally, demographic and economic data from 2021 were drawn from the American Community Survey, an annual survey conducted by the US Census Bureau that provides detailed estimates of social, economic, housing, and demographic characteristics at the state and local levels [29].

Study Cohort and Outcomes

The study cohort included all 50 states in the United States, the District of Columbia, and the Commonwealth of Puerto Rico, with 411 state maltreatment policy items assessed in each jurisdiction in calendar year 2021. Each policy item represents a specific child welfare policy provision (eg, a mandated reporting requirement, definition of child maltreatment, screening procedure, or investigation practice) and is organized within the SCAN database into predefined policy domains, including child maltreatment definitions, mandated reporting, screening, investigations, child welfare responses, and child welfare system context (Table S1 in Multimedia Appendix 1). These domains are part of the original SCAN database structure. Policy items vary in measurement format depending on the underlying policy provision and include binary, categorical, ordinal, and descriptive variables. Although the SCAN database organizes policy items into predefined domains, these domains were used only to describe the source variables. The subsequent analyses identified latent policy dimensions empirically using exploratory factor analysis (EFA) rather than relying on the original SCAN categorization.

Maltreatment outcomes were evaluated for individuals aged <18 years. While some states, such as Illinois, allow youth aged 18 years or older to remain in care, others do not. To ensure consistent denominators across jurisdictions, individuals aged 18 years or older were excluded.

The study examined three outcomes: (1) incidence of maltreatment, measured as the number of substantiated or indicated maltreatment cases per 1000 children aged <18 years in each state; (2) maltreatment recurrence, defined as the number of children previously identified as having experienced maltreatment who experienced subsequent incidents per 1000 children in the population; and (3) maltreatment fatalities, defined as the number of deaths resulting from maltreatment per 1,000,000 children. For the first 2 outcomes, only substantiated or indicated maltreatment recorded by state child protective services in NCANDS with report dates in 2021 were included. Although these outcomes represent related dimensions of child maltreatment burden, they capture distinct aspects of child welfare involvement and were analyzed separately rather than modeled as a combined outcome.

Quantification of State Maltreatment Policy Using NLP

Overview

The primary study variable was state maltreatment policy. To prepare these data, 411 state maltreatment policy items from 2021 across 52 US jurisdictions were processed using NLP models.

NLP Approach

This study applied NLP methods to classify and analyze 411 state maltreatment policy items from 2021 across 52 US jurisdictions. A zero-shot classification [30], which uses pretrained models to assign labels without task-specific fine-tuning, was selected because the policy data were not labeled for the specific classification tasks required in this study. Alternative approaches, such as supervised or few-shot classification, typically require curated training examples for each category, which were not available and would require substantial manual annotation. Zero-shot methods provide a scalable and consistent framework for applying predefined policy categories across heterogeneous statutory text while avoiding potential bias introduced through hand-labeled training data.

To evaluate the robustness of this approach, 6 NLP models were compared, representing distinct architectures and training paradigms, including encoder-based, sequence-to-sequence, and generative LLMs. Comparing these models allows assessment of whether classification results are consistent across fundamentally different modeling approaches, and whether more recent generative models provide advantages over traditional transformer-based classifiers in capturing the structure of policy text. The 6 models evaluated were as follows:

1. Bidirectional and auto-regressive transformer (BART) [31], a sequence-to-sequence transformer with a bidirectional encoder and an autoregressive decoder, designed for classification and text generation tasks;
2. Bidirectional encoder representations from transformers (BERT) [32], a bidirectional transformer encoder trained using masked-language modeling for deep contextual text understanding;
3. Robustly optimized BERT approach (RoBERTa) [33], an optimized version of BERT trained longer and on larger corpora to improve robustness and performance;
4. Decoding-enhanced BERT with disentangled attention (DeBERTa) [34], a transformer model that enhances BERT/RoBERTa with disentangled attention and improved position encoding;
5. Microsoft Copilot [35], generative transformer models adapted for instruction-following and classification;
6. Meta LLaMA 3.1 [36], an LLM in the LLaMA family, trained as an autoregressive transformer with strong general-purpose language understanding capabilities.

The models evaluated each policy item independently by assessing whether each jurisdiction's policy text supports or contradicts the policy item's title description. For example, for the policy item titled "How is reporting decentralized?," jurisdiction responses include "Each county has its own

XSL•FO
RenderX

reporting hotline," "Some counties have their own reporting hotlines," and "No county has its own reporting hotline." Each NLP model assigned a likelihood score from 0 to 1 to each response, indicating the degree to which it supported the policy title, where 1 represented the strongest support and 0 the weakest. For instance, the BART model assigned scores of 0.90, 0.81, and 0.59, respectively.

For LLMs, Microsoft Copilot and Meta LLaMA 3.1, the scores were generated by a structured prompt, which was applied separately to each policy item and took the following format:

"Please estimate the likelihood (on a scale from 0 to 1) that the response [A], [B], [C] supports or opposes the statement [Policy Title]. Stronger support should be closer to 1, a neutral or irrelevant stance should be around 0.5, and complete opposition should be close to 0. Provide likelihoods in this format: [A]: score1; [B]: score2; [C]: score3."

NLP Evaluation Metrics

The 6 models were evaluated using intrinsic and extrinsic metrics [37,38]. Intrinsic evaluation measured internal consistency and agreement by testing how well LLM embedding vectors capture the underlying characteristics of the texts, while extrinsic evaluation assessed the performance of key embedding features extracted from the policy documents based on their performance in downstream tasks and analyses [37]. Since later sections of the study applied EFA [39] and k-means clustering [40] to policy data, these techniques were used here as extrinsic evaluation metrics to determine how well different NLP models capture meaningful variation in policy content for the same set of features.

For intrinsic evaluation, 3 metrics were used. First, category consistency was assessed based on the relative ordering of classification scores across response categories. Specifically, items labeled as "Yes" were expected to receive higher scores than those labeled "No," while intermediate responses such as "Not Applicable," "Unknown," or "Missing" were expected to fall between "Yes" and "No." Second, for descriptive responses, semantic alignment was measured using 2 metrics. Because a true gold standard was unavailable for these responses, 2 proxy gold standards were used: (1) cosine similarity scores, where higher values indicate better semantic alignment, and (2) the average score across all methods. The correlation between each model's results and these proxy standards, where a higher correlation indicates better agreement, served as the metric to assess semantic alignment.

Extrinsic evaluation assessed EFA and clustering results derived from each NLP model-specific policy representation. In this stage, each model produced a complete set of scores for the 411 policy items across jurisdictions, which served as inputs for EFA and k-means clustering. Model performance was then assessed by comparing goodness-of-fit and clustering metrics to identify which NLP approaches most effectively capture meaningful structure in the policy data. EFA results were evaluated using the Bayesian information criterion (BIC) [41], where lower values indicate better model fit accounting for complexity, and the multiple R-squared of scores with factors (MRSF) [39], where higher values reflect better explanatory power of the factor scores. To explore different latent structures, EFA models with 2, 3, and 4 factors were examined. Clustering performance was analyzed using the Calinski-Harabasz score (CH) [42], where higher values indicate better cluster separation relative to complexity, and the between-cluster sum of squares to within-cluster sum of squares (BSS/WSS) ratio [43], where higher ratios represent more compact and well-separated clusters. Models with 2, 3, and 4 clusters were examined to evaluate clustering performance.

Qualitative Review

Finally, a qualitative review explored why certain methods perform better on specific metrics by analyzing model outputs, identifying recurring error patterns, and assessing consistency across similar policy items. This review considered factors such as sensitivity to nuanced language, handling of ambiguous or incomplete responses, and stability in scoring across different policy categories. This review aimed to gain an intuitive understanding of model behavior and to verify that the results aligned with expectations.

Overall model performance was determined by considering results across all intrinsic and extrinsic evaluation metrics rather than relying on a single metric or composite score. No weighting scheme was applied. Instead, models were compared based on their relative performance patterns across evaluation dimensions, with qualitative review used to further examine model behavior, identify potential sources of variation, and assess whether quantitative findings were consistent with expected policy interpretation.

Covariates

Demographic and economic characteristics were included as covariates to contextualize jurisdiction-level differences. Demographic covariates included age distribution, racial and ethnic composition, and gender. Economic conditions were measured by the proportion of the population living below 50%, 100%, 125%, and 400% of the federal poverty level, capturing varying levels of economic vulnerability.

Statistical Analysis

Exploratory Factor Analysis

Although the SCAN database organizes policy items into predefined domains for documentation purposes, these domains do not necessarily represent latent policy constructs. EFA was conducted on the 411 state maltreatment policy items to identify latent dimensions [39]. Consequently, policy items originating from the same SCAN domain were not constrained to load on the same factor; likewise, policy items from different SCAN domains could load on the same factor if they reflected similar underlying policy characteristics. The number of factors was determined using a scree plot, a graphical tool that helps identify the point where the rate of decrease in explained variance slows [44]. While this plot provides a helpful visual cue, identifying the "elbow" can be subjective and challenging [44,45]. In cases where the exact location of the "elbow" is ambiguous and several factor number choices were similarly justifiable, we opted for the smaller number of factors to ensure concise and easier interpretation.

XSL•FO
RenderX

To identify the underlying themes for each factor, policy items with high factor loadings were reviewed. These factors were then labeled based on the substantive patterns they capture. For each factor in each jurisdiction, factor scores were computed to quantify the degree to which each jurisdiction aligns with the identified policy dimensions [46]. The factor scores provided a reduced representation of the original 411 policy items and were used in subsequent analyses to characterize jurisdiction-level policy configurations.

The internal consistency of policy items contributing to each identified factor was assessed using McDonald ω [47]. ω was calculated based on policy items with absolute factor loadings above 0.4.

Clustering Analysis

K-means clustering was applied to the 411 state maltreatment policy items to identify groups of jurisdictions that share similar policy characteristics [40]. The optimal number of clusters was determined using the Silhouette method [48], which evaluated cluster cohesion and separation, along with interpretability considerations, to ensure the resulting groupings were meaningful and practically relevant. After the number of clusters is selected, each jurisdiction is assigned to a cluster based on the similarity of its policy item representations to the identified cluster profiles. To visually illustrate the geographic distribution of policy clusters, a US map is generated, with jurisdictions color-coded according to their assigned cluster. The policy characteristics of each cluster are identified by comparing average factor scores derived from the factor analysis.

Although k-means clustering and factor analysis are distinct techniques, they serve complementary purposes in this study. Factor analysis reduces the dimensionality of the policy items and identifies underlying policy domains, whereas clustering groups jurisdictions based on similarity across the full set of policy features. By conducting clustering on all items and using factor scores for interpretation, this approach preserves the full information available for grouping while enabling a more interpretable, lower-dimensional characterization of cluster differences. Consistency between cluster profiles and factor structures further supports the coherence and robustness of the observed policy patterns.

Demographic, Economic, and Outcome Comparisons

To characterize each cluster, policy factor scores were compared across clusters. Then, demographic, economic, and outcome differences were examined across jurisdiction clusters. Statistical differences across clusters were tested using ANOVA for continuous variables [49] and chi-square tests for categorical variables [50]. When significant differences are detected, post hoc pairwise comparisons are conducted to determine which clusters differ from each other. Given the small sample size of 52 US jurisdictions, a significance level of 0.10 is used to reduce the risk of type II errors and detect potential associations that may warrant further investigation [51,52].

Sensitivity Analyses

Leave-One-State-Out Sensitivity Analysis

Cluster robustness was assessed using a leave-one-jurisdiction-out sensitivity analysis. K-means clustering was repeated after sequentially removing each jurisdiction, and agreement between the resulting cluster assignments and the original clustering was evaluated using the adjusted Rand index (ARI) [53]. This analysis was conducted to evaluate whether the identified policy configurations were robust to the exclusion of individual jurisdictions and not driven by any single state's policy profile.

Sensitivity Analysis Including All NCANDS Maltreatment Records

To assess whether findings were sensitive to the definition of maltreatment outcomes, a sensitivity analysis was conducted using all screened-in maltreatment records reported in NCANDS in 2021, regardless of substantiation or indication status. The same outcome comparison procedures across policy clusters were applied to assess whether the observed associations between policy configurations and maltreatment outcomes remained consistent under an alternative maltreatment definition.

Subanalysis

To further examine policies specifically related to maltreatment fatalities, a separate and independent clustering analysis was conducted using only the 18 fatality-related policy items (Table S2 in Multimedia Appendix 1). Unlike the primary analysis, which clustered jurisdictions based on all 411 policy items, this subanalysis clustered jurisdictions solely according to fatality-related policies.

K-means clustering is applied to these 18 policy items to group jurisdictions with similar fatality-related policy characteristics. The maltreatment fatality rates across these clusters are then compared to identify any notable differences related to policy configurations. A violin plot—a visualization that combines a box plot with a density curve—is used to display the distribution of fatality rates by cluster. Statistical tests are conducted to determine whether the observed differences in fatality rates are statistically significant.

Ethical Considerations

This study was approved by the Northwestern University Institutional Review Board (STU00222024). The study consisted of a secondary analysis of deidentified administrative data obtained from the National Data Archive on Child Abuse and Neglect (NDACAN). Data access was granted by NDACAN under its Data Use Agreement, and all analyses were conducted in accordance with the terms of the Data Use Agreement and applicable institutional policies. Because the study involved retrospective analysis of existing deidentified data with no direct participant contact, the Northwestern University Institutional Review Board waived the requirement for informed consent.

The authors did not have access to information that could directly or indirectly identify individual participants during or after data collection. All datasets provided by NDACAN were deidentified prior to release, and analyses were conducted using

XSL•FO
RenderX

secure data management procedures in compliance with the NDACAN Data Use Agreement, which prohibits attempts to reidentify participants.

No participants were recruited for this secondary data analysis; therefore, no compensation was provided.

Results

Study Cohort Characteristics

The study included 52 jurisdictions in the United States: all 50 states, the District of Columbia, and the Commonwealth of Puerto Rico. In calendar year 2021, these jurisdictions had a combined population of 72,838,819 individuals aged <18 years, with 18,416,713 (25.3%) aged <5 years, and 54,422,106 (74.7%) between ages 5 and 17 years. In the same year, there were 3,774,528 maltreatment records, each representing a unique report and child combination, involving 3,134,446 unique children. Of these, 751,283 records were indicated, affecting 692,891 unique children. Although NCANDS contains individual-level child welfare records, the unit of analysis for this study was the jurisdiction. Individual-level records were aggregated to the jurisdiction level to derive maltreatment outcomes.

NLP Results

Table 1 presents the comparison of the 6 NLP models: BART, BERT, RoBERTa, DeBERTa, Copilot, and LLaMA 3.1. For intrinsic evaluation, category consistency scores indicate that DeBERTa and Copilot achieve the best performance (0.95). Semantic alignment results vary across metrics. Based on the similarity metric, LLaMA 3.1 performs best (0.21), followed by BART (0.19) and DeBERTa (0.14). In contrast, when evaluated using the average score metric, Copilot ranks highest (0.91).

For extrinsic evaluation, RoBERTa demonstrates the best performance in factor analysis across all models, achieving the lowest BIC values and the highest MRSF scores for different numbers of factors (2-factor: BIC=–601336; MRSF=0.98; 3-factor: BIC=–561327; MRSF=0.98; 4-factor: BIC=–535315; MRSF=0.98). DeBERTa follows closely with competitive results (2-factor: BIC=–581536; MRSF=0.98; 3-factor: BIC=–550143; MRSF=0.98; 4-factor: BIC=–526241; MRSF=0.97). For clustering, DeBERTa outperforms other models, achieving the highest CH and BSS/WSS ratios across all cluster solutions (2 clusters: CH=4.45; BSS/WSS=1.07; 3 clusters: CH=3.90; BSS/WSS=1.08; 4 clusters: CH=3.28; BSS/WSS=1.09).

The qualitative review showed key sources of error and model-specific challenges (Table S3 in Multimedia Appendix 1). For category consistency, RoBERTa had the lowest score due to struggles with negations (eg, “does not” or “failure to report”), leading to misclassifications, while LLaMA 3.1 often hallucinated artificial logic. In semantic alignment, LLM performance varied—while these models sometimes fabricated justifications instead of following response structures, they also demonstrated a better understanding of longer descriptive texts. However, evaluating semantic alignment was challenging, as similarity and average scores served as imperfect proxies. When comparing similarity and average score approaches, DeBERTa was neither the best nor the worst, performing well in some contexts but falling short in others. Finally, in extrinsic evaluation, DeBERTa outperformed other models, likely because it better preserves structural relationships and maintains meaningful variance in the data.

Based on intrinsic, extrinsic, and qualitative evaluations, DeBERTa was the best-performing model overall, achieving the highest category consistency, strong semantic alignment, the second-best factor analysis performance, and the best clustering result. Because a gold-standard annotation of policy items did not exist, a formal error rate was not calculated.

XSL•FO
RenderX

Table 1. Comparison of 6 natural language processing methods for quantifying state child maltreatment policies across 52 US jurisdictions using 2021 State Child Abuse & Neglect Policies Database data.

	BART[a]	BERT[b]	RoBERTa[c]	DeBERTa[d]	Copilot	LLaMA
Intrinsic						
Category consistency[e]	0.90	0.81	0.73	0.95	0.95	0.78
Semantic alignment—similarity[f]	0.19	0.04	0.10	0.14	0.10	0.21
Semantic alignment—average[g]	0.77	0.80	0.35	0.68	0.91	0.71
Extrinsic: factor analysis						
Factor analysis 2, BIC[h,i]	–464,493	–460,680	–601,336	–581,536	–457,318	–442,535
Factor analysis 2, MRSF[i]	0.97	0.97	0.98	0.98	0.96	0.97
Factor analysis 3, BIC[i]	–439,225	–439,606	–561,327	–550,143	–437,013	–425,035
Factor analysis 3, MRSF[i]	0.97	0.97	0.98	0.98	0.96	0.97
Factor analysis 4, BIC[i]	–420,815	–423,379	–535,315	–526,241	–419,011	–409,666
Factor analysis 4, MRSF[i]	0.97	0.97	0.98	0.97	0.96	0.96
Extrinsic: k-means						
K-means 2, CH[i]	3.57	2.81	2.67	4.45	2.69	2.81
K-means 2, BSS/WSS[i]	1.06	1.04	1.04	1.07	1.03	1.03
K-means 3, CH[i]	3.07	2.06	2.76	3.90	2.11	2.65
K-means 3, BSS/WSS[i]	1.07	1.03	1.05	1.08	1.03	1.05
K-means 4, CH[i]	2.80	2.49	2.74	3.28	2.04	2.79
K-means 4, BSS/WSS[i]	1.07	1.06	1.07	1.09	1.04	1.07

[a]BART: bidirectional and auto-regressive transformer.

[b]BERT: bidirectional encoder representations from transformers.

[c]RoBERTa: robustly optimized BERT approach.

[d]DeBERTa: decoding-enhanced BERT with disentangled attention.

[e]Category consistency: evaluates whether categorical descriptions align with expected patterns. Higher values indicate better consistency.

[f]Semantic alignment—similarity: using a similarity score as the proxy gold standard to evaluate semantic alignment. Higher values indicate better semantic alignment.

[g]Semantic alignment—average: using the average score across all methods as the proxy gold standard to evaluate semantic alignment. Higher values indicate better semantic alignment.

[h]BIC: Bayesian information criterion.

[i]Extrinsic evaluation metric row titles follow the format "[Method] [Number], [Metric]," where the number represents either the number of factors in exploratory factor analysis or the number of clusters in k-means clustering. The metrics include BIC, multiple R-squared of scores with factors (MRSF), Calinski-Harabasz score (CH), and the between-cluster sum of squares to within-cluster sum of squares ratio (BSS/WSS). Lower BIC values indicate a better fit, while higher values are preferable for MRSF, CH, and BSS/WSS.

EFA Results

The scree plot (Figure S1 in Multimedia Appendix 1) showed that although a clear "elbow" did not appear until over 50 factors—an impractically large number for interpretation—the decline in eigenvalues slowed around 3 factors, suggesting a point of diminishing returns. We interpreted this inflection as a minor elbow and selected a 3-factor solution to avoid the complexity of higher-dimensional models while still retaining meaningful structure.

The items with the highest absolute factor loadings for each factor (Table S4 in Multimedia Appendix 1) indicated that Factor 1 represented the definition domain, Factor 2 focused on the reporting domain, and Factor 3 captured the alternative response domain. Factor 1 (Definition) reflected the scope of maltreatment definitions, with higher scores corresponding to broader definitions that explicitly included emotional harm, abandonment, and sexual abuse, as well as more stringent safe haven requirements. Factor 2 (Reporting) covered policies related to mandated reporting, where higher scores were associated with a greater number of mandated reporters, more

XSL•FO
RenderX

stringent penalties for failure to report, and a reduced emphasis on reporter training. Factor 3 (Alternative Response) captured states' implementation of differential response systems, with higher scores indicating broader eligibility for alternative responses based on case characteristics such as risk level and maltreatment type.

The extracted policy dimensions demonstrated high internal consistency. McDonald ω values were 0.996, 0.959, and 0.938 for the 3 factors, indicating strong coherence among policy items contributing to each latent policy dimension.

Clustering Results

The k-means clustering analysis was used to identify distinct clusters based on policy characteristics. The optimal number of clusters was determined using the silhouette width, which peaked at 2 and 4 clusters (Figure S2 in Multimedia Appendix 1). To better capture the diversity and variability in the policy characteristics across jurisdictions, 4 clusters were chosen.

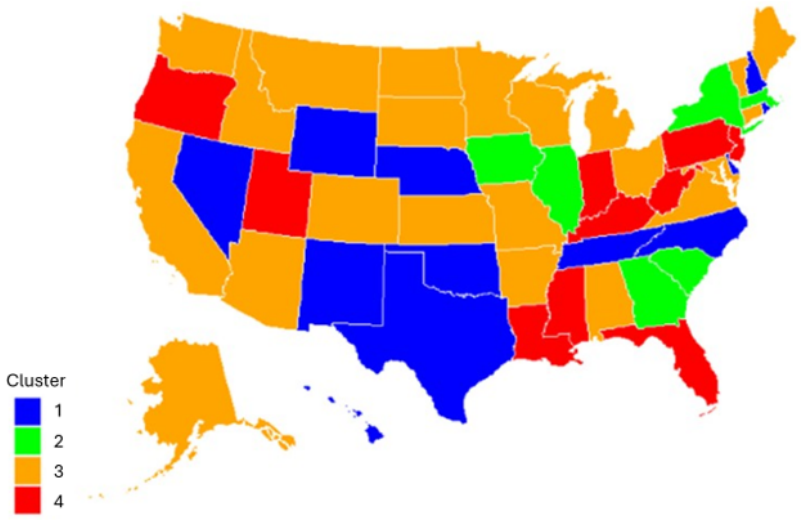
A total of 4 distinct policy clusters with varying geographic distributions across US states were identified (Figure 2). Cluster 1 (n=12) was concentrated in the South and parts of the West, including Texas, Oklahoma, and several Rocky Mountain states. Cluster 2 (n=6) was scattered, appearing in parts of the central, Northeast, and Southeast. Cluster 3 (n=23), the largest cluster, covered much of the central and western regions, as well as parts of the northern region, including the Midwest, Great Plains, and Mountain West. Cluster 4 (n=11) was primarily located in the southern and southeastern US, with a few states in the West also assigned to this group.

The policy characteristics of each cluster were assessed using factor scores (Table 2 and Figure S3 in Multimedia Appendix 1). Cluster 1 had the highest scores across all 3 domains (Definition: 0.51, Reporting: 1.05, Alternative Response: 0.74), reflecting broad maltreatment definitions, strict mandated reporting policies with strong penalties, and extensive use of differential response systems. Cluster 2 showed the lowest Reporting score (–1.37), indicating fewer mandated reporting requirements and weaker penalties. Cluster 3 had the lowest Definition score (–0.43), suggesting narrower maltreatment definitions and more lenient safe haven policies. Cluster 4 had the lowest Alternative Response score (–1.05), reflecting the limited implementation of differential response systems and a likely greater reliance on traditional case-handling approaches.

Regarding demographic and economic differences (Table 3), there were no significant differences in child population among clusters (cluster 1: 1.2 million; cluster 2: 2.1 million; cluster 3: 1.3 million; cluster 4: 1.4 million; P=.76). Demographically, there were minimal differences in age, gender, and race/ethnicity distributions across clusters. However, poverty status showed significant differences (Figure S4 in Multimedia Appendix 1), with cluster 4 showing a higher proportion of children below poverty levels (cluster 1: 0.17; cluster 2: 0.16; cluster 3: 0.15; cluster 4: 0.21; P=.048).

Regarding maltreatment outcomes (Table 4), cluster 2 exhibited the highest incidence of maltreatment, while cluster 1 had the lowest (cluster 1: mean 9.65, SD 4.10; cluster 2: mean 17.51, SD 6.13; cluster 3: mean 10.28, SD 6.47; cluster 4: mean 11.74, SD 6.47; P=.07). A similar trend was observed for maltreatment recurrence, with cluster 2 showing the highest rate (cluster 1: mean 0.52, SD 0.44; cluster 2: mean 1.74, SD 0.87; cluster 3: mean 0.66, SD 0.52; cluster 4: mean 0.68, SD 0.43; P<.001). In contrast, there were no significant differences in maltreatment fatalities across clusters (cluster 1: mean 21.63, SD 14.83; cluster 2: mean 25.74, SD 14.70; cluster 3: mean 25.80, SD 15.64; cluster 4: mean 23.20, SD 18.80; P=.89). Pairwise comparisons (Figure 3) revealed that cluster 2 had a significantly higher incidence of maltreatment than Clusters 1 and 4 and a significantly higher recurrence of maltreatment than Clusters 1, 3, and 4.

Figure 2. Geographic distribution of state child maltreatment policy clusters across 52 US jurisdictions based on 2021 State Child Abuse & Neglect Policies Database data.



XSL•FO
RenderX

Table 2. Comparison of latent child maltreatment policy factor scores across jurisdiction clusters identified through 2021 policy data from the State Child Abuse & Neglect Policies Database.

	Cluster 1 (n=12), mean (SD)	Cluster 2 (n=6), mean (SD)	Cluster 3 (n=23), mean (SD)	Cluster 4 (n=11), mean (SD)	*P* value[a]
Definition	0.51 (1.19)	0.00 (1.09)	–0.43 (0.58)	0.34 (1.15)	.03
Reporting	1.05 (0.42)	–1.37 (0.37)	–0.53 (0.44)	0.69 (0.83)	<.001
Alternative response	0.74 (0.76)	–0.03 (0.79)	0.12 (0.95)	–1.05 (0.23)	<.001

[a]*P* values are based on 1-way ANOVA tests.

Table 3. Comparison of demographic characteristics and poverty levels across state child maltreatment policy clusters among 52 US jurisdictions in 2021.

	Cluster 1 (n=12), mean (SD)	Cluster 2 (n=6), mean (SD)	Cluster 3 (n=23), mean (SD)	Cluster 4 (n=11), mean (SD)	*P* value[a]
Child population (in millions)	1.2 (2.0)	2.1 (1.2)	1.3 (1.7)	1.4 (1.1)	.76
Age (years)					
<5	0.25 (0.01)	0.25 (0.01)	0.26 (0.02)	0.25 (0.02)	.44
5-17	0.75 (0.01)	0.75 (0.01)	0.74 (0.02)	0.75 (0.02)	.44
Gender					
Male	0.50 (0.01)	0.49 (0.01)	0.50 (0.01)	0.49 (0.01)	.13
Female	0.50 (0.01)	0.51 (0.01)	0.50 (0.01)	0.51 (0.01)	.13
Race and ethnicity					
White alone	0.62 (0.19)	0.64 (0.12)	0.70 (0.15)	0.67 (0.18)	.56
Black alone	0.08 (0.07)	0.15 (0.10)	0.10 (0.11)	0.12 (0.11)	.52
Hispanic or Latino	0.18 (0.14)	0.12 (0.06)	0.11 (0.10)	0.19 (0.28)	.46
Poverty status					
Below 50% poverty level	0.06 (0.01)	0.06 (0.01)	0.06 (0.01)	0.08 (0.05)	.11
Below poverty level	0.17 (0.04)	0.16 (0.02)	0.15 (0.03)	0.21 (0.10)	.048
Below 125% poverty level	0.21 (0.05)	0.20 (0.03)	0.19 (0.03)	0.26 (0.12)	.04
Below 400% poverty level	0.13 (0.03)	0.13 (0.02)	0.12 (0.02)	0.16 (0.09)	.06

[a]*P* values are based on 1-way ANOVA tests.

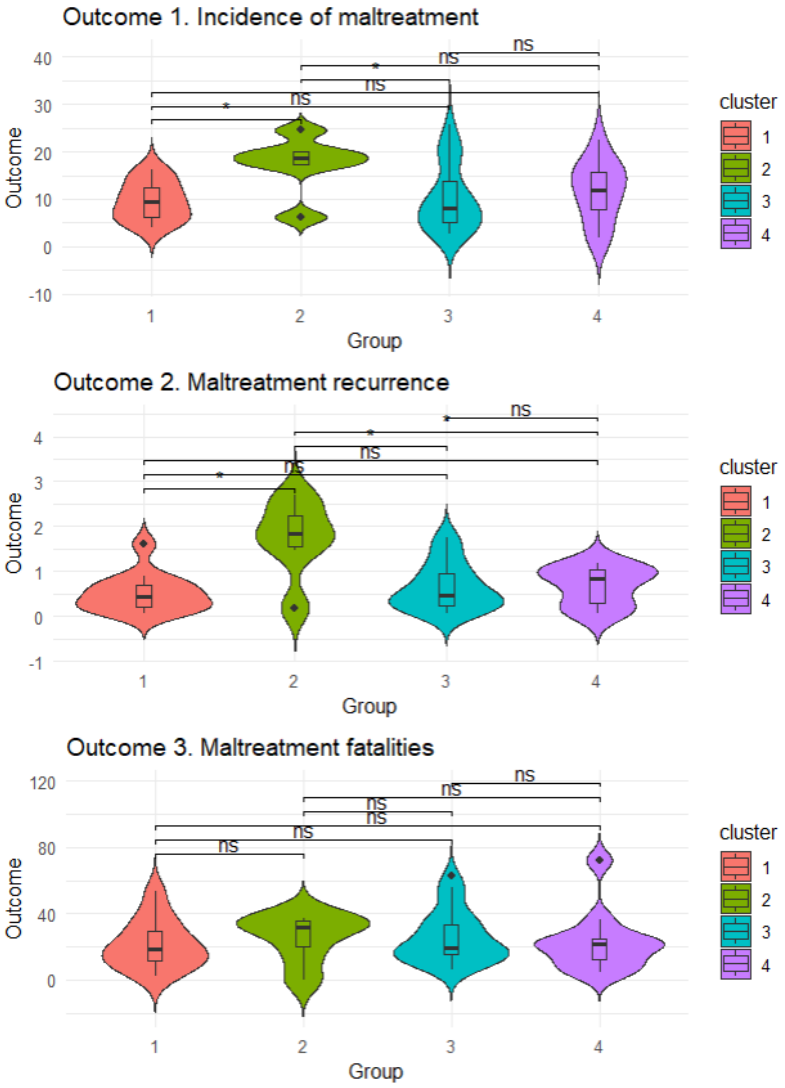
Table 4. Comparison of child maltreatment outcomes across state child maltreatment policy clusters among 52 US jurisdictions in 2021.

	Cluster 1 (n=12), mean (SD)	Cluster 2 (n=6), mean (SD)	Cluster 3 (n=23), mean (SD)	Cluster 4 (n=11), mean (SD)	*P* value[a]
Incidence of maltreatment	9.65 (4.10)	17.51 (6.13)	10.28 (6.47)	11.74 (6.47)	.07
Maltreatment recurrence	0.52 (0.44)	1.74 (0.87)	0.66 (0.52)	0.68 (0.43)	<.001
Maltreatment fatalities	21.63 (14.83)	25.74 (14.70)	25.80 (15.64)	23.20 (18.80)	.89

[a]*P* values are based on 1-way ANOVA tests.

XSL•FO
RenderX

Figure 3. Pairwise comparisons of maltreatment outcomes across state child maltreatment policy clusters among 52 US jurisdictions in 2021.



Sensitivity Analyses

The leave-one-jurisdiction-out sensitivity analysis demonstrated high cluster stability (mean ARI 0.984, range 0.726-1.000), suggesting that the identified policy configurations were not driven by individual jurisdictions.

A sensitivity analysis using all screened-in NCANDS maltreatment records in 2021, regardless of substantiation or indication status, showed similar findings. Cluster 2 continued to have the highest maltreatment incidence and recurrence, although only the difference between cluster 2 and cluster 3 remained statistically significant (Figure S5 and Table S5 in Multimedia Appendix 1).

Subanalysis Results

An independent clustering analysis based solely on the 18 fatality-related policy items (Figure S6 in Multimedia Appendix 1) identified 4 fatality-policy clusters, which are distinct from the 4 policy clusters reported in the primary analysis. Among these fatality-policy clusters, cluster 2 had the least clear definition for child fatalities and near-fatalities, with the lowest scores in “child fatalities definition includes injury,” “child fatalities definition includes death of children in foster care,” “child fatalities definition includes other,” “child near-fatalities definition includes general reference to condition/injury,” and “child near-fatalities definition includes specific injury or treatment” (Table S6 in Multimedia Appendix 1).

Regarding outcomes, there were no significant differences among fatality-related policy clusters in terms of incidence of maltreatment or maltreatment recurrence (Table 5). However, a significant difference was found in maltreatment fatalities (cluster 1: mean 22.64, SD 11.82; cluster 2: mean 47.91, SD 8.35; cluster 3: mean 25.63, SD 18.62; cluster 4: mean 19.03, SD 11.64; P=.03).

Pairwise comparisons showed that cluster 2 had significantly higher maltreatment fatalities compared to Clusters 1, 3, and 4 (Figure 4). The fatality distribution for cluster 2 was relatively narrow and centered around a higher median, indicating consistently elevated fatality rates. In contrast, cluster 3 showed a wider and more skewed distribution with greater variability, while clusters 1 and 4 showed more moderate and tightly clustered fatality rates with lower medians.

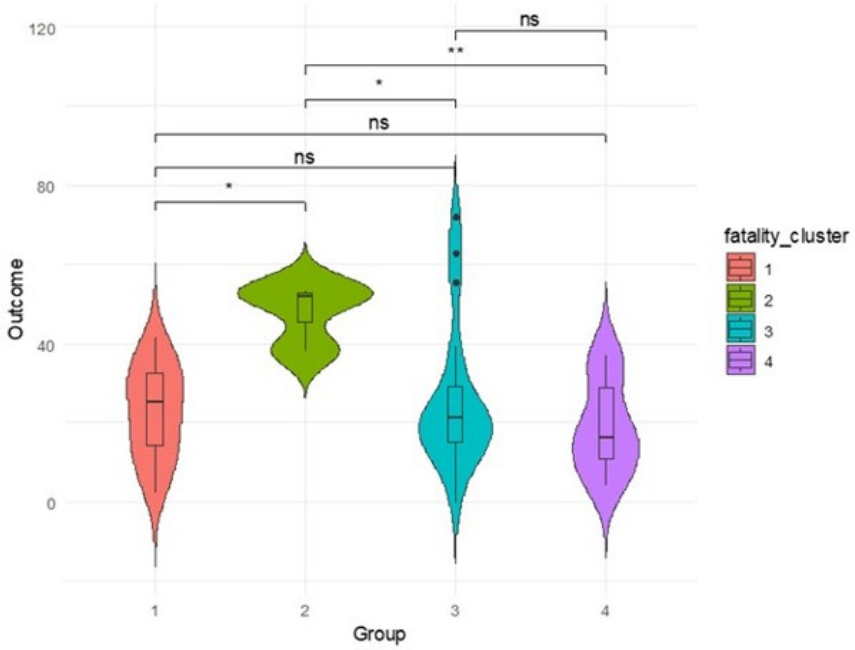
XSL•FO
RenderX

Table 5. Comparison of maltreatment outcomes across jurisdictions grouped by fatality-related policy clusters.

	Cluster 1 (n=15), mean (SD)	Cluster 2 (n=3), mean (SD)	Cluster 3 (n=20), mean (SD)	Cluster 4 (n=14), mean (SD)	*P* value[a]
Incidence of maltreatment	10.57 (7.11)	14.56 (6.21)	11.21 (5.63)	11.43 (7.35)	.82
Maltreatment recurrence	0.61 (0.55)	0.89 (0.21)	0.75 (0.57)	0.89 (0.84)	.68
Maltreatment fatalities	22.64 (11.82)	47.91 (8.35)	25.63 (18.62)	19.03 (11.64)	.03

[a]*P* values are based on 1-way ANOVA tests.

Figure 4. Pairwise outcome differences for subanalysis based on the 18 maltreatment fatality policy items.



Discussion

Summary of Findings

This study analyzed state child maltreatment policies and identified distinct policy clusters associated with maltreatment outcomes. Key findings indicated that states with the fewest mandated reporting requirements and weakest penalties (cluster 2) were associated with the highest substantiated or indicated maltreatment incidence and recurrence rates recorded in NCANDS, whereas states with the strongest reporting requirements, broadest maltreatment definitions, and most extensive use of alternative response (cluster 1) exhibited the lowest rates of substantiated or indicated maltreatment and recurrence. Additionally, states with broad maltreatment definitions and stringent reporting policies (cluster 4) had a high proportion of individuals below the poverty level, although maltreatment outcomes in these states were not significantly different.

Interpretation of Findings

Previous studies found that mandated reporting policies increased the number of maltreatment reports but did not have a significant impact on substantiated or indicated maltreatment incidence [7,54]. This study expanded upon them by suggesting that mandated reporting policies, when coupled with broad definitions of maltreatment and alternative response systems, were associated with lower rates of substantiated or indicated maltreatment and recurrence. These findings highlight the importance of considering the broader policy context in evaluating the role of mandated reporting.

While previous studies had linked child poverty to higher maltreatment reports [55,56], this study found that clusters with broader definitions and limited alternative response systems tended to be more economically disadvantaged. Although maltreatment outcomes in these states were slightly higher, differences were not statistically significant. These findings suggest the importance of considering how policy frameworks may intersect with socioeconomic conditions to shape child welfare involvement.

This relationship between child fatality definitions and maltreatment fatality rates has not been explicitly examined in prior research. These findings suggested that a lack of definitional specificity may influence reporting and intervention efforts, potentially contributing to higher fatality rates.

Policy Implications

The findings suggested several considerations for child maltreatment policy. First, states may benefit from integrating mandated reporting with coordinated prevention and response approaches. Strengthening reporting requirements while also implementing broad maltreatment definitions and alternative response systems was associated with lower rates of

XSL•FO
RenderX

substantiated or indicated maltreatment and recurrence recorded in NCANDS. Developing comprehensive frameworks that balance reporting with both preventive services and supportive response options, such as alternative response systems, could help improve child welfare outcomes [57-59].

Additionally, states with broad maltreatment definitions and stringent reporting policies had higher rates of poverty. While maltreatment outcomes in these states were not significantly different, the overlap between expansive definitions and economic disadvantage raises concerns that reports of maltreatment may, at times, reflect conditions of financial hardship rather than actual harm. This highlights the importance of refining definitions to ensure they do not conflate poverty with neglect [60,61].

Standardizing child fatality definitions across jurisdictions could also enhance the accuracy of reporting and effectiveness of intervention efforts. Establishing clear and consistent criteria for classifying and reporting child fatalities related to maltreatment may improve cross-state comparisons and inform more effective prevention strategies [62,63].

Lastly, the use of NLP and other advanced analytics in policy analysis demonstrated the potential for data-driven approaches in shaping more effective child welfare policies. Leveraging these technologies could provide valuable insights into policy development and evaluation.

Limitations and Future Directions

This study had several limitations. First, while the NLP models quantified policy characteristics, they may struggle with nuanced interpretations, particularly in complex policy language. Second, the study relied on state-reported NCANDS data, which may contain inconsistencies in reporting practices across jurisdictions. Third, the analysis was limited to policies in 2021, as more recent data were not available at the time of the study, potentially missing subsequent policy changes. Fourth, the cross-sectional design limited the ability to assess causal relationships between policy configurations and maltreatment outcomes. Finally, semantic alignment was challenging to evaluate due to the absence of a true gold standard, and reliance on proxy measures may not fully capture model accuracy.

Beyond these limitations, several sources of structural uncertainty are worth noting, as their exact magnitude remains largely unknown. Cross-state variation in reporting practices means observed outcome differences may partly reflect how jurisdictions define and record cases rather than underlying policy effects alone. Lag times between policy enactment and implementation add further uncertainty, since coded 2021 policies may not yet have been fully operationalized, while outcomes that year may partly reflect earlier policy environments. Tracking inconsistencies across state data systems also introduces measurement noise that is difficult to quantify but may affect the stability of cluster-outcome associations.

Future research should refine NLP approaches to enhance policy interpretation, explore longitudinal analyses to assess policy changes over time, and examine the impact of federal initiatives such as the Family First Prevention Services Act on state policy configurations and child welfare outcomes. With larger longitudinal datasets, future studies could also explore integrated machine learning or hierarchical modeling approaches to jointly model co-occurring policy configurations and multiple child welfare outcomes while accounting for temporal relationships and potential confounding factors. Recent advances in causal machine learning methods for policy evaluation [64] could further help address confounding in future work assessing the causal impact of specific policy configurations on maltreatment outcomes. Finally, parametric approaches integrating multiple co-occurring policies to predict population-level outcomes have also been developed in other domains, such as epidemic modeling of nonpharmacological interventions [65], and may offer a useful direction for future longitudinal extensions of this work.

Conclusions

This study identified distinct state-level child maltreatment policy configurations and examined their associations with maltreatment outcomes using an NLP-based analytical framework. This study advanced NLP-based analysis of child maltreatment policies across all US states, offering a scalable, reproducible approach to identify policy patterns. By integrating EFA and clustering, this study moved beyond descriptive comparisons and uncovered latent policy structures and their associations with demographic, economic, and child maltreatment outcomes. These findings offered a data-driven basis for comparing policies, identifying gaps, and informing reform.

Acknowledgments

Generative AI tools were not used in the generation, analysis, or interpretation of this manuscript. Grammarly was used for grammar and language editing assistance.

Funding

The authors declared no financial support was received for this work.

Data Availability

The datasets used in this study are available from the National Data Archive on Child Abuse and Neglect (NDACAN) but are not publicly available due to data use restrictions. Both the National Child Abuse and Neglect Data System (NCANDS) and the State Child Abuse & Neglect Policies (SCAN) dataset require submission of a data use agreement and approval by NDACAN.

XSL•FO
RenderX

Researchers may request access through NDACAN [66] and must comply with all data security and use requirements. Demographic and economic data from the American Community Survey are publicly available through the US Census Bureau.

Authors' Contributions

Conceptualization: ZL, RAE
Data curation: ZL
Formal analysis: ZL, NS, LNM, RAE, NJ
Investigation: ZL, NS, LNM, RAE, NJ
Methodology: ZL, NS, LNM
Project administration: ZL, RAE, NJ
Resources: RAE, NJ
Software: ZL
Supervision: RAE, NJ, NS, LNM
Validation: ZL, NS, LNM, RAE, NJ
Visualization: ZL
Writing—original draft: ZL
Writing—review and editing: ZL, RAE, NS, LNM, NJ

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary figures and tables providing factor analysis diagnostics, cluster determinations, geographic distributions, outcome comparisons, and SCAN policy item evaluations.
[DOCX File , 260 KB - publichealth_v12i1e99643_app1.docx]

References

1. Child maltreatment. Administration for Children and Families. 2025. URL: https://acf.gov/sites/default/files/documents/cb/cm2023.pdf [accessed 2026-09-25]
2. Norman RE, Byambaa M, De R, Butchart A, Scott J, Vos T. The long-term health consequences of child physical abuse, emotional abuse, and neglect: a systematic review and meta-analysis. PLoS Med 2012;9(11):e1001349 [FREE Full text] [doi: 10.1371/journal.pmed.1001349] [Medline: 23209385]
3. Merrick MT, Ford DC, Ports KA, Guinn AS. Prevalence of adverse childhood experiences from the 2011-2014 behavioral risk factor surveillance system in 23 states. JAMA Pediatr 2018;172(11):1038-1044 [FREE Full text] [doi: 10.1001/jamapediatrics.2018.2537] [Medline: 30242348]
4. Danese A, Widom CS. Objective and subjective experiences of child maltreatment and their relationships with psychopathology. Nat Hum Behav 2020;4(8):811-818. [doi: 10.1038/s41562-020-0880-3] [Medline: 32424258]
5. Petruccelli K, Davis J, Berman T. Adverse childhood experiences and associated health outcomes: a systematic review and meta-analysis. Child Abuse Negl 2019;97:104127. [doi: 10.1016/j.chiabu.2019.104127] [Medline: 31454589]
6. Klika JB, Rosenzweig J, Merrick M. Economic burden of known cases of child maltreatment from 2018 in each state. Child Adolesc Soc Work J 2020;37(3):227-234. [doi: 10.1007/s10560-020-00665-5]
7. Day E, Tach L, Mihalec-Adkins B. State child welfare policies and the measurement of child maltreatment in the United States. Child Maltreat 2022;27(3):411-422. [doi: 10.1177/10775595211006464] [Medline: 33832331]
8. Mandated reporting. Child Welfare Information Gateway. URL: https://www.childwelfare.gov/topics/safety-and-risk/mandated-reporting/?top=78 [accessed 2026-04-16]
9. Johnson-Motoyama M, Ginther DK, Phillips R, Beer OWJ, Merkel-Holguin L, Fluke J. Differential response and the reduction of child maltreatment and foster care services utilization in the U.S. from 2004 to 2017. Child Maltreat 2023;28(1):152-162. [doi: 10.1177/10775595211065761] [Medline: 35062827]
10. The Child Abuse Prevention and Treatment Act (CAPTA): background, programs, and funding. EveryCSReport. 2009. URL: https://www.everycrsreport.com/reports/R40899.html [accessed 2026-04-16]
11. LaBrenz C, Kim Y, Baiden P, Shipe SL, Littleton T, Choi M, et al. State child maltreatment policies and disparities in substantiation: a study of state-administered child welfare systems in the U.S. Child Maltreat 2023;28(4):700-712. [doi: 10.1177/10775595221143136] [Medline: 36458462]
12. Puls HT, Hall M, Anderst JD, Gurley T, Perrin J, Chung PJ. State spending on public benefit programs and child maltreatment. Pediatrics 2021;148(5):5. [doi: 10.1542/peds.2021-050685] [Medline: 34663680]
13. Gunes E, Florczak CK. Multiclass classification of policy documents with large language models. arXiv Preprint posted online on October 12, 2023 [FREE Full text]

XSL•FO
RenderX

14. Osnabrügge M, Vannoni M. Quality of legislation and compliance: a natural language processing approach. PSRM 2024;13(3):736-744. [doi: 10.1017/psrm.2024.21]
15. Swarnakar P, Modi A. NLP for climate policy: creating a knowledge platform for holistic and effective climate action. arXiv Preprint posted online on May 12, 2021. [doi: 10.48550/arXiv.2105.05621]
16. Planas J, Firebanks-Quevedo D, Naydenova G, Sharma R, Taylor C, Buckingham K. Beyond modeling: NLP pipeline for efficient environmental policy analysis. arXiv Preprint posted online on January 8, 2022 [FREE Full text]
17. Nay J. Natural language processing and machine learning for law and policy texts. SSRN. Rochester, NY; 2018. URL: https://papers.ssrn.com/abstract=3438276 [accessed 2024-08-22]
18. Rosholm M, Bodilsen ST, Michel B, Nielsen AS. Predictive risk modeling for child maltreatment detection and enhanced decision-making: evidence from Danish administrative data. PLoS One 2024;19(7):e0305974 [FREE Full text] [doi: 10.1371/journal.pone.0305974] [Medline: 38985689]
19. Ahn E, An R, Jonson-Reid M, Palmer L. Leveraging machine learning for effective child maltreatment prevention: a case study of home visiting service assessments. Child Abuse Negl 2024;151:106706. [doi: 10.1016/j.chiabu.2024.106706] [Medline: 38428267]
20. Negriff S, Lynch FL, Cronkite DJ, Pardee RE, Penfold RB. Using natural language processing to identify child maltreatment in health systems. Child Abuse Negl 2023;138:106090 [FREE Full text] [doi: 10.1016/j.chiabu.2023.106090] [Medline: 36758373]
21. Annapragada AV, Donaruma-Kwoh MM, Annapragada AV, Starosolski ZA. A natural language processing and deep learning approach to identify child abuse from pediatric electronic medical records. PLoS One 2021;16(2):e0247404 [FREE Full text] [doi: 10.1371/journal.pone.0247404] [Medline: 33635890]
22. Victor BG, Perron BE, Sokol RL, Fedina L, Ryan JP. Automated identification of domestic violence in written child welfare records: leveraging text mining and machine learning to enhance social work research and evaluation. J Soc Social Work Res 2021;12(4):631-655. [doi: 10.1086/712734]
23. Niu F, Dhamayanti M, Setiawati EP, Arisanti N. Artificial intelligence for early detection of child maltreatment in healthcare: a narrative review integrating technical, ethical, clinical, and governance perspectives. Children and Youth Services Review 2026;188:109106. [doi: 10.1016/j.childyouth.2026.109106]
24. Fox-Sowell S. Pennsylvania county taps natural language processing to help child welfare caseworkers. StateScoop. 2023. URL: https://statescoop.com/nlp-ai-washington-county-pennsylvania-child-welfare-caseworkers/ [accessed 2025-02-22]
25. Saxena D, Moon ESY, Chaurasia A, Guan Y, Guha S. Rethinking 'Risk' in algorithmic systems through a computational narrative analysis of casenotes in child-welfare. : Association for Computing Machinery; 2023 Presented at: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems; 2026 July 23; New York, NY, USA p. 1-19 URL: https://dl.acm.org/doi/10.1145/3544548.3581308
26. Weigensberg EC, Islam N, Knab J, Grider MA, Page J, Larson A. State Child Abuse and Neglect (SCAN) Policies Database 2019-2021. National Data Archive on Child Abuse and Neglect. 2022. URL: https://www.ndacan.acf.hhs.gov/datasets/dataset-details.cfm?ID=268 [accessed 2024-02-20]
27. National Child Abuse and Neglect Data System (NCANDS) Agency File. National Data Archive on Child Abuse and Neglect. 2024. URL: https://www.ndacan.acf.hhs.gov/datasets/datasets-list-ncands-state-agency-file.cfm [accessed 2024-08-08]
28. National Child Abuse and Neglect Data System (NCANDS) Child File. National Data Archive on Child Abuse and Neglect. 2024. URL: https://www.ndacan.acf.hhs.gov/datasets/datasets-list-ncands-child-file.cfm [accessed 2024-08-08]
29. American Community Survey. Census.gov. 2024. URL: https://www.census.gov/programs-surveys/acs [accessed 2025-02-25]
30. Yin W, Hay J, Roth D. Benchmarking zero-shot text classification: datasets, evaluation and entailment approach. arXiv Preprint posted online on August 31, 2019 [FREE Full text]
31. Lewis M, Liu Y, Goyal N, Ghazvininejad M, Mohamed A, Levy O. BART: denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. arXiv Preprint posted online on October 29, 2019 [FREE Full text]
32. Devlin J, Chang M, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. arXiv Preprint posted online on October 11, 2018 [FREE Full text]
33. Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D. RoBERTa: a robustly optimized BERT pretraining approach. arXiv Preprint posted online on July 26, 2019 [FREE Full text]
34. He P, Liu X, Gao J, Chen W. DeBERTa: decoding-enhanced BERT with disentangled attention. arXiv Preprint posted online on June 5, 2020 [FREE Full text]
35. Microsoft 365 Copilot. Microsoft. 2024. URL: https://copilot.microsoft.com [accessed 2025-02-28]
36. Meta Llama 3.1. Meta AI. 2024. URL: https://ai.meta.com/blog/meta-llama-3-1/ [accessed 2024-08-16]
37. Wang B, Wang A, Chen F, Wang Y, Kuo CCJ. Evaluating word embedding models: methods and experimental results. SIP 2019;8(1). [doi: 10.1017/atsip.2019.12]
38. Qiu Y, Li H, Li S, Jiang Y, Hu R, Yang L. Revisiting correlations between intrinsic and extrinsic evaluations of word embeddings. In: Sun M, Liu T, Wang X, Liu Z, Liu Y, editors. Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data. Cham: Springer International Publishing; 2018:209-921.

XSL•FO
RenderX

39. Gorsuch RL. Factor Analysis. 2nd ed. New York: Psychology Press; 1983:448.
40. Hartigan JA, Wong MA. Algorithm AS 136: A k-means clustering algorithm. J R Stat Soc Series C Appl Stat 1979;28(1):100-108. [doi: 10.2307/2346830]
41. Sakamoto Y, Ishiguro M, Kitagawa G. Akaike Information Criterion Statistics. Netherland: Springer; 1986.
42. Calinski T, Harabasz J. A dendrite method for cluster analysis. Commun Stat Simul Comput 1974;3(1):1-27. [doi: 10.1080/03610917408548446]
43. Kriegel HP, Schubert E, Zimek A. The (black) art of runtime evaluation: are we comparing algorithms or implementations? Knowl Inf Syst 2016;52(2):341-378. [doi: 10.1007/s10115-016-1004-2]
44. Cattell RB. The scree test for the number of factors. Multivariate Behav Res 1966;1(2):245-276. [doi: 10.1207/s15327906mbr0102_10] [Medline: 26828106]
45. Kaiser HF. A second generation little jiffy. Psychometrika 2025;35(4):401-415. [doi: 10.1007/BF02291817]
46. Grice JW. Computing and evaluating factor scores. Psychol Methods 2001;6(4):430-450. [Medline: 11778682]
47. McDonald RP. Test Theory: A Unified Treatment. New Jersey, USA: Lawrence Erlbaum Associates Publishers; 1999.
48. Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math 1987;20:53-65. [doi: 10.1016/0377-0427(87)90125-7]
49. Fisher RA. Statistical methods for research workers. In: Kotz S, Johnson NL, editors. Breakthroughs in Statistics: Methodology and Distribution. New York, NY: Springer; 1992:66-70.
50. Pearson KX. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. Lond Edinb Dubl Philos Mag 2009;50(302):157-175. [doi: 10.1080/14786440009463897]
51. Mudge JF, Baker LF, Edge CB, Houlahan JE. Setting an optimal α that minimizes errors in null hypothesis significance tests. PLoS One 2012;7(2):e32734 [FREE Full text] [doi: 10.1371/journal.pone.0032734] [Medline: 22389720]
52. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. New York: Routledge; 1988:567.
53. Hubert L, Arabie P. Comparing partitions. J Classif 1985;2(1):193-218. [doi: 10.1007/bf01908075]
54. Ho GWK, Gross DA, Bettencourt A. Universal mandatory reporting policies and the odds of identifying child physical abuse. Am J Public Health 2017;107(5):709-716. [doi: 10.2105/AJPH.2017.303667] [Medline: 28323475]
55. Kim H, Drake B. Has the relationship between community poverty and child maltreatment report rates become stronger or weaker over time? Child Abuse Negl 2023;143:106333 [FREE Full text] [doi: 10.1016/j.chiabu.2023.106333] [Medline: 37379728]
56. Dale M. Addressing the underlying issue of poverty in child-neglect cases. American Bar Association. 2014. URL: https://www.americanbar.org/groups/litigation/resources/newsletters/childrens-rights/addressing-underlying-issue-poverty-child-neglect-cases/ [accessed 2025-03-19]
57. Rochford HI, Zeiger KD, Peek-Asa C. State-level education policies: opportunities for secondary prevention of child maltreatment. Child Abuse Negl 2023;136:106018. [doi: 10.1016/j.chiabu.2022.106018] [Medline: 36630852]
58. Shusterman GR, Hollinshead D, Fluke JD, Yuan YT. Alternative Responses to Child Maltreatment: Findings From NCANDS. USA: U.S. Department of Health and Human Services; Office of the Assistant Secretary for Planning and Evaluation; 2005.
59. Font S, Maguire-Jack K. The organizational context of substantiation in child protective services cases. J Interpers Violence 2021;36(15-16):7414-7435 [FREE Full text] [doi: 10.1177/0886260519834996] [Medline: 30862238]
60. Saar-Heiman Y. Understanding the relationships among poverty, child maltreatment, and child protection involvement: perspectives of service users and practitioners. J Soc Social Work Res 2022;13(1):117-141. [doi: 10.1086/713999]
61. Lefebvre R, Fallon B, Van Wert M, Filippelli J. Examining the relationship between economic hardship and child maltreatment using data from the Ontario incidence study of reported child abuse and neglect-2013 (OIS-2013). Behav Sci (Basel) 2017;7(1):6 [FREE Full text] [doi: 10.3390/bs7010006] [Medline: 28208690]
62. Covington T, Collier A. Child maltreatment fatality reviews: learning together to improve systems that protect children and prevent maltreatment. National Center for Fatality Review and Prevention 2018 [FREE Full text]
63. Campbell KA, Wood JN, Lindberg DM, Berger RP. A standardized definition of near-fatal child maltreatment: results of a multidisciplinary Delphi process. Child Abuse Negl 2021;112:104893 [FREE Full text] [doi: 10.1016/j.chiabu.2020.104893] [Medline: 33373847]
64. Rehill P, Biddle N. Transparency challenges in policy evaluation with causal machine learning: improving usability and accountability. Data Policy 2024;6:e43. [doi: 10.1017/dap.2024.35]
65. Tutsoy O, Polat A, Colak S, Balikci K. Development of a multi-dimensional parametric model with non-pharmacological policies for predicting the COVID-19 pandemic casualties. IEEE Access 2020;8:225272-225283. [doi: 10.1109/access.2020.3044929]
66. National Data Archive on Child Abuse and Neglect. URL: https://www.ndacan.acf.hhs.gov/ [accessed 2026-08-27]

Abbreviations

ARI: adjusted Rand index
BART: bidirectional and auto-regressive transformer

XSL•FO
RenderX

BERT: bidirectional encoder representations from transformers
BIC: Bayesian information criterion
BSS/WSS: between-cluster sum of squares to within-cluster sum of squares
CAPTA: Child Abuse Prevention and Treatment Act
CH: Calinski-Harabasz score
DeBERTa: decoding-enhanced BERT with disentangled attention
EFA: exploratory factor analysis
LLM: large language model
MRSF: multiple R-squared of scores with factors
NCANDS: National Child Abuse and Neglect Data System
NDACAN: National Data Archive on Child Abuse and Neglect
NLP: natural language processing
RoBERTa: robustly optimized BERT approach
SCAN: State Child Abuse & Neglect Policies Database
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Mavragani, T Sanchez; submitted 27.Apr.2026; peer-reviewed by EL Thibodeau, M Lloyd Sieger, O Tutsoy; comments to author 20.Jul.2026; revised version received 03.Aug.2026; accepted 08.Aug.2026; published 16.Sep.2026.

Please cite as:
Luo Z, Epstein RA, Sambamoorthi N, Muhammad LN, Jordan N
Using Natural Language Processing to Examine State Child Maltreatment Policies and Associations With Outcomes: Multistate Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e99643
URL: https://publichealth.jmir.org/2026/1/e99643
doi:10.2196/99643
PMID:

©Zhidi Luo, Richard A Epstein, Nethra Sambamoorthi, Lutfiyya N Muhammad, Neil Jordan. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 16.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effectiveness of Online Training on Cardiopulmonary Resuscitation and Automated External Defibrillator Use in Simulations With Rural Populations: Quasi-Experimental Community Intervention Study

Elena Taverna-Llauradó[1,2,3], MD; Sara Martinez-Torres[2,3,4], PhD; Alba Roca-Biosca[5], PhD; Meritxell Pallejà-Millán[2,3,6], MSc; Francisco Martín-Luján[2,6], PhD; Cristina Rey-Reñones[2,3,6], PhD

[1]Primary Care Unit Camp de Tarragona, Institut Català de la Salut, Reus, Spain

[2]Primary Healthcare Research Support Unit Camp de Tarragona, Institut Universitari d'Investigació en Atenció Primària Jordi Gol, Reus, Spain

[3]TIC-AP Study Group, Primary Healthcare Research Support Unit Camp de Tarragona, Institut Universitari d'Investigació en Atenció Primària Jordi Gol, Reus, Catalunya, Spain

[4]Universitat Oberta de Catalunya, Barcelona, Spain

[5]Nursing Department, Universitat Rovira i Virgili, Tarragona, Spain

[6]School of Medicine and Health Sciences, Universitat Rovira i Virgili, C/St Llorenç 21, Reus, Spain

Corresponding Author:
Cristina Rey-Reñones, PhD
Primary Healthcare Research Support Unit Camp de Tarragona, Institut Universitari d'Investigació en Atenció Primària Jordi Gol, Reus, Spain

Abstract

Background: Sudden death due to cardiorespiratory arrest has a high mortality rate and often occurs outside hospital settings. Prompt initiation of cardiopulmonary resuscitation (CPR) by bystanders, along with the use of an automated external defibrillator (AED), has been shown to double survival rates. Given the challenges of ensuring timely emergency response in rural areas, implementing basic CPR training programs can help improve survival outcomes.

Objective: This study aimed to evaluate the effectiveness of online CPR-AED training delivered to participants from a rural area of Camp de Tarragona.

Methods: This quasi-experimental study consisted of 2 phases: phase 1, evaluation of the effectiveness of online training on CPR-AED knowledge (with pretest and posttest assessments), and phase 2, evaluation of the effectiveness of online training on CPR-AED maneuvers through simulation at 1 month and 6 months after online training. The sample of the study comprised residents in a rural area of Tarragona, Spain. A descriptive statistical analysis of the study population was conducted. For quantitative data with a nonnormal distribution, the median and IQR were presented. Categorical data were described as frequencies and percentages. A bivariate analysis was performed to compare the pretraining and posttraining quantitative variables using the Student 2-tailed *t* test.

Results: In total, 55 participants were included in the study. Of these, 74.5% (n=55) were women, the mean age was 41.5 (SD 9.1) years, and 94.5% (n=55) were employed. Overall, 52 participants completed the online training. The median time required to complete the course was 261.5 (IQR 935; range 125-327) minutes. In total, 51 participants took part in the first practical simulation, which was conducted 1 month after the theoretical training. The mean score obtained in this first simulation was 7.5 out of 10. Six months after the theoretical training, 46 participants completed a second simulation.

Conclusions: Online CPR-AED training is effective in improving CPR-AED knowledge and skills in a rural population in the short and medium term.

Trial Registration: ClinicalTrials.gov NCT05747495; https://clinicaltrials.gov/study/NCT05747495

(JMIR Public Health Surveill 2026;12:e80359) doi:10.2196/80359

KEYWORDS

out-of-hospital cardiac arrest; cardiopulmonary resuscitation; automated external defibrillator; emergency medical services; primary care; rural medicine

XSL•FO
RenderX

Introduction

Out-of-hospital cardiac arrest (OHCA) remains a major cause of mortality worldwide, with survival critically dependent on early recognition, prompt initiation of cardiopulmonary resuscitation (CPR), and rapid use of an automated external defibrillator (AED). Despite sustained public health initiatives, recent epidemiological data indicate that bystander CPR and AED deployment remain suboptimal in many regions, particularly in rural or geographically dispersed populations [1]. In Europe, OHCA represents a substantial public health burden, with an estimated annual incidence ranging from 53 cases per 100,000 inhabitants in Austria to 166 per 100,000 in Bosnia and Herzegovina [2].

Although advances in prehospital and emergency care have improved outcomes, overall survival remains limited. The EuReCa TWO registry, encompassing resuscitation attempts in 28 European countries, reported survival to hospital discharge in only 8% of patients [2].

In Spain, national registry data from 2022 documented an OHCA incidence of 24.2 cases per 100,000 inhabitants, with marked regional variation, with Catalonia reporting one of the highest incidence rates at 35.2 cases per 100,000. Overall survival to hospital discharge was 11.5%. These outcomes remain strongly associated with the timeliness and quality of bystander intervention, particularly CPR provision and AED accessibility. The “chain of survival”—early recognition, timely bystander CPR, rapid defibrillation, advanced life support, and postresuscitation care—continues to be central to improving OHCA outcomes [3].

Each minute of delay in chest compressions reduces survival by approximately 10%, whereas defibrillation within 3 to 5 minutes for shockable rhythms may increase survival to as much as 30% [4]. However, bystander CPR rates vary widely, from 26% in some regions to approximately 69% in communities with comprehensive public training programs.

Rural populations face distinct barriers to emergency response education, including limited access to certified instructors, long travel distances to training sites, and fewer opportunities for refresher courses [5]. These structural constraints likely contribute to lower CPR training rates and reduced AED awareness compared with urban communities, potentially widening disparities in OHCA outcomes [1]. In such settings, early intervention often depends on lay bystanders before the arrival of emergency medical services (EMS) [6-8].

Hands-only CPR has been demonstrated to be as effective as conventional CPR for most adult cardiac arrests, lowering barriers for lay responders and facilitating dissemination through community-based programs [6]. Nonetheless, substantial inequities persist between urban and rural areas. In rural contexts, EMS response times frequently exceed 10 to 15 minutes, AED availability is limited, and community CPR training coverage is lower. These challenges are often compounded by socioeconomic disadvantage, with underserved populations less likely to receive bystander CPR or early defibrillation [6].

Disparities in early resuscitation are particularly concerning in pediatric OHCA, where rapid intervention is essential. Globally, the annual incidence among children younger than 12 years ranges from 8 to 20 cases per 100,000 [9]. However, most schoolteachers report insufficient CPR training and limited confidence in performing or teaching resuscitation techniques. Although hands-only CPR is generally preferred by educators, overall preparedness remains inadequate [10].

Therefore, training schoolteachers in CPR may have a multiplier effect, enabling them to respond effectively in school environments while simultaneously transmitting life-saving skills to students and fostering a broader culture of preparedness within the community [11]. This approach is especially relevant in rural regions, where access to EMS is constrained and scalable digital solutions—such as online CPR courses—may represent practical and cost-effective alternatives.

Although traditional face-to-face CPR-AED training has proven effective, the COVID-19 pandemic accelerated the adoption of digital health education modalities, including blended and fully online programs. Emerging evidence suggests that online or hybrid CPR instruction combined with simulation-based practice can improve short-term knowledge retention and psychomotor skills, although findings remain heterogeneous and dependent on instructional design, learner characteristics, and assessment strategies [12].

Simulation-based assessment is increasingly used to evaluate CPR competence in standardized environments, enabling objective measurement of procedural accuracy, timing, and checklist-based performance [12,13]. However, improvements observed in simulated contexts should be interpreted cautiously, as such performance may not fully translate to real-world behavior during actual emergencies.

Despite growing interest in remote and community-based CPR education, evidence specifically examining online CPR-AED training for rural lay populations remains limited [1]. Most studies have focused on urban cohorts, health science students, or workplace settings, leaving an important knowledge gap regarding the feasibility and potential impact of digital training strategies in geographically isolated communities [13].

To address this gap, this study evaluates changes in theoretical knowledge and simulated CPR-AED performance among rural community participants before and after completion of an online training program supplemented by in-person simulation practice. Given the quasi-experimental pre-post design and the absence of a concurrent control group, the analysis focuses on observed improvements rather than causal inference.

Methods

Study Design, Setting, and Recruitment

This study was designed as a quasi-experimental community intervention without a control group. It was structured in two sequential phases:

- Phase 1: evaluation of the effectiveness of online training on CPR-AED theoretical knowledge

XSL•FO
RenderX

- Phase 2: evaluation of the effectiveness of online training on CPR-AED practical skills through in-person, simulation-based assessment

The study was carried out in 2 primary health care centers located in the municipalities of Borges del Camp and Cornudella de Montsant (province of Tarragona, Catalonia, Spain). The reference population consisted of residents from 3 rural municipalities in Catalonia.

A systematic community-based recruitment strategy was implemented. First, the research team contacted local authorities and institutional representatives to present the objectives and procedures of the study and to request collaboration. The study was subsequently disseminated through municipal associations, schools, and secondary education institutes, as well as through community notice boards and local communication channels. Informational meetings were organized when it was feasible to explain the study and resolve questions from potential participants.

Interested individuals contacted the research team or the primary health care centers directly. After receiving detailed information about the study, those willing to participate provided informed consent and were consecutively enrolled until the target sample size was reached. Recruitment procedures were identical across the participating municipalities to ensure consistency.

Participant Selection

Participants were selected based on specific eligibility criteria.

The inclusion criteria were as follows:

- Residence and/or employment in the Priorat or Baix Camp regions
- Age ≥18 years
- Access to the internet and compatible devices to complete the online training

The exclusion criteria were as follows:

- Presence of a language barrier (limited or no comprehension of Catalan)
- Presence of any disability preventing participation in the course and/or performance of CPR and AED procedures
- Anticipated inability to attend simulation sessions for evaluation

Prior CPR-AED training was not considered as an exclusion criterion.

Description of Intervention

The intervention consisted of an online training program based on the European Resuscitation Council (ERC) guidelines, as described in detail in the study protocol [10,14].

The primary objective of the training was to provide both theoretical and practical knowledge regarding emergency situations and the use of an AED.

Participants interested in enrolling were required to register on a Moodle platform, where all training materials were hosted. The course included interactive modules, audiovisual resources, and assessments to ensure the acquisition of both theoretical and practical competencies.

Evaluation of Intervention

Phase 1: Assessment of Knowledge Acquisition Following Online CPR-AED Training

To evaluate the effectiveness of the online training program, participants completed 2 standardized knowledge assessments: a pretraining test to establish baseline knowledge and a posttraining test to measure knowledge gained after course completion.

The assessment instrument was developed in accordance with the guidelines of the Canadian Council on Resuscitation and the ERC and consisted of 10 multiple-choice questions. Participants were required to achieve a perfect score (10/10) to pass. An unlimited number of attempts was permitted, and results were categorized dichotomously as pass or fail. Only participants who successfully completed the posttraining assessment were eligible to participate in the subsequent phase of the study.

Phase 2: Assessment of Practical Skill Retention Over Time

Participants who consented to continue in the study were invited to perform a simulated CPR-AED scenario at 2 follow-up intervals: 1 month (short term) and 6 months (medium term) after completing the online training.

During these sessions, participants were exposed to a standardized cardiac arrest scenario requiring the application of CPR and AED skills. Each simulation was independently assessed by 2 trained observers using a structured checklist to assess performance and retention of critical skills. The items evaluated in the checklist were (1) safe approach, (2) check response, (3) open the airway, (4) check for spontaneous breathing, (5) request external assistance, (6) ventilation-compressions ratio, (7) quality of chest compressions, (8) patch placement, (9) AED use, and (10) immediate compressions. At the end of the simulation, the participants were informed of the mistakes they made and, when applicable, how they should have proceeded.

Outcome Measures

Effectiveness of Online Training on CPR-AED Knowledge

The primary outcome measure of effectiveness was the difference in scores between the pretraining assessment and the first attempt of the posttraining assessment at the end of the course. Secondary variables included the number of attempts required to achieve a perfect score and the score obtained on the first posttraining attempt.

Effectiveness of Online Training on CPR-AED Performance in a Simulated Environment

In phase 2, the primary outcome was the performance score obtained during the in-person simulation. Success in the simulation was defined as the correct execution of at least 8 of

XSL•FO
RenderX

10 key actions listed on a standardized evaluation checklist [15,16].

To evaluate the effectiveness of the training in the simulation, the necessary sample size was calculated using the statistical software GRANMO (version 7.11) using the option of averages observed with respect to a reference sample. Accepting an α risk of .05 and a β risk of .2 for a bilateral contrast, 52 participants were needed to detect a difference of ≥0.5 points in the face-to-face evaluation of the online training, assuming a common SD of 1.07. The number of participants was increased by 30% to compensate for potential losses to follow-up [14].

Statistical Analyses

A descriptive analysis of the study population was conducted. Quantitative variables following a normal distribution are expressed as mean (SD), whereas nonnormally distributed variables are presented as median (IQR). Normality was assessed using the Kolmogorov-Smirnov test. Categorical variables are reported as frequencies and percentages.

Bivariate analyses were performed to compare pretraining and posttraining quantitative variables using the Student 2-tailed *t* test.

Ethical Considerations

The study was conducted in accordance with the principles of the Declaration of Helsinki and the guidelines of good clinical practice. The protocol was approved by the clinical research ethics committee of the Primary Care Research Institute IDIAP Jordi Gol (23/081-P) on June 15, 2023.

Data confidentiality was ensured in compliance with Spanish legislation on personal data protection (Ley Orgánica 03/2018 of December 5, on the Protection of Personal Data and Guarantee of Digital Rights) [17]. Participants received an information sheet explaining that their participation was voluntary, anonymous, and confidential and that they could withdraw from the study at any time prior to data verification. All participants were given the opportunity to ask questions and receive clarification from the research team before enrollment. Verbal and written informed consent for participation and audio recording were obtained from all participants.

Each participant was assigned a unique identification code upon inclusion. Study data were recorded using a purpose-designed electronic questionnaire and stored on a secure digital platform accessible only through the corporate intranet of the Catalan Health Institute in Tarragona. Access was password-protected and restricted to authorized investigators responsible for data entry. The dataset is not publicly available; however, access may be granted upon reasonable request to the principal investigator and with approval from the IDIAP Jordi Gol Research Ethics Committee.

Participants did not receive any financial compensation for their involvement in the study. Study findings will be disseminated to participating centers and research teams and shared with the scientific community through peer-reviewed publications and conference presentations. Dissemination in nonspecialist settings is also planned.

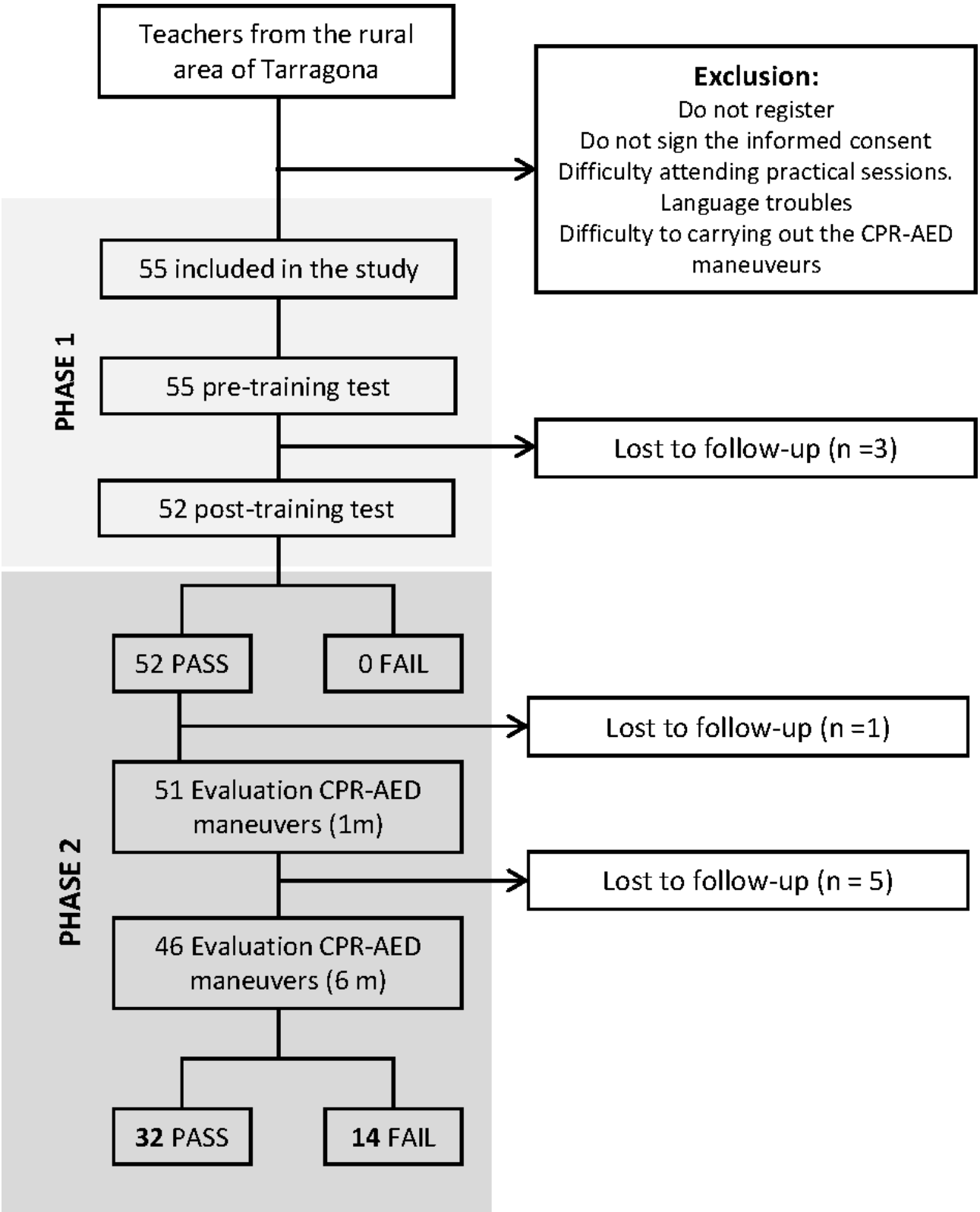
Results

Sociodemographic Characteristics of the Population

In total, 55 participants were included in the study (Figure 1). Of these, 74.5% (n=55) were women, 94.5% (n=55) belonged to the active working population, and 60% (n=33) held a university degree. Slightly more than half (n=32, 58.2%) of the participants had received previous training related to the topic under study. Overall, 7.3% (n=4) had previously witnessed a cardiac arrest; however, none reported knowing how to respond appropriately. Most participants resided in rural areas of the Baix Camp region. The mean age of the participants was 41.5 years. A summary of the main characteristics of the study population is presented in Table 1.

XSL•FO
RenderX

Figure 1. Flowchart of participant recruitment and inclusion in a quasi-experimental community intervention study evaluating the effectiveness of an online CPR-AED training program among rural community residents in Catalonia, Spain. The diagram shows the stages of eligibility assessment, enrollment, allocation to phase 1 (online theoretical training), participation in phase 2 (in-person simulation-based skills assessment), and final analysis. CPR-AED: cardiopulmonary resuscitation–automated external defibrillator.



XSL•FO
RenderX

Table . Main sociodemographic characteristics of the study population (N=55).

Variables	Participants, n (%)
Sex	
Female	41 (74.5)
Male	13 (23.6)
Nonbinary	1 (1.8)
Employment status	
Student	2 (3.6)
Unemployed	1 (1.8)
Employed	52 (94.5)
Educational level	
University diploma	33 (60)
Bachelor's degree or higher	17 (30.9)
Nonqualified	1 (1.8)
Skilled nonmanual	3 (5.5)
Other	1 (1.8)
Participants with previous training	32 (58.2)
Participants who had witnessed an arrest	4 (7.3)
Participants who did not know how to respond to the arrest	4 (100)
Place of residence: rural	39 (70.9)
Region	
Baix Camp	41 (74.5)
Priorat	6 (10.9)
Alt Camp	2 (3.6)
Tarragonès	5 (9.1)
Ribera d'Ebre	1 (1.8)

Phase 1: Evaluation of the Effectiveness of Online Training on CPR-AED Theoretical Knowledge

Of the 55 individuals enrolled in the study, 52 (94.5%) participants completed the online training. The median time required to complete the course was 201.5 minutes, with an IQR of 935 minutes.

The median number of pretest attempts was 1 (IQR 0). The mean pretest score was 6.72 (SD 2.06) points out of 10. The median number of posttest attempts was 2 (IQR 1), with a mean posttest score of 7.77 (SD 1.10) points out of 10.

The mean difference between pretraining and posttraining scores was 1.05 (SD 2.06) points, which was statistically significant (P=.01). These data are presented in Table 2.

Table . Main variables analyzed from the online training.

Variables	Values
Time taken to complete the course (minutes), median (IQR)	201.5 (935)
Number of pretest attempts, median (IQR)	1 (0)
Pretest score (out of 10), mean (SD)	6.72 (2.06)
Number of posttest attempts, median (IQR)	2 (1)
Posttest score (out of 10), mean (SD)	7.77 (1.10)

XSL•FO
RenderX

Phase 2: Evaluation of the Effectiveness of Online Training on CPR-AED Skills Through Simulation-Based Assessment

Of the 52 participants who completed the online training, 51 (98.1%) took part in the first practical simulation, which was carried out 1 month after the theoretical training. The average score obtained in this first simulation was 7.5 out of 10.

Six months after the theoretical training, 46 participants completed a second simulation. Among them, 42 achieved a passing score—defined as ≥8 points—representing 91.3% of those who participated in this second practical test. The average score was 9.0 out of 10. In contrast, 4 (8.7%) participants did not reach the minimum score to be considered competent, scoring ≤7 points.

A statistically significant improvement was observed between the first and final simulation scores, with a mean difference of 1.31 points (SD 2.05; $P<.001$), indicating a positive evolution in performance over time. Additionally, a paired-samples analysis comparing the scores from both simulations showed a significant increase in the second attempt, with an average difference of 0.89 points ($P=.01$), further supporting the sustained effect of the training over 6 months. The main results obtained in the simulation phases are shown in Table 3.

Table . Main results obtained in the simulation phases.

Simulation phases	First simulation (n=51), n (%)	Second simulation (n=46), n (%)
Safe approach	36 (70.6)	32 (69.6)
Check response	41 (80.4)	37 (80.4)
Open the airway	19 (37.2)	17 (37)
Check for spontaneous breathing	42 (82.4)	38 (82.6)
Request external assistance	31 (60.8)	33 (71.7)
Compressions-breathing relation	41 (80.4)	37 (80.4)
Quality of chest compressions	31 (60.8)	30 (65.2)
Patch placement	41 (80.4)	37 (80.4)
Automated external defibrillator use	40 (78.4)	36 (78.3)
Immediate compressions	44 (86.3)	39 (84.8)

An analysis of different variables revealed that, in the first simulation, there were no significant differences in scores based on participants' gender, place of residence (rural or urban), prior experience witnessing a cardiac arrest, or having received previous CPR or AED training. These factors did not appear to influence initial practical performance. However, in the second simulation, a significant association was found between having previous training and achieving a passing score ($P=.004$), suggesting that prior exposure to resuscitation concepts may enhance long-term skill retention. No significant differences were observed in the second simulation based on sex, place of residence, or previous experience witnessing a cardiac arrest. The differences in simulation scores according to participant characteristics are presented in Table 4.

Table . Differences between the punctuation of the simulation according to different variables.

Variables	First simulation, n (%)			Second simulation, n (%)		
	Pass	Fail	*P* value	Pass	Fail	*P* value
Gender			.58			.21
Woman	17 (51.5)	16 (48.5)		24 (75)	8 (25)	
Man	6 (50)	6 (50)		8 (61.5)	5 (38.5)	
Nonbinary	0 (0)	1 (100)		0 (0)	1 (100)	
Participants with previous training	14 (58.3)	10 (41.7)	.24	22 (8.5)	5 (18.5)	.04
Participants who had witnessed an arrest	1 (25)	3 (75)	.29	3 (75)	1 (25)	.81
Place of residence: rural	18 (52.9)	16 (47.1)	.47	23 (67.6)	11 (32.4)	.63

XSL•FO
RenderX

Discussion

This community-based study evaluated changes in theoretical knowledge and simulated CPR-AED performance following completion of an online training program supplemented with in-person simulation among lay participants—primarily primary and secondary schoolteachers—from rural areas in southern Catalonia.

Consistent with the study objectives, participants demonstrated significant improvements in posttraining knowledge scores and simulation performance compared with baseline. These findings suggest that a digitally delivered CPR-AED curriculum, when paired with structured hands-on practice, may enhance preparedness among non–health care professionals in geographically dispersed settings.

Out of the 55 individuals initially recruited, 52 completed the posttraining evaluation. Median knowledge scores increased from 6.72 (IQR 6-8.57) to 7.77 (IQR 7-8.67 out of 10), and participants required fewer attempts to complete the theoretical assessment, suggesting improved mastery of key resuscitation concepts. In addition, the improvement observed between the 2 simulated scenarios indicates retention of procedural skills, which is particularly relevant given that simulation assessments focused on core elements of basic life support and AED use. Although participants could repeat the test, the observed reductions in errors and increased procedural accuracy may reflect learning through feedback and iterative practice.

Unlike many community courses that focus solely on hands-only CPR, this program provided comprehensive basic life support training, including AED deployment. The online format offered logistical advantages and the potential to reach a broad audience, particularly in rural areas where access to face-to-face instruction is often limited.

All simulations were conducted individually at the participants' workplaces using manikins and standardized cardiac arrest scenarios under supervision by certified instructors and observers, following the Catalan Council of Resuscitation guidelines derived from the ERC. The observed gains are consistent with prior studies demonstrating the feasibility and educational value of virtual or blended CPR instruction for lay populations [18-20].

However, the characteristics of the study sample warrant cautious interpretation. Approximately 70% of participants were women, 60% held a university degree, and approximately 30% resided in urban areas while working in rural schools. These demographic features may limit generalizability to the broader rural population and suggest that participants may have been more motivated or educationally advantaged than the general community.

Several limitations should be considered. First, the quasi-experimental pre-post design without a control group precludes causal inference regarding the effectiveness of the intervention. Second, although simulation-based assessments provide standardized and objective measures of performance, they may not fully reflect behavior during real-life cardiac arrest events.

Technical barriers also affected implementation, as internet connectivity was inconsistent in some rural locations, potentially limiting access to the online materials. Additionally, some substitute teachers who participated in the first simulation were unable to complete follow-up assessments because they were no longer employed at the same schools, resulting in attrition between testing phases.

Finally, the demographic composition of the sample and partial inclusion of participants residing in urban settings restrict external validity and should be taken into account when extrapolating these findings to other rural contexts.

This study suggests that online CPR-AED training supplemented by workplace-based simulation may represent a scalable approach to improving resuscitation preparedness among laypeople in rural communities. Beyond immediate educational gains, such programs have the potential to strengthen community emergency response capacity by establishing local networks of trained individuals, particularly in school environments where teachers can serve both as first responders and multipliers of knowledge.

Participants' recommendations to integrate CPR education into primary and secondary school curricula, provide regular refresher training, and improve community awareness of public-access AEDs align with international calls for school-based resuscitation education as a cornerstone of population-wide preparedness. Therefore, future public health strategies should consider incorporating accessible digital training platforms alongside broader AED deployment to reduce disparities in OHCA outcomes, particularly in underserved rural areas.

Further research using controlled designs, longer follow-up periods, and more representative rural samples is needed to determine whether observed improvements in knowledge and simulated performance translate into sustained behavioral change and improved real-world clinical outcomes.

Acknowledgments

The authors thank the Primary Health Care Research Institute IDIAP Jordi Gol for providing assistance.

The research team would like to thank the study participants as well as the educational centers that facilitated its execution.

XSL•FO
RenderX

Funding

This research was funded by the Institut Universitari d'Investigació en Atenció Primària IDIAP Jordi Gol predoctoral fellowship (grant PREDOC-ECO-22/18). The article processing charges were funded by the Departament de Recerca i Universitats de la Generalitat de Catalunya to the Intervencions Sanitaries i Accions Comunitàries Research Group (2021 SGR 00884).

Data Availability

The datasets generated or analyzed during this study are not publicly available due to ethical considerations but are available upon reasonable request from the corresponding author.

Authors' Contributions

ET-L, SM-T, and CR-R coordinated the study, served as principal investigators, were responsible for the design of the study, and participated in the drafting of the manuscript. SM-T and MP-M conducted the formal analysis. ET-L and CR-R prepared the original draft. CR-R, SM-T, and FM-L supervised the final manuscript. All authors read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

References

1. Simmons KM, McIsaac SM, Ohle R. Impact of community-based interventions on out-of-hospital cardiac arrest outcomes: a systematic review and meta-analysis. Sci Rep 2023 Jun 23;13(1):10231. [doi: 10.1038/s41598-023-35735-y] [Medline: 37353542]
2. Gräsner JT, Wnent J, Herlitz J, et al. Survival after out-of-hospital cardiac arrest in Europe - results of the EuReCa TWO study. Resuscitation 2020 Mar 1;148:218-226. [doi: 10.1016/j.resuscitation.2019.12.042] [Medline: 32027980]
3. Ruiz Azpiazu JI, Fernández Del Valle P, Escriche MC, et al. Incidence, treatment, and factors associated with survival of out-of-hospital cardiac arrest attended by Spanish emergency services: report from the Out-of-Hospital Spanish Cardiac Arrest Registry for 2022 [Article in English, Spanish]. Emergencias 2024 Apr;36(2):131-139. [doi: 10.55633/s3me/014.2024] [Medline: 38597620]
4. Monsieurs KG, Nolan JP, Bossaert LL, et al. European Resuscitation Council guidelines for resuscitation 2015: section 1. Executive summary. Resuscitation 2015 Oct;95:1-80. [doi: 10.1016/j.resuscitation.2015.07.038] [Medline: 26477410]
5. Khatib A, Dababneh S, Ng T, et al. Short survey on cardiopulmonary resuscitation and automated external defibrillator training in rural British Columbia schools: preliminary findings and hypothesis-generating insights. CJC Open 2024;6(10):1240-1247. [doi: 10.1016/j.cjco.2024.07.006] [Medline: 39525343]
6. Liou FY, Lin KC, Chien CS, et al. The impact of bystander cardiopulmonary resuscitation on patients with out-of-hospital cardiac arrests. J Chin Med Assoc 2021 Dec 1;84(12):1078-1083. [doi: 10.1097/JCMA.0000000000000630] [Medline: 34610624]
7. Hsia RY, Huang D, Mann NC, et al. A US national study of the association between income and ambulance response time in cardiac arrest. JAMA Netw Open 2018 Nov 2;1(7):e185202. [doi: 10.1001/jamanetworkopen.2018.5202] [Medline: 30646394]
8. Hart J, Vakkalanka JP, Okoro U, Mohr NM, Ahmed A. Rural out-of-hospital cardiac arrest patients more likely to receive bystander CPR: a retrospective cohort study. Prehosp Emerg Care 2025;29(8):1104-1110. [doi: 10.1080/10903127.2024.2443478] [Medline: 39688931]
9. Katzenschlager S, Kelpanides IK, Ristau P, et al. Out-of-hospital cardiac arrest in children: an epidemiological study based on the German Resuscitation Registry identifying modifiable factors for return of spontaneous circulation. Crit Care 2023 Sep 7;27(1):349. [doi: 10.1186/s13054-023-04630-3] [Medline: 37679812]
10. Alwidyan MT, Alkhatib ZI, Alrawashdeh A, et al. Knowledge and willingness of schoolteachers in Jordan to perform CPR: a cross-sectional study. BMJ Open 2023 Aug 8;13(8):e073080. [doi: 10.1136/bmjopen-2023-073080] [Medline: 37553198]
11. Mollo A, Beck S, Degel A, Greif R, Breckwoldt J. Kids save lives: who should train schoolchildren in resuscitation? A systematic review. Resusc Plus 2024;20:100755. [doi: 10.1016/j.resplu.2024.100755] [Medline: 39282501]
12. Gino B, Siraj S, Peixoto M, Benson A, Dubrowski A. Comparing learning outcomes in cardiopulmonary resuscitation (CPR) and/or automated external defibrillator (AED) training for laypeople in face-to-face, online, and mixed training methods: an integrative literature review. Cureus 2023 May;15(5):e38489. [doi: 10.7759/cureus.38489] [Medline: 37273311]
13. Salehpoor-Emran M, Pashaeypoor S, Majdabadi ZA, Böttiger BW, Poortaghi S, Haghani S. The effect of online CPR training on the knowledge and practice of the Red Crescent Student Association Volunteers during the COVID-19 pandemic: a randomized clinical trial study. Resusc Plus 2025;25:101010. [doi: 10.1016/j.resplu.2025.101010] [Medline: 40703812]
14. Taverna-Llauradó E, Martínez-Torres S, Granado-Font E, et al. Online platform for cardiopulmonary resuscitation and automated external defibrillator training in a rural area: a community clinical trial protocol. BMJ Open 2024 Feb 7;14(2):e079467. [doi: 10.1136/bmjopen-2023-079467] [Medline: 38326271]

XSL•FO
RenderX

15. Consell Català de Resusscitació. Barcelona: Manual de l'alumne. In: Ressuscitació Cardiopulmonar: Consell Català de Resusscitació; 2010.
16. Basic life support. European Resuscitation Council. URL: https://www.erc.edu/education/for-professionals/course-types/bls/ [accessed 2026-05-12]
17. Organic law 3/2018, of 5 December, on the protection of personal data and guarantee of digital rights. Agencia Estatal Boletin Oficial del Estado. URL: https://www.boe.es/eli/es/lo/2018/12/05/3 [accessed 2026-05-08]
18. Castillo J, Rodríguez-Higueras E, Belmonte R, Rodríguez C, López A, Gallart A. Efficacy of virtual reality simulation in teaching basic life support and its retention at 6 months. Int J Environ Res Public Health 2023 Feb 24;20(5):4095. [doi: 10.3390/ijerph20054095] [Medline: 36901106]
19. Lin L, Ni S, Cheng J, et al. Effect of synchronous online vs. face-to-face cardiopulmonary resuscitation training on chest compression quality: a pilot randomized manikin study. Am J Emerg Med 2021 Dec;50:80-84. [doi: 10.1016/j.ajem.2021.07.009]
20. Han S, Park HJ, Nah S, et al. Instructor-led distance learning for training students in cardiopulmonary resuscitation: a randomized controlled study. PLoS One 2021;16(5):e0251277. [doi: 10.1371/journal.pone.0251277] [Medline: 33956873]

Abbreviations

AED: automated external defibrillator
CPR: cardiopulmonary resuscitation
EMS: emergency medical services
ERC: European Resuscitation Council
OHCA: out-of-hospital cardiac arrest

Edited by A Mavragani; submitted 09.Jul.2025; peer-reviewed by CJ Cunningham, N Mor; revised version received 27.Feb.2026; accepted 12.Mar.2026; published 22.May.2026.

Please cite as:
Taverna-Llauradó E, Martinez-Torres S, Roca-Biosca A, Pallejà-Millán M, Martín-Luján F, Rey-Reñones C
Effectiveness of Online Training on Cardiopulmonary Resuscitation and Automated External Defibrillator Use in Simulations With Rural Populations: Quasi-Experimental Community Intervention Study
JMIR Public Health Surveill 2026;12:e80359
URL: https://publichealth.jmir.org/2026/1/e80359
doi:10.2196/80359

© Elena Taverna-Llauradó, Sara Martinez-Torres, Alba Roca-Biosca, Meritxell Pallejà-Millán, Francisco Martín-Luján, Cristina Rey-Reñones. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 22.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Occupational Fatigue and Multidimensional Traffic Risk Outcomes Among Motorcycle-Based Food Delivery Workers: Cross-Sectional Study

Sookyung Kim[1], BSN, MPH, PhD; Chang Gi Park[2], PhD; Mi-So Shim[3], BSN, MPH, PhD

[1]School of Nursing, Soonchunhyang University, Cheonan, Republic of Korea

[2]Department of Population Health Nursing, College of Nursing, University of Illinois Chicago, Chicago, IL, United States

[3]College of Nursing, Keimyung University, 1095, Dalgubeol-daero, Dalseo-gu, Daegu, Republic of Korea

Corresponding Author:
Mi-So Shim, BSN, MPH, PhD
College of Nursing, Keimyung University, 1095, Dalgubeol-daero, Dalseo-gu, Daegu, Republic of Korea

Abstract

Background: Although motorcycle-based food delivery workers face a significant risk of accidents, previous research has primarily focused on traffic accidents, neglecting the multidimensional nature of safety, which includes perceived accident risk, near-miss experiences, and accident-related anxiety.

Objective: This study addresses this gap by investigating how occupational fatigue and health behaviors are associated with the Traffic Accident Risk Index (TARI) and its subdomains (near-miss experiences, self-rated accident anxiety, and other-rated accident anxiety).

Methods: A cross-sectional study was conducted. Data were collected from South Korean delivery workers via an online survey and analyzed using multiple linear regression.

Results: In total, 336 workers were included in the analysis. Occupational fatigue was positively associated with the overall risk index (*B*=0.017, 95% CI 0.012 to 0.021; *P*<.001) and all subdomains, including near-miss experiences (*B*=0.016, 95% CI 0.011 to 0.021; *P*<.001), other-rated accident anxiety (*B*=0.019, 95% CI 0.013 to 0.024; *P*<.001), and self-rated accident anxiety (*B*=0.016, 95% CI 0.011 to 0.021; *P*<.001). Inconsistent helmet use was associated with a higher TARI (*B*=0.413, 95% CI 0.030 to 0.796; *P*=.04), other-rated accident anxiety (*B*=0.495, 95% CI 0.025 to 0.966; *P*=.04), and self-rated accident anxiety (*B*=0.493, 95% CI 0.036 to 0.949; *P*=.04). Insufficient physical activity was associated with a higher TARI (*B*=0.368, 95% CI 0.016 to 0.719; *P*=.04) and self-rated accident anxiety (*B*=0.586, 95% CI 0.168 to 1.005; *P*=.006). Current smoking was associated with near-miss experiences (*B*=0.298, 95% CI 0.041 to 0.555; *P*=.02). Conversely, shorter break times were associated with lower accident risk (30 min to 1 h: *B*=−0.329, 95% CI −0.650 to −0.007; *P*=.045) and near-miss experiences (<30 min: *B*=−0.474, 95% CI −0.936 to −0.011; *P*=.045; 30 min to 1 h: *B*=−0.470, 95% CI −0.846 to −0.095; *P*=.01) than breaks exceeding 2 hours.

Conclusions: Occupational fatigue was associated with a higher overall perceived accident risk, more near-miss experiences, and greater accident-related anxiety. Modifiable health behaviors showed additional domain-specific associations. Prevention efforts may benefit from combining fatigue management with strategies to improve helmet use, increase physical activity, and support smoking cessation. Future research should refine the measurement of break time and establish evidence-based rest guidance for motorcycle-based delivery workers.

(JMIR Public Health Surveill 2026;12:e92667) doi:10.2196/92667

KEYWORDS

accidents; traffic; fatigue; occupational health; motorcycles; health risk behavior

Introduction

In recent years, the number of gig workers, defined as individuals who provide labor through digital platforms, has increased significantly worldwide [1]. The rapid expansion of food delivery applications has markedly increased the number of motorcycle-based food delivery workers (MFDWs) [2]. Because MFDWs deliver orders using motorcycles, they are routinely exposed to road hazards and face a disproportionately high risk of traffic accidents when compared with other occupational groups [3]. Furthermore, algorithmic management and piece-rate payment systems often compel MFDWs to operate under intense time pressure, potentially shaping risky riding behaviors and elevating crash risks [4]. Moreover, the concept of rest in the gig economy is often ambiguous; periods labeled as breaks may conflate restorative rest with unpaid

XSL•FO
RenderX

waiting times for algorithmically assigned tasks, potentially hindering effective recovery [5].

Although traffic accidents involving physical injury remain a primary concern [6], safety outcomes in this workforce encompass more than just accidents. Near-miss experiences and accident-related anxiety are critical yet understudied dimensions of traffic safety. Accident-related anxiety reflects not only a psychological response to danger but also a perceived vulnerability to crashes [7]. Accident-related anxiety, when combined with occupational fatigue, a state of exhaustion known to impair cognitive performance and elevate anxiety, may further compromise safety [8]. Importantly, from a preventative perspective, identifying early indicators before actual crash events is crucial for recognizing individuals at high risk of traffic accidents. Measures such as near-miss experiences and accident-related anxiety may provide valuable insight into preaccident risk and support the development of preventive interventions [9]. Therefore, understanding traffic safety requires a multidimensional approach that considers near-miss experiences and anxiety along with actual accidents among MFDWs.

Previous reviews have consistently reported sociodemographic characteristics and behavioral factors as key risk factors for adverse traffic safety outcomes among motorcycle riders. Sociodemographic factors, such as younger age and lower education level, along with behavioral characteristics, including alcohol consumption, smoking, nonhelmet use, and risky driving behaviors, have been consistently associated with traffic crashes and injuries [10,11]. Additionally, occupational status–related factors, such as long working hours and work schedules, are important correlates of traffic safety outcomes, and rider fatigue has been proposed as a significant risk factor [12]. Occupational fatigue has been identified as an important risk factor among motorcycle riders; however, it has rarely been assessed using validated instruments. Moreover, comprehensive analyses that simultaneously consider occupational fatigue, health-related behaviors, occupational status, and sociodemographic characteristics with adequate control for potential confounders remain limited.

Despite a growing literature on gig delivery workers, key gaps remain. Existing evidence has largely relied on crash- or injury-related outcomes, whereas near-miss or traffic-conflict events, widely discussed as proactive surrogate safety indicators, have been less commonly examined in this population [6,13]. Furthermore, the psychological aspects of safety, particularly traffic accident anxiety, have seldom been analyzed in relation to occupational fatigue and health behaviors. Additionally, existing studies lack comprehensive analyses that integrate occupational fatigue measured with validated instruments alongside health-related behaviors, occupational status, and sociodemographic characteristics, while adequately controlling for potential confounders. Addressing these gaps is vital to formulate prevention strategies that transcend regulatory enforcement. Therefore, this study aimed to bridge these limitations by investigating the work- and health-related determinants of traffic safety among MFDWs. Using a multidimensional framework, we assessed 3 distinct domains: Traffic Accident Risk Index (TARI), near-miss experiences, and accident-related anxiety. By integrating these experiential and psychological dimensions, this study aimed to provide comprehensive evidence to help occupational health policymakers improve the safety and well-being of platform delivery workers.

This study aimed to (1) assess the levels of traffic accident risk, including near-miss experiences, other-rated accident anxiety, self-rated accident anxiety, and occupational fatigue; (2) describe the demographic characteristics, occupational status, and health behaviors of MFDWs; (3) investigate the relationship between TARI and occupational fatigue; and (4) identify factors associated with TARI and its 3 subdomains among MFDWs.

Methods

Study Design

A cross-sectional descriptive study design was used. This study was reported in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines (Checklist 1).

Participants

The participants were MFDWs nationwide who were registered with platform-based delivery services and were recruited using convenience sampling. Given the fragmented nature of platform labor and the absence of a centralized registry for gig workers, MFDWs are difficult to reach through probability-based recruitment. Consequently, this sampling strategy was adopted to maximize access to actively working riders in real-world settings. Inclusion criteria were (1) being an adult and (2) having worked as an MFDW for more than 3 months. Individuals who could not understand or communicate in Korean were excluded.

Data Collection

The participants were recruited between August and October 2023 through offline and online channels. For offline recruitment, the researcher contacted multiple delivery agencies in 2 cities (Asan and Cheonan, Chungcheongnam-do, South Korea) and posted study notices at agencies that agreed to participate. MFDWs who expressed interest were directed to the study information sheet and consent form using QR codes. For online recruitment, a professional survey agency was contracted to distribute study invitations to MFDWs registered on major platform-based delivery services in South Korea. Participants began the survey after confirming their consent online. The study information sheet and informed consent form included details such as the right to withdraw from the study at any time during the study period; guarantees of anonymity; benefits, risks, inconveniences associated with participation; and confidentiality of personal information and records. In total, 336 participants voluntarily submitted consent forms and online survey responses.

Measures

Demographic Characteristics

Demographic characteristics were assessed using a questionnaire on age, gender, marital status, education, and socioeconomic status. Educational level was assessed based on the participants'

XSL•FO
RenderX

highest degree and categorized into 3 groups: middle school or below, high school, and college or above. Socioeconomic status was classified as high, moderate, or low.

Occupational Status

Occupational status was assessed using items related to career year, weekly working hours, breaks (including mealtimes), and helmet use. Career year was assessed by identifying the total duration that the participants had worked as professional riders. Break times, including mealtimes, were measured by asking the participants to report their daily break duration: 30 minutes or less, 30 minutes to less than 1 hour, 1 to less than 2 hours, or 2 hours or more. Finally, helmet use was assessed by asking the participants whether they always wear, usually wear, usually do not wear, or never wear. The responses were dichotomized into consistent helmet use (always or usually wear it) and inconsistent helmet use (usually do not wear it or never wear it).

Health-Related Characteristics

The following health-related characteristics were assessed: BMI, smoking status, alcohol problems, physical activity level, sleep duration, and the presence of disease. BMI was calculated from participants' self-reported height and weight by dividing weight (kg) by the square of height (m^2). The participants were then classified as underweight (<18.5 kg/m^2), normal (18.5 to <23.0 kg/m^2), overweight (23.0 to <25.0 kg/m^2), or obese (≥25.0 kg/m^2). Smoking status was assessed by asking whether the participant currently smoked ("yes" or "no"). Alcohol problems were assessed using the Alcohol Use Disorder Identification Test Concise, which consists of 3 items [14]. The participants were classified as having an alcohol problem if their total score was 6 or higher for males and 5 or higher for females [14]. Physical activity level was measured by asking the participants how many times they exercised per week. Sleep duration was classified into 7 to 9 hours and less than 7 or more than 9 hours based on adult recommendations and increased crash risk among drivers [15,16]. The presence of a disease was determined by asking a physician whether the participant had been diagnosed with any medical condition.

Occupational Fatigue

Fatigue was assessed using the Korean version of the Swedish Occupational Fatigue Inventory [17], originally developed by Ahsberg [18]. It consists of 20 items divided into 5 subscales (lack of energy, physical exertion, physical discomfort, lack of motivation, and sleepiness), each of which is scored on a 7-point Likert scale from 0 ("strongly disagree") to 6 ("strongly agree"). The sum of all the items ranged from 0 to 120 points, with a higher score indicating more severe fatigue. Cronbach α was in the range from 0.81 to 0.92 (lack of energy=0.92, physical exertion=0.87, physical discomfort=0.81, lack of motivation=0.92, and sleepiness=0.89) at development and was 0.95 in this study.

Traffic Accident Risk Index

Developed by Lee and Lee [9] and revised by Song and Lee [19], TARI captures perceived and self-reported dimensions of traffic risk, including near-miss experiences and accident-related anxiety, rather than objective accident incidence. The questionnaire was divided into 3 subdomains: 2 items assessed near-miss experiences as either perpetrators or victims, 2 items assessed other-rated accident anxiety (ie, the perceived likelihood of being involved in an accident as judged by people around the participant), and 2 items assessed self-rated accident anxiety. Each item was measured on a 5-point Likert scale ranging from 1 ("not at all") to 5 ("very much"). The overall score was calculated by averaging the scores of all items (range 1 - 5), with higher scores indicating a higher likelihood of traffic accidents. The score was treated as a continuous variable, and multiple linear regression was applied, as this approach is commonly used and appropriate for composite measures derived from Likert-type items. Cronbach α was 0.87 at development and 0.88 in this study. For the subdomains, Cronbach α values were 0.75 for near-miss experiences, 0.84 for other-rated accident anxiety, and 0.87 for self-rated accident anxiety.

Data Analysis

Data were analyzed using Stata software (version 18.0; StataCorp LP). The TARI and occupational fatigue levels, along with the participants' demographic characteristics, occupational status, and health-related characteristics, were analyzed using frequencies, percentages, means, and SDs. Correlations between occupational fatigue, TARI, and its domains were analyzed using Pearson correlation coefficient. Multiple regression analyses were conducted separately for TARI and each of its 3 domains to identify associated risk factors. To assess the potential for common method bias, Harman single-factor test was conducted using 26 items from the 2 primary psychological constructs (occupational fatigue and TARI).

Ethical Considerations

This study was approved by the Institutional Review Board of Soonchunhyang University (202305-SB-065) in accordance with the Declaration of Helsinki. Informed consent was obtained from all participants prior to their participation in the survey. All collected data were anonymized and deidentified to protect participants' privacy and confidentiality. Participants received a coupon worth KRW 10,000 (≈US $7) as a token of appreciation.

Results

General Characteristics

General characteristics of the participants are shown in Table 1. Approximately 90.5% (304/336) of the participants were male, with a mean age of 30.78 (SD 6.30) years. Among them, 77.7% (261/336) were single, 16.4% (55/336) had graduated from middle school, and 5.4% (18/336) reported a high socioeconomic status. Regarding occupational status, 9.8% (33/336) had less than 1 year of career experience; 29.2% (98/336) worked less than 40 hours per week; 25.3% (85/336) had breaks of less than 30 minutes, including mealtimes; and 93.2% (313/336) reported consistent helmet use (always or usually wearing a helmet). In terms of health-related characteristics, 47.0% (158/336) had a normal BMI, 36.9% (124/336) were smokers, 26.5% (89/336) reported alcohol problems, 10.4% (35/336) engaged in physical activity almost

XSL•FO
RenderX

every day, the mean sleep duration was 7.79 (SD 1.77) hours, 37.2% (125/336) reported short or long sleep duration, and 14.6% (49/336) reported the presence of disease.

XSL•FO
RenderX

Table . General characteristics of motorcycle-based food delivery workers (N=336).

Variables and characteristics	Values
Demographics	
Age (y)	
Mean (SD)	30.78 (6.30)
18 - 29, n (%)	172 (51.7)
30 - 39, n (%)	130 (39.0)
≥40, n (%)	31 (9.3)
Gender, n (%)	
Female	32 (9.5)
Male	304 (90.5)
Marital status, n (%)	
Single	261 (77.7)
Married	63 (18.7)
Divorced/widowed	12 (3.6)
Education, n (%)	
≤Middle school	55 (16.4)
High school	177 (52.7)
≥College	104 (30.9)
Socioeconomic status, n (%)	
High	18 (5.4)
Moderate	184 (54.7)
Low	134 (39.9)
Occupational status	
Career year, n (%)	
≤1	33 (9.8)
>1 to 3	166 (49.6)
>3 to 5	110 (32.8)
≥5	26 (7.8)
Working hours per week	
Mean (SD)	45.91 (15.81)
<40, n (%)	98 (29.2)
40 - 52, n (%)	138 (41.6)
>52, n (%)	96 (28.9)
Breaks including mealtimes, n (%)	
<30 min	85 (25.3)
30 min to 1 h	131 (39.0)
1 to 2 h	82 (24.4)
≥2 h	38 (11.3)
Helmet use, n (%)	
Consistent	313 (93.2)
Inconsistent	23 (6.9)
Health-related characteristics	
BMI	

XSL•FO
RenderX

Variables and characteristics	Values
Mean (SD)	23.35 (3.01)
Normal, n (%)	158 (47.0)
Underweight, n (%)	11 (3.3)
Overweight, n (%)	97 (29.0)
Obese, n (%)	69 (20.6)
Smoking status, n (%)	
Current smoker	124 (36.9)
Nonsmoker	212 (63.1)
Alcohol problems, n (%)	
Yes	89 (26.5)
No	247 (73.5)
Physical activity, n (%)	
Almost every day	35 (10.4)
3 - 4 times per week	69 (20.5)
1 - 2 times per week	91 (27.1)
Almost never	141 (42.0)
Sleep duration	
Mean (SD)	7.79 (1.77)
<7 or >9, n (%)	125 (37.2)
7 - 9, n (%)	198 (61.3)
Presence of disease, n (%)	
Yes	49 (14.6)
No	287 (85.4)

Traffic Accident Risk and Occupational Fatigue Levels

The average TARI score was 2.65 (SD 0.97), with scores ranging from 1 to 5. The means and SDs of the 3 domains of TARI (near-miss experiences, other-rated accident anxiety, and self-rated accident anxiety) were 3.03 (SD 1.08), 2.56 (SD 1.21), and 2.38 (SD 1.11), respectively. The mean total score for occupational fatigue was 48.46 (SD 26.58), with scores ranging from 0 to 113. Table 2 also presents the mean scores for the 5 occupational fatigue subscales.

Table . Characteristics of the occupational fatigue and the Traffic Accident Risk Index among motorcycle-based food delivery workers (N=336).

Variables	Mean (SD; range[a])
Occupational fatigue	48.46 (26.58; 0-113)
Lack of energy	10.91 (5.93; 0-24)
Physical exertion	7.91 (5.87; 0-20)
Physical discomfort	9.74 (6.15; 0-24)
Lack of motivation	9.01 (5.59; 0-24)
Sleepiness	10.89 (5.85; 0-24)
Traffic Accident Risk Index	2.65 (0.97; 1-5)
Near-miss experiences	3.03 (1.08;1-5)
Other-rated accident anxiety	2.56 (1.21; 1-5)
Self-rated accident anxiety	2.38 (1.11;1-5)

[a]Minimum-maximum.

Exploratory univariate comparisons suggested that TARI and subdomain scores varied across several demographic,

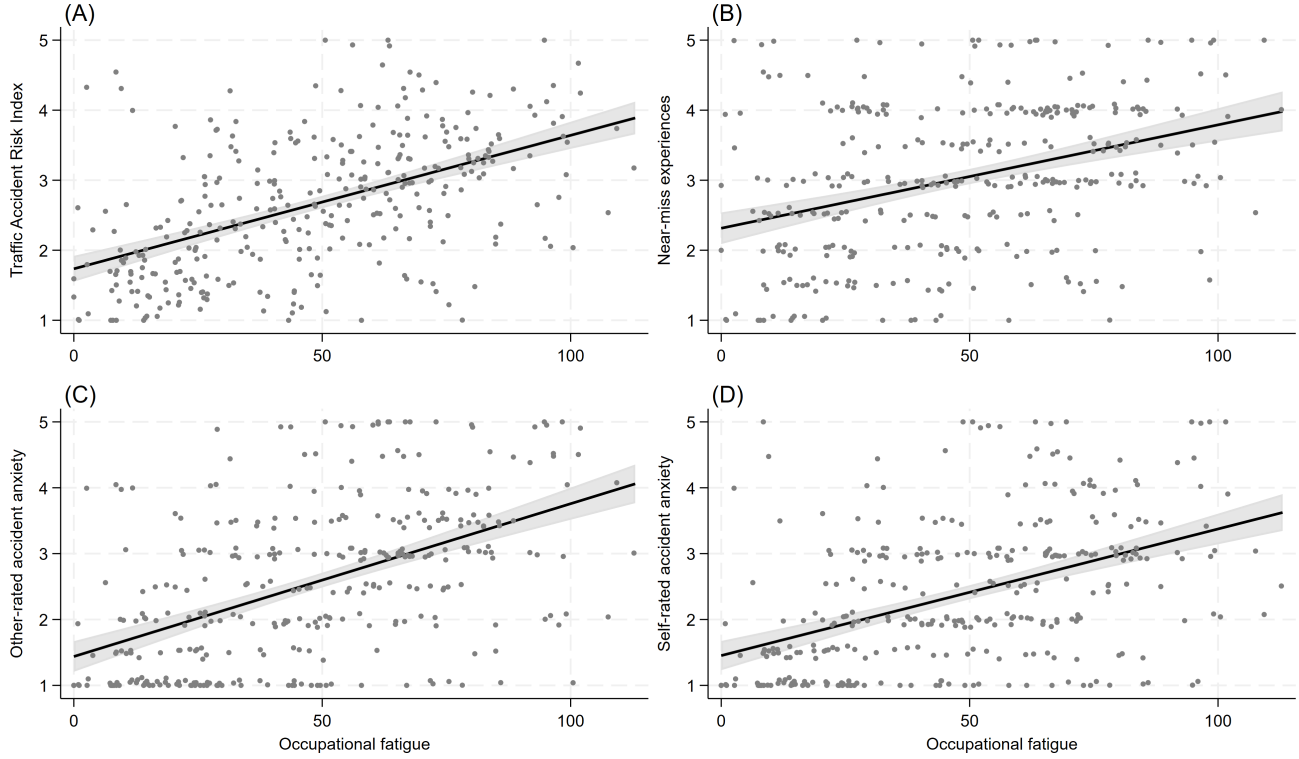
XSL•FO
RenderX

occupational, and health-related characteristics (Multimedia Appendix 1). These results were used to contextualize and inform the multivariable models.

Correlation Between Traffic Accident Risk Outcomes and Occupational Fatigue

Occupational fatigue was positively correlated with the overall TARI and its subdomains, including near-miss experiences and accident-related anxiety (r=0.36-0.52, all P<.001; Figure 1). These unadjusted patterns were consistent with the multivariable models (Table 3). The full correlation matrix is provided in Multimedia Appendix 2.

Figure 1. Unadjusted associations between occupational fatigue and traffic accident risk outcomes. Scatterplots show bivariate relationships between occupational fatigue and (A) Traffic Accident Risk Index, (B) near-miss experiences, (C) other-rated accident anxiety, and (D) self-rated accident anxiety. Solid lines indicate fitted values from linear regression, and shaded bands represent 95% CIs. Multivariable-adjusted estimates are reported in Table 3.



XSL•FO
RenderX

Table . Factors affecting Traffic Accident Risk Index and its 3 domains among motorcycle-based food delivery workers (N=336).

Variables	TARI[a]		Near-miss experiences		Other-rated accident anxiety		Self-rated accident anxiety	
	B (SE)	95% CI	*B* (SE)	95% CI	*B* (SE)	95% CI	*B* (SE)	95% CI
Age (y) (ref[b]: 18 - 29 y)								
30 - 39	0.119 (0.114)	−0.104 to 0.343	0.148 (0.133)	−0.113 to 0.409	0.053 (0.139)	−0.220 to 0.327	0.157 (0.135)	−0.109 to 0.423
≥40	0.208 (0.217)	−0.219 to 0.635	0.349 (0.254)	-0.150 to 0.849	0.215 (0.266)	−0.309 to 0.738	0.060 (0.258)	−0.448 to 0.569
Sex (ref: female)								
Male	0.155 (0.172)	−0.184 to 0.493	−0.081 (0.201)	−0.477 to 0.315	0.353 (0.211)	−0.062 to 0.769	0.192 (0.205)	−0.212 to 0.595
Marital status (ref: single)								
Married	0.022 (0.148)	−0.269 to 0.314	−0.030 (0.173)	−0.371 to 0.311	−0.085 (0.182)	−0.442 to 0.273	0.181 (0.177)	−0.167 to 0.529
Widowed/divorced	−0.362 (0.330)	−1.011 to 0.288	−0.348 (0.386)	−1.107 to 0.411	−0.458 (0.405)	−1.254 to 0.339	−0.280 (0.393)	−1.054 to 0.494
Education (ref: ≥College)								
≤Middle school	0.209 (0.203)	−0.190 to 0.608	−0.287 (0.237)	−0.753 to 0.180	0.525[c] (0.249)	0.036 to 1.014	0.389 (0.242)	−0.086 to 0.865
High school	−0.028 (0.106)	−0.237 to 0.181	−0.239 (0.124)	−0.484 to 0.006	0.012 (0.130)	−0.245 to 0.268	0.143 (0.127)	−0.107 to 0.392
SES[d] (ref: High)								
Moderate	−0.049 (0.246)	−0.533 to 0.436	-0.353 (0.288)	−0.919 to 0.213	−0.085 (0.302)	−0.679 to 0.509	0.292 (0.293)	−0.285 to 0.869
Low	−0.005 (0.244)	−0.486 to 0.476	−0.253 (0.286)	−0.815 to 0.310	−0.066 (0.300)	−0.656 to 0.523	0.303 (0.291)	−0.270 to 0.876
Career year (ref: ≤1 y)								
>1 to 3	0.012 (0.173)	−0.328 to 0.352	0.104 (0.202)	−0.293 to 0.502	−0.006 (0.212)	−0.423 to 0.411	−0.062 (0.206)	−0.467 to 0.343
>3 to 5	0.154 (0.187)	−0.214 to 0.522	0.189 (0.218)	−0.241 to 0.619	0.252 (0.229)	−0.199 to 0.703	0.022 (0.223)	−0.416 to 0.460
≥5	−0.119 (0.239)	−0.590 to 0.352	0.040 (0.280)	−0.511 to 0.591	−0.269 (0.294)	−0.847 to 0.309	−0.128 (0.285)	−0.689 to 0.434
Working hours per week (ref: <40 h)								
40 - 52	−0.202 (0.141)	−0.479 to 0.075	−0.300 (0.164)	−0.623 to 0.024	0.001 (0.172)	−0.339 to 0.340	−0.308 (0.168)	−0.637 to 0.022
>52	−0.078 (0.150)	−0.373 to 0.218	−0.044 (0.176)	−0.390 to 0.301	0.108 (0.184)	−0.255 to 0.471	−0.297 (0.179)	−0.649 to 0.056
Break time (ref: ≥2 h)								
≤30 min	−0.281 (0.201)	−0.677 to 0.115	−0.474[c] (0.235)	−0.936 to −0.011	−0.228 (0.247)	−0.714 to 0.257	−0.141 (0.240)	−0.613 to 0.331
30 min to 1 h	−0.329[c] (0.163)	−0.650 to −0.007	−0.470[c] (0.191)	−0.846 to −0.095	−0.317 (0.200)	−0.711 to 0.077	−0.198 (0.194)	−0.581 to 0.184
1 to 2 h	−0.237 (0.175)	−0.581 to 0.108	−0.254 (0.205)	−0.657 to 0.149	−0.211 (0.215)	−0.634 to 0.212	−0.245 (0.209)	−0.656 to 0.165
Helmet use (ref: consistent)								
Inconsistent	0.413[c] (0.195)	0.030 to 0.796	0.251 (0.228)	−0.197 to 0.699	0.495[c] (0.239)	0.025 to 0.966	0.493[c] (0.232)	0.036 to 0.949
BMI (ref: underweight to normal)								

XSL•FO
RenderX

Variables	TARI[a]		Near-miss experiences		Other-rated accident anxiety		Self-rated accident anxiety	
	B (SE)	95% CI	*B* (SE)	95% CI	*B* (SE)	95% CI	*B* (SE)	95% CI
Over-weight/Obese	−0.096 (0.105)	−0.302 to 0.110	0.009 (0.122)	−0.232 to 0.249	−0.135 (0.128)	−0.387 to 0.118	−0.162 (0.125)	−0.407 to 0.083
Smoking (ref: nonsmoker)								
Current smoker	0.008 (0.112)	−0.212 to 0.227	0.298[c] (0.130)	0.041 to 0.555	−0.172 (0.137)	−0.441 to 0.098	−0.104 (0.133)	−0.366 to 0.158
Alcohol problem (ref: no)								
Yes	0.135 (0.110)	−0.082 to 0.352	0.170 (0.129)	−0.084 to 0.424	0.085 (0.135)	−0.182 to 0.351	0.151 (0.131)	−0.108 to 0.410
Physical activity (ref: almost everyday)								
3 - 4 times/wk	0.246 (0.199)	−0.145 to 0.637	0.222 (0.232)	−0.235 to 0.679	0.160 (0.244)	−0.320 to 0.639	0.356 (0.237)	−0.110 to 0.822
1 - 2 times/wk	0.247 (0.189)	−0.124 to 0.618	0.336 (0.221)	−0.098 to 0.770	0.064 (0.231)	−0.391 to 0.520	0.340 (0.225)	−0.102 to 0.782
Almost never	0.368[c] (0.179)	0.016 to 0.719	0.340 (0.209)	−0.071 to 0.751	0.177 (0.219)	−0.254 to 0.607	0.586[e] (0.213)	0.168 to 1.005
Sleep duration (ref: 7 - 9 h)								
<7 or >9 h	−0.028 (0.099)	−0.223 to 0.167	0.017 (0.116)	−0.211 to 0.245	−0.096 (0.122)	−0.335 to 0.143	−0.006 (0.118)	−0.239 to 0.226
Presence of disease (ref: no)								
Yes	−0.017 (0.149)	−0.310 to 0.275	0.100 (0.174)	−0.242 to 0.442	−0.001 (0.182)	−0.360 to 0.358	−0.151 (0.177)	−0.500 to 0.198
Occupational fatigue	0.017[f] (0.002)	0.012 to 0.021	0.016[f] (0.003)	0.011 to 0.021	0.019[f] (0.003)	0.013 to 0.024	0.016[f] (0.003)	0.011 to 0.021

[a]TARI: Traffic Accident Risk Index.

[b]ref: reference group.

[c]*P*<.05.

[d]SES: socioeconomic status.

[e]*P*<.01.

[f]*P*<.001.

Factors Associated With TARI and the 3 Subdomains Among MFDWs

The results of the multiple linear regression analyses for TARI and its 3 subdomains, including near-miss experiences, other-rated accident anxiety, and self-rated accident anxiety, are presented in Table 3. All 4 regression models were statistically significant, with adjusted R^2 values of 29.9%, 21.6%, 29.2%, and 23.5% for the TARI, near-miss experiences, other-rated accident anxiety, and self-rated accident anxiety models, respectively. Multicollinearity was assessed across all regression models, and all variance inflation factors were less than 10, indicating no significant multicollinearity. When TARI was the dependent variable, the statistically significant factors included break time, helmet use, physical activity, and occupational fatigue. Compared with participants with break times of 2 hours or more, those with break times of 30 minutes to 1 hour had lower perceived traffic accident risk (*B*=−0.329, 95% CI −0.650 to −0.007; *P*=.045). Additionally, inconsistent helmet use (*B*=0.413, 95% CI 0.030 to 0.796; *P*=.04), lack of physical activity (*B*=0.368, 95% CI 0.016 to 0.719; *P*=.04), and higher levels of occupational fatigue (*B*=0.017, 95% CI 0.012 to 0.021; *P*<.001) were associated with higher TARI scores.

In the model with near-miss experiences as the dependent variable, the statistically significant factors included break times, smoking, and occupational fatigue. Compared with participants with break times of 2 hours or more, those with break times of less than 30 minutes (*B*=−0.474, 95% CI −0.936 to −0.011; *P*=.045) or between 30 minutes and 1 hour (*B*=−0.470, 95% CI −0.846 to −0.095; *P*=.01) were associated with lower levels of near-miss experiences. Additionally, smoking (*B*=0.298, 95% CI 0.041 to 0.555; *P*=.02) and higher levels of occupational fatigue (*B*=0.016, 95% CI 0.011 to 0.021; *P*<.001) were associated with increased near-miss experiences.

In the model with other-rated accident anxiety as the dependent variable, statistically significant factors included educational level, helmet use, and occupational fatigue. Participants with an educational level of middle school or lower (*B*=0.525, 95% CI 0.036 to 1.014; *P*=.04), compared with those with a college education or higher, those with inconsistent helmet use (*B*=0.495, 95% CI 0.025 to 0.966; *P*=.04), and those with higher

XSL•FO
RenderX

occupational fatigue (*B*=0.019, 95% CI 0.013 to 0.024; *P*<.001) were associated with greater other-rated anxiety about accidents.

Finally, in the model with self-rated accident anxiety as the dependent variable, statistically significant factors included helmet use, physical activity, and occupational fatigue. Participants with inconsistent helmet use (*B*=0.493, 95% CI 0.036 to 0.949; *P*=.04), those who did not engage in physical activity (B=0.586, 95% CI 0.168 to 1.005; *P*=.006) compared with those who exercised almost daily, and those with higher levels of occupational fatigue (*B*=0.016, 95% CI 0.011 to 0.021; *P*<.001) were associated with higher self-rated accident anxiety.

Discussion

This study contributes to the growing body of literature on MFDWs by adopting a multidimensional view of traffic safety beyond crash-based outcomes. We examined TARI, near-miss experiences, and 2 forms of accident-related anxiety (self-rated and other-rated anxiety). This distinction is meaningful because behavioral precursors to accidents (near-miss experiences) and the psychological dimensions of safety (accident anxiety) may not share identical correlates. By modeling these outcomes separately, our findings offer a more granular understanding of how occupational fatigue and health behaviors are related to safety precursors and perceived vulnerability in platform-based delivery work.

Across all the models, occupational fatigue emerged as the most consistent correlate of adverse safety outcomes. Higher fatigue was associated with a higher overall traffic accident risk, more near-miss experiences, and greater accident-related anxiety. This pattern is broadly consistent with prior evidence indicating that fatigue is linked to degraded cognitive performance and safety-critical decision-making [20,21]. Additionally, fatigue has been associated with reduced situational awareness and slower reaction times, which have been linked to near-miss events and judgment errors in complex traffic environments [22,23]. The consistency of the association across multiple safety domains suggests that fatigue is a salient marker of increased safety risks and psychological strain in MFDWs. Although causal inference was not warranted in this cross-sectional study, the consistency of fatigue associations across all safety domains suggests that fatigue management warrants prioritization in this workforce. At the platform level, immediate steps could include in-app advisory alerts recommending rest after extended delivery periods, as well as the integration of fatigue-related considerations into the delivery system design, followed by rider-facing educational interventions that raise awareness of fatigue-inducing factors and promote monitoring. Looking further ahead, the development of cumulative fatigue monitoring systems integrated into delivery platforms, as well as policy-level requirements for platforms to disclose algorithmic workload indicators to occupational health authorities, should be pursued.

Several modifiable health behaviors were independently associated with specific safety domains, even after adjusting for fatigue and other covariates. The inconsistent use of helmets was associated with a higher overall accident risk and greater levels of self-rated and other-rated accident anxiety, consistent with the evidence that helmet use remains a key protective factor for motorcyclists [10,24]. Insufficient physical activity was associated with higher overall risk and self-rated accident anxiety. This finding aligns with broader evidence linking physical activity to lower anxiety symptoms and improved mental well-being [25] and may plausibly relate to self-regulation or coping capacity relevant to safety, although the underlying mechanisms in this occupational context remain unclear and warrant further investigation. Smoking was specifically associated with higher near-miss experiences, which may reflect distraction-related or broader risk-prone patterns [11,26-28]. Together, these findings provide preliminary evidence supporting a multicomponent view of safety promotion, suggesting that fatigue management may be most effective when considered alongside targeted health-behavior strategies.

The multivariable models explained a significant proportion of the variance in the outcomes (adjusted R^2=29.9% for TARI, 21.6% for near-miss experiences, 29.2% for other-rated accident anxiety, and 23.5% for self-rated accident anxiety), suggesting that fatigue and the included behavioral and occupational covariates capture important correlates of multidimensional safety. At the same time, the remaining unexplained variance indicates that additional work- and context-related determinants may also contribute to safety outcomes and should be incorporated in future studies.

We also observed a counterintuitive inverse association between longer self-reported break duration and safety outcomes, with riders reporting breaks of 2 hours or more showing higher accident risk and near-miss experiences. This counterintuitive finding should not be interpreted as evidence that shorter breaks are protective or that reducing rest improves safety among MFDWs. This finding should be regarded as exploratory and hypothesis generating, as alternative explanations cannot be excluded, including reverse causation and potential sensitivity to how break duration was categorized. One plausible hypothesis raised by this study is that the observed pattern may reflect a structural measurement limitation inherent in algorithmically managed delivery. In platform-based food delivery, periods that MFDWs label as “breaks” may frequently coincide with unpaid on-duty waiting time, during which riders remain logged into the app and available for dispatch [5]. If so, those reporting the longest “breaks” may not represent a well-rested MFDW, but rather one with prolonged on-duty availability that sustains cognitive load rather than enabling genuine recovery. Recovery research suggests that the restorative value of rest depends not only on its duration but also on psychological detachment from work demands [29], and such detachment may be structurally difficult to achieve within algorithmically managed gig work. Previous studies further suggested that prolonged idle periods can induce compensatory time pressure once orders resume, which may be associated with riskier riding behavior [4,30]. However, these explanations remain speculative, and this hypothesis requires direct empirical testing in future studies. Future studies should employ more precise measures that distinguish restorative rest from on-duty waiting, ideally complemented by objective indicators such as platform activity logs or ecological momentary assessments.

XSL•FO
RenderX

This study had some limitations. First, its cross-sectional design precluded causal inferences. However, given that MFDWs represent a particularly vulnerable occupational group in which traffic incidents frequently result in serious injuries [3], identifying strong correlates such as occupational fatigue provides critical and immediate evidence for prioritizing interventions, even in the absence of longitudinal confirmation. Second, self-reported measures may have introduced recall or reporting bias. In particular, break times may not clearly distinguish actual rest from idle waiting times. Nevertheless, it is important to note that psychological constructs, such as "accident anxiety" and "near-miss experiences," are inherently subjective. These internal vulnerability states cannot be captured through objective accident records alone, making self-reporting an appropriate and valid assessment method. As the 2 primary psychological constructs, occupational fatigue and TARI, were assessed using a single self-report survey, the possibility of common method bias cannot be fully excluded. To examine this issue, we conducted Harman single-factor test, which showed that the first factor accounted for 45.78% of the total variance, below the commonly used 50% threshold. Third, several important unmeasured variables may limit the interpretation of our findings. Objective exposure indicators such as riding distance, number of deliveries per shift, shift timing, route type, urban traffic exposure, weather conditions, incentive pressure, and platform algorithm demands were not captured. Given that these factors may vary considerably across MFDWs and could independently influence safety outcomes, the findings regarding TARI and its subdomains should be interpreted with caution. Furthermore, residual confounding cannot be ruled out, as health behaviors such as smoking and physical inactivity may partly reflect unmeasured risk-prone tendencies, and unmeasured work-context factors may have contributed to the observed associations. Finally, because convenience sampling was employed, the findings should be generalized with caution. In particular, South Korea's food delivery market is characterized by a predominance of algorithmic dispatch systems, peak-hour demand concentration, and highly developed road infrastructure, which may differ substantially from delivery contexts in other countries [31,32]. Therefore, findings should be interpreted with caution when applied to other national settings. Consequently, future studies should use longitudinal designs and more precise measurements of rest and workload, along with objective exposure indicators, to validate these findings. Real-time approaches, such as ecological momentary assessments and complementary objective indicators, may help clarify temporal relationships and support more actionable guidance on fatigue management and recovery for MFDWs working under algorithmically managed piece-rate conditions.

This study contributes to the literature on MFDWs by conceptualizing traffic safety as a multidimensional construct encompassing accident risk, near-miss experiences, and accident-related anxiety. Occupational fatigue was consistently associated with all the outcomes, suggesting its utility as an important indicator of elevated safety risks and accident-related anxiety. Modifiable health behaviors, including inconsistent helmet use, insufficient physical activity, and smoking, were also associated with specific safety domains, indicating the potential value of considering fatigue management alongside health-behavior approaches. The inverse association observed between longer self-reported break duration and safety outcomes warrants cautious interpretation, as break time may partly reflect on-duty waiting rather than restorative rest. Future studies using longitudinal designs and more precise measures of rest and exposure are required to clarify temporality and inform context-sensitive strategies for supporting recovery during platform delivery.

Acknowledgments

We thank all participants for their time and valuable contributions to this study. No generative artificial intelligence tools were used in the preparation of this manuscript.

Funding

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean Government (MSIT; RS-2023-00211574) and the Soonchunhyang University Research Fund. The funder was not involved in the study design, data collection, analysis, interpretation, or writing of the manuscript.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: SK, MSS, CGP
Data curation: SK, MSS
Formal analysis: MSS
Funding acquisition: SK
Investigation: SK
Methodology: SK, CGP
Project administration: SK
Resources: SK

XSL•FO
RenderX

Supervision: SK, CGP
Validation: MSS
Visualization: MSS, SK
Writing – original draft: SK, MSS
Writing – review & editing: SK, MSS, CGP

Conflicts of Interest

None declared.

Multimedia Appendix 1
Exploratory univariate comparisons of the Traffic Accident Risk Index and its three domains by participants' characteristics (N=336).
[DOCX File, 35 KB - publichealth_v12i1e92667_app1.docx]

Multimedia Appendix 2
Pearson correlations between traffic accident risk outcomes and occupational fatigue among motorcycle-based food delivery workers (N=336).
[DOCX File, 23 KB - publichealth_v12i1e92667_app2.docx]

Checklist 1
STROBE checklist.
[PDF File, 86 KB - publichealth_v12i1e92667_app3.pdf]

References

1. OECD/ILO/European Union. Handbook on measuring digital platform employment and work. : OECD Publishing; 2023 URL: https://www.oecd.org/en/publications/handbook-on-measuring-digital-platform-employment-and-work_0ddcac3b-en.html#related-publications [accessed 2026-05-28]
2. Online food delivery - worldwide. Statista. URL: https://www.statista.com/outlook/emo/online-food-delivery/worldwide [accessed 2025-10-07]
3. Kwangsukstith S, Surawattanasakul V, Mahakkanukrauh C, et al. Major injuries and associated factors in traffic accidents among motorcycle food delivery riders during the COVID-19 pandemic in Thailand. Heliyon 2024 Oct 30;10(20):e39032. [doi: 10.1016/j.heliyon.2024.e39032] [Medline: 39640680]
4. Christie N, Ward H. Delivering hot food on motorcycles: a mixed method study of the impact of business model on rider behaviour and safety. Saf Sci 2023 Feb;158:105991. [doi: 10.1016/j.ssci.2022.105991]
5. Realizing decent work in the platform economy ILC113/v(1). : International Labour Organization; 2024 URL: https://www.ilo.org/sites/default/files/2024-07/ILC113-V%281%29-%5BWORKQ-231121-002%5D-Web-EN.pdf [accessed 2025-11-17]
6. Useche SA, Traficante S, Llamazares FJ, Marin C. The human cost of fast deliveries: a systematic literature review of occupational risks and safety outcomes in last-mile delivery workers. J Transp Health 2025 Oct;44:102133. [doi: 10.1016/j.jth.2025.102133]
7. Clapp JD, Olsen SA, Danoff-Burg S, et al. Factors contributing to anxious driving behavior: the role of stress history and accident severity. J Anxiety Disord 2011 May;25(4):592-598. [doi: 10.1016/j.janxdis.2011.01.008] [Medline: 21377829]
8. Useche SA, Robayo S, Orozco-Fontalvo M. The hidden cost of your "too fast food": stress-related factors and fatigue predict food delivery riders' occupational crashes. Int J Occup Saf Ergon 2024 Sep;30(3):825-834. [doi: 10.1080/10803548.2024.2356997] [Medline: 38853658]
9. Lee SY, Lee SC. A validation study and confirmatory factor analysis of the Traffic Accident Risk Index (TARI). Korean J Ind Organ Psychol 2010 Mar;23(1):75-87. [doi: 10.24230/ksiop.23.1.201003.75]
10. Kiwango G, Katopola D, Francis F, Möller J, Hasselberg M. A systematic review of risk factors associated with road traffic crashes and injuries among commercial motorcycle drivers. Int J Inj Contr Saf Promot 2024 Jun;31(2):332-345. [doi: 10.1080/17457300.2024.2319628] [Medline: 38385344]
11. Konlan KD, Hayford L. Factors associated with motorcycle-related road traffic crashes in Africa, a scoping review from 2016 to 2022. BMC Public Health 2022 Apr 5;22(1):649. [doi: 10.1186/s12889-022-13075-2] [Medline: 35382791]
12. Amoadu M, Ansah EW, Sarfo JO. Psychosocial work factors, road traffic accidents and risky driving behaviours in low-and middle-income countries: a scoping review. IATSS Research 2023 Jul;47(2):240-250. [doi: 10.1016/j.iatssr.2023.03.005]
13. Zheng Y, Ma Y, Guo L, Cheng J, Zhang Y. Crash involvement and risky riding behaviors among delivery riders in China: the role of working conditions. Transp. Res. Rec 2019 Apr;2673(4):1011-1022. [doi: 10.1177/0361198119841028]
14. Lee JH, Kong KA, Lee DH, Choi YH, Jung KY. Validation and proposal for cut-off values of an abbreviated version of the Alcohol Use Disorder Identification Test using the Korean National Health and Nutrition Examination Survey. Clin Exp Emerg Med 2018 Jun;5(2):113-119. [doi: 10.15441/ceem.17.228] [Medline: 29973036]

XSL•FO
RenderX

15. Acute sleep deprivation and risk of motor vehicle crash involvement. : AAA Foundation for Traffic Safety; 2016 URL: https://aaafoundation.org/research/acute-sleep-deprivation-risk-motor-vehicle-crash-involvement [accessed 2026-05-28]
16. Watson NF, Badr MS, Belenky G, et al. Recommended amount of sleep for a healthy adult: a joint consensus statement of the American Academy of Sleep Medicine and Sleep Research Society. Sleep 2015 Jun 1;38(6):843-844. [doi: 10.5665/sleep.4716] [Medline: 26039963]
17. Lee S, Seong S, Park S, et al. Korean version of the Swedish Occupational Fatigue Inventory among construction workers: cultural adaptation and psychometric evaluation. Int J Environ Res Public Health 2021 Apr 18;18(8):4302. [doi: 10.3390/ijerph18084302] [Medline: 33919602]
18. Ahsberg E. Dimensions of fatigue in different working populations. Scand J Psychol 2000 Sep;41(3):231-241. [doi: 10.1111/1467-9450.00192] [Medline: 11041305]
19. Song H, Lee S. Changes in traffic accident risk degree according to driving workload rate and moderate effect of ego-resilience. Korean J Ind Organ Psychol 2011;24(1):29-50. [doi: 10.24230/kjiop.v24i1.29-50]
20. Techera U, Hallowell M, Stambaugh N, Littlejohn R. Causes and consequences of occupational fatigue: meta-analysis and systems model. J Occup Environ Med 2016 Oct;58(10):961-973. [doi: 10.1097/JOM.0000000000000837] [Medline: 27525527]
21. Williamson A, Lombardi DA, Folkard S, Stutts J, Courtney TK, Connor JL. The link between fatigue and safety. Accid Anal Prev 2011 Mar;43(2):498-515. [doi: 10.1016/j.aap.2009.11.011] [Medline: 21130213]
22. Dawson D, Reynolds AC, Van Dongen HPA, Thomas MJW. Determining the likelihood that fatigue was present in a road accident: a theoretical review and suggested accident taxonomy. Sleep Med Rev 2018 Dec;42:202-210. [doi: 10.1016/j.smrv.2018.08.006] [Medline: 30274744]
23. Zhang M, Murphy LA, Fang D, Caban-Martinez AJ. Influence of fatigue on construction workers' physical and cognitive function. Occup Med (Lond) 2015 Apr;65(3):245-250. [doi: 10.1093/occmed/kqu215] [Medline: 25701835]
24. Lee G, Park JH, Lee SC, Lee S, Song KJ, Shin SD. The epidemiology and outcomes of severe road traffic injury by road user type in Korea: a nationwide retrospective observational study. Injury 2024 Oct;55(10):111732. [doi: 10.1016/j.injury.2024.111732] [Medline: 39084036]
25. McDowell CP, Dishman RK, Gordon BR, Herring MP. Physical activity and anxiety: a systematic review and meta-analysis of prospective cohort studies. Am J Prev Med 2019 Oct;57(4):545-556. [doi: 10.1016/j.amepre.2019.05.012] [Medline: 31542132]
26. de Rome L, Brown J, Baldock M, Fitzharris M. Near-miss crashes and other predictors of motorcycle crashes: findings from a population-based survey. Traffic Inj Prev 2018;19(sup2):S20-S26. [doi: 10.1080/15389588.2018.1536822] [Medline: 30540505]
27. Jomnonkwao S, Champahom T, Se C, Hantanong N, Ratanavaraha V. Contributing factors to near-miss experiences of motorcyclists in Thailand: a random parameter probit model approach. Heliyon 2023 Dec;9(12):e22625. [doi: 10.1016/j.heliyon.2023.e22625] [Medline: 38090010]
28. Kim S, Woo MS, Hong S. Health and health management among motorcycle-based food delivery workers in South Korea: a qualitative interview study. Int J Qual Stud Health Well-being 2026 Dec 31;21(1):2613971. [doi: 10.1080/17482631.2026.2613971] [Medline: 41520275]
29. Sonnentag S, Fritz C. The Recovery Experience Questionnaire: development and validation of a measure for assessing recuperation and unwinding from work. J Occup Health Psychol 2007 Jul;12(3):204-221. [doi: 10.1037/1076-8998.12.3.204] [Medline: 17638488]
30. Hsu CK, Lee TC, Wong KI. Time is money: effects of time-sensitive incentive on risky driving among on-demand food-delivery motorcyclists in Taiwan. Saf Sci 2024 Aug;176:106549. [doi: 10.1016/j.ssci.2024.106549]
31. Chung N. Thinking food delivery platforms infrastructurally: the practices and politics of Baemin's infrastructuralisation in Seoul. New Technol Work Employ 2025 Jul;40(2):214-241. [doi: 10.1111/ntwe.12311]
32. Moszoro M, Soto M. Road quality and mean speed score. IMF Working Papers. Preprint posted online on May 20, 2022. [doi: 10.5089/9798400210440.001]

Abbreviations

MFDW: motorcycle-based food delivery worker
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
TARI: Traffic Accident Risk Index

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 02.Feb.2026; peer-reviewed by S Saita, TAK Manne; revised version received 02.May.2026; accepted 20.May.2026; published 12.Jun.2026.

Please cite as:
Kim S, Park CG, Shim MS
Occupational Fatigue and Multidimensional Traffic Risk Outcomes Among Motorcycle-Based Food Delivery Workers: Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e92667
URL: https://publichealth.jmir.org/2026/1/e92667
doi:10.2196/92667

© Sookyung Kim, Chang Gi Park, Mi-So Shim. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 12.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Population-Level Impact of the Enterovirus A71 Vaccination Program on Hand, Foot, and Mouth Disease: Ecological Time-Series Study

Ye Tong[1*], PhD; Huan Fan[1*], MSc; Jing Wang[1], MSc; Xinru Zeng[1], MSc; Xiaoqing Cheng[1], MSc; Changjun Bao[1], MSc; Liguo Zhu[1,2], PhD; Hong Ji[1], PhD; Xiang Huo[1], PhD

[1]Department of Acute Infectious Disease Control and Prevention, Jiangsu Provincial Center for Disease Control and Prevention, 172 Jiangsu Road, Nanjing, China

[2]School of Public Health, Nanjing Medical University, Nanjing, China

[*]these authors contributed equally

Corresponding Author:
Hong Ji, PhD
Department of Acute Infectious Disease Control and Prevention, Jiangsu Provincial Center for Disease Control and Prevention, 172 Jiangsu Road, Nanjing, China

Abstract

Background: Hand, foot, and mouth disease (HFMD), a common childhood illness caused by various enteroviruses, poses a significant public health threat in the Asia-Pacific region, where severe cases associated with enterovirus A71 (EV71) are a major concern. The EV71 vaccination program was introduced in China in late 2016. Although randomized controlled trials have established the robust efficacy and safety of these vaccines, assessing their real-world performance remains crucial. Subsequent studies have evaluated its real-world effectiveness in several provinces, including Zhejiang and Guangdong. However, evidence on its real-world impact in reducing EV71-associated HFMD in Jiangsu Province remains limited.

Objective: This study aimed to describe HFMD epidemiological characteristics and to evaluate the population-level effectiveness of the EV71 vaccination program in Jiangsu Province.

Methods: We characterized the temporal distribution of EV71-related HFMD cases based on longitudinal surveillance data collected through the National Notifiable Diseases Surveillance System in Jiangsu Province from 2012 to 2019 and assessed the impact of vaccination using a Bayesian structural time series model under a counterfactual framework. The intervention effect of the EV71 vaccine was estimated by comparing the observed data with predictions from a counterfactual model scenario without vaccination.

Results: A total of 932,274 HFMD cases were reported in Jiangsu from 2012 to 2019, including 5190 (0.56%) severe cases. An estimated 140,876 (15.11%) cases were attributed to EV71. EV71-associated HFMD cases showed a marked decline, with estimated numbers of 22,303, 9796, and 3900 in 2017, 2018, and 2019, respectively. We estimated that 30,117 EV71 cases (95% credible interval −1777 to 93,601) were prevented by the vaccination program from 2017 to 2019, corresponding to a reduction of 45.55% (95% credible interval −5.19% to 72.22%). The effectiveness of vaccination increased annually, with reductions of −1.03% (95% credible interval −94.85% to 48.29%) in 2017, 55.54% (95% credible interval 13.94%-77.56%) in 2018, and 82.28% (95% credible interval 65.77%-90.89%) in 2019. Furthermore, we observed that children younger than 4 years obtained greater benefits compared with those older than 4 years, with the greatest reduction of 57.68% (95% credible interval 13.04%-79.65%) in the 3- to 4-year age group, followed by a 48.09% (95% credible interval 28.00%-62.98%) reduction in the 0- to 2-year age group. In contrast, the reduction was markedly lower, at 16.75 % (95% credible interval −96.53% to 65.47%) in children older than 4 years during the 2017 to 2019 period.

Conclusions: EV71 vaccination is an effective measure to prevent HFMD. The sharp decline in EV71-associated HFMD cases following the implementation of the EV71 vaccination program suggests a potential causal relationship. Therefore, strengthening vaccine coverage remains a public health priority.

(JMIR Public Health Surveill 2026;12:e85604) doi:10.2196/85604

KEYWORDS

hand, foot, and mouth disease; enterovirus A71; vaccine efficacy; Chinese population; counterfactual model; real-world study

XSL•FO
RenderX

Introduction

Hand, foot, and mouth disease (HFMD) is a highly contagious viral illness that is endemic to China and poses a significant public health burden, particularly among children younger than 5 years [1]. Surveillance data from the Chinese Center for Disease Control and Prevention identified 7,200,092 HFMD cases from 2008 to 2012, resulting in 2457 deaths [2]. HFMD can be caused by various enteroviruses, such as enterovirus A71 (EV71) and coxsackieviruses, including serotypes A16, A6, A10, B1, B2, and others. Among the various enteroviruses responsible for HFMD, EV71 has been identified as a major pathogen associated with severe and fatal cases [3]. Its neurotropic nature allows it to cross the blood-brain barrier, potentially causing serious neurological complications, such as poliomyelitis-like paralysis, encephalitis, and fatal cardiopulmonary failure in young children [4,5]. Moreover, EV71 is highly contagious and can be transmitted directly via respiratory droplets, saliva, or the fecal-oral route [6], as well as indirectly through fomites on various surfaces such as metal, plastic, and paper [7].

Epidemiological surveillance data have highlighted the substantial disease burden caused by EV71. In 2016, EV71 was estimated to be responsible for approximately 70% of severe HFMD cases and more than 90% of HFMD-associated deaths in China [8,9]. In response to this threat, several inactivated EV71 vaccines were developed by Chinese pharmaceutical companies and were granted regulatory approval in 2015 for use in preventing EV71-related HFMD [10-13]. The national EV71 vaccination program was subsequently launched in 2016, with large-scale promotion and mobilization efforts taking place from 2017 onward [14].

While randomized controlled trials [11-13] have demonstrated the high efficacy and safety of these vaccines in controlled clinical settings [15], it is essential to evaluate their performance under real-world conditions. Postmarketing surveillance and population-level evaluations provide critical evidence to guide public health policy and vaccination strategies. A few recent studies have provided encouraging real-world data. For instance, Xiao et al [16] reported a 41.4% (95% credible interval 37.1%-43.7%) reduction in EV71-associated HFMD cases in Guangdong Province between 2017 and 2019 following the introduction of the vaccine. Similarly, Zheng et al [17] observed a 29% (95% credible interval 24%-34%) decline in EV71-related HFMD cases in Zhejiang Province. Additionally, research by Head et al [18] reported a 60% (95% credible interval 41%-72%) reduction in Chengdu City of Sichuan Province between 2017 and 2018. Recently, a nationwide study further estimated that EV71 vaccination was associated with a reduction of approximately 297,946 EV71-related cases during 2017 to 2018, corresponding to an overall decline of nearly 17% [19]. However, the effectiveness of the EV71 vaccination program varies across regions in China, exhibiting considerable spatial heterogeneity that warrants further investigation.

Despite these findings, there remains a paucity of evidence from other key regions in China. Jiangsu Province, located in eastern China, is one of the most densely populated and economically developed provinces in the country, with high mobility and urbanization that could influence disease transmission dynamics [20]. However, the population-level effectiveness of the EV71 vaccination program in Jiangsu has not been thoroughly investigated.

In this study, we aimed to generate population-level evidence on the real-world effectiveness of the EV71 vaccination program in Jiangsu Province. By leveraging high-quality, longitudinal surveillance data and using an advanced Bayesian structural time series (BSTS) model, we sought to quantify the vaccine's impact on EV71-associated HFMD and provide insights to inform future vaccination strategies and public health decision-making in other endemic regions.

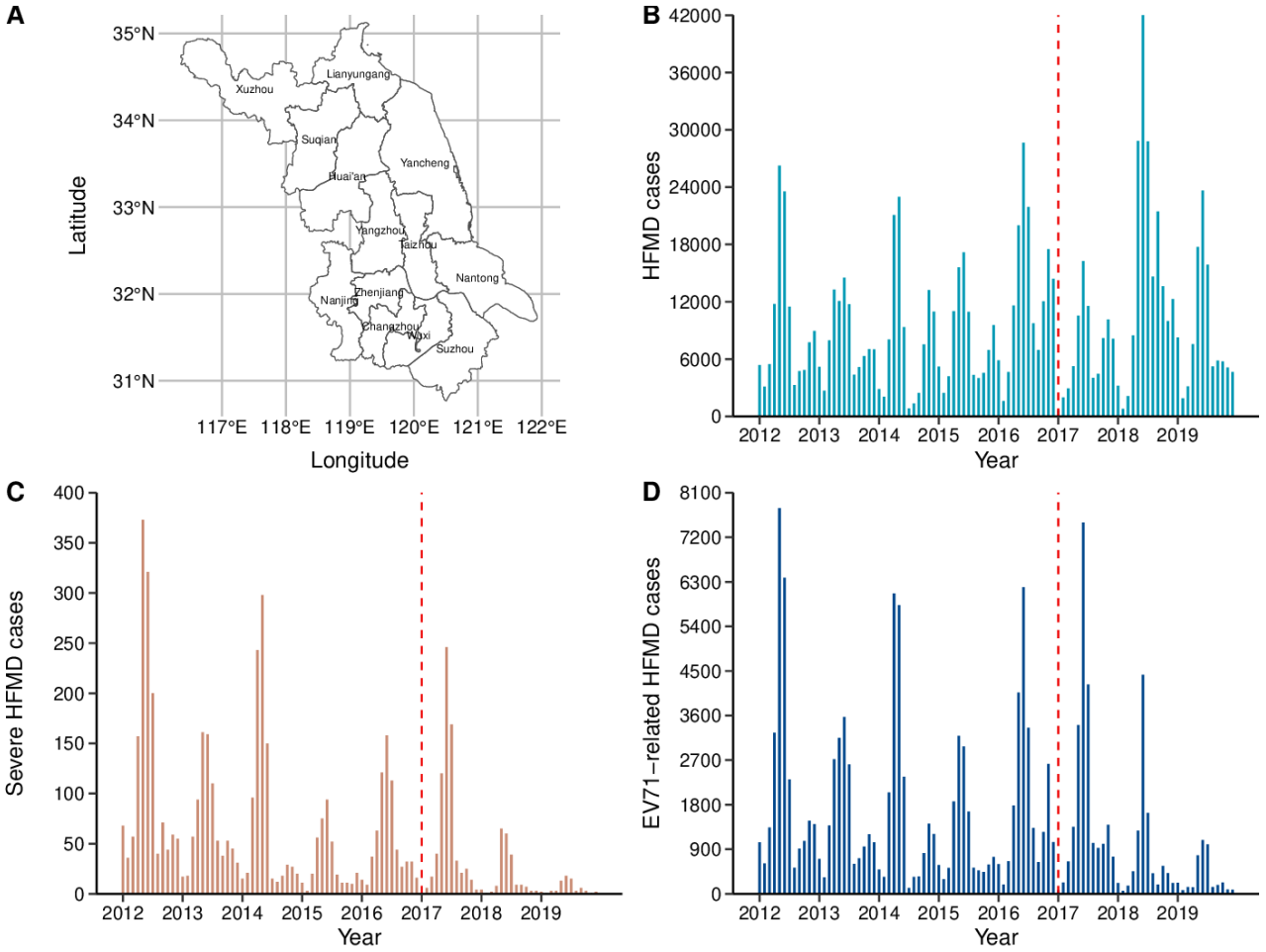
Methods

Study Population

We conducted this longitudinal observational study in Jiangsu, located in eastern China, which comprises 13 cities, such as Nanjing, Suzhou, and Changzhou (Figure 1A). This study followed the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist (Checklist 1). The study population consisted of children aged 0 to 14 years, totaling approximately 11 million in 2017 [21].

XSL•FO
RenderX

Figure 1. The red dotted line denotes the initiation of large-scale promotion and mobilization for vaccination. (A) Study area comprising 13 cities in Jiangsu Province, China; (B) temporal trends in total hand, foot, and mouth disease (HFMD) cases; (C) severe HFMD cases; and (D) enterovirus A71 (EV71)–related HFMD cases.



Data Collection of HFMD Surveillance Data

We included reported cases of HFMD among children aged 0 to 14 years reported between 2012 and 2019. HFMD is a notifiable infectious disease in China. Clinically diagnosed cases are required to be reported through the National Notifiable Diseases Surveillance System, a nationwide direct network reporting platform. Demographic and diagnostic information for each HFMD case is included in the report. Physicians must complete and submit a case report card within 24 hours of diagnosis. The information on the report card is subsequently reviewed by the Centers for Disease Control and Prevention at the county, municipal, and provincial levels.

Etiological data on HFMD cases from 2012 to 2019 were obtained from the HFMD etiological surveillance system in Jiangsu Province. According to the surveillance protocol, each county-level hospital was required to collect specimens from at least 5 first-time medical visits for HFMD each month; if fewer than 5 cases were reported, specimens from all available cases were collected whenever possible. In the event of an HFMD outbreak, a minimum of 5 specimens were required for pathogen testing. Additionally, all severe and fatal cases were mandatorily sampled.

Specimens were tested for EV71, coxsackievirus A16, and other enteroviruses using polymerase chain reaction for nucleic acid detection. Because not all reported HFMD cases underwent etiological testing, we used a statistical method to estimate the monthly number of EV71-related HFMD cases:

EV71 cases = (number of general HFMD cases × EV71-positive rate among general cases) + number of EV71-positive cases among severe HFMD cases.

General cases were defined as fever accompanied by rashes on the hands, feet, mouth, or buttocks (with some cases presenting without fever), while severe cases were defined as those with neurological involvement or respiratory or circulatory dysfunction.

Statistical Analysis

Within a counterfactual framework, we used a BSTS model to evaluate the intervention effects of the EV71 vaccination. The BSTS model is a sophisticated state-space model that integrates Bayesian inference to analyze time-series data [22,23]. It decomposes a time series into interpretable components, such as trend, seasonality, and external covariates, while also providing not only point predictions but also credible intervals, making it a versatile tool in epidemiology.

First, we developed BSTS models based on training sets including EV71-induced cases from 2012 to 2016. Under the counterfactual scenario assuming the absence of the EV71

XSL•FO
RenderX

vaccination program, the model was developed to simulate the monthly number of EV71-related cases. The BSTS model was defined as follows:

$$(1)\log(y_t)=Z_t T\alpha_t+\varepsilon_t$$

$$(2)\alpha_{t+1}=T_t\alpha_t+R_t\eta_t$$

The equation (1) was the observation equation, which related the logarithmically transformed EV71-induced monthly incidence y_t to a vector of latent "state" variables *at*. In our study, the at contained long-term trend (year) and seasonality (12 months each year). The transition equation (2) specifies the dynamics of latent variables as they change over time. Moreover, η_t and ε_t were independently and normally distributed error terms, and Z_t, T_t, and R_t were model matrices mixed with known values (0 or 1) and other unknown parameters [24].

The spike-and-slab prior was applied for the optimal variable selection, and the Kalman filter was used for time-series prediction [24]. We used the Markov chain Monte Carlo algorithm with 10,000 iterations to simulate the posterior distribution of predictions, and then, Bayesian model averaging was used to smooth over a large number of potential results. The model fitting goodness was evaluated by root mean square error, mean absolute percentage error, and R^2 [25].

Subsequently, we used the fitted model to predict the EV71-associated HFMD cases from 2017 to 2019. The relative difference (RD) between predicted and observed cases in the BSTS model was used to estimate the EV71 vaccination effectiveness. The RD was calculated by subtracting the observed cases from the predicted cases and then dividing by the predicted cases. Additionally, we performed age-stratified analyses to evaluate the effectiveness of EV71 vaccination program among children aged 0 to 2, 3 to 4, and 5 to 14 years.

Ethical Considerations

The study received ethical exemption approval from the Ethics Committee of Jiangsu Provincial Center for Disease Control and Prevention (SL2025-B031-01), and the requirement for informed consent was waived by the ethics committee. The HFMD surveillance data used in this study were collected as part of legally mandated infectious disease prevention and control activities. These data were obtained from compliant sources and stored securely, and their secondary use fully adhered to ethical standards. All analyses were performed on deidentified datasets, ensuring that no harm or infringement of rights occurred to data providers. Given that the study is beneficial to HFMD prevention and control and presents minimal risk, it meets the criteria for exemption from ethical review.

Results

HFMD Cases

This analysis included 932,274 cases of HFMD in Jiangsu during the period from 2012 to 2019, which exhibited seasonal and biannual patterns (Figure 1B). Among the reported cases, 5190 (0.56%) were classified as severe (Figure 1C), and an estimated 140,876 (15.11%) cases were attributed to EV71.

Notably, we observed that the proportion of EV71 cases declined sharply after the vaccine implementation, with estimated case counts of 22,303 in 2017, 9796 in 2018, and 3900 in 2019. For instance, the number of EV71 cases peaked at nearly 8000 in May of 2012, whereas in 2019, it was almost below 1100 (Figure 1D).

The Fitting Goodness of BSTS Models

On the basis of the training sets, we constructed the BSTS model and fitted the monthly cases from 2012 to 2016 (Multimedia Appendix 1). In summary, we observed that the performance of the model was accurate and robust, with a root mean square error of 0.33, a mean absolute percentage error of 3.46%, and an R^2 of 78.76%.

The Effectiveness of the EV71 Vaccination Program During 2017 to 2019

Overall, during the EV71 vaccination period (2017 to 2019), the observed number of EV71 cases was significantly lower than the number forecasted by the model (Figure 2 and Table 1). We estimated that 30,117 (95% credible interval −1777 to 93,601) EV71 cases were prevented by EV71 vaccination, corresponding to a relative reduction of 45.55% (95% credible interval −5.19% to 72.22%).

XSL•FO
RenderX

Figure 2. Monthly trends of predicted counterfactual enterovirus A71 (EV71)–related cases (blue) and observed cases (red) in Jiangsu Province during 2017 to 2019. HFMD: hand, foot, and mouth disease.

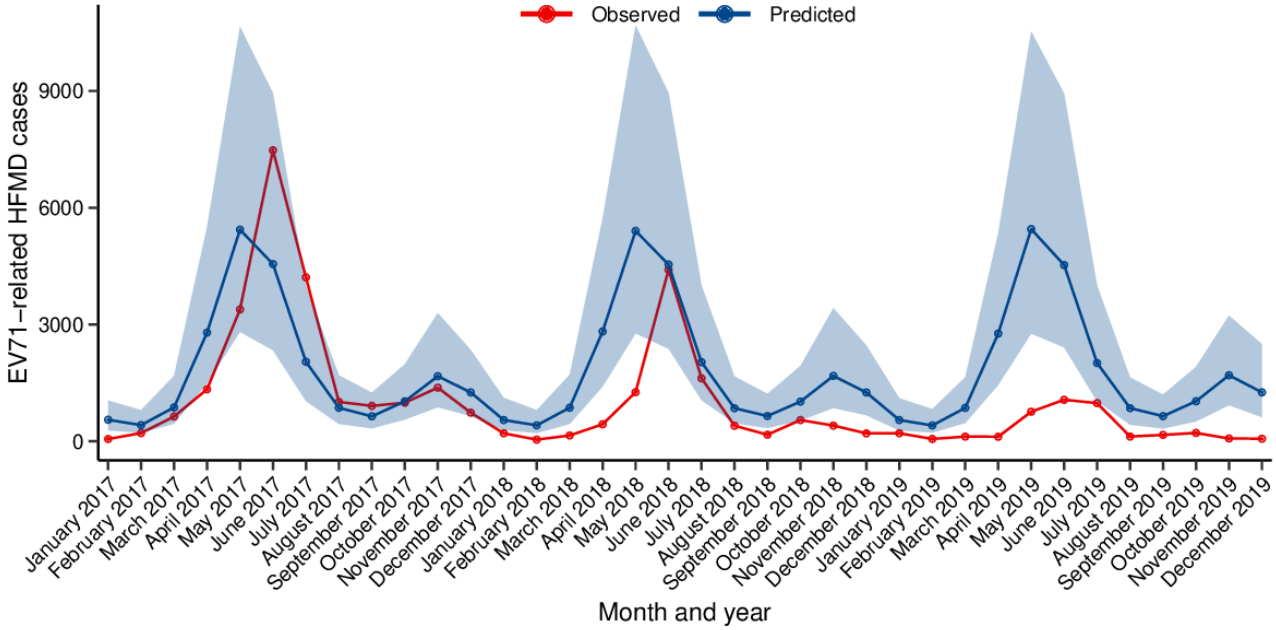


Table . The observed, predicted, and prevented enterovirus A71–induced hand, foot, and mouth disease cases and relative reductions among children in Jiangsu Province, 2017 to 2019.

Year	Observed cases, n	Predicted cases, n (95% credible interval)	Prevented cases[a], n (95% credible interval)	Relative reduction[b] (%; 95% credible interval)
2017 - 2019	35,999	66,116 (34,222 to 129,600)	30,117 (−1777 to 93,601)	45.55 (−5.19 to 72.22)
2017	22,303	22,077 (11,446 to 43,131)	−226 (−10,857 to 20,828)	−1.03 (−94.85 to 48.29)
2018	9796	22,032 (11,382 to 43,651)	12,236 (1586 to 33,855)	55.54 (13.94 to 77.56)
2019	3900	22,008 (11,394 to 42,818)	18,108 (7494 to 38,918)	82.28 (65.77 to 90.89)

[a]Prevented cases were calculated as predicted cases under the counterfactual scenario without vaccination minus observed cases.

[b]Relative reduction values were calculated by dividing the number of prevented cases by the number of predicted cases.

Furthermore, the magnitude of the decline exhibited a progressive increase over years. We observed that −226 (95% credible interval −10,857 to 20,828) EV71-infected cases were prevented in 2017, whereas the number of prevented cases increased to 12,236 (95% credible interval 1586-33,855) in 2018 and further to 18,108 (95% credible interval 7494-38,918) in 2019, with the relative reduction of −1.03% (95% credible interval −94.85% to 48.29%), 55.54% (95% credible interval 13.94%-77.56%), and 82.28% (95% credible interval 65.77%-90.89%) for the respective years.

The Effectiveness of the EV71 Vaccination Program by Age Group

The relative reductions of EV71-related HFMD cases by age group from 2017 to 2019 are shown in Figure 3 and Table 2. Among age subgroups, the RD exhibited an increasing trend over the years, with the highest RD observed in the children aged 3 to 4 years. The relative reduction during 2017 to 2019 was 48.09% (95% credible interval 28.00%-62.98%) for children aged 0 to 2 years, 57.68% (95% credible interval 13.04%-79.65%) for children aged 3 to 4 years, and 16.75% (95% credible interval −96.53% to 65.47%) for children older than 4 years.

XSL•FO
RenderX

Figure 3. Monthly trends of predicted counterfactual enterovirus A71 (EV71)–related cases (blue) and observed cases (red) among different age groups in Jiangsu Province during 2017 to 2019. HFMD: hand, foot, and mouth disease. (A) Children aged 0-2 years; (B) children aged 3-4 years; (C) children older than 4 years.

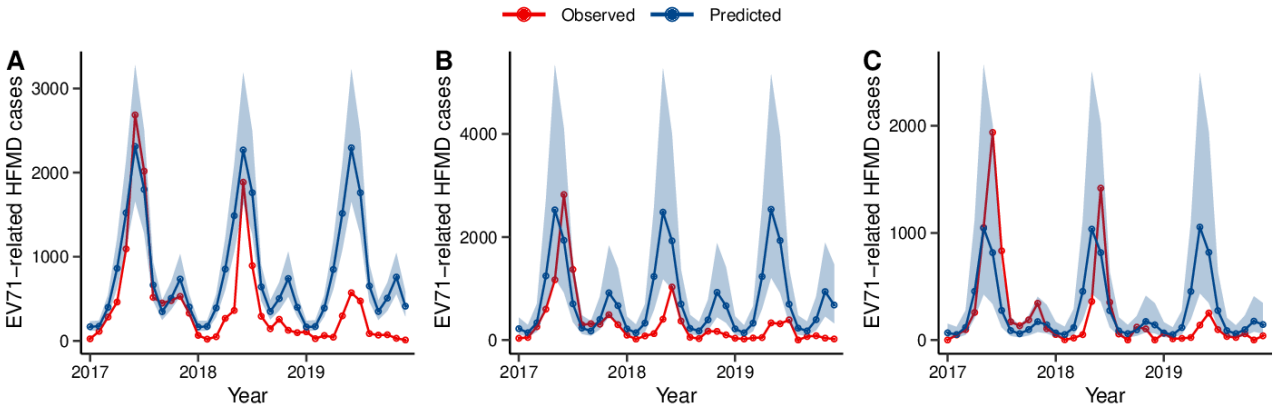


Table . Age-stratified relative reduction of enterovirus A71–induced hand, foot, and mouth disease cases among children in Jiangsu Province, 2017 to 2019.

Age group (years)	Relative reduction in 2017 - 2019, (%; 95% credible interval)	Relative reduction in 2017, (%; 95% credible interval)	Relative reduction in 2018, (%; 95% credible interval)	Relative reduction in 2019 (%; 95% credible interval)
0-2	48.09 (28.00 to 62.98)	9.16 (−25.83 to 35.23)	54.20 (36.43 to 67.48)	81.20 (73.91 to 86.53)
3-4	57.68 (13.04 to 79.65)	15.70% (−72.40 to 59.97)	72.03 (42.94 to 86.55)	85.44 (69.72 to 92.91)
>4	16.75 (−96.53 to 65.47)	−52.51 (−260.67 to 7.33)	24.82 (−76.16 to 69.09)	77.80 (47.28 to 90.62)

Discussion

Principal Findings

In this study, we used an advanced BSTS model to conduct the first population-based assessment of EV71 vaccine effectiveness using real-world surveillance data from Jiangsu Province, China. Our analysis revealed a stable and consistent seasonal fluctuation in EV71 cases from 2012 to 2016. However, a progressive decline in case numbers over subsequent years was observed from 2017 to 2019. On the basis of a counterfactual analysis framework, we estimated that the EV71 vaccination program led to a 45.55% reduction in EV71-associated HFMD cases. Furthermore, we identified age-related differences in EV71 vaccine effectiveness, with varying levels of protection observed across different age groups. These findings provide robust empirical evidence supporting the real-world effectiveness of the EV71 vaccination program in Jiangsu Province and offer valuable data to inform future public health strategies targeting EV71-related HFMD.

Although EV71 cases declined after vaccine introduction in 2017, the overall HFMD cases showed a temporary increase in 2018. This may be explained by intensified surveillance and reporting efforts following vaccine rollout, as well as the circulation of non-EV71 enterovirus serotypes during that year. Similar postvaccination increases in total HFMD notifications have also been reported in other provinces [16,17,26,27]. Given the geographical proximity and demographic similarities between Jiangsu and Zhejiang Provinces, our findings demonstrate remarkable consistency with the results reported by Zheng et al [17] in Zhejiang, further supporting the robustness of our findings.

However, the protective efficacy observed in our study was lower than that reported in Guangdong Province and particularly in Chengdu City of Sichuan Province, where vaccine effectiveness was estimated to be as high as 60% during 2017 to 2018 [18]. These regional differences may be attributable to variations in age distribution and, more importantly, vaccination coverage rates [17]. For instance, Guangdong Province achieved a median coverage rate of 35.8% in 2019 [16], whereas a study conducted in Chengdu City among children aged 6 to 59 months reported a coverage rate of 54.3% in 2018 [18]. In contrast, Jiangsu Province reported markedly lower vaccination coverage rates. In 2019, vaccination coverage was only 10.6% in Nanjing, the provincial capital, and an estimated 12.7% in Lianyungang [28,29], underscoring the suboptimal vaccination uptake across Jiangsu Province.

Our longitudinal analysis also demonstrated a progressive increase in the relative reduction rate of EV71 cases, which may reflect the incremental improvements in vaccine coverage in Jiangsu Province. For example, in Nanjing, the coverage rate rose from 5.7% in 2017 to 12.6% in 2018, before slightly declining to 10.6% in 2019. Additionally, we found that the vaccine effectiveness was highest among children aged 3 to 4 years, compared with those younger than 3 years or older than 4 years, during the 2017 to 2019 period. This age-specific trend aligns with findings from previous studies, yet it failed to attain statistical significance [17,30]. The superior benefit observed in children aged 3 to 4 years is likely attributable to their more mature immune systems compared with younger children, enabling a stronger and more durable vaccine-induced immune response [31]. In contrast, among children older than 4 years, a substantial proportion may have already acquired natural

XSL•FO
RenderX

immunity through prior EV71 infection, thereby reducing the incremental benefit of vaccination. Additionally, differences in vaccination coverage across age groups may also have contributed to the observed pattern.

Several limitations of this study should be acknowledged. First, our findings may be susceptible to unmeasured confounding factors, such as socioeconomic conditions, differences in vaccination coverage across age groups, or concurrent changes in health care policy. However, the consistency of our results with prior studies, suggesting no substantial unmeasured confounding, strengthens confidence in the robustness of these estimates of EV71 vaccination effectiveness. Second, as an ecological study, we evaluated vaccine effectiveness at the population level and lacked access to individual-level data, which would be necessary to account for potential confounding variables and assess individual-level determinants of vaccine response. Third, although estimates based on EV71-positive rates may involve uncertainty due to incomplete testing, the routine surveillance sampling approach supports their validity. Finally, as with other similar evaluations, our study was limited by the relatively short postvaccination observation period. The wide credible intervals indicate substantial uncertainty in the estimated number of cases prevented. Ongoing long-term surveillance will be essential to fully understand the durability and long-term effectiveness of the EV71 vaccination program.

Conclusions

Our study provides a robust population-level evaluation of the effect of the EV71 vaccination program in Jiangsu Province during the period of 2017 to 2019. We observed a substantial reduction in EV71-associated HFMD cases following the implementation of the vaccination policy, with the magnitude of the decline increasing progressively in line with improvements in vaccine coverage. Additionally, vaccine effectiveness was notably higher among children aged 3 to 4 years compared with other age groups. These findings provide strong real-world evidence supporting the utility of the EV71 vaccination program in reducing disease burden and offer a scientific basis for optimizing population-level intervention strategies to control EV71-related HFMD.

Acknowledgments

The authors are grateful to all staff involved in the enterovirus surveillance and monitoring program for their valuable contributions and dedicated work.

Funding

This work was supported by the National Key Research and Development Program of China (2024YFC2311302 and 2024YFC2310403).

Data Availability

The datasets generated or analyzed during this study are not publicly available due to privacy restrictions but are available from the corresponding author on reasonable request and with permission from the Jiangsu Provincial Center for Disease Control and Prevention.

Authors' Contributions

LZ, HJ, and XH serve as co-corresponding authors and contributed equally to study supervision, coordination, and intellectual input into the interpretation and contextualization of the findings. YT was responsible for data collection and drafted the manuscript. HF was responsible for laboratory detection and quality control. JW, XZ, XC, and CB contributed to data interpretation and provided methodological support.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Fitting performance of the Bayesian structural time series model for monthly cases from 2012 to 2016.
[PNG File, 12 KB - publichealth_v12i1e85604_app1.png]

Checklist 1
STROBE checklist for observational studies.
[PDF File, 288 KB - publichealth_v12i1e85604_app2.pdf]

References

1. Ji H, Fan H, Lu PX, et al. Surveillance for severe hand, foot, and mouth disease from 2009 to 2015 in Jiangsu province: epidemiology, etiology, and disease burden. BMC Infect Dis 2019 Jan 22;19(1):79. [doi: 10.1186/s12879-018-3659-7] [Medline: 30669973]

XSL•FO
RenderX

2. Nayak G, Bhuyan SK, Bhuyan R, Sahu A, Kar D, Kuanar A. Global emergence of enterovirus 71: a systematic review. Beni Suef Univ J Basic Appl Sci 2022;11(1):78. [doi: 10.1186/s43088-022-00258-4] [Medline: 35730010]
3. Jin Y, Zhang R, Wu W, Duan G. Antiviral and inflammatory cellular signaling associated with enterovirus 71 infection. Viruses 2018 Mar 28;10(4):155. [doi: 10.3390/v10040155] [Medline: 29597291]
4. Mao Q, Wang Y, Bian L, Xu M, Liang Z. EV-A71 vaccine licensure: a first step for multivalent enterovirus vaccine to control HFMD and other severe diseases. Emerg Microbes Infect 2016 Jul 20;5(7):e75. [doi: 10.1038/emi.2016.73] [Medline: 27436364]
5. Chan LG, Parashar UD, Lye MS, et al. Deaths of children during an outbreak of hand, foot, and mouth disease in Sarawak, Malaysia: clinical and pathological characteristics of the disease. For the Outbreak Study Group. Clin Infect Dis 2000 Sep;31(3):678-683. [doi: 10.1086/314032] [Medline: 11017815]
6. Omaña-Cepeda C, Martínez-Valverde A, del Mar Sabater-Recolons M, Jané-Salas E, Marí-Roig A, López-López J. A literature review and case report of hand, foot and mouth disease in an immunocompetent adult. BMC Res Notes 2016 Mar 15;9:165. [doi: 10.1186/s13104-016-1973-y] [Medline: 26975350]
7. Solomon T, Lewthwaite P, Perera D, Cardosa MJ, McMinn P, Ooi MH. Virology, epidemiology, pathogenesis, and control of enterovirus 71. Lancet Infect Dis 2010 Nov;10(11):778-790. [doi: 10.1016/S1473-3099(10)70194-8] [Medline: 20961813]
8. Ji T, Han T, Tan X, et al. Surveillance, epidemiology, and pathogen spectrum of hand, foot, and mouth disease in mainland of China from 2008 to 2017. Biosaf Health 2019 Jun;1(1):32-40. [doi: 10.1016/j.bsheal.2019.02.005]
9. Xing W, Liao Q, Viboud C, et al. Hand, foot, and mouth disease in China, 2008-12: an epidemiological study. Lancet Infect Dis 2014 Apr;14(4):308-318. [doi: 10.1016/S1473-3099(13)70342-6] [Medline: 24485991]
10. Li R, Liu L, Mo Z, et al. An inactivated enterovirus 71 vaccine in healthy children. N Engl J Med 2014 Feb 27;370(9):829-837. [doi: 10.1056/NEJMoa1303224] [Medline: 24571755]
11. Nguyen TT, Chiu CH, Lin CY, et al. Efficacy, safety, and immunogenicity of an inactivated, adjuvanted enterovirus 71 vaccine in infants and children: a multiregion, double-blind, randomised, placebo-controlled, phase 3 trial. Lancet 2022 Apr 30;399(10336):1708-1717. [doi: 10.1016/S0140-6736(22)00313-0] [Medline: 35427481]
12. Zhu FC, Meng FY, Li JX, et al. Efficacy, safety, and immunology of an inactivated alum-adjuvant enterovirus 71 vaccine in children in China: a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial. Lancet 2013 Jun 8;381(9882):2024-2032. [doi: 10.1016/S0140-6736(13)61049-1] [Medline: 23726161]
13. Zhu F, Xu W, Xia J, et al. Efficacy, safety, and immunogenicity of an enterovirus 71 vaccine in China. N Engl J Med 2014 Feb 27;370(9):818-828. [doi: 10.1056/NEJMoa1304923] [Medline: 24571754]
14. Li ZQ, Qin ZQ, Tan HF, et al. Analysis of the coverage of inactivated enterovirus 71 (EV71) vaccine and adverse events following immunization with the EV71 vaccine among children from 2016 to 2019 in Guangzhou. Expert Rev Vaccines 2021 Jul;20(7):907-918. [doi: 10.1080/14760584.2021.1933451] [Medline: 34036862]
15. Chong P, Liu CC, Chow YH, Chou AH, Klein M. Review of enterovirus 71 vaccines. Clin Infect Dis 2015 Mar 1;60(5):797-803. [doi: 10.1093/cid/ciu852] [Medline: 25352588]
16. Xiao J, Zhu Q, Yang F, et al. The impact of enterovirus A71 vaccination program on hand, foot, and mouth disease in Guangdong, China: a longitudinal surveillance study. J Infect 2022 Oct;85(4):428-435. [doi: 10.1016/j.jinf.2022.06.020] [Medline: 35768049]
17. Zheng D, Shen L, Wen W, et al. The impact of EV71 vaccination program on hand, foot and mouth disease in Zhejiang Province, China: a negative control study. Infect Dis Model 2023 Dec;8(4):1088-1096. [doi: 10.1016/j.idm.2023.09.001] [Medline: 37745754]
18. Head JR, Collender PA, Lewnard JA, et al. Early evidence of inactivated enterovirus 71 vaccine impact against hand, foot, and mouth disease in a major center of ongoing transmission in China, 2011-2018: a longitudinal surveillance study. Clin Infect Dis 2020 Dec 15;71(12):3088-3095. [doi: 10.1093/cid/ciz1188] [Medline: 31879754]
19. Liu Y, Song Y, Liu F, et al. Effectiveness of the enterovirus A71 vaccine on hand, foot, and mouth disease: a real-world study in China. Clin Microbiol Infect 2025 Feb;31(2):258-265. [doi: 10.1016/j.cmi.2024.09.020] [Medline: 39343096]
20. Lessani MN, Li Z, Jing F, et al. Human mobility and the infectious disease transmission: a systematic review. Geo Spat Inf Sci 2024;27(6):1824-1851. [doi: 10.1080/10095020.2023.2275619] [Medline: 40046953]
21. National Bureau of Statistics of China. China Statistical Yearbook: China Statistics Press; 2017.
22. Kalinina I, Bidyuk P, Gozhyj A. Construction of forecast models based on Bayesian structural time series. Presented at: 2022 IEEE 17th International Conference on Computer Sciences and Information Technologies (CSIT); Nov 10-12, 2022. [doi: 10.1109/CSIT56902.2022.10000484]
23. Zhang R, Mi H, He T, et al. Trends and multi-model prediction of hepatitis B incidence in Xiamen. Infect Dis Model 2024 Dec;9(4):1276-1288. [doi: 10.1016/j.idm.2024.08.001] [Medline: 39224908]
24. Almarashi AM, Khan K. Bayesian structural time series. Nanosci Nanotechnol Lett 2020 Jan;12(1):54-61. [doi: 10.1166/nnl.2020.3083]
25. Hodson TO. Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. Geosci Model Dev 2022;15(14):5481-5487. [doi: 10.5194/gmd-15-5481-2022]

XSL•FO
RenderX

26. Han Y, Chen Z, Zheng K, et al. Epidemiology of hand, foot, and mouth disease before and after the introduction of enterovirus 71 vaccines in Chengdu, China, 2009-2018. Pediatr Infect Dis J 2020 Oct;39(10):969-978. [doi: 10.1097/INF.0000000000002745] [Medline: 32433221]
27. Wang J, Zhang S. Epidemiological characteristics and trends of hand-foot-mouth disease in Shanghai, China from 2011 to 2021. Front Public Health 2023 May 30;11:1162209. [doi: 10.3389/fpubh.2023.1162209] [Medline: 37325298]
28. Liang Y, Xu L, Ma M. Analysis on the coverage of enterovirus 71 (EV71) vaccine from 2017 to 2019 in Nanjing City, Jiangsu Province. J Med Health 2021(1):282-284.
29. Xie B, Pan Y, Chen Y, Xu L. Analysis of EV71 vaccination status in Lianyungang City: a study from 2012-2021 birth cohort. South China J Prev Med 2023;49(8):981-985.
30. Cui Y, Yang YN, Zheng RR, et al. Epidemiological characteristics of hand, foot, and mouth disease clusters during 2016-2020 in Beijing, China. J Med Virol 2022 Oct;94(10):4934-4943. [doi: 10.1002/jmv.27906] [Medline: 35655366]
31. Kloc M, Ghobrial RM, Kuchar E, Lewicki S, Kubiak JZ. Development of child immunity in the context of COVID-19 pandemic. Clin Immunol 2020 Aug;217:108510. [doi: 10.1016/j.clim.2020.108510] [Medline: 32544611]

Abbreviations

BSTS: Bayesian structural time series
EV71: enterovirus A71
HFMD: hand, foot, and mouth disease
RD: relative difference
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Mavragani, T Sanchez; submitted 12.Oct.2025; peer-reviewed by H Qi, S Zhu; revised version received 10.Dec.2025; accepted 31.Dec.2025; published 10.Mar.2026.

Please cite as:
Tong Y, Fan H, Wang J, Zeng X, Cheng X, Bao C, Zhu L, Ji H, Huo X
Population-Level Impact of the Enterovirus A71 Vaccination Program on Hand, Foot, and Mouth Disease: Ecological Time-Series Study
JMIR Public Health Surveill 2026;12:e85604
URL: https://publichealth.jmir.org/2026/1/e85604
doi:10.2196/85604

© Ye Tong, Huan Fan, Jing Wang, Xinru Zeng, Xiaoqing Cheng, Changjun Bao, Liguo Zhu, Hong Ji, Xiang Huo. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 10.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Impact of US Food and Drug Administration's Advisory and Enforcement Actions on US Sales of Esco Bars Products in 2023: Synthetic Control Approach

Hyungsik Shin, PhD; Jonathan Hannings, MA; Zachary Cahn, PhD; Xin Xu, PhD; Robin Toblin, PhD; Matthew Farrelly, PhD; Brian A King, PhD

Center for Tobacco Products, US Food and Drug Administration, 10903 New Hampshire Avenue, Silver Spring, MD, United States

Corresponding Author:
Hyungsik Shin, PhD
Center for Tobacco Products, US Food and Drug Administration, 10903 New Hampshire Avenue, Silver Spring, MD, United States

Abstract

Background: In 2023, Esco Bars was the second most commonly reported e-cigarette brand used among US middle and high school student e-cigarette users. These products have not been authorized for sale in the United States by the US Food and Drug Administration (FDA). On May 12, 2025, and May 25, 2023, the FDA issued an import alert and a warning letter, respectively, to the manufacturer requiring them to immediately remove these unauthorized products from the market. On June 22, 2023, and July 27, 2023, the FDA also issued warning letters to US retailers and distributors, respectively, regarding the illicit sale of these unauthorized products. This study evaluated the impact of these advisory and enforcement actions on retail sales of Esco Bars in the United States.

Objective: We evaluated the impact of these 2023 FDA advisory and enforcement actions on illicit sales of unauthorized Esco Bars to better inform future regulatory decision-making.

Methods: This study used a synthetic control method to evaluate the impact of 2023 FDA advisory and enforcement actions on Esco Bars sales based on weekly sales data from the NielsenIQ Retail Measurement Service between January 1, 2023, and December 30, 2023. Data came from the 48 contiguous states and were not available for internet and vape shop sales. First, we standardized the data by transforming the weekly sales into *z* scores. Second, we created a synthetic comparison for Esco Bars sales based on sales of other e-cigarettes for which the FDA did not take action during the analytical period. Finally, we calculated the difference in sales between Esco Bars and the comparison group before and after the FDA actions. To assess the robustness of the findings, we used other standardization methods for sensitivity analyses.

Results: By the last week of 2023, actual Esco Bars' weekly sales were 68.5% lower than the estimated sales in the comparison group (P=.02). Over the 5-month period following FDA advisory and enforcement actions, Esco Bars product sales were reduced by approximately 1.7 million equivalized units (P=.06). These findings were robust across sensitivity analyses.

Conclusions: The FDA's advisory and enforcement actions substantially reduced sales of Esco Bars products in the United States, with a sustained impact noted over a 5-month period. These findings underscore the importance and impact of FDA actions against unauthorized tobacco products as part of a comprehensive regulatory approach.

(JMIR Public Health Surveill 2026;12:e81033) doi:10.2196/81033

KEYWORDS

unauthorized e-cigarettes; illicit sales of e-cigarettes; US Food and Drug Administration; FDA; FDA advisory and enforcement actions; synthetic control method; SCM; difference-in-differences; DiD

Introduction

E-cigarettes have been the most commonly used tobacco product among youth in the United States since 2014, surpassing conventional cigarettes [1]. E-cigarettes can generally have a lower risk than conventional cigarettes and have the potential to benefit adult smokers who transition completely [2]. However, youth uptake and use of these products is problematic, including because e-cigarettes contain nicotine, which is a highly addictive chemical compound. Nicotine exposure can be associated with multiple adverse health consequences, including harming adolescent brain development and priming the brain for addiction to other drugs [3-5].

Youth e-cigarette use prevalence increased rapidly in the early 2010s and peaked in 2019 at nearly 28%, largely driven by flavored cartridge-based systems [6]. Since then, the market landscape has continued to evolve, and flavored disposable

XSL•FO
RenderX

products have become the most popular type of e-cigarettes among youth [7]. According to the 2023 National Youth Tobacco Survey, approximately 60.7% of youth reporting current e-cigarette use used disposable e-cigarettes, and Esco Bars was the second most commonly reported brand used among youth who currently use e-cigarettes [7].

In the United States, tobacco products not on the market prior to 2007, including e-cigarettes, are required to submit a premarket tobacco product application (PMTA) and receive authorization by the US Food and Drug Administration (FDA) before the product can be marketed. The manufacturer of Esco Bars submitted a PMTA. However, in 2022, the FDA issued a refuse to accept (RTA) letter to the manufacturer of Esco Bars, finding that its submitted premarket tobacco product application did not provide sufficient information to permit substantive review per regulatory requirements. Receipt of an RTA letter indicates that a manufacturer does not have authorization to legally market its products in the United States. Nevertheless, the manufacturer continued to produce and sell these products.

As a result of common youth use of Esco Bars, and following the manufacturer's lack of compliance with regulatory review decisions, the FDA took targeted advisory and enforcement actions as a part of a comprehensive approach to reduce youth use of tobacco products [8]. The FDA can issue untitled letters and warning letters as advisory actions to encourage voluntary compliance when responding to illicit tobacco product marketing. Warning letters provide instructions to a company (ie, manufacturer, distributor, or retailer) in violation of FDA regulations, including directions for prompt action to address the violations and a timeline for a response [9]. FDA may also pursue enforcement measures including import alerts, civil money penalties, seizures, and injunctions [10,11]. Import alerts signal that violative products may be detained upon entry into the country [12].

To underscore the FDA's commitment to protecting youth against unauthorized e-cigarettes, the FDA issued an import alert for Esco Bars products on May 12, 2023 [13]. Additionally, the FDA issued a warning letter to the manufacturer of Esco Bars products on May 25, 2023, as the firm had been manufacturing, distributing, and/or importing these unauthorized tobacco products in the United States despite prohibitions stated in the RTA letter [14]. Following these actions, FDA further issued warning letters to retailers and distributors of e-cigarette products on June 22, 2023, and July 27, 2023, respectively [15,16], for selling and/or distributing unauthorized Esco Bars products.

Existing studies have analyzed FDA warning letters on e-cigarettes issued to retailers and manufacturers between 2018 and 2021 by violation type and company characteristic [17,18], including one study that also assessed compliance among online retailers [17]. However, none of these studies have examined the impact of these advisory actions on e-cigarette sales. To our knowledge, this is the first study evaluating the impact of FDA advisory and enforcement actions on illicit tobacco product sales following noncompliance with regulatory review decisions. To address this gap in the scientific literature, this study explores the association between the FDA's targeted advisory and enforcement actions in 2023 and subsequent changes in Esco Bars product sales by the end of 2023.

Methods

Overview

Data

The FDA's Center for Tobacco Products licenses NielsenIQ Retail Measurement Service (RMS) data, which includes weekly tobacco product sales from 48 contiguous states (not including Hawaii and Alaska). NielsenIQ RMS data come from several sources, including electronic point-of-sale data from stores through product bar code checkout scanners at retailer registers, coding of retail circulars (eg, in-store flyers and advertisements promoting products), and in-store audits (ie, field auditors who capture in-store display information promoting products). NielsenIQ uses proprietary statistical methods to project nationally representative sales of nicotine and tobacco products by universal product code.

Unit Sales

The FDA's Center for Tobacco Products receives NielsenIQ tobacco product sales data in both raw and equivalized units. The equivalized unit was created by NielsenIQ to provide a standardized volume measure for products with varying package sizes. Equivalized unit can vary by tobacco product type. For e-cigarettes, 1 equivalized unit is equal to 1 disposable device, prefilled pod, starter kit, or e-liquid bottle. For this study, we used weekly e-cigarette equivalized unit sales from January 1, 2023, to December 30, 2023.

Retail Channel

Weekly e-cigarette equivalized sales were assessed by combining two primary NielsenIQ retail channels: (1) convenience stores, which encompass smaller stores with a limited selection of grocery products, including chain, franchise, and independent convenience stores; and (2) Extended All Outlets Combined, which is a combination of food and drug stores plus select mass merchandiser (eg, Target), club store (eg, Sam's Club), military commissary (eg, Army & Air Force Exchange Service), and dollar store (eg, Dollar General) accounts [19].

Sample

NielsenIQ reports e-cigarette brands as listed on external product packages. We identified 4 Esco Bars-related sub-brands in NielsenIQ RMS data in 2023, including Esco Bars, Esco Bars Carsonator, Esco Bars Fruitia, and Esco Bars Ripe Collection. According to the data, these 4 sub-brands are produced by the same manufacturer and share the major brand, Esco Bars, in their names. Therefore, we reported all 4 sub-brands as Esco Bars products for the purposes of this study. Following the same approach, we combined other sub-brands that are produced by the same manufacturers and share similar major brand names. Details about these sub- and major brands are available upon request. We included only brands that provided full sales data for the entire study period. In addition, we excluded e-cigarette products that do not contain nicotine, that is, accessories (eg, e-cigarette battery, coils, and carrying cases), as they are not

XSL•FO
RenderX

applicable comparisons for Esco Bars for this study. These excluded products represented very few e-cigarette equivalized unit sales (<0.1%) during the analytical timeframe.

Study Design

We used the synthetic control method (SCM) to draw inferences about the likelihood that the intervention had a causal effect on Esco Bars' sales. This is similar to a conventional difference-in-difference (DiD) approach, as both methods compare changes over time between a group exposed to the intervention and an unexposed comparison group [20]. Although DiD models have been widely used to assess potential effects of policy interventions, their validity critically relies on the parallel trends assumption (ie, that the treatment and control groups would have changed in the same way over time if the intervention had not occurred, and thus changes during the postintervention period are due to the treatment itself) [21]. DiD was not suitable for this study because all potential comparison e-cigarette brands failed parallel trends tests. Therefore, we used the SCM, which provides a systematic way of creating an appropriate comparison when the parallel trends assumption cannot be met [22].

In this study, the intervention group was Esco Bars, which was subject to FDA advisory and enforcement actions in 2023. The intervention period was between May 14, 2023, and July 29, 2023. Although the FDA first issued an import alert for Esco Bars products on May 12, 2023, we used May 14, 2023, as the first date of intervention because NielsenIQ weekly sales are tracked from Sunday through Saturday. May 12, 2023, was a Friday falling in the week ending May 13, 2023. For the same reason, we chose July 29, 2023, as the last day of the intervention because July 27, 2023, (the date when the FDA took the last advisory action, issuing warning letters to distributors, against Esco Bars product sales in 2023) was a Thursday falling in the week ending on July 29, 2023. We assumed that adjusting the intervention period by one or two days based on NielsenIQ weekly sales would not meaningfully affect our findings.

To be conservative and fully evaluate the combined impact of multiple FDA advisory and enforcement actions together, we defined the postintervention period as 5 months between July 30 and December 30, 2023, and reported our major findings accordingly. In Figure S1 in Multimedia Appendix 1, we present additional estimates, which include the impact of advisory and enforcement actions during the intervention period, starting from May 14, 2023.

The synthetic comparison of Esco Bars was generated from a set of e-cigarettes that did not encounter any FDA advisory and enforcement actions, including import alerts, warning letters, refuse to file letters, RTAs, or marketing denial orders during the analytical period.

Following these criteria, we identified a total of 51 e-cigarette brands for the "donor pool," that is, all brands that were eligible to be part of the comparison, from the 2023 NielsenIQ RMS data. Among these 51 brands, weights for the synthetic comparison were optimized based on their closeness of fit to the actual Esco Bars' sales during the preintervention period, that is, from January 1, 2023, to May 13, 2023. In the main model, the synthetic comparison consisted of a weighted average of 10 selected brands. Details about these brands and their weights are reported in Table S1 in Multimedia Appendix 1.

A proper execution of the SCM requires that variables in the data are rescaled to correct for differences in size between units [23]. Weekly unit sales varied substantially by brand in the final sample, ranging from less than 1000 equivalized units to a little more than 5 million equivalized units each week. To address the issue, following existing studies [24,25], we rescaled weekly equivalized unit sales using a *z* score standardization method. For a given week, we subtracted each brand's preintervention averages from its weekly sales and divided the result by its pre-intervention standard deviation. In this case, a 1-unit change in standardized sales for a given brand is equivalent to a change of 1 preintervention standard deviation. To assess the sensitivity of our findings to the choice of standardization methods, we conducted sensitivity analyses using two alternatives: (1) min-max normalization—subtracting the preintervention minimum and dividing by the preintervention range, and (2) robust rescaling—subtracting the preintervention median and dividing by the preintervention interquartile range. A more detailed methodological approach for each standardization method is provided in Multimedia Appendix 1.

Statistical Analyses

The application of conventional statistical approaches to assess significance is not available for the SCM approach [22,26]. Therefore, we estimated the *P* value using the concept of the Fisher exact test [27]. First, we calculated differences in sales between Esco Bars and its synthetic comparison during the postintervention period. This difference was calculated as the postintervention mean squared error (MSE). Then we calculated the ratio of the postintervention MSE to the preintervention MSE to determine the relative magnitude of the intervention effect. Second, we constructed synthetic estimates for every comparison brand and calculated intervention effects for each brand. These differences for the comparison brands are false intervention effects because these brands should not have been affected by the intervention (ie, FDA advisory and enforcement actions against illicit sales of Esco Bars). Then, we compared the intervention effect for Esco Bars to the false intervention effects of all 51 brands. Finally, we obtained a *P* value as the share of false intervention effects greater than or equal to the actual intervention effect [28]. If the ratio is sufficiently small, we can reasonably conclude that the intervention effect is unlikely to be a chance result. Technical details of the SCM and Fisher exact test are presented in Multimedia Appendix 1.

All data analyses were independently conducted using Stata version 16.1 (StataCorp LLC), R version 4.3.1 (R Foundation for Statistical Computing), and Python version 3.11.5 (Python Software Foundation) together with Anaconda version 23.10.0 (Anaconda, Inc).

Ethical Considerations

We used deidentified aggregated retail sales data licensed from NielsenIQ for this study. All sales data were fully anonymized. Therefore, this study did not involve human participants, human

XSL•FO
RenderX

tissue, or identifiable personal data, and ethical approval was not required in accordance with institutional and national guidelines.

Results

Findings

Of the 52 brands in our final sample, including Esco Bars, 67.3% (35/52) of brands represented only disposable e-cigarettes, 1.9% (1/52) represented only prefilled cartridges, and 30.8% (16/52) represented multiple types of e-cigarette products. The average weekly sales were estimated at approximately 114,000 equivalized units per brand (SD 662,000), with the median at 1400 units per week (IQR 10,000). The large difference between the mean and median reflects the disproportionate share of sales captured by the most popular brands in our sample, providing additional supporting evidence for using standardized sales prior to implementing the SCM.

Figure 1 presents weekly standardized unit sales of Esco Bars and the synthetic comparison in 2023. Unit sales of Esco Bars were very close to that of the comparison, 0.03% lower on average, during the preintervention period. The high similarity provides strong evidence for the parallel trends assumption to support the validity of the SCM model. Our findings also show that Esco Bars sales started to decline the week ending May 21, 2023, immediately after the FDA advisory and enforcement actions, while estimated sales of the synthetic comparison remained largely steady compared to the preintervention period, following a similar moderate downward trend throughout 2023. As a result, during the 5-month postintervention period, between the week ending on July 29, 2023, and the week ending on December 30, 2023, total Esco Bars unit sales were 1.7 million units lower than the synthetic comparison, which represented estimated sales in the absence of FDA intervention. Weekly sales of Esco Bars were 68.5% lower than that of the synthetic comparison by the week ending December 30, 2023.

Figure 1. Weekly *z* score standardized unit sales of Esco Bars and the synthetic comparison before and after the US Food and Drug Administration (FDA)'s advisory and enforcement actions. Comparison of weekly observed (actual) versus modeled (synthetic) Esco Bars sales in 2023 in *z* score standardized units [2,3]. The red area represents the difference between the standardized actual and synthetic Esco Bars sales during the intervention period, while the blue area represents the same difference for the postintervention period. Week 1 and week 52 correspond to the week ending January 7, 2023, and the week ending December 30, 2023, respectively. Dashed vertical lines denote the first and last weeks during which the specified FDA advisory and enforcement actions occurred: Week 20 corresponds to the week ending on May 20, 2023, denoting the first week after the FDA's import alert for Esco Bars was issued on May 12, 2023. Week 30 corresponds to the week ending on July 29, 2023, denoting the week of the FDA's last advisory action against unauthorized Esco Bars in 2023, warning letters issued to distributors on July 27, 2023.

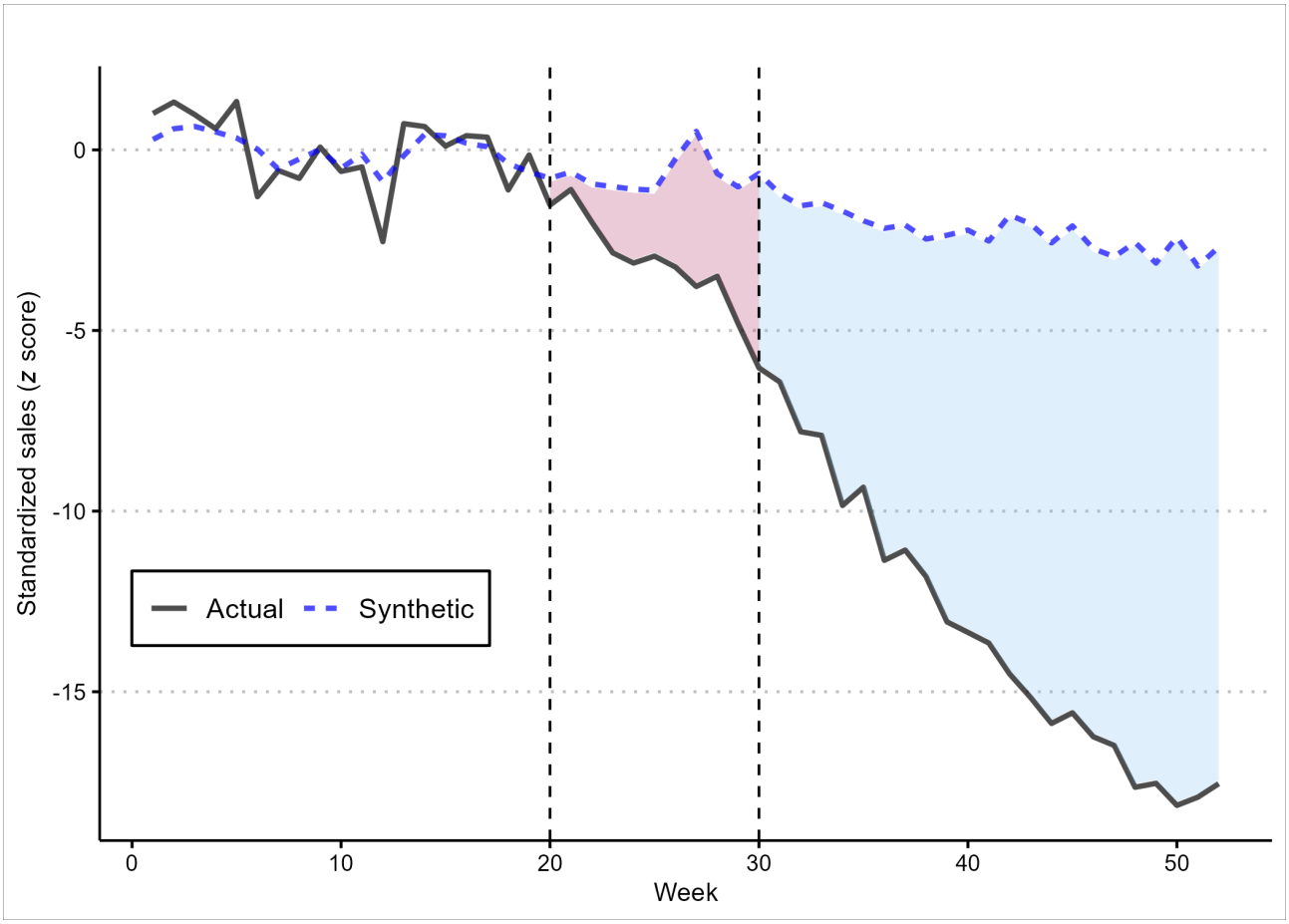
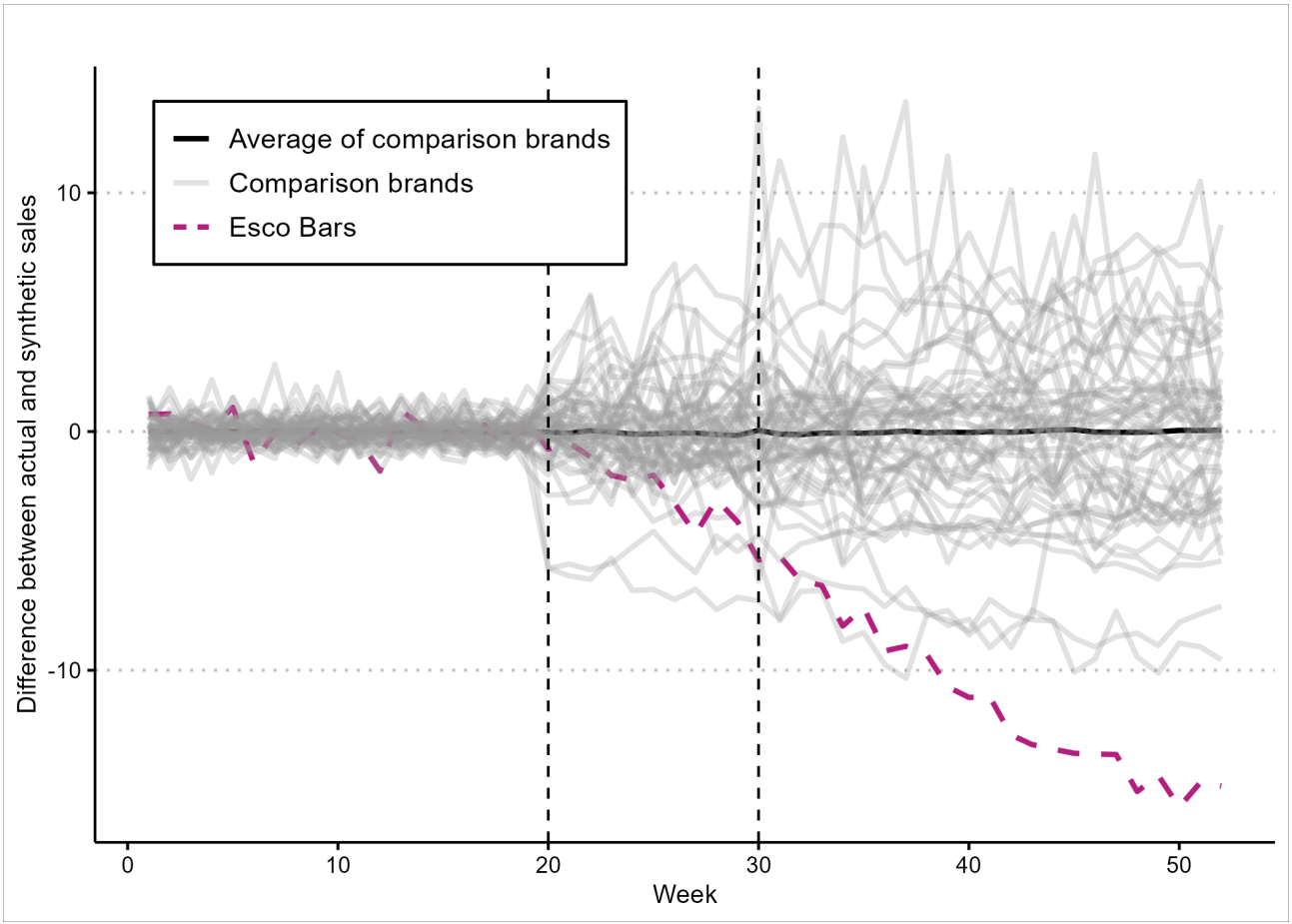


Figure 2 presents the intervention effect of Esco Bars and false intervention effects of all 51 e-cigarette brands in the "donor pool." In all cases, the effects illustrate the difference between weekly sales of a brand and estimated sales from that brand's synthetic comparison. Figures 1 and 2 show that Esco Bars sales started to decline immediately following FDA's first enforcement action, when the import alert for Esco Bars was issued on May 12, 2023. In addition, it also suggests that,

XSL•FO
RenderX

compared to all other brands included in the analysis, Esco Bars had the largest sales reduction during the postintervention period (a decline of 14.8 SDs vs 9.5 SDs from Avata, the brand having the second largest sales decline). The findings in Figure 2 further reinforce the causal connection between FDA advisory and enforcement actions subsequent to noncompliance with regulatory review decisions and Esco Bars sales reductions.

Figure 2. Estimated intervention effect: differences between actual weekly sales for an e-cigarette brand and corresponding synthetic sales. Gray and purple lines represent differences between weekly observed sales for a brand and modeled sales estimated using a corresponding synthetic comparison. Week 1 and week 52 correspond to the week ending January 7, 2023, and the week ending December 30, 2023, respectively. Dashed vertical lines denote the first and last weeks during which the specified US Food and Drug Administration (FDA) advisory and enforcement actions occurred. Week 20 corresponds to the week ending on May 20, 2023, denoting the first week after FDA's import alert for Esco Bars was issued on May 12, 2023. Week 30 corresponds to the week ending on July 29, 2023, denoting the week of FDA's last advisory action against unauthorized Esco Bars in 2023, warning letters were issued to distributors on July 27, 2023.



We estimated *P* values based on the concept of the Fisher exact test, which depends on the ranking of the ratio of sales changes between the post- and preintervention periods. Esco Bars were ranked third out of the 52 brands in the analysis and thus had a *P* value of .06 (3/52, 5.8%). Two other brands, Juice Head and Vodo Bar, were ranked higher than Esco Bars based on the Fisher exact tests. We further investigated these 2 brands with higher significance (Figure 3). Among all brands included in this analysis, Esco Bars had the largest postintervention MSE by a substantive margin, indicating that in the postintervention period, Esco Bars sales diverged from its synthetic comparison more than any other e-cigarette brand. However, the MSE ratios for Juice Head and Vodo Bar were ranked higher because they had the lowest MSEs in the preintervention period, indicating their sales diverged from their synthetic comparisons less than any other brand in the period. This context suggests that the higher MSE ratios of Juice Head and Vodo Bar do not preclude our finding that the FDA advisory and enforcement actions had significant impacts on Esco Bars sales during the postintervention period.

XSL•FO
RenderX

Figure 3. Investigation of esco bars ranking among all e-cigarette brands. Postintervention mean-squared error (MSE) and preintervention MSE were presented for each e-cigarette brand included in this study. MSE measures how close each brand's actual sales were to modeled sales of the corresponding synthetic comparison. Labels reflect the 5 e-cigarette brands from our sample with the highest ratios of postintervention MSE to preintervention MSE.

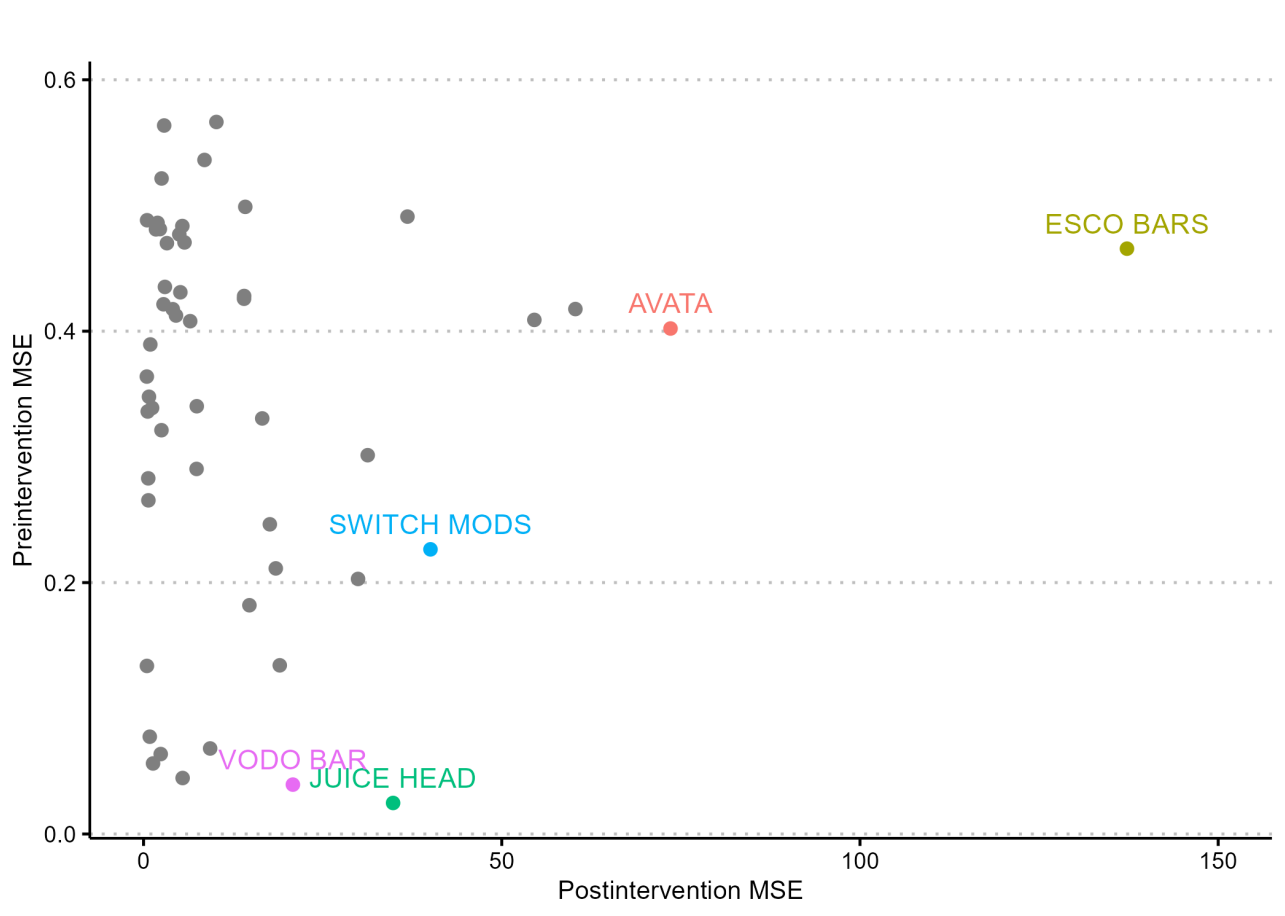
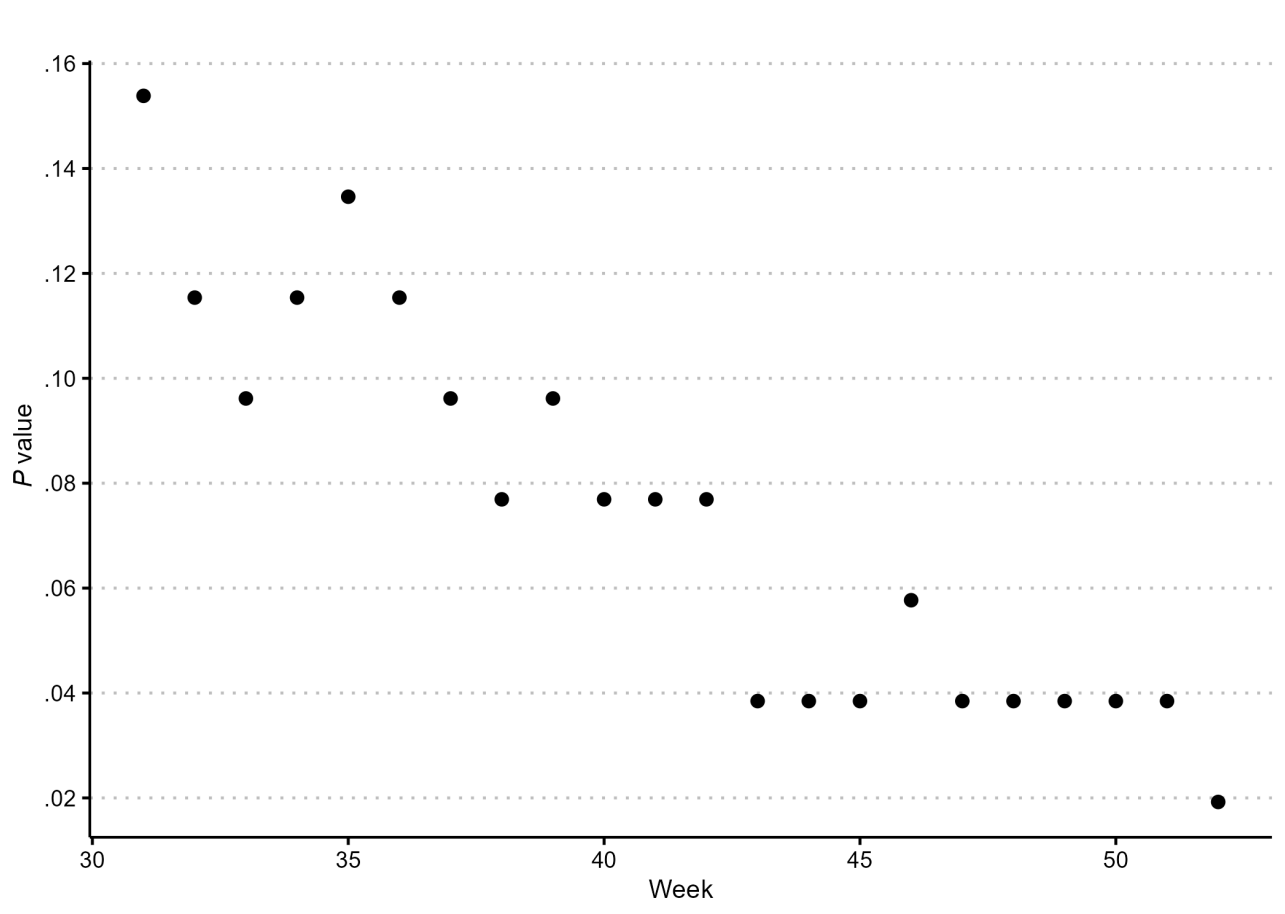


Figure 4 presents *P* values of weekly intervention effects of Esco Bars sales during the postintervention period. In addition to the overall *P* value (.06), week-specific *P* values provide additional evidence on the variations of intervention effects over time. Our findings show that week-specific *P* values decreased rapidly during the postintervention period. The *P* values of the intervention effect immediately after FDA advisory and enforcement actions were .15 (8/52, 15.3%) on the week ending on May 21, 2023, and .02 (1/52, 2%) on the week ending on December 30, 2023. By the 17th week after FDA advisory and enforcement actions, the intervention effect for Esco Bars was statistically significant based on the conventional threshold ($P<.05$). This indicates that these FDA actions were associated with a reduction in Esco Bars sales.

XSL•FO
RenderX

Figure 4. Week-specific *P* values for intervention effect on Esco Bars during the postintervention period. This figure depicts week-specific *P* values for the intervention effect on sales of Esco Bars in the postintervention period in 2023. Week 30 corresponds to the week ending on July 29, 2023, denoting the week of FDA's last advisory action against unauthorized Esco Bars in 2023, and warning letters were issued to the distributors on July 27, 2023. Week 52 corresponds to the week ending on December 30, 2023, the end of our postintervention period. Following the Fisher exact test, the *P* value for a given week represents the total share of intervention effects (placebo plus actual) that is greater than or equal to the actual intervention effect in a given week. By the final week in the study, the actual intervention effect was the highest out of 52 total intervention effects (1/52, 2%; *P*=.02).



Sensitivity Analysis

Table 1 presents the results based on alternative data standardization methods. These sensitivity analyses confirmed that our findings are independent of standardization methods for weekly sales. Our major findings remained consistent across all 3 methods, suggesting that the total Esco Bars sales fell by 1.7 million EQ units during the 5 months after the FDA's advisory and enforcement actions. By the last week of 2023, weekly sales of Esco Bars were reduced by almost 70%, and *P* values of week-specific intervention effects ranged from .02 to .04, reflecting the significant sales changes during the postintervention period.

Table . Sensitivity analysis results by the standardization method. Three standardization methods were used in this sensitivity analysis: (1) a *z* score method in the main model, which subtracts each brand's preintervention average from its weekly sales and divides the result by its preintervention SD; (2) a min-max normalization method, which subtracts the preintervention minimum and divides by the preintervention range; and (3) a robust rescaling method, which subtracts the preintervention median and divides by the preintervention IQR. The min-max normalization and robust rescaling methods are used for sensitivity analyses.

	z score (main model)	Min-max normalization	Robust rescaling
Esco Bars rank among 52 brands	3rd	6th	4th
Overall *P* value	.06	.12	.08
P value at the final week in 2023	.02	.04	.04
Total reductions in Esco Bars sales within 5 months (equivalized units)	1.7 million	1.7 million	1.7 million
Weekly Esco Bars sales reduction by the final week of 2023 (%)	68.50	68.50	68.70

XSL•FO
RenderX

Discussion

Principal Findings

Our findings demonstrate that the FDA's advisory and enforcement actions against illicit sales of unauthorized Esco Bars, including the import alert and warning letters that occurred from May to July 2023, resulted in a substantial reduction in product sales within 5 months. These actions not only reduced Esco Bars weekly sales by nearly 70% but also had sustained effects through the end of 2023 in the United States. As a result, total Esco Bars sales were reduced by approximately 1.7 million EQ units over the 5-month period. Notably, our conclusions were consistent and robust across sensitivity analyses.

Our findings confirmed that the FDA's import alerts and warning letters against Esco Bars products, which were illegally marketed following FDA's regulatory review decisions, were successful deterrents and had sustained impacts on illicit sales and distribution of these products across the supply chain. It is possible that the manufacturer, distributors, and retailers of these unauthorized products also took voluntary actions to restrain sales and availability of this product after receiving warning letters and/or being placed on import alert. These actions, although voluntary, were under the influence of FDA advisory and enforcement actions, and thus, should be considered as part of the impacts of these actions. These findings reinforce that the FDA's targeted advisory and enforcement actions are effective against sales and availability of other unauthorized commercial tobacco and nicotine products in the US market.

Our findings also generally align with concurrent patterns of youth cigarette use [7,29]. For example, data from the 2023 and 2024 National Youth Tobacco Survey found substantial declines in youth use of Esco Bars products, dropping from an estimated 440,000 students in 2023 to 160,000 in 2024, with the brand falling out of the top 5 most popular among youth by 2024 [7,29]. Although the focus of these studies was not specifically on FDA advisory and enforcement actions against illicit sales of unauthorized Esco Bars products in 2023, the findings on youth behaviors provide indirect supporting evidence for the effectiveness of these actions.

However, it is also important to note that, although Esco Bars' sales fell significantly after FDA regulatory review decisions and advisory and enforcement actions, these unauthorized products continued to be available illegally in the United States throughout 2023. This shows that there are actors in the marketplace who are intent on marketing unauthorized products, and thus regulatory vigilance needs to be an active and continual effort.

Accordingly, the FDA has continued to issue additional warning letters and updated import alerts and has pursued enforcement actions against firms to address the continued marketing, distribution, and sales of unauthorized products, including youth-appealing e-cigarettes, beyond 2023. As of October 2024, the FDA had issued more than 700 warning letters to firms for manufacturing, selling, and/or distributing unauthorized new tobacco products including e-cigarettes; issued more than 690 warning letters to retailers, including brick-and-mortar and online retailers, for the sales of unauthorized tobacco products; and filed civil money penalty complaints against more than 75 manufacturers and more than 150 retailers for distribution and/or sales of unauthorized tobacco products [30]. In addition, the Department of Justice (DOJ), in collaboration with the FDA, has filed a total of 8 injunctions against firms that continued to market unauthorized e-cigarette products [31]. Our findings in this analysis suggest that such actions are likely to have similarly successful impacts on regulating the marketplace and enforcing FDA decisions requiring the removal of illegal products from the market.

The complexity of the ever-evolving US e-cigarette marketplace has been exacerbated by the flow of unauthorized e-cigarettes entering the country. The availability and illicit sales of these unauthorized e-cigarette products, many of which are popular among young people, pose serious threats to the health and development of US youth and young adults [3-5,29,32,33]. To address these unauthorized products, the FDA has developed and implemented a comprehensive strategy across the supply chain including with other federal enforcement partners, such as the US Customs and Border Protection and the DOJ.

For example, in collaboration with US Customs and Border Protection, the FDA has made multiple seizures of unauthorized e-cigarettes, including (1) the seizure announced on December 14, 2023, of approximately 1.4 million units of unauthorized e-cigarettes with an estimated retail value of more than US $18 million;[34,35] (2) the seizure announced on April 30, 2024, of more than 45,000 units of unauthorized e-cigarette products with an estimated retail value of US $0.7 million;[36] (3) the seizure announced on May 22, 2025, of nearly 2 million units of unauthorized e-cigarette products with an estimated retail value of US $33.8 million [37]; and (4) the seizure announced on September 10, 2025, of approximately 4.7 million units of unauthorized e-cigarette products with an estimated retail value of US $86.5 million to accelerate further enforcement actions [38]. More recently, as announced on September 25, 2025, the FDA, DOJ, the Bureau of Alcohol, Tobacco, Firearms and Explosives, and the US Marshals Service seized more than 2.1 million unauthorized e-cigarettes taken from 5 distributors and 6 retailers across 7 different states [39].

These enforcement actions are part of the FDA's comprehensive enforcement strategy across the entire supply chain to increasingly address illicit sales of unauthorized e-cigarette products, particularly for products popular among youth. In addition to these enforcement strategies, the agency continues to educate the public about the risks of these products and to increase voluntary compliance by educating those in the supply chain (eg, retailers) about which products are legal to sell.

Continued surveillance of commercial tobacco and nicotine products available in the United States can provide data on unauthorized product sales and better support FDA efforts to protect the public health, while additional evaluation studies of application review decisions, warning letters, and import alerts can provide further evidence of the effectiveness of the FDA's regulatory authorities, including product review and advisory and enforcement actions. In addition, given the extent of illicit

XSL•FO
RenderX

tobacco product sales globally, including e-cigarette sales, continued surveillance could also help inform regulatory actions across the world, as appropriate.

Limitations

This study has some limitations. First, the NielsenIQ RMS data did not include e-cigarette sales from Alaska or Hawaii. The data also did not include e-cigarette sales from certain retail outlets, including tobacco specialty stores such as vape shops, liquor stores, and online stores. However, studies based on nationally representative data have shown that less than 25% of current youth e-cigarette users obtained their products from vape shops or online stores [40,41]. Second, this analysis did not account for state or local advisory and enforcement actions against unauthorized e-cigarettes during the analytical time period. However, it is reasonable to assume that these actions would have broadly applied to all unauthorized e-cigarettes and thus were unlikely to differentially affect the estimated impact of this analysis, which was focused on Esco Bars. While this study was not designed to assess the impact of FDA advisory and enforcement actions against unauthorized Esco Bars products on disposable e-cigarette sales, other data have shown that disposable e-cigarette sales have also declined from June 18, 2023, throughout the rest of the year [42]. Third, our findings may underestimate the overall impact of FDA advisory and enforcement actions against Esco Bars, as our analytical period stops at the end of 2023, and the FDA took additional actions in 2024. In addition, since the study was based on retail sales data, we were unable to assess potential behavioral changes among Esco Bars users. However, economic studies have shown that sales and consumption tend to be highly correlated [43-45], and recent studies have shown that Esco Bars has become less popular among youth between 2023 and 2024 [7,29]. Additional research using population-based survey data could present direct behavioral changes associated with the FDA's tobacco and nicotine product advisory and enforcement actions.

Conclusions

In response to the illicit sale of Esco Bars e-cigarette products, FDA issued multiple advisory and enforcement actions against firms that marketed, sold, and/or distributed unauthorized tobacco products. Our findings demonstrate that these actions (ie, warning letters and an import alert) substantially reduced sales of Esco Bars products in the United States, with a sustained impact noted over a 5-month period. Specifically, these actions reduced sales by approximately 1.7 million EQ units within 5 months. These findings underscore the importance and impact of FDA actions against unauthorized tobacco products as part of a comprehensive regulatory approach.

Acknowledgments

The authors would like to thank Esther Salazar and Deborah Neveleff from the US Food and Drug Administration for their insightful feedback. BAK's work on the study was undertaken while he was employed at the US Food and Drug Administration's Center for Tobacco Products. The authors confirm that no generative artificial intelligence tools (such as ChatGPT or other large language models) were used in the writing, editing, data analysis, or creation of content for this manuscript. All work presented is the original creation of the authors.

The findings and conclusions in this study are those of the authors and do not necessarily represent the official position of the US Food and Drug Administration. The author's own analyses, calculations, and conclusions were informed in part by NielsenIQ's Retail Measurement Service data for the e-cigarettes category from January 1, 2023, to December 30, 2023, for Total US Extended All Outlets Combined and convenience stores and are those of the US Food and Drug Administration and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein, or in developing, reviewing, or confirming the research approaches used in connection with this report. NielsenIQ Retail Measurement Service data consist of weekly purchase and pricing data generated from participating retail store point-of-sale systems in all US markets. See the NielsenIQ website for more information [46].

Funding

This study is funded by the US Food and Drug Administration.

Data Availability

The data cannot be shared publicly because they contain proprietary information and are subject to contractual restrictions.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Online appendix: model specifications.
[DOCX File, 150 KB - publichealth_v12i1e81033_app1.docx]

References

XSL•FO
RenderX

1. US Centers for Disease Control and Prevention, Public Health Service, Office of the Surgeon General. E-cigarette use among youth and young adults: a report of the surgeon general. : US Department of Health and Human Services; 2016 URL: https://stacks.cdc.gov/view/cdc/44007 [accessed 2026-03-09]
2. The relative risks of tobacco products. US Food and Drug Administration. URL: https://www.fda.gov/tobacco-products/health-effects-tobacco-use/relative-risks-tobacco-products#:~:text=as%20a%20whole.-,Are%20E%2DCigarettes%20a%20Lower%2DRisk%20Alternative%20to%20Cigarettes?,the%20statutory%20public%20health%20standard [accessed 2025-03-11]
3. England LJ, Bunnell RE, Pechacek TF, Tong VT, McAfee TA. Nicotine and the developing human: a neglected element in the electronic cigarette debate. Am J Prev Med 2015 Aug;49(2):286-293. [doi: 10.1016/j.amepre.2015.01.015] [Medline: 25794473]
4. Happer JP, Courtney KE, Baca RE, et al. Nicotine use during late adolescence and young adulthood is associated with changes in hippocampal volume and memory performance. Front Neurosci 2024;18:1436951. [doi: 10.3389/fnins.2024.1436951] [Medline: 39221006]
5. Leslie FM. Unique, long-term effects of nicotine on adolescent brain. Pharmacol Biochem Behav 2020 Oct;197:173010. [doi: 10.1016/j.pbb.2020.173010] [Medline: 32738256]
6. King BA. Flavors are a major driver of the youth e-cigarette epidemic. Am J Public Health 2020 Jun;110(6):773-774. [doi: 10.2105/AJPH.2020.305670] [Medline: 32374696]
7. Birdsey J, Cornelius M, Jamal A, et al. Tobacco product use among U.S. middle and high school students—National Youth Tobacco Survey, 2023. MMWR Morb Mortal Wkly Rep 2023 Nov 3;72(44):1173-1182. [doi: 10.15585/mmwr.mm7244a1] [Medline: 37917558]
8. FDA seeks $20K+ fines against retailers selling unauthorized youth-appealing e-cigarettes. US Food and Drug Administration. 2024. URL: https://www.fda.gov/tobacco-products/ctp-newsroom/fda-seeks-20k-fines-against-retailers-selling-unauthorized-youth-appealing-e-cigarettes#:~:text=According%20to%20the%202023%20National,Bars%20products%20during%20the%20period [accessed 2025-03-19]
9. About warning and close-out letters. US Food and Drug Administration. 2024. URL: https://www.fda.gov/inspections-compliance-enforcement-and-criminal-investigations/warning-letters/about-warning-and-close-out-letters [accessed 2024-04-23]
10. Advisory and enforcement actions against industry for unauthorized tobacco products. US Food and Drug Administration. 2024. URL: https://www.fda.gov/tobacco-products/compliance-enforcement-training/advisory-and-enforcement-actions-against-industry-unauthorized-tobacco-products#4 [accessed 2024-04-23]
11. Advisory action letters. US Food and Drug Administration. 2023. URL: https://www.fda.gov/animal-veterinary/compliance-enforcement/advisory-action-letters [accessed 2025-09-08]
12. Improt alerts. US Food and Drug Administration. 2024. URL: https://www.fda.gov/industry/actions-enforcement/import-alerts [accessed 2024-04-23]
13. Import alert 98-06. US Food and Drug Administration. 2023. URL: https://www.accessdata.fda.gov/cms_ia/importalert_1163.html [accessed 2024-02-08]
14. FDA puts firms responsible for Esco Bars and Breeze—two popular disposable e-cigarette brands—on notice. US Food and Drug Administration. 2023. URL: https://www.fda.gov/tobacco-products/ctp-newsroom/fda-puts-firms-responsible-esco-bars-and-breeze-two-popular-disposable-e-cigarette-brands-notice [accessed 2024-01-31]
15. FDA inspection blitz leads to more than 180 warning letters to retailers for the illegal sale of youth-appealing Elf Bar and Esco Bars e-cigarettes. US Food and Drug Administration. 2023. URL: https://www.fda.gov/news-events/press-announcements/fda-inspection-blitz-leads-more-180-warning-letters-retailers-illegal-sale-youth-appealing-elf-bar [accessed 2024-01-31]
16. FDA puts distributors on notice for illegal e-cigarettes popular with youth, including Elf Bar/EB design and Esco Bars. US Food and Drug Administration. 2023. URL: https://www.fda.gov/tobacco-products/ctp-newsroom/fda-puts-distributors-notice-illegal-e-cigarettes-popular-youth-including-elf-bar/eb-design-and-esco-bars#:~:text=FDA%2C%20EC%20Supply%2C%20Inc [accessed 2024-01-31]
17. Nguyen H, Dennehy CE, Tsourounis C. Violation of US regulations regarding online marketing and sale of e-cigarettes: FDA warnings and retailer responses. Tob Control 2020 Dec;29(e1):e4-e9. [doi: 10.1136/tobaccocontrol-2019-055106] [Medline: 32123138]
18. Schillo BA, Bertrand A, Briggs J, et al. Analysis of e-cigarette warning letters issued by the Food and Drug Administration in 2020 and 2021. Tob Control 2024 Feb 20;33(2):247-251. [doi: 10.1136/tc-2022-057359] [Medline: 36229228]
19. Trigger S, Xu X, Malarcher A, Salazar E, Shin H, Babb S. Trends in over-the-counter nicotine replacement therapy sales, US, 2017 2020. Am J Prev Med 2023 May;64(5):650-657. [doi: 10.1016/j.amepre.2022.12.008] [Medline: 36682916]
20. Fredriksson A, Oliveira GD. Impact evaluation using difference-in-differences. RAUSP Manag J 2019 Oct 14;54(4):519-532. [doi: 10.1108/RAUSP-05-2019-0112]
21. Dimick JB, Ryan AM. Methods for evaluating changes in health care policy: the difference-in-differences approach. JAMA 2014 Dec 10;312(22):2401-2402. [doi: 10.1001/jama.2014.16153] [Medline: 25490331]

XSL•FO
RenderX

22. Abadie A, Diamond A, Hainmueller J. Comparative politics and the synthetic control method. Am J Pol Sci 2015 Feb;59(2):495-510. [doi: 10.1111/ajps.12116]
23. Abadie A. Using synthetic controls: feasibility, data requirements, and methodological aspects. J Econ Lit 2021 Jun 1;59(2):391-425. [doi: 10.1257/jel.20191450]
24. Sun L, Ben-Michael E, Feller A. Using multiple outcomes to improve the synthetic control method. Rev Econ Stat 2025:1-29. [doi: 10.1162/rest_a_01592]
25. Pinkney S. An improved and extended bayesian synthetic control. arXiv. Preprint posted online on Mar 30, 2021. [doi: 10.48550/arXiv.2103.16244]
26. Abadie A, L'Hour J. A penalized synthetic control estimator for disaggregated data. J Am Stat Assoc 2021 Oct 2;116(536):1817-1834. [doi: 10.1080/01621459.2021.1971535]
27. Upton GJG. Fisher's exact test. J R Stat Soc Ser A Stat Soc 1992;155(3):395-402. [doi: 10.2307/2982890]
28. Abadie A, Diamond A, Hainmueller J. Synthetic control methods for comparative case studies: estimating the effect of California's tobacco control program. J Am Stat Assoc 2010 Jun;105(490):493-505. [doi: 10.1198/jasa.2009.ap08746]
29. Park-Lee E, Jamal A, Cowan H, et al. Notes from the field: e-cigarette and nicotine pouch use among middle and high school students—United States, 2024. MMWR Morb Mortal Wkly Rep 2024 Sep 5;73(35):774-778. [doi: 10.15585/mmwr.mm7335a3] [Medline: 39236021]
30. FDA warns firms illegally selling e-cigarettes resembling products with smart technology, including phones and gaming devices. US Food and Drug Administration. 2024. URL: https://www.fda.gov/tobacco-products/ctp-newsroom/fda-warns-firms-illegally-selling-e-cigarettes-resembling-products-with-smart-technology-including-phones#:~:text=%E2%80%8C%20%20%20

46. NielsenIQ. URL: http://www.NielsenIQ.com/global/en/ [accessed 2026-03-11]

Abbreviations

DiD: difference-in-difference
DOJ: Department of Justice
FDA: US Food and Drug Administration
MSE: mean squared error
RMS: Retail Measurement Service
RTA: refuse to accept
SCM: synthetic control method

Edited by A Mavragani, T Sanchez; submitted 05.Aug.2025; peer-reviewed by A Bertrand, JM Martin-Alvarez; accepted 10.Dec.2025; published 12.Mar.2026.

Please cite as:
Shin H, Hannings J, Cahn Z, Xu X, Toblin R, Farrelly M, King BA
The Impact of US Food and Drug Administration's Advisory and Enforcement Actions on US Sales of Esco Bars Products in 2023: Synthetic Control Approach
JMIR Public Health Surveill 2026;12:e81033
URL: https://publichealth.jmir.org/2026/1/e81033
doi:10.2196/81033

© Hyungsik Shin, Jonathan Hannings, Zachary Cahn, Xin Xu, Robin Toblin, Matthew Farrelly, Brian A King. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 12.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Uncovering the Reasons Behind Maternal Care Dropout in Bangladesh: Cross-Sectional Study

Syeda Saima Alam[1], MSc; Plabon Sarkar[2], MSc; M A Rifat[3], MSc, MPH; Sumaiya Jahan[2], MSc; Rokibul Islam[2], MSc; Israt Jahan[1], MSc; Sanjib Saha[4], PhD

[1]Noakhali Science and Technology University, Noakhali, Bangladesh
[2]University of Dhaka, Dhaka, Bangladesh
[3]BRAC University, Dhaka, Bangladesh
[4]Lund University, BMC, Sölvegatan 19, Lund, Sweden

Corresponding Author:
Sanjib Saha, PhD
Lund University, BMC, Sölvegatan 19, Lund, Sweden

Abstract

Background: Utilization of the maternal continuum of care (CoC)—comprising adequate antenatal care (ANC), skilled birth attendance, and postnatal care (PNC)—is critical for improving maternal and child health outcomes. However, dropout from the CoC remains substantial in Bangladesh, with women discontinuing services at different stages of pregnancy, delivery, and postpartum care.

Objective: This study aimed to quantify maternal dropout at each stage of the CoC and identify socioeconomic and demographic factors associated with discontinuity, comparing two nationally representative survey rounds.

Methods: Data were drawn from the Bangladesh Demographic and Health Surveys (BDHS) 2017‑2018 and 2022. Women aged 15 to 49 years with a live birth in the preceding 2 to 3 years were included. Completion of full CoC was defined as receiving at least 4 ANC visits, delivering with a skilled birth attendant, and obtaining at least 1 PNC contact within 48 hours of delivery. Predisposing (age, education, parity, religion, and division), enabling (wealth index, media exposure, health care access, and residence), and need factors (terminated pregnancy and desired pregnancy status) were identified using the Andersen Behavioral Model. Survey-weighted multivariable logistic regression models were fitted for each CoC component and overall CoC completion, with interaction terms to assess whether associations differed between survey rounds.

Results: Among 8424 mothers, 27.9% (n=2350) failed to complete all components of the maternal CoC. Dropout was highest at the ANC stage (n=4962, 55.7%), followed by PNC (n=3976, 47.2%) and skilled birth attendant–assisted delivery (n=3378, 40.1%). Between survey rounds, overall CoC dropout decreased significantly from 31.9% (BDHS 2017‑2018) to 22.4% (BDHS 2022), reflecting modest improvements in service continuity. Factors significantly associated with higher odds of CoC dropout included lower maternal education (adjusted odds ratio [AOR] 2.70, 95% CI 1.94‑3.77; $P<.001$), higher parity (AOR 2.73, 95% CI 2.12‑3.50; $P<.001$), lower wealth quintiles (AOR 4.04, 95% CI 3.02‑5.41; $P<.001$), and rural residence (AOR 1.40, 95% CI 1.18‑1.67; $P<.001$). Protective factors included older maternal age at delivery (AOR 0.56, 95% CI 0.42‑0.74; $P<.001$) and history of ever-terminated pregnancy (AOR 0.74, 95% CI 0.63‑0.86; $P<.001$). Significant temporal interactions (all $P<.05$) indicated that the strength of associations for education, parity, religion, wealth, media exposure, health care access barriers, residence, and pregnancy desire differed between survey rounds, reflecting changing determinants of CoC engagement amid policy reforms and pandemic disruptions.

Conclusions: Maternal, socioeconomic, and geographic factors are strongly associated with discontinuity along the maternal health care continuum in Bangladesh. Statistically significant temporal variations underscore the impact of evolving health policies and system disruptions on maternal service utilization patterns. Targeted, area-specific interventions addressing these determinants across all CoC components are essential to improve maternal health care retention and achieve better maternal and child health outcomes.

(JMIR Public Health Surveill 2026;12:e85875) doi:10.2196/85875

KEYWORDS

maternal dropout; continuum of care; antenatal care; skilled birth attendants; postnatal care; demographic and health survey; Bangladesh

XSL•FO
RenderX

Introduction

The maternal continuum of care (CoC) consists of 4 or more antenatal care (ANC) visits, childbirth by a skilled birth attendant (SBA), and a postnatal care (PNC) visit by skilled health care providers within 48 hours of delivery [1]. When a woman and child pair fulfills all 3 criteria, they are considered to have attained a complete CoC [1]. The CoC encompasses integrated service delivery for mothers and children, ensuring that women have access to quality care throughout pregnancy, childbirth, and the postpartum period [2]. It is critical for improved health outcomes, efficient and effective delivery of nutritional interventions, increased service utilization, and addressing underlying factors [3], with recent multicountry evidence underscoring that completion of the CoC remains suboptimal, particularly in the low- and middle-income countries [4].

The completion of CoC demonstrates a marked reduction in undernutrition indicators, for example, stunting, wasting, and underweight among children alongside a decline in maternal and child mortality rates [3,5]. Empirical research indicates that adequate continuity of care can avert half a million maternal fatalities, 4 million neonatal fatalities, and 6 million pediatric fatalities worldwide [6-8]. Consequently, withdrawal from the CoC signifies noncompliance with prescribed maternal care measures [9], and emerging findings from Bangladesh indicate that the use of the full maternal CoC is positively associated with optimal complementary feeding practices among children aged 6 to 23 months [10].

In Bangladesh, around 53% of pregnant women drop out from receiving 4 or more ANC visits, 47% of the childbirths are not supported by SBAs, and 48% of women do not receive PNC visits within 2 days after the delivery [11,12]. The dropouts of women and children from the CoC can severely impact maternal and child nutrition. For instance, lower ANC and PNC visits result in poor infant and young child feeding practices that increase the risk of malnutrition and impaired cognitive development, which may result in long-term adverse educational and economic outcomes for children [13]. In addition, insufficient ANC visits and nutrition counseling can lead to maternal malnutrition, such as anemia and pregnancy complications, while a lack of SBA during delivery and PNC visits might pose a risk to maternal mortality and poor reproductive health [13]. Considering the significance of the CoC components, they are being evaluated for consistency to safeguard the survival and welfare of both mother and newborn. This, in turn, may serve as a crucial strategy for achieving the Sustainable Development Goals' target of a maternal mortality ratio under 70 per 100,000 live births and a neonatal mortality rate under 12 per 1000 live births by 2030 [6,13].

The CoC is essential for enhancing maternal and child nutrition outcomes, as interruptions in one service frequently cause disruptions in others. For example, data from Bangladesh indicate that women who cease ANC are more inclined to also abandon skilled birth attendance and PNC [13,14]. It is important to understand the many factors that lead to dropout in both individual and combined parts of the CoC to create complete and effective interventions. Key factors—such as maternal age, socioeconomic status, education level, place of residence, religion, parity, and occupation—significantly influence adherence to and gaps within the continuum, affecting the success of maternal and child nutrition initiatives [13,15]. To create targeted and comprehensive services that can remove barriers, make sure services continue, and ultimately improve the health and nutrition of mothers and children in Bangladesh, it is important to identify and analyze these factors in a strong way.

This study aims to comprehensively quantify and compare maternal dropouts at each stage of the CoC (ANC, delivery, and PNC) and from the overall CoC between Bangladesh Demographic and Health Surveys (BDHS) 2017-2018 and BDHS 2022, and to identify and contrast the socioeconomic and demographic factors associated with these discontinuities. It is hypothesized that the magnitude and pattern of maternal dropouts differ significantly between BDHS 2017‑2018 and BDHS 2022 across the stages of the CoC and for the overall CoC, and that the socioeconomic and demographic determinants of discontinuity at different stages of the CoC differ in strength and direction between the 2 survey rounds.

By exploring where and why women disengage from critical services and by comparing 2 distinct survey rounds, the research provides crucial insights into gaps that undermine maternal and child health outcomes. The findings from this study will help inform targeted strategies and policy interventions to reduce dropouts, improve service uptake, and ultimately contribute to achieving maternal and child health goals nationwide.

Methods

Study Design, Data Sources, and Sampling

This study used data from BDHS 2017‑2018 and BDHS 2022 to determine how CoC utilization and its covariates have changed over time. Both surveys are nationally representative, are cross-sectional in design, and employed 2-stage stratified random sampling. The overall response rate for both surveys was 98%. The detailed methods are outlined elsewhere [11,16].

A total of 8424 ever-married women aged 15 to 49 years who had a live birth within the last 3 and 2 years prior to the surveys, respectively, for BDHS 2017‑2018 and BDHS 2022, were retrieved. However, the study only analyzed samples with complete information regarding maternal utilization of CoC and coupled with respective covariates.

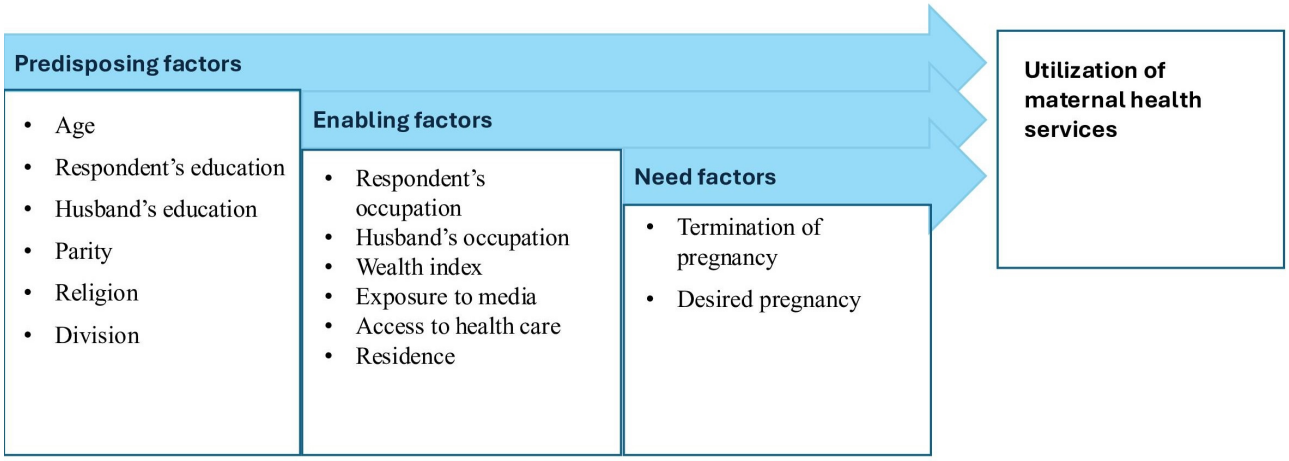
Conceptual Model

This study utilized the Andersen Behavioral Model of Health Care Services to systematically identify factors affecting the utilization of CoC services. The Andersen model is a widely used conceptual framework for understanding health care utilization, positing that service use is shaped by 3 key domains: predisposing, enabling, and need factors [17,18]. Predisposing factors encompass demographic and social attributes, including age, gender, education, and health beliefs, that affect an individual's propensity to pursue care. This study identified maternal age, respondent's education, husband's education, parity, religion, and division as predisposing factors. Maternal

XSL•FO
RenderX

age and parity are recognized factors influencing obstetric risk and health-seeking behavior, whereas maternal education affects health literacy, awareness of maternal health risks, and autonomy in decision-making. The husband's education is a significant predisposing factor, as it indicates social status, influences household health beliefs, and affects decision-making authority concerning maternal health care utilization [19]. Enabling factors include the means and resources that make it easier or harder to get services, such as community resources, household wealth, and health care availability [20]. This study identified the respondent's occupation, the husband's occupation, wealth index, media exposure, health care access, and residence as enabling factors. The household wealth index indicates the economic ability to access services, while exposure to mass media functions as a crucial enabling factor by disseminating maternal health information, shaping health-seeking behaviors, and addressing knowledge deficiencies, especially in resource-constrained environments. Access to health care and residence encompasses geographic and structural impediments to service utilization. Need factors encompass both perceived and assessed health status, indicating the actual or subjective requirement for medical attention [21]. Termination of pregnancy and desired pregnancy were chosen as need factors because termination of pregnancy indicates obstetric history that heightens the risks of future pregnancy complications, whereas pregnancy intention affects the motivation to pursue and complete the continuum of maternal care. In this study, the classification and definitions of analytical variables adhered closely to the established schema, ensuring conceptual coherence. Figure 1 shows how these areas interact with each other, as described in the Andersen Behavioral Model. It shows how combinations of predisposing, enabling, and need factors affect the use of CoC services and, in the end, maternal and child health outcomes.

Figure 1. Theoretical framework based on the Andersen Health Care Utilization Model.



Outcome Variable

The outcome variable is maternal dropout from CoC, which consists of dropping out from three integral components, which are (1) at least 4 ANC visits by skilled care providers during pregnancy, (2) delivery assisted by SBAs, and (3) at least 1 PNC visit by skilled care providers within 48 hours of delivery [14]. The skilled care providers in this aspect are designated as doctors, nurses, midwives, paramedics, community SBAs, and subassistant community medical officers. A mother was considered to have dropped out from the full CoC if she had not received any of these components. The mothers were assigned code 1 if they received all 3 components of the CoC; otherwise, code 0. Similarly, mothers who utilized any specific component of the CoC were assigned a code 1, otherwise 0, and analyzed separately.

Covariates

We considered covariates, including age at delivery (<19 y, 19-30 y, and 31 - 49 y); education level (no education, primary, secondary, and higher); husband's education level (no education, primary, secondary, and higher); occupation (not working and working); husband's occupation (not working and working); parity (1, 2 - 3, and >3); ever-terminated pregnancy (never/no and ever/yes); desired pregnancy (yes and no); wealth index (poorest, poorer, middle, richer, and richest); any exposure to television, radio, or news (yes and no); facing problems such as getting permission and money and long distance to access health care (big problem and not a big problem); religion (Muslim and other); place of residence (rural and urban); division of residence (Barisal, Chittagong, Dhaka, Khulna, Mymensingh, Rajshahi, Rangpur, and Sylhet); and survey round (BDHS 2017 - 2018 and BDHS 2022). These covariates were categorized based on how they align with the Andersen Behavioral Model (Figure 1).

Statistical Analysis

The distribution of covariates by maternal dropouts from CoC was examined using cross-tabulations in a frequency distribution table, and the differences among 4 individual components were observed using the χ^2 test.

The association between prevalence of dropout from maternal CoC and covariates was analyzed using logistic regression models. Four models were utilized based on the outcome variables, including dropout from (1) 4 ANC, (2) delivery assisted by SBAs, (3) PNC within 48 hours of delivery, and (4) all CoC components. Given that the selection of variables was guided by a theoretical framework, the Andersen Behavioral Model of Health Services, no covariate was excluded in the

XSL•FO
RenderX

multivariable models, irrespective of the level of significance of the association in the univariable models.

Data from BDHS 2017 - 2018 and BDHS 2022 were pooled, and a binary indicator for survey round was included in all models. We fitted survey-weighted logistic regression models for each stage of the CoC (completion of adequate ANC, institutional delivery, and PNC contact) as well as overall CoC completion, with socioeconomic and demographic characteristics as predictors. To assess whether the patterns of dropout and the associations between maternal characteristics and CoC engagement differed between survey rounds, we included interaction terms between survey round and each key predictor (eg, wealth quintile, education level, and place of residence). Equality of the survey-specific effects across socioeconomic and demographic factors was evaluated using joint Wald tests of the interaction coefficients; statistical significance indicated that the log odds ratios (ie, the strength and direction of associations) differed between BDHS 2017 - 2018 and BDHS 2022 [22].

To ensure the accuracy of the estimates, the analyses were adjusted for sampling weight, strata, and primary sampling units. SEs of the odds ratio and variance inflation factors were observed to examine the multicollinearity. An SE of the odds ratio >2 was considered to indicate multicollinearity among the covariates [23]. All the statistical tests were conducted considering $P<.05$ as the significance level. The software STATA 17 (StataCorp, College Station) was used for data analysis. The study is reported in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines for observational studies (Checklist 1).

Ethical Considerations

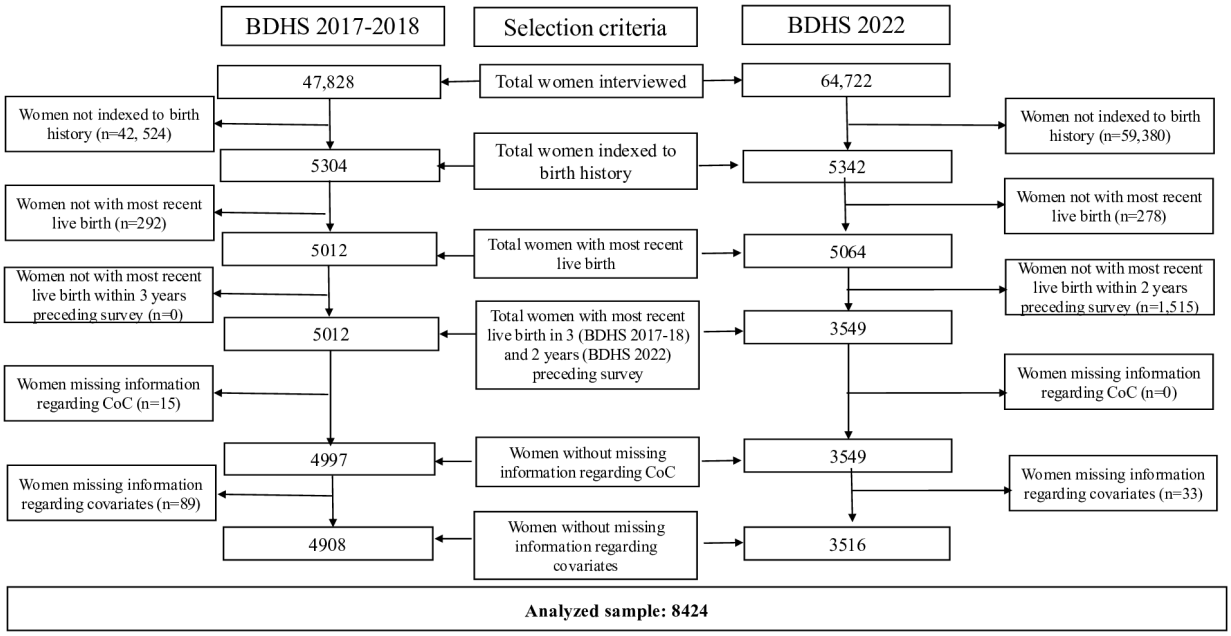
The Demographic and Health Survey received approval from the ICF Macro Institutional Review Board, United States, which complies with all the criteria of 45 CFR 46 “Protection of Human Subjects.” As the data were collected in Bangladesh, the BDHS received ethical approval from the National Research Ethics Committee of the Bangladesh Medical Research Council, Dhaka, Bangladesh. Before being enrolled in the study, verbal informed agreement was obtained from each participant and their intimate partners (all ever-married women aged 15 - 49 y). If the respondents were incapable of reading, verbal consent was considered the most appropriate method to confirm participation.

All survey data were anonymized before being made available for analysis, and no financial or other incentives were offered to participants for their participation or for the use of their data in this study. The manuscript and its supplementary materials do not include any information or images that could lead to the identification of individual respondents. Comprehensive descriptions of the ethical procedures and informed consent process are provided in the BDHS documentation and official reports.

Results

The sample selection process for this study is illustrated in Figure 2. In this study, 8424 mothers with complete information about outcome and covariates were analyzed. In the analyzed sample, maternal dropouts from having ≥4 ANC, delivery assisted by SBAs, PNC within 48 hours of delivery, and full CoC were 55.7% (n=4962), 40.1% (n=3378), 47.2% (n=3976), and 27.9% (n=2350), respectively (Figure 3).

Figure 2. The sample selection process. BDHS: Bangladesh Demographic and Health Surveys; CoC: continuum of care.



XSL•FO
RenderX

Figure 3. Maternal dropout from the continuum of care (CoC) and its component. ANC: antenatal care; BDHS: Bangladesh Demographic and Health Surveys; PNC: postnatal care; SBA: skilled birth attendant.

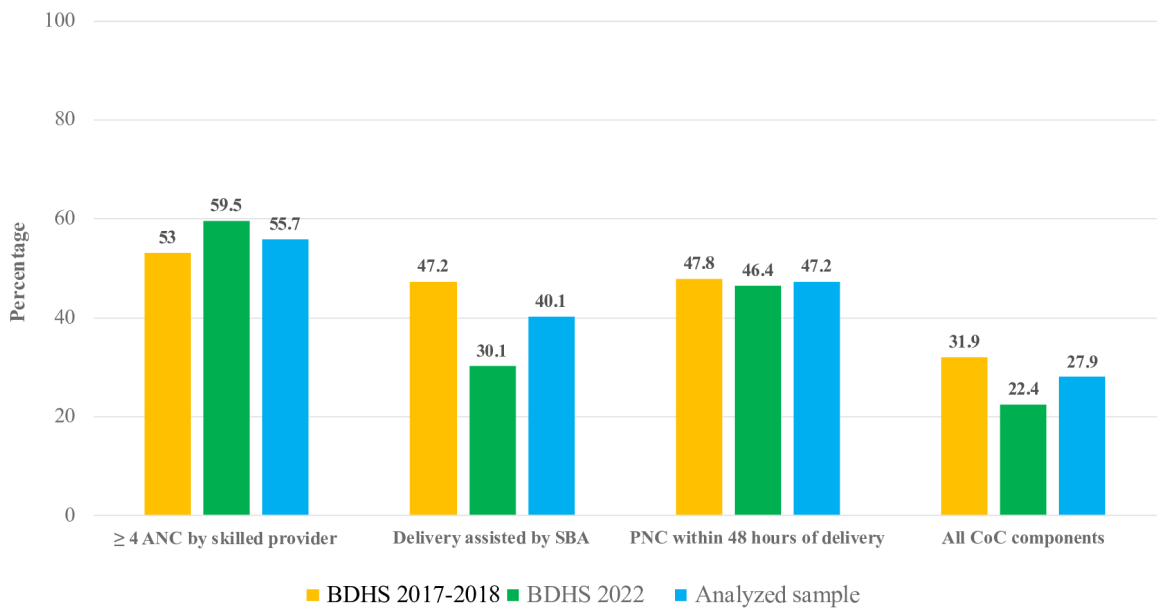


Table 1 demonstrates the difference in dropouts from individual and all CoC components across BDHS 2017‑2018 and 2022 and shows clear changes over time. Across both BDHS rounds, dropouts from the maternal CoC remained high, with particularly large attrition among socioeconomically disadvantaged and rural women. The dropout at each stage (≥4 ANC visits, SBA-assisted delivery, and ≥4 PNC contacts) was consistently higher among women with little or no education, from poorer households, without media exposure, reporting major problems in accessing care, and residing in rural or underserved divisions. Educational and wealth gradients were steep: women with no schooling or in the poorest quintile had roughly doubled the dropout from full CoC compared with highly educated or richest women in both surveys. Although some absolute levels improved between 2017‑2018 and 2022, the pattern of persistent socioeconomic and geographic inequities in CoC coverage remained the dominant feature of the data.

XSL•FO
RenderX

Table . Association between the coverage of ≥4 antenatal care (ANC), postnatal care (PNC), and delivery by skilled birth attendants (SBAs), and full continuum of care (CoC) by maternal characteristics in Bangladesh according to the Bangladesh Demographic and Health Surveys (BDHS) survey 2017 - 2018 (N=4908) and 2022 (N=3516).

Covariates	Dropout from ANC		Dropout from SBA-assisted delivery		Dropout from PNC		Dropout from all CoC components	
	BDHS 2017 - 2018, n (%)	BDHS 2022, n (%)	BDHS 2017 - 2018, n (%)	BDHS 2022, n (%)	BDHS 2017 - 2018, n (%)	BDHS 2022, n (%)	BDHS 2017 - 2018, n (%)	BDHS 2022, n (%)
Predisposing factors								
Age at delivery, y	*P*=.03	*P*=.002	*P*=.49	*P*=.23	*P*=.72	*P*=.18	*P*=.03	*P*=.03
<19	425 (52)	290 (63)	385 (47)	142 (31)	385 (47)	214 (46)	248 (30)	1060 (31)
19-30	1745 (59)	1442 (59)	1586 (46)	711 (29)	1831 (47)	1120 (46)	1060 (31)	1060 (31)
31-49	371 (56)	317 (53)	319 (49)	195 (32)	339 (52)	300 (50)	1060 (31)	1060 (31)
Education level	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001
No education	241 (79)	125 (77)	223 (73)	88 (54)	222 (73)	109 (67)	184 (61)	70 (43)
Primary	898 (66)	590 (73)	897 (66)	378 (47)	900 (66)	489 (61)	648 (48)	292 (36)
Secondary	1138 (48)	1097 (59)	1019 (43)	523 (28)	1023 (44)	823 (44)	634 (27)	371 (20)
Higher	264 (30)	237 (34)	151 (17)	59 (9)	164 (18)	213 (31)	78 (9)	42 (6)
Husbands' education	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001
No education	475 (70)	373 (71)	469 (69)	261 (50)	468 (69)	343 (65)	358 (53)	205 (39)
Primary	1017 (62)	714 (60)	970 (59)	413 (40)	973 (59)	554 (54)	660 (40)	309 (30)
Secondary	783 (48)	686 (57)	682 (42)	291 (24)	941 (58)	497 (41)	437 (27)	206 (17)
Higher	266 (28)	276 (36)	169 (18)	83 (11)	773 (19)	240 (31)	89 (9)	55 (7)
Parity	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001
1	821 (44)	733 (54)	643 (35)	276 (20)	656 (35)	806 (60)	399 (22)	194 (14)
2 - 3	1286 (52)	1082 (59)	1229 (50)	592 (32)	1236 (50)	874 (47)	808 (33)	445 (24)
>3	434 (73)	234 (72)	418 (70)	180 (56)	417 (70)	215 (67)	337 (56)	136 (42)
Religion	*P*=.003	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001
Muslim	2355 (52)	1917 (59)	2136 (48)	1004 (31)	2342 (50)	1545 (48)	1450 (32)	744 (23)
Others	186 (45)	132 (47)	154 (37)	44 (16)	257 (62)	89 (32)	94 (23)	31 (11)
Division	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001
Barisal	314 (60)	240 (64)	262 (50)	124 (33)	265 (51)	181 (49)	203 (39)	101 (27)
Chittagong	485 (59)	339 (58)	423 (52)	180 (31)	400 (49)	266 (45)	295 (36)	138 (23)
Dhaka	336 (46)	285 (53)	271 (37)	130 (24)	283 (39)	247 (46)	180 (25)	100 (19)
Khulna	202 (40)	224 (56)	175 (35)	50 (13)	173 (34)	115 (29)	90 (18)	37 (9)
Mymensingh	308 (52)	229 (51)	326 (55)	186 (41)	326 (55)	253 (56)	204 (34)	118 (26)
Rajshahi	257 (50)	211 (60)	221 (43)	88 (25)	221 (43)	144 (41)	140 (27)	61 (17)
Rangpur	209 (38)	258 (62)	262 (48)	135 (33)	264 (48)	206 (50)	138 (25)	87 (21)
Sylhet	430 (63)	263 (66)	378 (56)	155 (39)	377 (55)	222 (55)	294 (43)	133 (33)
Enabling factors								

XSL•FO
RenderX

Covariates	Dropout from ANC		Dropout from SBA-assisted delivery		Dropout from PNC		Dropout from all CoC components	
	BDHS 2017‑2018, n (%)	BDHS 2022, n (%)	BDHS 2017‑2018, n (%)	BDHS 2022, n (%)	BDHS 2017‑2018, n (%)	BDHS 2022, n (%)	BDHS 2017‑2018, n (%)	BDHS 2022, n (%)
Occupation	*P*=.31	*P*=.21	*P*<.001	*P*<.001	*P*<.001	*P*=.07	*P*<.001	*P*=.002
Not working	1518 (51)	1554 (58)	1217 (41)	731 (27)	1233 (42)	1229 (46)	842 (28)	561 (21)
Working	1023 (53)	495 (60)	1073 (55)	317 (39)	1076 (56)	405 (49)	702 (36)	214 (26)
Husbands' occupation	*P*=.15	*P*=.10	*P*=.02	*P*=.02	*P*=.05	*P*=.22	*P*=.49	*P*=.19
Not working	24 (63)	40 (49)	11 (29)	15 (19)	12 (32)	43 (53)	10 (26)	13 (16)
Working	2517 (52)	2009 (58)	2279 (47)	1033 (30)	2297 (47)	1591 (46)	1534 (32)	762 (22)
Wealth index	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001
Poorest	736 (69)	540 (75)	772 (73)	388 (54)	769 (72)	471 (66)	572 (54)	297 (41)
Poorer	610 (62)	504 (69)	594 (60)	279 (38)	598 (60)	386 (53)	406 (41)	212 (29)
Middle	459 (52)	442 (62)	471 (54)	199 (28)	409 (46)	320 (45)	253 (29)	143 (20)
Richer	452 (46)	345 (50)	356 (37)	126 (18)	368 (38)	252 (37)	225 (23)	91 (13)
Richest	284 (28)	218 (33)	159 (16)	56 (8)	165 (17)	205 (31)	88 (9)	32 (5)
Media exposure[a]	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001
No exposure	1186 (67)	1080 (70)	1138 (65)	626 (40)	1129 (64)	847 (55)	852 (48)	498 (32)
Any exposure	1355 (43)	969 (49)	1152 (37)	422 (21)	1180 (37)	787 (40)	692 (22)	277 (14)
Accessing health care[b]	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001
Not a problem	875 (42)	740 (52)	766 (37)	358 (25)	781 (37)	608 (43)	469 (23)	253 (18)
Big problem	1666 (59)	1309 (63)	1524 (54)	690 (33)	1528 (54)	1026 (49)	1075 (38)	522 (25)
Residence	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001
Urban	687 (41)	503 (44)	561 (33)	205 (18)	575 (34)	432 (38)	359 (21)	147 (13)
Rural	1854 (58)	1546 (65)	1729 (54)	843 (36)	1734 (54)	1202 (51)	1185 (37)	628 (27)
Need factors								
Terminated pregnancy	*P*=.02	*P*=.002	*P*=.02	*P*=.07	*P*=.009	*P*=.33	*P*=.09	*P*=.09
No/never	2138 (53)	1712 (59)	2142 (53)	877 (30)	2122 (52)	1349 (47)	1302 (32)	655 (23)
Yes/ever	403 (48)	337 (53)	476 (57)	171 (27)	477 (57)	285 (45)	242 (29)	120 (19)
Desired pregnancy	*P*<.001	*P*=.02	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001	*P*<.001
Yes	1975 (50)	1618 (57)	1718 (44)	781 (28)	1734 (45)	1556 (55)	1130 (29)	580 (21)
No	616 (59)	431 (62)	572 (55)	267 (38)	575 (55)	326 (47)	414 (40)	195 (28)

[a]Media exposure: newspaper or television or radio.

[b]Problem to access health care: permission to go, monetary constraints, or distance to health facilities.

Table 2 shows the adjusted odds ratios (aORs) for factors that are linked to dropping out of all parts of the maternal CoC. It compares BDHS 2017‑2018 and BDHS 2022. Joint Wald tests of the survey-round×covariate interactions indicated significant

XSL•FO
RenderX

disparities in the associations for various key predictors between the 2 survey rounds.

XSL•FO
RenderX

Table . Factors associated with dropouts from the continuum of care among the mothers in Bangladesh from 2017 to 2022.

Covariates	BDHS[a] 2017 - 18		BDHS 2022		Pooled		
	AOR[b] (95% CI)	*P* value	AOR (95% CI)	*P* value	AOR (95% CI)	*P* value	*P* value[c]
Predisposing factors							
Age at delivery (years) (Reference: <19 y)	1	—[d]	1	—	1	—	<.001
19-30	0.72 (0.57-0.90)	.004	0.82 (0.57-1.18)	.29	0.75 (0.62-0.91)	.004	
31-49	0.49 (0.34-0.69)	<.001	0.68 (0.41-1.14)	.15	0.56 (0.42-0.74)	<.001	
Education level (Reference: higher)	1	—	1	—	1	—	<.001
Secondary	1.49 (1.10-2.01)	.009	1.65 (1.10-2.47)	.02	1.52 (1.19-1.94)	.001	
Primary	2.17 (1.52-3.10)	<.001	2.28 (1.46-3.55)	<.001	2.17 (1.64-2.87)	<.001	
No education	3.20 (2.12-4.82)	<.001	2.16 (1.22-3.81)	.008	2.70 (1.94-3.77)	<.001	
Husband's education (Reference: higher)	1	—	1	—	1	—	<.001
Secondary	1.63 (1.19-2.23)	.003	1.47 (1.01-2.15)	.04	1.59 (1.25-2.03)	<.001	
Primary	1.84 (1.34-2.53)	<.001	1.98 (1.32-2.96)	.001	1.91 (1.48-2.44)	<.001	
No education	1.20 (1.38-2.87)	<.001	2.21 (1.46-3.34)	<.001	2.09 (1.59-2.76)	<.001	
Parity (Reference: 1)	1	—	1	—	1	—	<.001
2 - 3	1.60 (1.30-1.95)	<.001	1.73 (1.36-2.20)	<.001	1.64 (1.40-1.91)	<.001	
>3	2.79 (2.04-3.82)	<.001	2.61 (1.70-4.00)	<.001	2.73 (2.12-3.50)	<.001	
Religion (Reference: others)	1	—	1	—	1	—	<.001
Muslim	1.50 (1.08-2.07)	.02	2.42 (1.39-4.23)	.002	1.78 (1.33-2.38)	<.001	
Division (Reference: Dhaka)	1	—	1	—	1	—	<.001
Barisal	1.03 (0.71-1.49)	.88	1.20 (0.78-1.86)	.40	1.11 (0.83-1.48)	.49	
Chittagong	1.34 (0.96-1.85)	.08	1.09 (0.75-1.59)	.66	1.23 (0.96-1.57)	.10	
Khulna	0.47 (0.32-0.71)	<.001	0.40 (0.25-0.66)	.000	0.46 (0.34-0.62)	<.001	
Mymensingh	0.85 (0.57-1.24)	.39	0.83 (0.55-1.26)	.39	0.84 (0.63-0.11)	.23	
Rajshahi	0.78 (0.56-1.09)	.15	0.82 (0.54-1.25)	.36	0.79 (0.61-1.02)	.08	
Rangpur	0.51 (0.34-0.77)	.002	0.71 (0.45-1.10)	.13	0.59 (0.44-0.80)	.001	
Sylhet	1.26 (0.85-1.86)	.25	1.34 (0.92-1.96)	.13	1.31 (0.98-1.74)	.07	
Enabling factors							
Occupation (Reference: working)	1	—	1	—	1	—	.41
Not working	0.94 (0.80-1.10)	.42	0.91 (0.72-1.16)	.46	0.93 (0.81-1.06)	.30	
Husbands' occupation (Reference: working)	1	—	1	—	1	—	.95

XSL•FO
RenderX

Covariates	BDHS[a] 2017 - 18		BDHS 2022		Pooled		
	AOR[b] (95% CI)	*P* value	AOR (95% CI)	*P* value	AOR (95% CI)	*P* value	*P* value[c]
Not working	1.30 (0.59-2.85)	.51	1.04 (0.47-2.32)	.91	1.07 (0.60-1.92)	.81	
Wealth index (Reference: richest)	1	—	1	—	1	—	<.001
Poorest	4.08 (2.82-5.89)	<.001	3.93 (2.40-6.42)	<.001	4.04 (3.02-5.41)	<.001	
Poorer	3.38 (2.39-4.78)	<.001	3.10 (2.01-5.09)	.001	3.32 (2.53-4.35)	<.001	
Middle	2.41 (1.71-3.38)	<.001	2.22 (1.40-3.54)	.001	2.34 (1.79-3.06)	<.001	
Richer	2.16 (1.56-2.99)	<.001	1.83 (1.13-2.96)	.01	2.05 (1.57-2.67)	<.001	
Media exposure[e] (Reference: any exposure)	1	—	1	—	1	—	<.001
No exposure	1.61 (1.33-1.95)	<.001	1.60 (1.32-1.95)	<.001	1.58 (1.38-1.82)	<.001	
Accessing health care[f] (Reference: big problem)	1	—	1	—	1	—	.02
Not a big problem	0.84 (0.71-1.00)	.05	0.92 (0.75-1.11)	.36	0.87 (0.76-0.99)	.04	
Residence (Reference: urban)	1	—	1	—	1	—	<.001
Rural	1.40 (1.12-1.72)	.003	1.44 (1.08-1.94)	.01	1.40 (1.18-1.67)	<.001	
Need factors							
Terminated pregnancy (Reference: no/never)	1	—	1	—	1	—	<.001
Yes/ever	0.79 (0.65-0.95)	.02	0.67 (0.51-0.87)	.003	0.74 (0.63-0.86)	<.001	
Desired pregnancy (Reference: yes/then)	1	—	1	—	1	—	.07
No	1.36 (1.12-1.66)	.002	1.12 (0.87-1.42)	.34	1.27 (1.09-1.48)	.002	
Survey round (Reference: BDHS 2017 - 2018)	—	—	—	—	1	—	—
BDHS 2022	—	—	—	—	0.59 (0.51-0.69)	<.001	

[a]BDHS: Bangladesh Demographic and Health Surveys.

[b]AOR: adjusted odds ratio.

[c]*P* value obtained from joint Wald tests of the interaction coefficients between survey years and covariates.

[d]Not applicable.

[e]Media exposure: newspaper, television, or radio.

[f]Problem to access health care: permission to go, monetary constraints, or distance to health facilities.

The impacts of age at delivery, maternal education, husband's education, parity, religion, division, wealth index, media exposure, health care access, residence, terminated pregnancy, and desired pregnancy exhibited significant variations between survey rounds. Conversely, the associations with maternal occupation (*P*=.40) and husband's occupation (*P*=.95) exhibited no significant differences between BDHS 2017 - 2018 and

XSL•FO
RenderX

BDHS 2022, indicating consistent patterns for these variables over time.

Some important patterns that were specific to the survey are as follows: the protective effect of older age at delivery was stronger in 2017 - 2018 (AOR 0.49, 95% CI 0.34-0.69 for 31–49 years) than in 2022 (AOR 0.68, 95% CI 0.41-1.14); the aOR for Muslim religion went up from 1.50 (95% CI 1.08-2.07) in 2017 - 2018 to 2.42 (95% CI 1.39-4.23) in 2022; and women with no education were at a greater disadvantage in 2017 - 2018 (AOR 3.20, 95% CI 2.12-4.82) than in 2022 (AOR 2.16, 95% CI 1.22-3.81), which suggests that the education gradient is getting smaller. The survey round main effect shows that the odds of dropping out are much lower in BDHS 2022 than in 2017 - 2018 (AOR 0.59, 95% CI 0.51 - 0.69, $P<.001$). This means that the completion of the CoC has improved over time.

In the supplementary materials, we presented the factors associated with dropouts from ANC, SBA, and PNC, separately (see Tables 1, 2, and 3 in Multimedia Appendix 1). In all the regression models, SEs of the odds ratio for all covariates were <2, indicating an acceptable level of multicollinearity among the independent variables. Furthermore, the variance inflation factor of the covariates was less than 3 for both the adjusted and unadjusted models (see Table 4 in Multimedia Appendix 2).

Discussion

Principal Findings

This study investigated the trends in the CoC over the past 5 years and analyzed the factors contributing to dropouts from CoC in Bangladesh, which include ANC, skilled birth attendance, and PNC. Using the Andersen Healthcare Utilization Model, we discovered that predisposing factors (age, education level of the respondent and her husband, parity, desired pregnancy, religion, and survey round), enabling factors (occupation of the respondent and her husband, wealth index, and media exposure), and need factors (desired pregnancy and terminated pregnancy) increased the likelihood of maternal dropout from CoC. From BDHS 2017 - 2018 (31.9%) to BDHS 2022 (22.4%), there was a significant drop in the overall dropout from CoC. This suggests that maternal health service utilization has gotten a little better.

The statistically significant differences noted in the associations between covariates and maternal CoC dropouts across the 2 survey rounds ($P<.05$ in the interaction test) illustrate Bangladesh's evolving maternal health landscape from 2017 to 2022, characterized by policy reforms and unprecedented disruptions in the health system. During this time between surveys, Bangladesh put the National Strategy for Maternal Health 2019 - 2030 into action. This plan focused on making it easier for everyone to get maternal health services and improving community-based care delivery [24]. The COVID-19 pandemic (2020 - 2022) significantly disrupted maternal health service provision and utilization, with studies indicating approximately 30% reductions in ANC, institutional deliveries, and postnatal care in 2020, followed by an incomplete recovery by 2021 [25]. This disruption had a bigger impact on vulnerable groups, such as women from lower wealth quintiles, people who live in rural areas, and people with little education. This could make existing inequalities worse. The substantial interaction effects for enabling factors, including wealth index ($P<.001$), media exposure ($P<.001$), health care access barriers ($P=.02$), and residence ($P<.001$), indicate that the protective or risk-enhancing roles of these determinants fluctuated between survey rounds. This likely reflects the health system's varying capacity to serve different population subgroups during the pandemic and the inconsistent pace of postpandemic recovery across socioeconomic strata [26-28].

These changes in CoC dropout patterns over time can be explained by a number of interconnected barriers. Lower maternal education is frequently correlated with diminished financial independence and decision-making abilities, which subsequently adversely impacts the continuity of health care services for both the mother and her children [29]. The COVID-19 pandemic made financial problems worse. Studies showed that household income dropped by 19% to 20%, and access to maternal health care was disrupted [25,30]. As the cost of living went up, there were more demands on limited household resources. For example, maintaining continuity through the CoC requires multiple visits to health care facilities. This requirement competes with time for household duties and leads to high indirect medical costs, such as transportation and medications [26]. Women from lower socioeconomic backgrounds in Bangladesh bear a disproportionate share of health care financing [31]. This makes the burden of out-of-pocket costs even heavier for them.

To address dropout from maternal health services, Bangladesh has implemented social health protection schemes, such as Shasthyo Surokhsha Karmasuchi and Maternal Health Voucher Scheme, targeting disadvantaged populations [32]. However, evidence suggests that these programs typically fail to monitor the complete maternal CoC, resulting in women discontinuing service uptake after initial contact with 1 component of the continuum [33]. This programmatic limitation highlights the critical gap between initiating maternal health care and maintaining continuity across all phases of pregnancy, delivery, and postpartum care.

Moreover, the notable temporal fluctuations in need factors, especially desired pregnancy status ($P=.07$) and history of terminated pregnancies ($P<.001$), may indicate evolving reproductive health behaviors and family planning dynamics amid this era of economic instability and service interruption [34]. The temporal fluctuations in the strength of associations concerning predisposing factors, such as age at delivery, parity, religion, and division, signify the changing socioeconomic and geographic determinants of maternal health care utilization. The significant interaction for division ($P<.001$) highlights enduring geographic disparities in maternal health care access, with regions such as Khulna and Rangpur exhibiting improved outcomes, whereas others encountered heightened obstacles during the intersurvey period. These ongoing geographic disparities, along with the ongoing difficulty of keeping CoC continuity across all parts, show that Bangladesh needs customized, area-specific interventions to make the full maternal health care pathway stronger [35].

XSL•FO
RenderX

Moreover, the high rate of unintended pregnancies, which are frequently identified late, along with related effects, such as anxiety and depression, discourages women from obtaining CoC in Bangladesh. In addition, late detection inhibits mothers from commencing ANC punctually, resulting in inadequate access to the prescribed number of ANC visits. Our finding that women with unwanted pregnancies had higher odds of dropping out of the CoC is consistent with previous literature linking unintended pregnancy to lower utilization of ANC, skilled delivery, and postnatal services in Bangladesh and other low- and middle-income countries [36,37]. This pattern does not merely indicate a lack of motivation; it likely signifies the presence of multiple intersecting disadvantages. Women with unintended pregnancies are more prone to encounter financial constraints, restricted decision-making authority, stigma or fear of judgment from health care providers, and symptoms of depression or stress, all of which can hinder timely and sustained engagement with services [36].

The study utilized the latest 2 consecutive nationally representative surveys, with high response rates. Therefore, these datasets allow for reliable conclusions about maternal health care utilization patterns across the country within the defined time. In addition, the selection of factors is guided by the Andersen Behavioral Model of Health Care Services, providing a more precise estimate. The evaluation of outcomes and factors followed validated measurements, minimizing the risk of assessment bias. In contrast to the strengths of the study, there remain some challenges to interpreting the findings. First, the cross-sectional design of the study limits establishing a causal nature of the association found between identified factors and CoC components. Second, the possibility of recall bias may affect the accuracy of findings, as the recall period spanned up to 3 years in the datasets. Third, the analyses were restricted to women with complete information on CoC utilization and all covariates, which introduces the possibility of selection bias if women with missing data differ systematically from those included in the analytic sample. Finally, the study did not explore health care system–related factors, such as health care providers' attitudes, patient-provider relationships, availability of workforce, quality of care, and factors underlying the decision-making dynamics of women to uptake CoC services. These contextual elements are critical to fully understanding health care utilization.

This study offers crucial insights into the utilization of CoC in maternal health care, emphasizing the importance of policy support to improve the maternal utilization of CoC, the development of relevant interventions, and the provision of services.

Conclusion

In this study of 8424 mothers, more than 1 in 4 failed to complete the continuum of maternal care, with dropout most pronounced at the ANC stage. Dropout was strongly associated with lower education, higher parity, lower wealth, rural residence, and younger maternal age and was also influenced by media exposure, pregnancy intention, and religious affiliation, underscoring the complex interplay of sociodemographic and experiential factors in maternal health service utilization.

Funding

The authors declared no financial support was received for this work.

Data Availability

Publicly accessible Bangladesh Demographic and Health data were used in this study. The data can be accessed upon reasonable request to the corresponding author.

Authors' Contributions

Conceptualization: SSA, PS, MAR
Data curation: SSA, PS, MAR, RI, IJ
Formal analysis: MAR, SSA, PS
Investigation: SSA, PS, MAR, RI, IJ
Methodology: SSA, PS, MAR, SS
Project administration: SSA, PS, MAR
Resources: SSA, PS, MAR
Software: SSA, PS, MAR, RI, IJ
Supervision: SS, MAR
Validation: SS, MAR, SSA, PS, RI, IJ
Visualization: SS, MAR
Writing — original draft: SSA, PS, MAR, RI, IJ
Writing — review and editing: SS

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Factors associated with dropout from antenatal care, skilled birth attendant delivery, and postnatal care.
[DOCX File, 54 KB - publichealth_v12i1e85875_app1.docx]

Multimedia Appendix 2
Fitness of regression models.
[DOCX File, 32 KB - publichealth_v12i1e85875_app2.docx]

Checklist 1
STROBE checklist.
[DOCX File, 29 KB - publichealth_v12i1e85875_app3.docx]

References

1. Atnafu A, Kebede A, Misganaw B, et al. Determinants of the continuum of maternal healthcare services in Northwest Ethiopia: findings from the primary health care project. J Pregnancy 2020;2020:4318197. [doi: 10.1155/2020/4318197] [Medline: 32908704]
2. Standards for improving quality of maternal and newborn care in health facilities. : World Health Organization; 2016 URL: https://cdn.who.int/media/docs/default-source/mca-documents/qoc/quality-of-care/standards-for-improving-quality-of-maternal-and-newborn-care-in-health-facilities_1a22426e-fdd0-42b4-95b2-4b5b9c590d76.pdf?sfvrsn=3b364d8_4 [accessed 2026-03-11]
3. Saito A, Kondo M. Continuum of care for maternal and child health and child undernutrition in Angola. BMC Public Health 2024 Mar 4;24(1):680. [doi: 10.1186/s12889-024-18144-2] [Medline: 38439029]
4. Ahmed R, Gebre S, Demelash M, et al. The continuum of care for maternal health in Africa: a systematic review and meta-analysis. PLOS ONE 2024;19(7):e0305780. [doi: 10.1371/journal.pone.0305780] [Medline: 39024369]
5. Zelka MA, Yalew AW, Debelew GT. Effectiveness of a continuum of care in maternal health services on the reduction of maternal and neonatal mortality: systematic review and meta-analysis. Heliyon 2023 Jun;9(6):e17559. [doi: 10.1016/j.heliyon.2023.e17559] [Medline: 37408879]
6. Kerber KJ, de Graft-Johnson JE, Bhutta ZA, Okong P, Starrs A, Lawn JE. Continuum of care for maternal, newborn, and child health: from slogan to service delivery. The Lancet 2007 Oct;370(9595):1358-1369. [doi: 10.1016/S0140-6736(07)61578-5]
7. Tinker A, ten Hoope-Bender P, Azfar S, Bustreo F, Bell R. A continuum of care to save newborn lives. The Lancet 2005 Mar;365(9462):822-825. [doi: 10.1016/S0140-6736(05)71016-3]
8. The World Health Report 2005: make every mother and child count. : World Health Organization; 2005 URL: https://iris.who.int/server/api/core/bitstreams/f20eb1c5-e257-4f6b-b91b-62ab2ac3901c/content [accessed 2026-03-11]
9. Sserwanja Q, Mukunya D, Nabachenje P, et al. Continuum of care for maternal health in Uganda: a national cross-sectional study. PLoS ONE 2022;17(2):e0264190. [doi: 10.1371/journal.pone.0264190] [Medline: 35202413]
10. Rifat MA, Islam R, Didar RB, et al. Effects of maternal use of the continuum of care on complementary feeding practices in Bangladesh: cross-sectional study. JMIR Public Health Surveill 2025 Oct 7;11:e76666. [doi: 10.2196/76666] [Medline: 41060072]
11. Bangladesh demographic and health survey 2017-18. : National Institute of Population Research and Training (NIPORT) and ICF; 2020 URL: https://dhsprogram.com/pubs/pdf/FR344/FR344.pdf [accessed 2026-03-11]
12. Shill LC, Tithi SH, Jahan I, Mokbul M, Saha P, Alam SS. Factors associated with utilization of antenatal care services among women in Noakhali District, Bangladesh. Int J Child Health Nutr 2024;13(4):244-250. [doi: 10.6000/1929-4247.2024.13.04.6]
13. Parvin N, Rahman M, Islam MJ, Haque SE, Sarkar P, Mondal MNI. Socioeconomic inequalities in the continuum of care across women's reproductive life cycle in Bangladesh. Sci Rep 2022 Sep 16;12(1):15618. [doi: 10.1038/s41598-022-19888-w] [Medline: 36114413]
14. Rahman R, Rahman M, Haque SE. Empowerment dimensions and their relationship with continuum care for maternal health in Bangladesh. Sci Rep 2021 Sep 21;11(1):18760. [doi: 10.1038/s41598-021-98181-8] [Medline: 34548545]
15. Muluneh AG, Kassa GM, Alemayehu GA, Merid MW. High dropout rate from maternity continuum of care after antenatal care booking and its associated factors among reproductive age women in Ethiopia, evidence from Demographic and Health Survey 2016. PLOS ONE 2020;15(6):e0234741. [doi: 10.1371/journal.pone.0234741] [Medline: 32530948]
16. Bangladesh demographic and health survey 2022. : The DHS Program, ICF; 2024 URL: https://www.dhsprogram.com/pubs/pdf/FR386/FR386.pdf [accessed 2026-03-11]
17. Andersen RM. Revisiting the behavioral model and access to medical care: does it matter? J Health Soc Behav 1995 Mar;36(1):1-10. [Medline: 7738325]
18. Kabir MR. Adopting Andersen's behavior model to identify factors influencing maternal healthcare service utilization in Bangladesh. PLOS ONE 2021;16(11):e0260502. [doi: 10.1371/journal.pone.0260502] [Medline: 34843566]

XSL•FO
RenderX

19. Amin R, Shah NM, Becker S. Socioeconomic factors differentiating maternal and child health-seeking behavior in rural Bangladesh: a cross-sectional analysis. Int J Equity Health 2010 Apr 3;9(1):9. [doi: 10.1186/1475-9276-9-9] [Medline: 20361875]
20. Methun MIH, Ahinkorah BO, Roy S, et al. Inequalities in adequate maternal healthcare opportunities: evidence from Bangladesh Demographic and Health Survey 2017-2018. BMJ Open 2023 Oct 10;13(10):e070111. [doi: 10.1136/bmjopen-2022-070111] [Medline: 37816563]
21. Rahman AE, Perkins J, Islam S, et al. Knowledge and involvement of husbands in maternal and newborn health in rural Bangladesh. BMC Pregnancy Childbirth 2018 Jun 18;18(1):247. [doi: 10.1186/s12884-018-1882-2] [Medline: 29914410]
22. Allison PD. Comparing logit and probit coefficients across groups. Sociol Methods Res 1999 Nov;28(2):186-208. [doi: 10.1177/0049124199028002003]
23. Mahfuzur MR, Billah MA, Liebergreen N, et al. Exploring spatial variations in level and predictors of unskilled birth attendant delivery in Bangladesh using spatial analysis techniques: findings from nationally representative survey data. PLOS ONE 2022;17(10):e0275951. [doi: 10.1371/journal.pone.0275951] [Medline: 36282794]
24. Hossain AT, Hazel EA, Rahman AE, et al. Effective multi-sectoral approach for rapid reduction in maternal and neonatal mortality: the exceptional case of Bangladesh. BMJ Glob Health 2024 May 6;9(Suppl 2):e011407. [doi: 10.1136/bmjgh-2022-011407] [Medline: 38770805]
25. Hossain AT, Akter E, Manna RM, et al. Impact of COVID-19 on the utilisation of maternal health services in Bangladesh: a division-level analysis. J Glob Health 2024 Nov 22;14:05040. [doi: 10.7189/jogh.14.05040] [Medline: 39575613]
26. Methun MIH, Haq I, Uddin MSG, et al. Socioeconomic correlates of adequate maternal care in Bangladesh: analysis of the Bangladesh Demographic and Health Survey 2017-18. Biomed Res Int 2022;2022:8027712. [doi: 10.1155/2022/8027712] [Medline: 36398069]
27. Alam A, Alam S, Mobasshira K, et al. Exploring urban-rural inequalities of maternal healthcare utilization in Bangladesh. Heliyon 2025 Jan 30;11(2):e41945. [doi: 10.1016/j.heliyon.2025.e41945] [Medline: 39911430]
28. Aziz S, Basit A, Sultana S, Homer CSE, Vogel JP. Inequalities in women's utilization of postnatal care services in Bangladesh from 2004 to 2017. Sci Rep 2022 Feb 17;12(1):2747. [doi: 10.1038/s41598-022-06672-z] [Medline: 35177728]
29. Haider MR, Qureshi ZP, Khan MM. Effects of women's autonomy on maternal healthcare utilization in Bangladesh: evidence from a national survey. Sex Reprod Healthc 2017 Dec;14:40-47. [doi: 10.1016/j.srhc.2017.09.002] [Medline: 29195633]
30. Mhajabin S, Hossain AT, Nusrat N, et al. Indirect effects of the early phase of the COVID-19 pandemic on the coverage of essential maternal and newborn health services in a rural subdistrict in Bangladesh: results from a cross-sectional household survey. BMJ Open 2022 Feb 3;12(2):e056951. [doi: 10.1136/bmjopen-2021-056951] [Medline: 35115357]
31. Huq NM, Al-Amin AQ, Howlader SR, Kabir MA. Paying out of pocket for healthcare in Bangladesh - a burden on poor? Iran J Public Health 2015 Jul;44(7):1024-1025. [Medline: 26576387]
32. Sultana N, Hossain A, Das H, et al. Is the maternal health voucher scheme associated with increasing routine immunization coverage? Experience from Bangladesh. Front Public Health 2023;11:963162. [doi: 10.3389/fpubh.2023.963162] [Medline: 36817885]
33. Mahmood SS, Amos M, Hoque S, et al. Does healthcare voucher provision improve utilisation in the continuum of maternal care for poor pregnant women? Experience from Bangladesh. Glob Health Action 2019;12(1):1701324. [doi: 10.1080/16549716.2019.1701324] [Medline: 31825301]
34. Gipson JD, Hossain MB, Koenig MA. Measurement of and trends in unintended birth in Bangladesh, 1983-2000. J Health Popul Nutr 2011 Aug;29(4):400-405. [doi: 10.3329/jhpn.v29i4.8457] [Medline: 21957679]
35. Rahman M, Rabby MF, Sakib MMH, Hossain MM. Examining regional disparities in maternal and child health in Bangladesh using cluster analysis of MICS 2019 data. Sci Rep 2025;15(1):32741. [doi: 10.1038/s41598-025-03003-w]
36. Khan MN, Harris ML, Huda MN, Loxton D. A population-level data linkage study to explore the association between health facility level factors and unintended pregnancy in Bangladesh. Sci Rep 2022 Sep 7;12(1):15165. [doi: 10.1038/s41598-022-19559-w] [Medline: 36071170]
37. Mekonnen BD, Vasilevski V, Bali AG, Sweet L. Effect of pregnancy intention on completion of maternity continuum of care in Sub-Saharan Africa: systematic review and meta-analysis. BMC Pregnancy Childbirth 2024 Nov 28;24(1):802. [doi: 10.1186/s12884-024-06998-8] [Medline: 39609727]

Abbreviations

ANC: antenatal care
AOR: adjusted odds ratio
BDHS: Bangladesh Demographic and Health Surveys
CoC: continuum of care
PNC: postnatal care
SBA: skilled birth attendant
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 14.Oct.2025; peer-reviewed by E Kendal, S Jabeen; revised version received 21.Jan.2026; accepted 25.Feb.2026; published 01.Apr.2026.

Please cite as:
Alam SS, Sarkar P, Rifat MA, Jahan S, Islam R, Jahan I, Saha S
Uncovering the Reasons Behind Maternal Care Dropout in Bangladesh: Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e85875
URL: https://publichealth.jmir.org/2026/1/e85875
doi:10.2196/85875

© Syeda Saima Alam, Plabon Sarkar, M A Rifat, Sumaiya Jahan, Rokibul Islam, Israt Jahan, Sanjib Saha. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 1.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring Motives for Reducing Alcohol Consumption Among Users of an Alcohol Reduction App: Content Analysis

Lucille P Seppi[1]; Olivia M Maynard[1], PhD; Dimitra Kale[2], PhD; John A Cunningham[3,4], PhD; Claire Garnett[1,2], PhD

[1]School of Psychological Science, University of Bristol, 12a Priory Road, Bristol, United Kingdom

[2]Department of Behavioural Science and Health, University College London, London, United Kingdom

[3]Department of Addictions Sciences, King's College London, London, United Kingdom

[4]Centre for Addiction and Mental Health, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Claire Garnett, PhD
School of Psychological Science, University of Bristol, 12a Priory Road, Bristol, United Kingdom

Abstract

Background: It is important to understand the motives behind why people want to reduce their alcohol consumption to inform messaging for public health campaigns.

Objective: This study aimed to identify the motives for reducing alcohol consumption among users of the Drink Less app in the United Kingdom.

Methods: Content analysis of Drink Less app users' responses to the prompt "I want to drink less because..." Users were aged 18 years or older, lived in the United Kingdom, and had downloaded the app between May 20, 2016, and June 24, 2024 (n=2520). Inductive content analysis was conducted to analyze users' motives to drink less, and the frequencies were stratified by age, sex, and Alcohol Use Disorders Identification Test (AUDIT) score categories.

Results: The most common motives to drink less were wanting to improve their physical health (1329/2520, 52.7%), feel better in their body (823/2520, 32.7%), improve their mental well-being (567/2520, 22.5%), regain agency (483/2520, 19.2%), live a different life (321/2520, 12.7%), and have better relationships (309/2520, 12.3%). The motives for drinking less,"improve physical health" and "feel better in their body," had a lower prevalence among users in higher AUDIT risk zones compared with low-risk, while wanting to "improve their mental well-being," "regain agency," "live a different life," and "have better relationships" had a higher prevalence among users in higher AUDIT risk zones.

Conclusions: Users of an alcohol reduction app in the United Kingdom most commonly reported wanting to improve their physical health, feel better in their bodies, and improve their mental well-being as their motives for drinking less alcohol. The pattern of prevalence of different motives varied by AUDIT risk zones, highlighting the importance of tailoring interventions based on what is most likely to be motivating for individuals.

(JMIR Public Health Surveill 2026;12:e88992) doi:10.2196/88992

KEYWORDS

alcohol; motives to reduce; drinkers; United Kingdom; content analysis

Introduction

Alcohol consumption is a leading risk factor for early mortality, ill health, and disability in the United Kingdom [1], as well as contributing to health inequalities [2]. There is a dose-response relationship between the volume of alcohol consumed and the risk of alcohol-related harm (eg, stroke) [3], meaning anyone who drinks alcohol can benefit from reducing their consumption. It is important to understand the motives behind why people want to reduce their alcohol consumption to inform messaging for public health campaigns and create targeted guidance based on high-risk groups.

As of December 2025, around a third of risky drinkers had made an attempt in the past year to cut down or stop drinking, and a third believed they should or wanted to cut down [4]. Motivation is an important part of behavior change [5], and there is evidence that high motivation levels predict whether someone makes an attempt to reduce their consumption [6]. Motivation is commonly addressed in guidance documents and treatment manuals for brief alcohol interventions, and in the Cochrane Review of brief alcohol interventions in the form of identifying reasons for wanting to reduce alcohol consumption [7]. Having personally relevant motives may help to increase the level of motivation and lead more people to make a reduction attempt. Therefore, it is important to understand, among those who are motivated to make a reduction attempt, what the motive behind this was. This could be used to inform public health campaigns,

XSL•FO
RenderX

with common motives highlighted to encourage others to make a reduction attempt as well.

Previous research has looked at understanding people's motives for reducing their alcohol consumption. Adults in England reported motives to improve fitness, weight loss, future health concerns, social factors, cost, advice from a health care professional, and fewer social occasions [8-10]. Adults in Australia reported that health, lifestyle, and social motives were most common [11], alongside fear of the negative consequences among younger age groups [12,13]. People drinking at heavier levels reported motives such as major life changes, weighing up pros and cons of drinking and not drinking, health and financial concerns, and being affected by seeing someone drunk or high [14,15].

However, there are some gaps in the previous research. Some studies only included risky drinkers, so no one who had reduced their drinking to low-risk levels was asked about their motives [8,9]. In the cases of population surveys [8-11], very heavy drinkers tend to be underrepresented [16] despite being at greater risk of alcohol-related harm. In most previous research, participants were asked to select their motives for drinking less from a list of response options [8-12,14], and the process of selecting or deriving the response options was unclear. And finally, the timing of the data collection in these studies was either retrospective, introducing scope for recall bias [8-11], or participants were asked to report why they were considering changing their drinking behavior [15].

The Drink Less app is a theory-based intervention to support users in reducing their alcohol consumption in the United Kingdom that has been developed and refined in a systematic way involving user feedback [17,18] with evidence for its effectiveness [19]. It is widely used [20] and popular among its users, with a 4.6-star rating on the UK Apple App Store as of February 2026 [21]. Drink Less is one of the top apps for alcohol-related searches and has over 90,000 unique users [21]. The app collects anonymous data on users' motives for drinking less in the form of an open-text response to the prompt "I want to drink less because..." These data can provide insight on motives to reduce alcohol consumption among UK drinkers at the moment at which they are making such an attempt. The aim of this study was to identify the Drink Less users' motives for reducing their alcohol consumption.

Methods

Study Design

Qualitative analysis of previously collected short-text responses to the prompt "I want to drink less because..." among users of the Drink Less app. The study protocol was preregistered on the Open Science Framework (osf.io/gnb8p) prior to data analysis starting.

Participants

The Drink Less app is freely available on the UK Apple App Store [21]. App users had to agree to the app's "Privacy Policy" and "Terms and Conditions" for their anonymous data to be used for scientific research purposes. In addition to this, for this study, users had to be aged 18 years or older, live in the United Kingdom, and have downloaded the Drink Less app between May 18, 2016, and July 20, 2024 (there was no minimum engagement required) to be included in this study. In cases of duplicate device IDs, only the first download was included.

The Drink Less App

The Drink Less app consists of 8 evidence-based modules to help users change their drinking behavior: Goal Setting, Self-Monitoring and Feedback, Action Planning, Normative Feedback, Cognitive Bias Re-Training, Insights, Behavioral Substitution, and Information About Antecedents [17,18]. The app was designed to help people reduce their alcohol consumption though there was no explicit user motive for reducing alcohol consumption that it was designed for or targeted on. Users could access all parts of the app at all times. The onboarding process (when first downloading the app) involves users agreeing to the app's "Privacy Policy" and "Terms and Conditions" (they can opt out at any time, and if they do not agree, they can still use the app but not have their data used for scientific research purposes). During the time period of this study, research studies were conducted within the app (eg, [19,22]) though none of these affected the functionality or visibility of the "I want to drink less because..." prompt.

Measures

Motive to Drink Less

All users received the prompt "I want to drink less because..." on the app dashboard and when setting goals though did not have to complete it. There was an open-text box for users to respond. If users changed their original response, the most recent version was used. There was no time limit on when the response had to be made by.

Demographic Characteristics

During app onboarding, users input their sex (male/female) and their year of birth (which was used to calculate their age). There were no other additional sociodemographic characteristics measured consistently across the time period to minimize participant burden and increase engagement with the app.

Alcohol Use Disorders Identification Test (AUDIT) Score

Users completed the AUDIT, a 10-item questionnaire with total scores ranging from 0 to 40 categorized into 4 risk zones [23]: zone I (score 0 - 7, low-risk); zone II (score 8 - 15, increasing risk); zone III (score 16 - 19, higher risk); and zone IV (score 20 - 40, at risk of alcohol dependence). This was completed during onboarding (ie, at the time of app download).

Data Analysis

An inductive content analysis was conducted to analyze the users' motives to drink less, allowing for large amounts of textual data to be explored qualitatively, identifying categories which can be quantified in terms of frequency [24]. This allows for patterns and meanings to be reported in a simpler manner with the aim of representing new insights and knowledge practically, guiding future action [25].

Researcher LPS familiarized themselves with the data and coded it inductively in Microsoft Excel using an iterative process that led to the development of *subthemes*, which came together to

XSL•FO
RenderX

form *themes* and ultimately into *categories*. The criteria for selection were broad in that everything mentioned was considered as a motivation and not discounted for not fitting the previous literature, considering that our aim was to inductively explore motivations to drink less. Furthermore, the analysis was limited to a descriptive level so as to not become too abstract or look for underlying meaning due to both the volume of data and the short-form text (which cannot provide context or insight into potential underlying meanings in the way long-form text or interviews can) [26]. Each user response could be coded multiple times. Once an initial codebook came together, LPS met with CG and OMM to discuss it, for example, whether anything had been overlooked or something may have more than one meaning, after which LPS returned to the coding individually, shaped by the discussion. As all available data on motives to drink less were coded, inductive thematic saturation (when new codes or themes are rarely or never generated) [27] was not considered as a marker at which to stop coding, though it was reached before the end of the coding process. Each response was coded by its content, up until the last response. The prevalence of users reporting each subtheme, theme, and category was quantified.

The frequency of categories and themes was also stratified by age group, sex, and AUDIT risk zones. There was a change to the preregistered protocol in that subthemes were not stratified by age group, sex, and AUDIT risk zones due to very small cell sizes.

Ethical Considerations

Ethics approval for the use of the Drink Less data for research purposes has been granted by University College London (CEHP/2020/579) and the University of Bristol (#23313) ethics committees. This particular study was approved by the School of Psychological Science Research Ethics Committee at the University of Bristol (#22068). The app has a privacy notice, and users provided consent for their anonymized data to be used for academic research purposes after downloading the app. Users can opt out from their data being used for academic research purposes within the app at any point. Participants were not compensated for taking part in this study.

Results

Overview

There were 93,640 unique users of the Drink Less app between May 20, 2016, and June 24, 2024, and 2550 (2.7%) of these users had responded to the prompt “I want to drink less because...” Of these, 30 responses were excluded due to being uninterpretable (eg, incoherent, not in English) (n=26), the user being under 18 years (n=3), or having missing demographic data (n=1).

The final sample consisted of 2520 unique user responses. The majority of users were between the ages of 35 and 54 years (n=1488, 59.1%) and identified as female (n=1594, 63.3%). Fewer than 5% (n=123) of users were categorized as a low-risk drinker, 57.2% (n=1441) as an increasing or higher risk drinker, and 37.9% (n=956) as at risk of alcohol dependence (see Table 1). The sample (ie, users completing the prompt) had a higher proportion of users who were younger (less than 54 years old), a higher proportion of female users, and a higher proportion of users at higher risk or risk of alcohol dependence compared with those not completing the prompt (see Table 1).

XSL•FO
RenderX

Table . Characteristics of users of the Drink Less app (2016-2024), including all users (n=93,640) and sample users (n=2520).

Characteristics	All users (n=93,640)	Users not completing the prompt (n=91,120)	Sample users (n=2520)	Chi-square (*df*)	*P* value
Age (years), n (%)				51.1 (5)	<.001
18 - 24	4124 (4.4)	4038 (4.4)	86 (3.4)		
25 - 34	15,045 (16.1)	14,614 (16.0)	431 (17.1)		
35 - 44	24,143 (25.8)	23,405 (25.7)	738 (29.3)		
45 - 54	27,295 (29.1)	26,545 (29.1)	750 (29.8)		
55 - 64	17,170 (18.3)	16,748 (18.4)	422 (16.7)		
65+	5863 (6.3)	5770 (6.3)	93 (3.7)		
Sex, n (%)				202.6 (1)	<.001
Male	47,522 (50.7)	46,596 (51.1)	926 (36.7)		
Female	46,118 (49.3)	44,524 (48.9)	1594 (63.3)		
AUDIT[a] risk zone, n (%)				290.4 (3)	<.001
I (score 0-7; low risk)	10,679 (11.4)	10,556 (11.6)	123 (4.9)		
II (score 8-15; increasing risk)	41,614 (44.4)	40,701 (44.7)	913 (36.2)		
III (score 16-19; higher risk)	17,062 (18.2)	16,534 (18.1)	528 (21.0)		
IV (score 20-40; at risk of alcohol dependence)	24,285 (25.9)	23,329 (25.6)	956 (37.9)		

[a]AUDIT: Alcohol Use Disorders Identification Test.

Of the sample users, the time difference between downloading the app and the date of the last update to users' drink less motive had a median value of 0.9 (IQR 0.6 - 6.6), indicating that over half of users reported their drink less motive within a day and three-quarters had reported their motive within 7 (mean 16.3, SD 57.8) days.

There were 19 categories, 29 themes, and 96 subthemes. On average, each response was coded 2.1 times, resulting in 5362 codes. The frequency of each category and theme is reported in Table 2, and the frequency of each category stratified by age, sex, and AUDIT risk zone is reported in Table 3. The frequency of users reporting each category and theme, and stratified by age, sex, and AUDIT risk zone, is reported in Multimedia Appendix 1. The frequency of subthemes is reported in Multimedia Appendix 2.

XSL•FO
RenderX

Table . Frequency of categories and themes of motives to drink less reported by users of the Drink Less app (2016-2024; sample users: n=2520).

Categories and themes[a]	Sample users (n=2520), n (%)
I want to improve my physical health	1329 (52.7)
I want to live a healthier lifestyle	1129 (85.0)
I am worried about future health problems	180 (13.6)
I have other health problems	72 (5.4)
I want to feel better in my body	823 (32.7)
Help with weight loss	718 (87.2)
Improve my fitness	113 (13.7)
I want to improve how I look	48 (5.8)
I want to improve the features on my body	32 (3.9)
I want to feel more attractive/sexy	3 (0.4)
I want to improve my mental well-being	567 (22.5)
I want to improve my overall well-being	346 (61.0)
Improve mental health	198 (34.9)
I want to feel clearer-headed, more alert, focused/be present	87 (15.3)
I want to regain agency	483 (19.2)
I want to be in control of my drinking	330 (68.3)
I want to be in control (of my actions, life...)	114 (23.6)
I do not want drinking to define me	93 (19.3)
I want to live a different life	321 (12.7)
I want to make space in my life for other things	136 (42.4)
I want to live (longer)	91 (28.3)
I want to be more...	83 (25.9)
I want to improve my quality of life	46 (14.3)
To have better relationships with the people in my life	309 (12.3)
I am doing it for someone	261 (84.5)
Drinking affects my relationships/social life	48 (15.5)
My drinking affects the way I treat and impact people	21 (6.8)
I want to improve my sex life	4 (1.3)
A decision that drinking was too expensive	228 (9.1)
Restore my energy	212 (8.4)
I do not want to (re)experience the side effects/consequences	162 (6.4)
Drinking hurts me	94 (3.7)
I want to have a better relationship with alcohol	64 (2.5)
I want to moderate my drinking, not abstain	61 (95.3)
Temporary abstinence (eg, Dry January)	4 (6.3)
I have changed my attitude about drinking	43 (1.7)
To support or encourage people in my life to drink less/have a better relationship with alcohol	43 (1.7)
Improve work life	40 (1.6)
I am planning for my future (eg, starting a family/finding a partner/planning for retirement)	36 (1.4)
Future family planning	19 (52.8)

XSL•FO
RenderX

Categories and themes[a]	Sample users (n=2520), n (%)
Planning for retirement/old age	17 (47.2)
Other	30 (1.2)
Someone or something inspired me to	28 (1.1)
Advice or concern	26 (92.3)
Media	2 (7.1)
I have seen the effects drinking has had on other people	17 (0.7)
I do not want to end up like someone I know	17 (100.0)
A significant holiday or event (eg, birthday or upcoming trip)	17 (0.7)

[a]The percentage of users for all categories and for themes within each category may be greater than 100% because each response could be coded multiple times.

XSL•FO
RenderX

Table . Frequency of categories of motives to drink less reported by users of the Drink Less app, and stratified by sociodemographic (age and sex) and drinking characteristics (AUDIT[a] risk zone) (sample users: n=2520)[b].

		Age in years, n (%)						Sex, n (%)		AUDIT risk zone[c], n (%)			
Category	All users (n=2520), n (%)	18 - 24 (n=86)	25 - 34 (n=431)	35 - 44 (n=738)	45 - 54 (n=750)	55 - 64 (n=422)	65+ (n=93)	Male (n=926)	Female (n=1594)	I (n=123)	II (n=913)	III (n=528)	IV (n=956)
I want to improve my physical health	1329 (52.7)	30 (34.9)	194 (45.0)	398 (53.9)	391 (52.1)	257 (60.9)	59 (63.4)	480 (51.8)	849 (53.3)	76 (61.8)	505 (55.3)	291 (55.1)	457 (47.8)
I want to feel better in my body	823 (32.7)	19 (22.1)	116 (26.9)	247 (33.5)	271 (36.1)	151 (35.8)	19 (20.4)	251 (27.1)	572 (35.9)	48 (39.0)	322 (35.3)	182 (34.5)	271 (28.3)
I want to improve my mental well-being	567 (22.5)	23 (26.0)	124 (28.8)	181 (24.5)	177 (23.6)	49 (11.6)	13 (14.0)	171 (18.5)	396 (24.8)	13 (10.6)	185 (20.3)	126 (23.9)	243 (25.4)
I want to regain agency	483 (19.2)	24 (27.9)	93 (21.6)	130 (17.6)	139 (18.5)	80 (19.0)	17 (18.3)	146 (15.8)	337 (21.1)	11 (8.9)	142 (15.6)	98 (18.6)	232 (24.3)
I want to live a different life	321 (12.7)	8 (9.3)	67 (15.5)	117 (15.9)	83 (11.1)	39 (9.2)	7 (7.5)	119 (12.9)	202 (12.7)	6 (4.9)	94 (10.3)	69 (13.1)	152 (15.9)
To have better relationships with the people in my life	309 (12.3)	8 (9.3)	60 (13.9)	133 (18.0)	81 (10.8)	21 (5.0)	6 (6.4)	136 (14.7)	173 (10.9)	6 (4.9)	59 (6.5)	65 (12.3)	179 (18.7)
A decision that drinking was too expensive	228 (9.1)	8 (9.3)	44 (10.2)	80 (10.8)	61 (8.1)	32 (7.6)	3 (3.2)	90 (9.7)	138 (8.7)	8 (6.5)	68 (7.4)	58 (11.0)	94 (9.8)
Restore my energy	212 (8.4)	5 (5.8)	32 (7.4)	61 (8.3)	77 (10.3)	30 (7.1)	7 (7.5)	51 (5.5)	161 (10.1)	9 (7.3)	90 (9.9)	47 (8.9)	66 (6.9)
I do not want to (re)experience the side effects/consequences	162 (6.4)	7 (8.1)	30 (7.0)	52 (7.0)	45 (6.0)	21 (5.0)	7 (7.5)	35 (3.8)	127 (8.0)	5 (4.1)	42 (4.6)	43 (8.1)	72 (7.5)
Drinking hurts me	94 (3.7)	5 (5.8)	16 (3.7)	28 (3.8)	25 (3.3)	12 (2.8)	8 (8.6)	36 (3.9)	58 (3.6)	3 (2.4)	21 (2.3)	20 (3.8)	50 (5.2)

XSL•FO
RenderX

		Age in years, n (%)						Sex, n (%)		AUDIT risk zone[c], n (%)			
Category	All users (n=2520), n (%)	18 - 24 (n=86)	25 - 34 (n=431)	35 - 44 (n=738)	45 - 54 (n=750)	55 - 64 (n=422)	65+ (n=93)	Male (n=926)	Female (n=1594)	I (n=123)	II (n=913)	III (n=528)	IV (n=956)
I want to have a better relationship with alcohol	64 (2.5)	3 (3.5)	9 (2.1)	21 (2.8)	17 (2.3)	11 (2.6)	3 (3.2)	20 (2.2)	44 (2.8)	4 (3.3)	23 (2.5)	13 (2.5)	24 (2.5)
I have changed my attitude about drinking	43 (1.7)	2 (2.3)	4 (0.9)	15 (2.0)	18 (2.4)	1 (0.2)	3 (3.2)	19 (2.1)	24 (1.5)	3 (2.4)	14 (1.5)	11 (2.1)	15 (1.6)
To support or encourage people in my life to drink less/have a better relationship with alcohol	43 (1.7)	0 (0.0)	5 (1.2)	20 (2.7)	17 (2.3)	1 (0.2)	0 (0.0)	13 (1.4)	30 (1.9)	0 (0.0)	18 (2.0)	13 (2.5)	12 (1.3)
Improve work life	40 (1.6)	3 (3.5)	12 (2.8)	15 (2.0)	9 (1.2)	1 (0.2)	0 (0.0)	17 (1.8)	23 (1.4)	1 (0.8)	8 (0.9)	12 (2.3)	19 (2.0)
I am planning for my future (eg, starting a family/finding a partner/planning for retirement)	36 (1.4)	1 (1.2)	7 (1.6)	14 (1.9)	8 (1.1)	6 (1.4)	0 (0.0)	13 (1.4)	23 (1.4)	1 (0.8)	16 (1.8)	7 (1.3)	12 (1.3)
Other	30 (1.2)	1 (1.2)	6 (1.4)	5 (0.7)	11 (1.5)	6 (1.4)	1 (1.1)	17 (1.8)	13 (0.8)	1 (0.8)	15 (1.6)	8 (1.5)	6 (0.6)
Someone or something inspired me to	28 (1.1)	2 (2.3)	6 (1.4)	6 (0.8)	11 (1.5)	2 (0.5)	1 (1.1)	12 (1.3)	16 (1.0)	3 (2.4)	7 (0.8)	5 (0.9)	13 (1.4)
I have seen the effects drinking has had on other people	17 (0.7)	2 (2.3)	3 (0.7)	2 (0.3)	9 (1.2)	1 (0.2)	0 (0.0)	4 (0.4)	13 (0.8)	1 (0.8)	3 (0.3)	4 (0.8)	9 (0.9)

XSL•FO
RenderX

		Age in years, n (%)						Sex, n (%)		AUDIT risk zone[c], n (%)			
Category	All users (n=2520), n (%)	18 - 24 (n=86)	25 - 34 (n=431)	35 - 44 (n=738)	45 - 54 (n=750)	55 - 64 (n=422)	65+ (n=93)	Male (n=926)	Female (n=1594)	I (n=123)	II (n=913)	III (n=528)	IV (n=956)
A significant holiday or event (eg, birthday or upcoming trip)	17 (0.7)	0 (0.0)	8 (1.9)	3 (0.4)	5 (0.7)	1 (0.2)	0 (0.0)	5 (0.5)	12 (0.8)	0 (0.0)	6 (0.7)	5 (0.9)	6 (0.6)

[a]AUDIT: Alcohol Use Disorders Identification Test.

[b]The percentage of users for all categories may be greater than 100% because each response could be coded multiple times.

[c]Zone I (score 0-7), zone II (score 8-15), zone III (score 16-19), and zone IV (score 20-40).

Category 1: I Want to Improve My Physical Health

Over half of users reported wanting to improve physical health (1329/2520, 52.7%), and most responses were under the theme "wanting to live a healthier lifestyle" (n=1129, 85%; such as unspecified goals to take care of one's general health or specific goals like lowering one's cholesterol), in contrast with being "worried about future health problems" (n=180, 13.6%; specific concerns such as future liver failure or cancer). The least common theme was having other health problems, which directly inspired a reduction attempt, including motives to stop the exacerbation of health problems caused through or impacted on by alcohol. This theme also described motives where alcohol reduction was not the primary goal, such as reducing consumption to aid recovery from something else.

The "improving physical health" category was reported by users across all age groups, increasing in prevalence from 34.9% (30/86) among 18 - 24 year olds to 63.4% (59/93) among those over 65 years old. This category had similar prevalence among male and female users. The prevalence of users reporting wanting to improve physical health was higher among those in AUDIT risk zone I (76/123, 61.8%), with prevalence at a lower level (457/956, 47.8%) among those in risk zone IV.

Category 2: I Want to Feel Better in My Body

Nearly a third of users (823/2520, 32.7%) reported wanting to feel better in their body. The most common theme was "Help with weight loss" (718/823, 87.2%), followed by "Improve my fitness" (113/823, 13.7%), consisting of either a general motive to do so or wanting to improve in a specific sport or exercise. The other themes were "I want to improve how I look" (48/823, 5.8%), "I want to improve the features on my body" (32/823, 3.9%), mainly referring to the skin, and "I want to feel more attractive/sexy" (3/823, 0.4%).

All age groups reported wanting to feel better in their body (20.4% to 36.1%), with prevalence highest across age groups 35 - 64 years old. More female users reported wanting to feel better in their body (572/1594, 35.9%) compared with male users (251/926, 27.1%). The prevalence of reporting wanting to feel better in their body was highest among AUDIT risk zone I (48/123, 39%) and lowest in prevalence among risk zone IV (271/956, 28.3%).

Category 3: I Want to Improve My Mental Well-Being

Around a quarter of users reported wanting to improve their mental well-being (567/2520, 22.5%), with the most common theme being to improve their overall well-being (346/567, 61%), mainly wanting to feel better or not wanting drinking to impact on their mood. Some also specified wanting a better relationship with themselves as well as to feel less irritable, stressed, or frustrated. The other common theme was improving mental health (198/567, 34.9%), which specified motives to reduce the symptoms of mental health conditions such as anxiety and depression, and a general desire to improve mental health.

It was more common to report wanting to improve mental well-being as a motive to drink less among users in the age groups of 18 - 54 years (over 23.6%). The prevalence of reporting wanting to improve mental well-being was slightly higher among female users (396/1594, 24.8%) compared with male users (171/926, 18.5%). Reporting wanting to improve mental well-being was highest among users in AUDIT risk zone IV (243/956, 25.4%) and lowest in prevalence among risk zone I (13/123, 10.6%).

Category 4: I Want to Regain Agency

Wanting to regain agency was reported by 19.2% (483/2520) of users, with the most common theme being "I want to be in control of my drinking" (330/483, 68.3%), which mainly consisted of a fear of becoming or being dependent as well as feeling guilty or ashamed about one's drinking or expressing concern about usually drinking more than intended. A small percentage of users wanted to make themselves or others proud or prove that they could reduce their consumption, whereas others felt that they "have to" or "need to" reduce for unspecified reasons.

Within this category, a quarter of users (114/483, 23.6%) expressed wanting to be in control over more than just their drinking, such as their actions or life. This consisted of users reporting wanting to be more responsible/reliable and the feeling that alcohol prevents them from making "good" choices. Lastly,

XSL•FO
RenderX

some users mentioned that they did not want drinking to define them (93/483, 19.3%), which mainly referred to issues of identity (ie, feeling more like oneself without drink), but also not wanting their drinking to embarrass them or others, and not wanting others to worry about them.

Over a quarter of users aged 18 - 24 years reported this category of wanting to regain agency (24/86, 27.9%) compared with slightly lower prevalence (17.6% to 21.6%) among other age groups. More female users (337/1594, 21.1%) reported wanting to regain agency as a motive to drink less compared with 15.8% (146/926) of male users. Reporting wanting to regain agency was highest among users in AUDIT risk zone IV (232/956, 24.3%) and lowest in risk zone I (11/123, 8.9%).

Category 5: I Want to Live a Different Life

Wanting to live a different life was reported by 12.7% (321/2520) of users. The most common theme was "I want to make space in my life for other things" (136/321, 42.4%), then "I want to live (longer)" (91/321, 28.3%) and "I want to be more" (83/321, 25.9%) of a certain quality, such as productive, successful, or resilient.

This category was more prevalent among users aged 25 - 44 years (~15%) compared with other age groups and was similarly prevalent between male (119/926, 12.9%) and female users (202/1594, 12.7%). The prevalence of wanting to live a different life was highest among users in the AUDIT risk zone IV (152/956, 15.9%) and lowest among those users in risk zone I (6/123, 4.9%).

Category 6: To Have Better Relationships With the People in My Life

Wanting to have better relationships with the people in their life as a motive to drink less was reported by 12.3% (309/2520) of users. The main theme of this category was "I am doing it for someone" (261/309, 84.5%), such as family, partners, friends, or someone unspecified. Other themes included "Drinking affects my relationships/social life" (48/309, 15.5%) and "My drinking affects the way I treat and impact people" (21/309, 6.8%).

The prevalence of wanting to have better relationships with the people in their life was higher among the 25 - 34 and 35 - 44 years age groups (60/431 [13.9%] and 133/738 [18%], respectively) and lower among other age groups. A higher proportion of male users reported this motive to drink less (136/926, 14.7%) compared with female users (173/1594, 10.9%). The prevalence of wanting to have better relationships with the people in their life was highest among users in AUDIT risk zone IV (179/956, 18.7%) and lowest among users in risk zone I (6/123, 4.9%).

Other Categories

All other categories for motives to drink less were mentioned by fewer than 10% of users: deciding that drinking was too expensive (228/2520, 9.1%), restoring energy (212/2520, 8.4%), not wanting to experience the side effects/consequences (162/2520, 6.4%), drinking hurts me (94/2520, 3.7%), wanting to have a better relationship with alcohol (64/2520, 2.5%), changed attitude about drinking (43/2520, 1.7%), supporting others to drink less (43/2520, 1.7%), improve work life (40/2520, 1.6%), planning for future (36/2520, 1.4%), other (30/2520, 1.2%), being inspired to (28/2520, 1.1%), seen the effects drinking has on others (17/2520, 0.7%), and a significant event (17/2520, 0.7%).

Discussion

Main Findings

The most common motives to reduce drinking among users of the Drink Less app in the United Kingdom were wanting to improve their physical health, feel better in their body, improve their mental well-being, and regain agency. These participants, by the nature of the sample, had all taken some action to download and start using an evidence-based digital intervention with evidence for its effectiveness [19], suggesting that these individuals were in either the "action" stage of change or perhaps in "preparation" [28], having downloaded the app but may not have planned on using it immediately. Wanting to improve their physical health was reported as a motive to drink less by over half of users, and this was present across all age groups though increased with age. Nearly a third of users reported wanting to feel better in their body (eg, weight loss and improved fitness), and this was more common among female users. Interestingly, there were clear patterns for the prevalence of motives across users' AUDIT risk zones. "Improving physical health" and "wanting to feel better in their body" were highest among users drinking at low-risk levels and decreased in prevalence with higher AUDIT risk zones (though in absolute terms, these motives were still higher than for other motives for drinking less).

In contrast, motives to "improve their mental well-being," "regain agency," "live a different life," and "have better relationships" all had the highest prevalence among users at risk of alcohol dependence and the lowest among those drinking at low risk. About a quarter of users reported wanting to improve their mental well-being, and this motive was reported more among users aged between 18 and 54 years and was slightly more common among female users. Around a fifth of users reported wanting to regain agency, mainly reporting wanting to be in control of their drinking and being worried about dependency. About a tenth reported wanting to live a different life and to have better relationships with the people in their life.

Comparison With Other Studies

The findings of this study that physical health and wanting to feel better in their body were common motives for drinking less is consistent with previous research [8-11,14,15,29], including that these motives were more commonly reported by older age groups. This study found that improving mental health was a relatively common motive to drink less though this motive is not typically seen across previous research, where there tends to be a focus on physical health, or health more generally. There were also some motives that came up in previous research, such as fewer social occasions [10] and treatment or therapy [15], that were not common in this study. This could be due to differences in the samples with participants in the current study being users of an alcohol reduction app and therefore consciously trying to reduce their alcohol consumption. The

XSL•FO
RenderX

motive to drink less due to worries around dependency tended not to be included as a response option in previous research but came up in this study and another study where participants could give free responses [15]. The samples in these two studies both included heavier drinkers, highlighting the importance of sampling heavier drinkers when assessing motives for drinking less as an important group to consider for tailoring public health interventions.

This study found motives to drink less that were about seeking a positive outcome (eg, wanting to feel happier and have better relationships with themselves) and not just avoiding a negative outcome (eg, not wanting to experience the consequences and worries about dependency). Another study asking participants about their motives to drink less without using preset response options found similarly positive reasons for reducing alcohol consumption [15].

A previous study found differences in motives based on the level of severity of alcohol use with "consequence-related" reasons (such as health problems and financial concerns) more common among those with more severe alcohol use, and "drifting out" reasons (changes in circumstances such as having children) more common among those with less severe alcohol use [30]. While the current study saw different patterns in motives to drink less based on AUDIT risk zone, it did not follow the same pattern as the previous study with physical health and feeling better in their body (both "consequence-related" reasons) being more common among app users with low-risk AUDIT zones.

Strengths and Limitations

A major strength of this study was that data collection was remote and in a naturalistic setting, meaning these findings are likely to have high external validity. Participants were also those that were currently trying to reduce their alcohol consumption and prompted to enter their motive (as free text) for drinking less when they downloaded the app, avoiding issues around recall bias. The app is also a stand-alone intervention, so no researcher was present, avoiding potential fear of judgment or social conformity [31]. Data were also collected from users with a range of ages and AUDIT scores, therefore, capturing data from heavier drinkers and people at risk of alcohol dependence, which is sometimes underrepresented in household surveys [16].

This study collected data from 2016 to 2024, though we did not adjust for time periods such as the COVID-19 pandemic, which are known to have affected the prevalence of attempts to reduce alcohol consumption in the United Kingdom [32]. Previous research has looked at time trends in motives for attempts to reduce alcohol consumption from a nationally representative cross-sectional survey and saw increases in attempts motivated by health concerns and those motivated by cost and social factors [9]. Future research could look to identify any trends in the prevalence of different alcohol reduction motives among this user group over time to see if this reflects what was seen across the general population.

The app users included in this study, who reported their motive for drinking less in the app, were only a small proportion of total app users (fewer than 3%) and differed by their age, sex, and AUDIT risk zone. Therefore, this sample is self-selected and not generalizable to all users of the Drink Less app. However, not completing the prompt is unlikely to have a bearing on users' readiness to change given the level of motivation already likely present; participants had downloaded and completed the onboarding process for an app named "Drink Less," which has a description in the UK App Store of "Are you looking to cut down how much you drink? If so, we can help." Also, there is previous evidence that the extent of behavioral engagement (eg, time spent on the app) with Drink Less does not have a mediating effect on its effectiveness, suggesting that these users who have greater behavioral engagement with the app and report their motive for drinking less are not more or less likely to be successful in reducing their alcohol consumption [33]. Therefore, we believe these findings provide useful information on what actually motivates an attempt to reduce alcohol consumption even though they are from a self-selected sample of users, and these may generalize to other people motivated to reduce their alcohol consumption but not using the app. Further to this, we only analyzed data on motives to drink less alcohol from one app. While the Drink Less app is a widely used, popular [21] and evidence-based app [19], future research could explore whether similar motives to drink less are reported in other alcohol reduction apps.

The vast majority of motives to drink less were reported within 1 week of downloading the app suggesting that these motives related to why they had made an attempt to reduce their drinking (by downloading the Drink Less app). While we do not know if users sustained their motive throughout their attempt, they were able to update it at any point. A quarter of users last updated their motive to drink less over a week after downloading the app, meaning that their AUDIT score (recorded on the day of app download) may have changed since onboarding. However, while the AUDIT-C (the first 3 items) is responsive to change [34], the full 10-item AUDIT is likely to be less responsive to change given the nature of the remaining 7 items. Further to this, in this study, we categorized users by their AUDIT risk zone (ie, scores ranging from 0 to 40 grouped into 4 categories), meaning a greater change in AUDIT score would be required for the user to change AUDIT risk zone.

Another limitation of this study was that the content analysis was conducted manually by one researcher, with no interrater reliability calculated.

Future Research

Understanding people's motivations to reduce their alcohol consumption has important implications for how to motivate risky drinkers to make reduction attempts given that identifying these motives is commonly addressed across guidance documents and treatment manuals for brief alcohol interventions [7]. We found different patterns in motives to drink less for people at either end of the AUDIT risk zone, which could be further explored in other populations and potentially used to inform tailoring of public health campaigns and interventions. This study's findings can be used to inform materials or public health campaigns encouraging risky drinkers to use the Drink Less app with examples of other users' motives, rather than to

XSL•FO
RenderX

only consume less alcohol given the skepticism and denial of national alcohol guidelines that some people have (eg, [35]).

By the nature of the sample, all the study participants were using an evidence-based digital intervention with evidence for its effectiveness for helping users reduce 2 UK units a week over a 6-month period [19], and therefore, their motives to drink less had all resulted in action to do something. However, we do not know whether there were certain types of motives which are more successful in reducing alcohol consumption than others, and this is something that we will look to investigate in the future.

The findings of this study also informed the response options for the question on motives for cutting down on alcohol consumption included in the Alcohol Toolkit Study [36] (with 2 new response options "Improve my mental health or well-being" and "To avoid bad experiences when drinking (aggressive behavior, hangovers, risk of injury, etc)" included).

Conclusions

Users of an alcohol reduction app, Drink Less, in the United Kingdom most commonly reported improving their physical health, feeling better in their body, improving their mental well-being, regaining agency, living a different life, and having better relationships with the people in their life as their motives for reducing their alcohol consumption. Wanting to improve mental well-being, regain agency, live a different life, and to have better relationships were all motives that had the highest prevalence among people at risk of alcohol dependence and the lowest among those drinking at low-risk levels. Understanding the motives behind why people are attempting to reduce their alcohol consumption, particularly among heavier drinkers, could inform how to encourage more people to make an attempt to reduce their drinking.

Funding

The development, refinement, and evaluation of Drink Less has been funded by the National Institute for Health and Care Research (NIHR), NIHR School for Public Health Research (NIHR SPHR), UK Centre for Tobacco and Alcohol Studies (UKCTAS), Society for the Study of Addiction (SSA), and Cancer Research UK (CRUK). This study is part of and contributing to a larger programme of research funded by NIHR (NIHR302923). The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

Data Availability

The data used in this study cannot be shared openly as the app users were not asked to consent for their data to be made open access.

Authors' Contributions

Conceptualization: CG
Data curation: CG
Formal analysis: LPS
Funding acquisition: CG
Supervision: CG
Supervision: OMM
Writing – original draft: LPS, CG
Writing – review & editing: LPS, OMM, DK, JAC, CG

Conflicts of Interest

LPS, OMM, DK, and JAC have no conflicts of interest to declare. CG worked on a project funded by Alcohol Change UK (ACUK) that ran from September 2021 to March 2023, with a no cost extension running until March 2025. The authors have since become aware that ACUK received <0.6% of its funds in 2024‐25 from Lucky Saint, an organization that produces and sells nonalcoholic drinks, and owns a pub that sells standard alcoholic drinks. In March 2025, Lucky Saint became an associate member of The Portman Group, a self-regulatory organization that is fully funded and controlled by the alcohol industry.

Multimedia Appendix 1
Frequency of categories and themes of motives to drink less reported by app users, stratified by sociodemographic and drinking characteristics (n=2520).
[DOCX File, 39 KB - publichealth_v12i1e88992_app1.docx]

Multimedia Appendix 2
Frequency of categories, themes, and subthemes reported by the Drink Less app users (n=2520).
[DOCX File, 33 KB - publichealth_v12i1e88992_app2.docx]

XSL•FO
RenderX

References

1. The public health burden of alcohol and the effectiveness and cost-effectiveness of alcohol control policies: an evidence review. : Public Health England; 2016 Dec URL: https://assets.publishing.service.gov.uk/media/5b6c5703ed915d3119112af6/alcohol_public_health_burden_evidence_review_update_2018.pdf [accessed 2026-04-18]
2. Boyd J, Sexton O, Angus C, Meier P, Purshouse RC, Holmes J. Causal mechanisms proposed for the alcohol harm paradox—a systematic review. Addiction 2022 Jan;117(1):33-56. [doi: 10.1111/add.15567] [Medline: 33999487]
3. Charlet K, Heinz A. Harm reduction—a systematic review on effects of alcohol reduction on physical and mental symptoms. Addict Biol 2017 Sep;22(5):1119-1159. [doi: 10.1111/adb.12414] [Medline: 27353220]
4. Buss V, Kock L, Brown J. Latest trends on alcohol consumption in England from the Alcohol Toolkit Study. Alcohol in England. URL: https://www.alcoholinengland.info/graphs/monthly-tracking-kpi [accessed 2026-04-18]
5. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6(1):42. [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
6. de Vocht F, Brown J, Beard E, et al. Motivation to reduce alcohol consumption and subsequent attempts at reduction and changes in consumption in increasing and higher-risk drinkers in England: a prospective population survey. Addiction 2018 May;113(5):817-827. [doi: 10.1111/add.14132] [Medline: 29368391]
7. Michie S, Whittington C, Hamoudi Z, Zarnani F, Tober G, West R. Identification of behaviour change techniques to reduce excessive alcohol consumption. Addiction 2012 Aug;107(8):1431-1440. [doi: 10.1111/j.1360-0443.2012.03845.x] [Medline: 22340523]
8. Beard E, Brown J, Kaner E, West R, Michie S. Predictors of and reasons for attempts to reduce alcohol intake: a population survey of adults in England. PLoS One 2017;12(3):e0173458. [doi: 10.1371/journal.pone.0173458] [Medline: 28278218]
9. Kale D, Buss V, Oldham M, Brown J, Shahab L, Jackson S. Trends in motives for attempts to reduce alcohol consumption among risky adult drinkers in England: a representative population survey, 2017-2024. Drug Alcohol Depend Rep 2025 Jun;15:100340. [doi: 10.1016/j.dadr.2025.100340] [Medline: 40470512]
10. Britton A, Bell S. Reasons why people change their alcohol consumption in later life: findings from the Whitehall II Cohort Study. PLoS One 2015;10(3):e0119421. [doi: 10.1371/journal.pone.0119421] [Medline: 25756213]
11. Pennay A, Callinan S, Livingston M, et al. Patterns in reduction or cessation of drinking in Australia (2001-2013) and motivation for change. Alcohol Alcohol 2019 Jan 1;54(1):79-86. [doi: 10.1093/alcalc/agy072] [Medline: 30346513]
12. Epler AJ, Sher KJ, Piasecki TM. Reasons for abstaining or limiting drinking: a developmental perspective. Psychol Addict Behav 2009 Sep;23(3):428-442. [doi: 10.1037/a0015879] [Medline: 19769427]
13. Stritzke WGK, Butt JCM. Motives for not drinking alcohol among Australian adolescents: development and initial validation of a five-factor scale. Addict Behav 2001;26(5):633-649. [doi: 10.1016/s0306-4603(00)00147-7] [Medline: 11676375]
14. Matzger H, Kaskutas LA, Weisner C. Reasons for drinking less and their relationship to sustained remission from problem drinking. Addiction 2005 Nov;100(11):1637-1646. [doi: 10.1111/j.1360-0443.2005.01203.x] [Medline: 16277625]
15. Cunningham JA, Blomqvist J, Koski-Jännes A, Cordingley J. Current heavy drinkers' reasons for considering change: results from a natural history general population survey. Addict Behav 2005 Mar;30(3):581-584. [doi: 10.1016/j.addbeh.2004.05.028] [Medline: 15718075]
16. Rehm J, Kilian C, Rovira P, Shield KD, Manthey J. The elusiveness of representativeness in general population surveys for alcohol. Drug Alcohol Rev 2021 Feb;40(2):161-165. [doi: 10.1111/dar.13148] [Medline: 32830351]
17. Garnett C, Perski O, Michie S, et al. Refining the content and design of an alcohol reduction app, Drink Less, to improve its usability and effectiveness: a mixed methods approach. F1000Res 2021;10:511. [doi: 10.12688/f1000research.51416.2] [Medline: 34646502]
18. Garnett C, Crane D, West R, Brown J, Michie S. The development of Drink Less: an alcohol reduction smartphone app for excessive drinkers. Transl Behav Med 2019 Mar 1;9(2):296-307. [doi: 10.1093/tbm/iby043] [Medline: 29733406]
19. Oldham M, Beard E, Loebenberg G, et al. Effectiveness of a smartphone app (Drink Less) versus usual digital care for reducing alcohol consumption among increasing-and-higher-risk adult drinkers in the UK: a two-arm, parallel-group, double-blind, randomised controlled trial. EClinicalMedicine 2024 Apr;70:102534. [doi: 10.1016/j.eclinm.2024.102534] [Medline: 38685934]
20. Garnett C, Perski O, Beard E, Michie S, West R, Brown J. The impact of celebrity influence and national media coverage on users of an alcohol reduction app: a natural experiment. BMC Public Health 2021 Jan 6;21(1):30. [doi: 10.1186/s12889-020-10011-0] [Medline: 33407283]
21. Drink Less. App Store for iPhone. URL: https://apps.apple.com/gb/app/drink-less/id1020579244 [accessed 2026-04-18]
22. Bell L, Garnett C, Bao Y, et al. How notifications affect engagement with a behavior change app: results from a micro-randomized trial. JMIR Mhealth Uhealth 2023 Jun 9;11:e38342. [doi: 10.2196/38342] [Medline: 37294612]
23. Babor TF, Higgins-Biddle JC, Saunders JB, Monteiro MG. AUDIT: The Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Care, 2nd edition: World Health Organization; 2001. URL: https://www.who.int/publications/i/item/WHO-MSD-MSB-01.6a [accessed 2026-04-18]
24. Mayring P. Qualitative content analysis. Forum Qual Soc Res 2000;1(2) [FREE Full text]
25. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]

XSL•FO
RenderX

26. Lindgren BM, Lundman B, Graneheim UH. Abstraction and interpretation during the qualitative content analysis process. Int J Nurs Stud 2020 Aug;108(2):103632. [doi: 10.1016/j.ijnurstu.2020.103632] [Medline: 32505813]
27. Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907. [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
28. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997;12(1):38-48. [doi: 10.4278/0890-1171-12.1.38] [Medline: 10170434]
29. Davies EL, Conroy D, Winstock AR, Ferris J. Motivations for reducing alcohol consumption: an international survey exploring experiences that may lead to a change in drinking habits. Addict Behav 2017 Dec;75:40-46. [doi: 10.1016/j.addbeh.2017.06.019] [Medline: 28692953]
30. Cunningham JA, Blomqvist J, Koski-Jännes A, Cordingley J. Maturing out of drinking problems: perceptions of natural history as a function of severity. Addict Res Theory 2005 Feb;13(1):79-84. [doi: 10.1080/16066350512331328131]
31. Cialdini RB, Goldstein NJ. Social influence: compliance and conformity. Annu Rev Psychol 2004;55(1):591-621. [doi: 10.1146/annurev.psych.55.090902.142015] [Medline: 14744228]
32. Jackson SE, Garnett C, Shahab L, Oldham M, Brown J. Association of the COVID-19 lockdown with smoking, drinking and attempts to quit in England: an analysis of 2019-20 data. Addiction 2021 May;116(5):1233-1244. [doi: 10.1111/add.15295] [Medline: 33089562]
33. Garnett C, Dinu LM, Oldham M, et al. Do engagement and behavioural mechanisms underpin the effectiveness of the Drink Less app? NPJ Digit Med 2024 Jun 29;7(1):174. [doi: 10.1038/s41746-024-01169-7] [Medline: 38951560]
34. Bradley KA, McDonell MB, Bush K, Kivlahan DR, Diehr P, Fihn SD. The AUDIT alcohol consumption questions: reliability, validity, and responsiveness to change in older male primary care patients. Alcohol Clin Exp Res 1998 Nov;22(8):1842-1849. [doi: 10.1111/j.1530-0277.1998.tb03991.x] [Medline: 9835306]
35. Crawford J, Cooke R, Shorter GW, Bendtsen M. Personal perceptions of risky drinking and alcohol guidelines – a qualitative analysis. BMC Public Health 2025 Sep 16;25(1):3049. [doi: 10.1186/s12889-025-24296-6] [Medline: 40958106]
36. Beard E, Brown J, West R, et al. Protocol for a national monthly survey of alcohol use in England with 6-month follow-up: “the Alcohol Toolkit Study”. BMC Public Health 2015 Mar 7;15(1):230. [doi: 10.1186/s12889-015-1542-7] [Medline: 25884652]

Abbreviations

AUDIT: Alcohol Use Disorders Identification Test

Edited by A Mavragani, T Sanchez; submitted 05.Dec.2025; peer-reviewed by DS Falk, M Bendtsen; revised version received 10.Mar.2026; accepted 31.Mar.2026; published 29.Apr.2026.

Please cite as:
Seppi LP, Maynard OM, Kale D, Cunningham JA, Garnett C
Exploring Motives for Reducing Alcohol Consumption Among Users of an Alcohol Reduction App: Content Analysis
JMIR Public Health Surveill 2026;12:e88992
URL: https://publichealth.jmir.org/2026/1/e88992
doi:10.2196/88992

© Lucille P Seppi, Olivia M Maynard, Dimitra Kale, John A Cunningham, Claire Garnett. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 29.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Mediation Analysis of the Relationship Between Health Literacy and the French General Population's Opinions on Hepatitis B Vaccination: Representative Cross-Sectional Survey of the SLAVACO Project

Bakary Cissé[1], MPH, MD; Sylvie Boyer[1], MSc, PhD; Jeremy K Ward[2], MSc, PhD; Julien Mancini[1,3], MD, PhD

[1]Aix Marseille Univ, INSERM, IRD, ISSPAM, SESSTIM, Sciences Economiques & Sociales de la Santé & Traitement de l'Information Médicale, Marseille, France

[2]CERMES3 (INSERM, CNRS, EHESS, Université de Paris), Paris, France

[3]APHM, Hop Ste Marguerite, BioSTIC, Biostatistique et Technologies de l'Information et de la Communication, Marseille, France

Corresponding Author:
Julien Mancini, MD, PhD
Aix Marseille Univ, INSERM, IRD, ISSPAM, SESSTIM, Sciences Economiques & Sociales de la Santé & Traitement de l'Information Médicale
Institut Paoli-Calmettes, 232 Bd Ste Marguerite, BP 156, 13273 Marseille Cedex 9
Marseille
France
Phone: 33 491384394
Email: julien.mancini@univ-amu.fr

Abstract

Background: In France, reluctance toward hepatitis B vaccination remains high, despite the availability of a safe and effective vaccine to prevent this infection. To boost vaccination coverage, it is therefore essential to identify the factors that are likely to encourage a more favorable opinion of this vaccine. Health literacy (HL) is one such factor. It refers to the individual ability to access, understand, critically appraise, and apply health information to make informed decisions about health issues for oneself and for others.

Objective: This study explored the mechanisms through which HL might affect opinions about hepatitis B vaccination, both directly and indirectly, by relevant factors, including opinions about vaccination in general, trust in government health agencies, and trust in medical doctors.

Methods: The analysis used data from the SLAVACO-Wave 3 (Suivi Longitudinal des Attitudes à l'Égard d'un Vaccin Contre la COVID-19) survey, conducted in December 2021 among a representative sample of French adults (N=1932). Favorable and unfavorable opinions of hepatitis B vaccination were measured using a 5-point Likert scale, while HL was assessed using the HLS_{19}-Q12 questionnaire (12-item general health literacy questionnaire used in Health Literacy Survey 2019-2021). A structural equation model examined the relationship between HL and hepatitis B vaccination opinions, taking into account the potential mediating role of trust in the health care system (ie, government health agencies and medical doctors).

Results: Findings showed that individuals with a favorable opinion of hepatitis B vaccination (1437/1932, 74.4%) had a higher HL level than those with a negative or neutral opinion (62.6 vs 57.0, $P<.001$). The association between HL and hepatitis B vaccination opinions was fully mediated by trust in the health care system. The indirect effect of HL was estimated at 0.068 (95% CI 0.042-0.093), accounting for 52.4% (0.068/0.1297) of the total effect. This effect was particularly pronounced in people over 50 years (0.084, 95% CI 0.042-0.126, accounting for 0.084/0.1306, 64.3% of the total effect). Goodness-of-fit indicators were satisfactory.

Conclusions: Enhancing HL might positively influence hepatitis B vaccination opinions and uptake through greater trust in the health care system. From a public health perspective, strategies should go beyond providing clear information and access to vaccines and actively work to strengthen trust in health care institutions and professionals. National campaigns correcting misconceptions about hepatitis B vaccination could be complemented by targeted interventions for groups most likely to hold negative opinions. Repeating this survey in the post–COVID-19 context could also reveal different trends, given evolving public perceptions of vaccines and health authorities.

XSL•FO
RenderX

(JMIR Public Health Surveill 2026;12:e82496) doi:10.2196/82496

KEYWORDS

health literacy; hepatitis B vaccination; vaccine hesitancy; trust; SLAVACO survey; France

Introduction

Hepatitis B is a communicable infectious disease associated with an increased risk of cirrhosis and liver cancer [1]. In 2022, the World Health Organization (WHO) estimated that 254 million people were living with chronic hepatitis B infection and that approximately 1.1 million died as a result of the disease, mainly from cirrhosis or hepatocellular carcinoma [2]. Although safe and effective vaccines have been developed, 1.5 million new hepatitis B virus (HBV) infections occur annually worldwide [3,4]. The WHO aims to reduce the incidence of chronic hepatitis infection by 90% and the annual number of deaths due to chronic hepatitis by 65% before 2030 [5]. To achieve these targets, the organization's guidelines [6] include systematic treatment for people with chronic HBV infection, the prevention of perinatal and infant infection through vaccination, and catch-up vaccinations for both high-risk populations (eg, injection drug users, men who have sex with men, and sex workers) and health care professionals. As of 2023, hepatitis B vaccination for infants was mandatory in 190 countries, and global coverage was estimated at 83% [7].

Hepatitis B is slightly endemic in France; in 2016, the estimated prevalence in the general population in metropolitan France was 0.3% [8]. Although this prevalence may seem low in the European context—it is higher than in Germany (0.2%) but lower than in Southern and Eastern European countries, where it varies between 0.5% and 1.5% [9]—hepatitis B remains a major public health problem due to its chronic progression and severe complications. Moreover, the burden is marked by unequal distribution, with significant socioeconomic disparities disproportionately affecting the most vulnerable populations [10]. Its impact is also reflected in mortality; from 2005 to 2020, a total of 2133 deaths occurred in patients hospitalized for related complications (cirrhosis or hepatocellular carcinoma) in metropolitan France, representing a lethality rate of 6% [11]. Beyond morbidity and mortality, hepatitis B also generates substantial direct and indirect health care costs, particularly in the stages of cirrhosis or hepatocellular carcinoma [12]. To combat this infection, starting in the late 1980s, the French National Authority for Health recommended vaccination for individuals at high risk of exposure and for all newborns. Moreover, a catch-up program for children and adolescents younger than 15 years was implemented in 1995 [13]. In the context of the WHO's goal to eliminate viral hepatitis by 2030, primary vaccination became mandatory for all infants born on or after January 1, 2018 [14]. The French primary childhood immunization schedule recommends 3 doses: the first at 2 months of age, the second at 4 months, and a booster at 11 months [15]. Full immunization is considered complete once the 3-dose series has been administered. As the majority of HBV infections in France in adults occur through sexual transmission [16], vaccination is strongly encouraged for persons older than 18 years of age. The immunization schedule consists of 3 doses, administered at 0, 1, and 6 months [15]. According to the 2021 LaboHEP survey, which was conducted among all public and private bioanalytical testing laboratories to estimate hepatitis screening activity, the rate of positive HBV diagnosis in metropolitan France was 54/100,000, representing a 10% increase compared to 2016 [17]. In addition, although hepatitis B vaccination coverage among infants has gradually increased over time, reaching over 90% for children aged 21 months since 2018 [18], it was below 50% among adolescents and adults in 2015 [19,20]; this reflects a clear generational gap in hepatitis B immunization.

Health Barometer surveys conducted between 2010 and 2023 show that in terms of vaccine hesitancy rankings in metropolitan France, the hepatitis B vaccine placed third after COVID-19 and influenza vaccines [21]. Despite this, no recent study has investigated the determinants of hesitancy over this vaccine. In the present analysis, we formulate the hypothesis that health literacy (HL), defined by Sørensen et al [22] as a concept "linked to literacy and that entails people's knowledge, motivation and competences to access, understand, appraise, and apply health information to make judgments and take decisions in everyday life concerning healthcare, disease prevention and health promotion to maintain or improve quality of life during the life course," might explain hepatitis B vaccine hesitancy in France. In general, higher HL is associated with greater vaccine acceptance and lower hesitancy [23]. This relationship may be influenced by the type of vaccine and the population studied. In the case of hepatitis B vaccination, controversies surrounding a potential link to demyelinating diseases [24] may contribute to differences or similarities compared to other vaccines. These specificities highlight that vaccine hesitancy is not uniform but rather a complex and multifactorial challenge to public health [25-27], frequently rooted as much in a lack of trust as in limited access to, and limited understanding of, vaccine-related information. Trust in the health care system—understood here as health agencies and medical doctors—is well-documented as a central determinant of vaccine acceptance [28,29]. However, it is often difficult to create strategies to build this trust. This gap highlights the relevance of exploring HL as a sustainable mediating lever to support trust [30]. Additionally, given that vaccine hesitancy does not necessarily regard specific vaccines but vaccination more broadly [31], any examination of opinions on a particular vaccine should also cover opinions on vaccination in general.

In this context, to enhance our understanding of hepatitis B vaccine opinions in France, this study explored the mechanisms by which HL influences the general public's opinion on hepatitis B vaccination. We used a structural equation model [32] to investigate the direct and indirect effects of HL on these opinions. A model with 5 hypotheses was analyzed as follows: (1) a higher HL is associated with a more favorable opinion of hepatitis B vaccination (direct effect represented by the coefficient "A"); (2) a higher HL contributes positively to trust

XSL•FO
RenderX

in the health care system and this trust is associated with a more favorable opinion of hepatitis B vaccination (indirect effect via "D" and "C"); (3) a higher HL has a positive influence on opinions about vaccination in general, which in turn are associated with a more favorable opinion of hepatitis B vaccination (indirect effect via "E" and "B"); (4) together, trust in the health care system and positive opinions of vaccination in general are mediators in the relationship between HL and opinions on hepatitis B vaccination (indirect effect via "D, F, and B"); and (5) the expected effect of HL on opinions of hepatitis B vaccination, in hypotheses 1 to 4, might vary according to gender, age, and financial deprivation.

Methods

Recruitment Procedure and Study Sample

For the present analysis, we used data from SLAVACO-Wave 3 (Suivi Longitudinal des Attitudes à l'Égard d'un Vaccin Contre la COVID-19), which was conducted between December 2 and 17, 2021. The SLAVACO project was a multiwave longitudinal survey conducted across metropolitan France. Its primary objective was to study the evolution of public attitudes toward different aspects of COVID-19 vaccination and attitudes toward vaccines more generally. Data were collected by the French Provence-Alpes-Côte d'Azur Regional Health Observatory using online self-administered questionnaires, lasting approximately 15 minutes, sent to 25,800 French adults selected by random sampling in an online panel of over 750,000 French households (Bilendi panel). The quota sampling method was then used to obtain a final sample of 2022 respondents corresponding to the adult French population in terms of gender (male and female), age (18-24, 25-34, 35-49 50-64, and 75+ years), type of employment (farmers, craftsmen, executives, intermediate professions, employees, workers, retirees, and other inactives), and population density (<2000, 2000-20,000, 20,000-100,000, and >100,000 inhabitants) in respondents' region of residence (Alsace, Aquitaine, Auvergne, Burgundy, Brittany, Center, Île-de-France, Languedoc, Nord-Pas-de-Calais, Normandy, Pays de la Loire, and Provence-Alpes-Côte d'Azur) [33]. Final adjustments were applied by weighting the data with the raking ratio and the macro Calmar program of Statistical Analysis System. The latter was designed using census data from the French INSEE (National Institute for Statistics and Economic Studies) [34].

Data Collected

After the participants' consent was obtained, we assessed their opinions about vaccination in general using a question taken from the national Health Barometer survey; this survey has been conducted regularly in France over the past 2 decades [35]. The question was as follows: "Are you strongly, moderately, not really or not at all in favour of vaccination in general?" with the following 5 answer options: "yes, strongly," "yes, moderately," "I don't have an opinion," "no, not really," and "no, not at all." Opinions specifically on hepatitis B vaccination were also assessed in SLAVACO using a similar question as follows: "Are you strongly, moderately, not really or not at all in favour of hepatitis B vaccination?" with the same 5-option response scale. Respondents who did not answer both of these 2 questions were excluded from the present analysis.

SLAVACO also assessed HL using the European Health Literacy Survey 2019-2021 Questionnaire, HLS_{19}-Q12 (12-item general health literacy questionnaire used in Health Literacy Survey 2019-2021). This questionnaire has been translated, applied, and validated in 17 countries [36,37], including France. Based on Sorensen's matrix, the HLS_{19}-Q12 comprises 12 items that measure HL in 4 cognitive dimensions (accessing, understanding, evaluating, and applying health information) and in 3 health contexts (health care, disease prevention, and health promotion) [22]. The answers to each question are selected on a 4-point scale from 0 (very difficult) to 3 (very easy). After summing the scores for all 12 items, we standardized the total score, ranging from 0 to 100, where a higher score reflected a higher level of HL. Internal consistency was excellent with a Cronbach α coefficient of 0.91.

To measure trust in the health care system, we used the methodology used in the 2021 Political Confidence Barometer, the main French longitudinal study on public trust in politics [38]. We measured 2 dimensions: trust in government health agencies and trust in medical doctors. A 5-point scale ranging from 0 (no trust at all) to 4 (complete trust) was used for each measure.

Our analysis also included self-reported socioeconomic variables. Specifically, gender was reported using 3 categories (man, woman, and other), and age was categorized into the following groups: 18-24, 25-34, 35-49, 50-64, 65-74, and 75+ years, in line with the INSEE's age categories [39]. Educational attainment was measured using 6 categories: no educational qualification, lower secondary school certificate, upper secondary school certificate, bachelor's degree, master's degree, and doctorate, while employment status was recorded as currently employed, not employed, or retired. Financial deprivation was determined using the question "How easy or difficult is it for you to pay all your bills at the end of the month?" Respondents who answered "difficult" or "very difficult" were classified as having financial difficulties. Finally, the presence of one or more chronic diseases was assessed using a "yes" or "no" question.

Statistical Analyses

Continuous variables were reported in terms of the mean and SD, while categorical variables were reported in terms of frequency and percentages. We performed descriptive analyses to explore covariables associated with opinions on hepatitis B vaccination. To do this, we treated these opinions and opinions on vaccination in general as binary variables: a favorable opinion was defined as an answer of "yes, strongly" or "yes, moderately." An unfavorable opinion was defined as responding "I don't have an opinion," "no, not really," or "no, not at all." For the bivariate and mediation analyses, a 0 to 4 scale was used ("no, not at all"=0; "no, not really"=1; "I don't have an opinion"=2; "yes, moderately"=3; and "yes, strongly"=4).

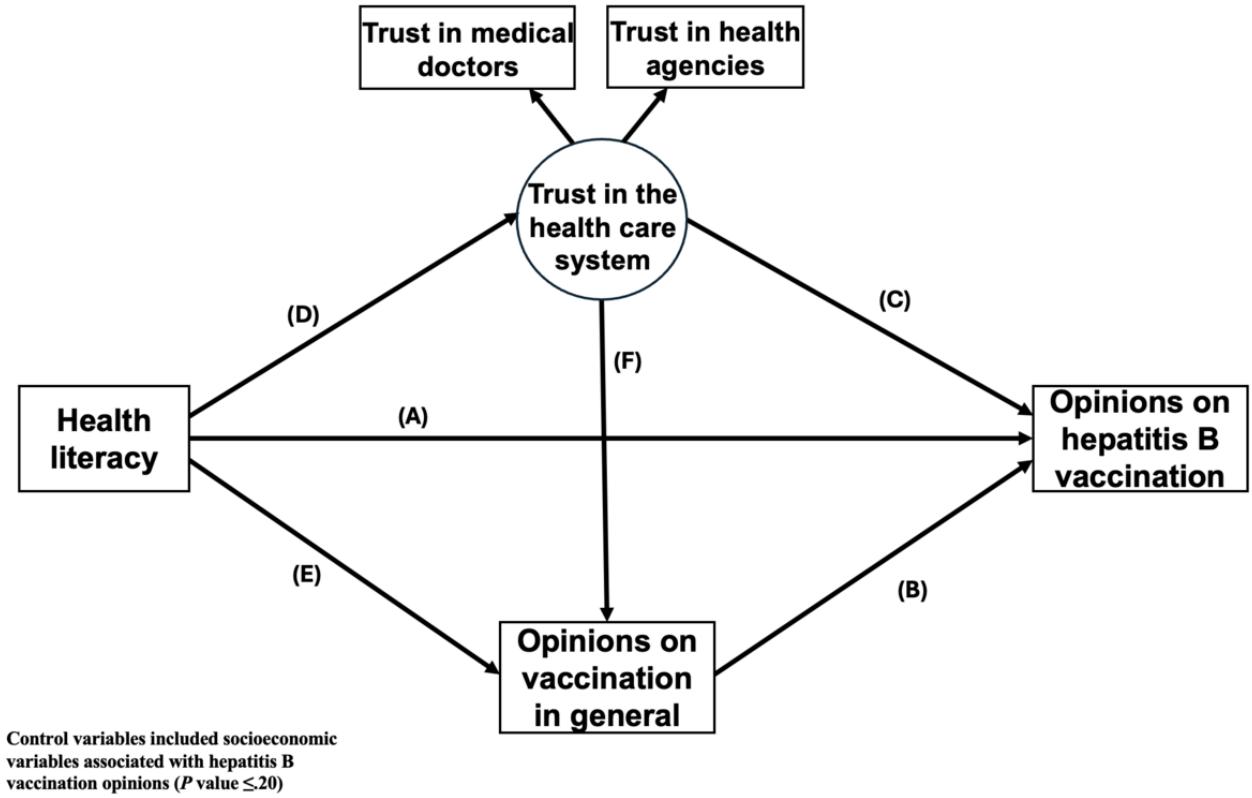
We conducted a Pearson correlation analysis to investigate the bivariate association between the 5 variables of interest (ie, opinions on vaccination in general, opinions on hepatitis B

XSL•FO
RenderX

vaccination, HL level, trust in government health agencies, and trust in medical doctors).

We performed a mediation analysis to explore whether favorable opinions toward vaccination in general and trust in the health care system mediated the association between HL and favorable opinions toward hepatitis B vaccination (Figure 1). Trust in the health care system was modelled using a latent variable based on 2 observed indicators: trust in government health agencies and trust in medical doctors. Possible mediation was then examined after stratification by gender (man vs woman), age (<50 years vs ≥50 years), and financial deprivation (yes vs no). The control variables included socioeconomic variables associated with hepatitis B vaccination opinions (*P* value ≤.20). The 95% CI for the direct, indirect, and total effects of the mediation models were estimated using 5000 bootstrap samples. The effects of HL were estimated per 10-point increase in the HL score. We used a series of indicators to assess the goodness-of-fit of the model: Tucker-Lewis Index (considered excellent above 0.95); comparative fit index (considered excellent above 0.95); and root-mean-square error of approximation (RMSEA, considered excellent if below 0.05).

Figure 1. Hypothesized mediation model for health literacy's effect on hepatitis B vaccination opinions (A-F are the coefficients from the Introduction).



All analyses were performed on the weighted database. Analyses were carried out with statistical R software (version 4.4.1; R Foundation), and the significance level was set to 5% for all 2-sided tests.

Ethical Considerations

This study was approved in France by the Ethics Evaluation Committee of the National Institute of Health and Medical Research (IRB00003888) on February 9, 2021 (21-770). After receiving information about SLAVACO-Wave 3, informed consent was obtained from all the participants before they started the survey. In accordance with standards practice for web-based surveys, we did not have access to any data that could identify respondents. Participants received no compensation.

Results

Sample Characteristics

Among the 2022 individuals who participated in SLAVACO-Wave 3, we excluded 90 (4.4%) because of missing answers to the questions on opinions on hepatitis B vaccination and/or vaccination in general. Of the 1932 adults comprising this study's sample, 52.6% (n=1016) were women and 34.1% (n=658) were retirees (Table 1). Most of the participants (n=1184, 61.4%) had at least an upper secondary school certificate, and most (n=1210, 62.6%) did not report financial difficulties. Half of the participants (n=952, 49.3%) reported one or more chronic diseases. A quarter (n=495, 25.6%) expressed an unfavorable opinion about hepatitis B vaccination, and a fifth (n=375, 19.4%) towards vaccination in general. Among those hesitant toward hepatitis B vaccination, 58.6% (290/495) were nevertheless favorable to vaccination in general, while 18.6% (290/1557) of those favorable to vaccination in general expressed hesitancy toward hepatitis B vaccination.

XSL•FO
RenderX

Table 1. Description of this study's sample and factors associated with opinions on hepatitis B vaccination in France (N=1932).

Variables	Overall (N=1932)	Opinions on hepatitis B vaccination		*P* value
		Favorable (n=1437, 74.4%)	Unfavorable (n=495, 25.6%)	
Age group (years), n (%)				.14
18-24	201 (10.4)	157 (78.4)	43 (21.6)	
25-34	282 (14.6)	216 (76.5)	66 (23.5)	
35-49	460 (23.8)	345 (75)	115 (25)	
50-64	466 (24.1)	324 (69.5)	142 (30.5)	
65-74	279 (14.4)	211 (75.8)	68 (24.2)	
75+	244 (12.6)	183 (75.1)	61 (24.9)	
Gender, n (%)				<.001
Man	916 (47.4)	739 (80.7)	177 (19.3)	
Woman	1016 (52.6)	698 (68.7)	318 (31.3)	
Employment status, n (%)				.30
Currently employed	802 (41.5)	611 (76.1)	191 (23.9)	
Unemployed	472 (24.4)	339 (71.9)	133 (28.1)	
Retired	658 (34.1)	487 (74)	171 (26)	
Educational attainment, n (%)				.01
No diploma	121 (6.3)	88 (72.7)	33 (27.3)	
Lower secondary school certificate	626 (32.4)	452 (72.3)	174 (27.7)	
Upper secondary school certificate	405 (21)	283 (69.8)	123 (30.2)	
Bachelor's degree	347 (18)	266 (76.6)	81 (23.4)	
Master's degree	208 (10.8)	171 (82)	37 (18)	
Doctorate	224 (11.6)	177 (78.8)	48 (21.2)	
Financial deprivation, n (%)				<.001
Yes	722 (37.4)	503 (69.7)	219 (30.3)	
No	1210 (62.6)	934 (77.2)	276 (22.8)	
Chronic diseases, n (%)				.02
Yes	952 (49.3)	730 (76.7)	202 (23.3)	
No	980 (50.7)	707 (72.1)	274 (27.9)	
Health literacy score (continuous), mean (SD)	61.1 (17.0)	62.6 (16.9)	57.0 (16.8)	<.001
Opinion on vaccination in general, n (%)				<.001
Favorable	1557 (80.6)	1267 (81.4)	290 (18.6)	
Unfavorable	375 (19.4)	170 (45.3)	205 (54.7)	
"Do you trust government health agencies?" n (%)				<.001
Not at all (0)	358 (18.5)	197 (54.9)	161 (45.1)	
Not really (1)	571 (29.5)	414 (72.5)	157 (27.5)	
No opinion (2)	80 (4.1)	46 (57.9)	33 (42.1)	
Moderately (3)	805 (41.7)	678 (84.2)	127 (15.8)	
Completely (4)	118 (6.1)	102 (86.3)	16 (13.7)	
"Do you trust medical doctors?" n (%)				<.001
Not at all (0)	50 (2.6)	30 (60.6)	20 (39.4)	
Not really (1)	148 (7.6)	90 (60.6)	58 (39.4)	
No opinion (2)	40 (2.1)	22 (55.4)	18 (44.6)	

XSL•FO
RenderX

Variables	Overall (N=1932)	Opinions on hepatitis B vaccination		*P* value
		Favorable (n=1437, 74.4%)	Unfavorable (n=495, 25.6%)	
Moderately (3)	1232 (63.8)	901 (73.1)	331 (26.9)	
Completely (4)	462 (23.9)	393 (85.2)	69 (14.8)	

Approximately half the sample (n=923, 47.8%) reported trusting government health agencies, while 87.7% (n=1694) reported trusting medical doctors. The mean HL score in the sample was 61.1 (SD 17.0). Individuals with a favorable opinion of hepatitis B vaccination had a higher HL score than those with an unfavorable opinion (62.6 vs 57.0, *P*<.001).

Correlation Analyses

The correlation matrix for all variables in the mediation model is reported in Table 2. A favorable opinion on hepatitis B vaccination was associated with a higher HL level (*r*=0.18, *P*<.001), a favorable opinion on vaccination in general (*r*=0.44, *P*<.001), a higher level of trust in government health agencies (*r*=0.27, *P*<.001), and a higher level of trust in medical doctors (*r*=0.22, *P*<.001). A higher HL level was positively and significantly correlated with a positive opinion on vaccination in general (*r*=0.20, *P*<.001), with a higher level of trust in government health agencies (*r*=0.27, *P*<.001), and with a higher level of trust in medical doctors (*r*=0.27, *P*<.001). A favorable opinion on vaccination in general was positively associated with a higher level of trust in government health agencies (*r*=0.27, *P*<.001), and with a higher level of trust in medical doctors (*r*=0.27, *P*<.001).

Table 2. Correlation matrix of mediation model variables.

Variables	Health literacy (A)	Opinion on hepatitis B vaccination (B)	Opinion on vaccination in general (C)	Trust in government health agencies (D)	Trust in medical doctors (E)
Health literacy (A)					
r	1	0.18	0.20	0.27	0.27
P value	—[a]	<.001	<.001	<.001	<.001
Opinion on hepatitis B vaccination (B)					
r	0.18	1	0.44	0.27	0.22
P value	<.001	—	<.001	<.001	<.001
Opinion on vaccination in general (C)					
r	0.20	0.44	1	0.27	0.27
P value	<.001	<.001	—	<.001	<.001
Trust in government health agencies (D)					
r	0.27	0.27	0.27	1	0.34
P value	<.001	<.001	<.001	—	<.001
Trust in medical doctors (E)					
r	0.27	0.22	0.27	0.34	1
P value	<.001	<.001	<.001	<.001	—

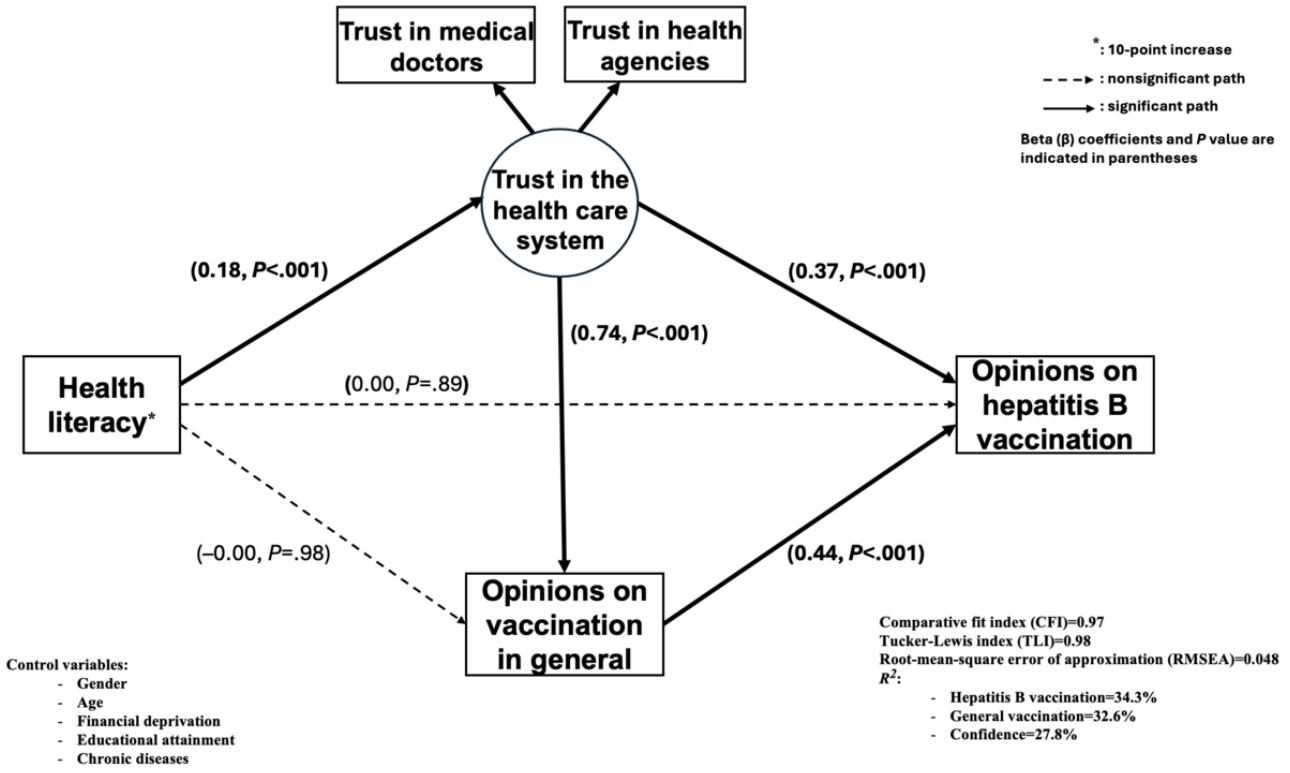
[a]Not available.

Mediation Analyses

The mediation model for all participants, with 2 mediators and a direct path, is presented in Figure 2. A significant and positive relationship was observed between trust in the health care system and a favorable hepatitis B vaccination opinion (β=0.37; *P*<.001). There was also a significant positive relationship between a favorable opinion on vaccination in general and a favorable opinion on hepatitis B vaccination (β=0.44; *P*<.001). A higher level of trust in the health care system had a significant and strong positive association with a favorable opinion on vaccination in general (β=0.74; *P*<.001). A higher HL level was positively and significantly associated with a higher level of trust in the health care system (β=0.18; *P*<.001), but not associated with a favorable opinion on vaccination in general and with a favorable opinion on hepatitis B vaccination (β=0.00; *P*=.89 and β=–0.00; *P*=.98, respectively). All estimations were adjusted for educational attainment and the presence of chronic diseases.

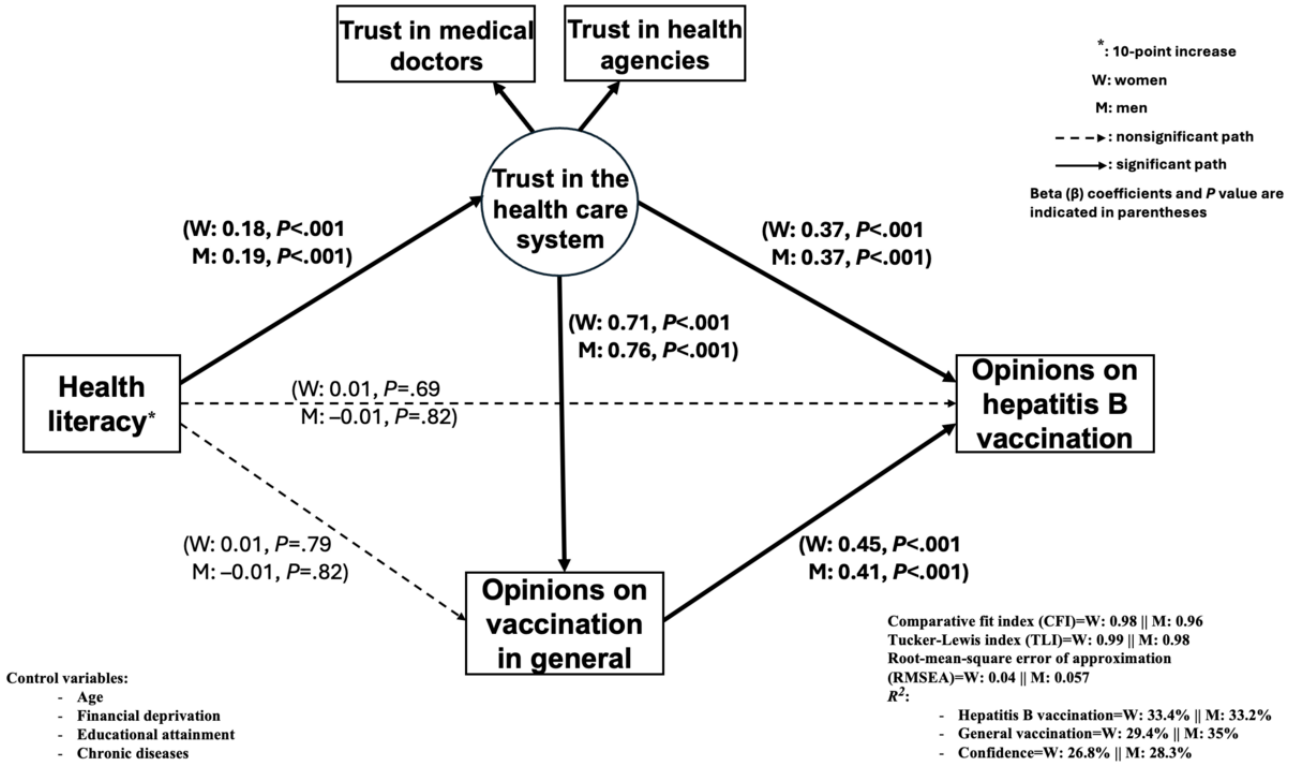
XSL•FO
RenderX

Figure 2. Mediation model of health literacy on opinions about hepatitis B vaccination (N=1932).



After stratifying by gender (Figure 3), we found that a higher level of trust in the health care system was positively associated with a favorable opinion on hepatitis B vaccination in women and men (β=0.37; *P*<.001 for both). A higher level of trust in the health care system was also significantly associated with a favorable opinion about vaccination in general for both genders (women: β=0.71; *P*<.001 and men: β=0.76; *P*<.001). A higher HL level was positively associated with a higher level of trust in the health care system for both genders (women: β=0.18; *P*<.001 and men: β=0.19; *P*<.001) but had no direct effect on a favorable opinion about vaccination in general or about hepatitis B vaccination.

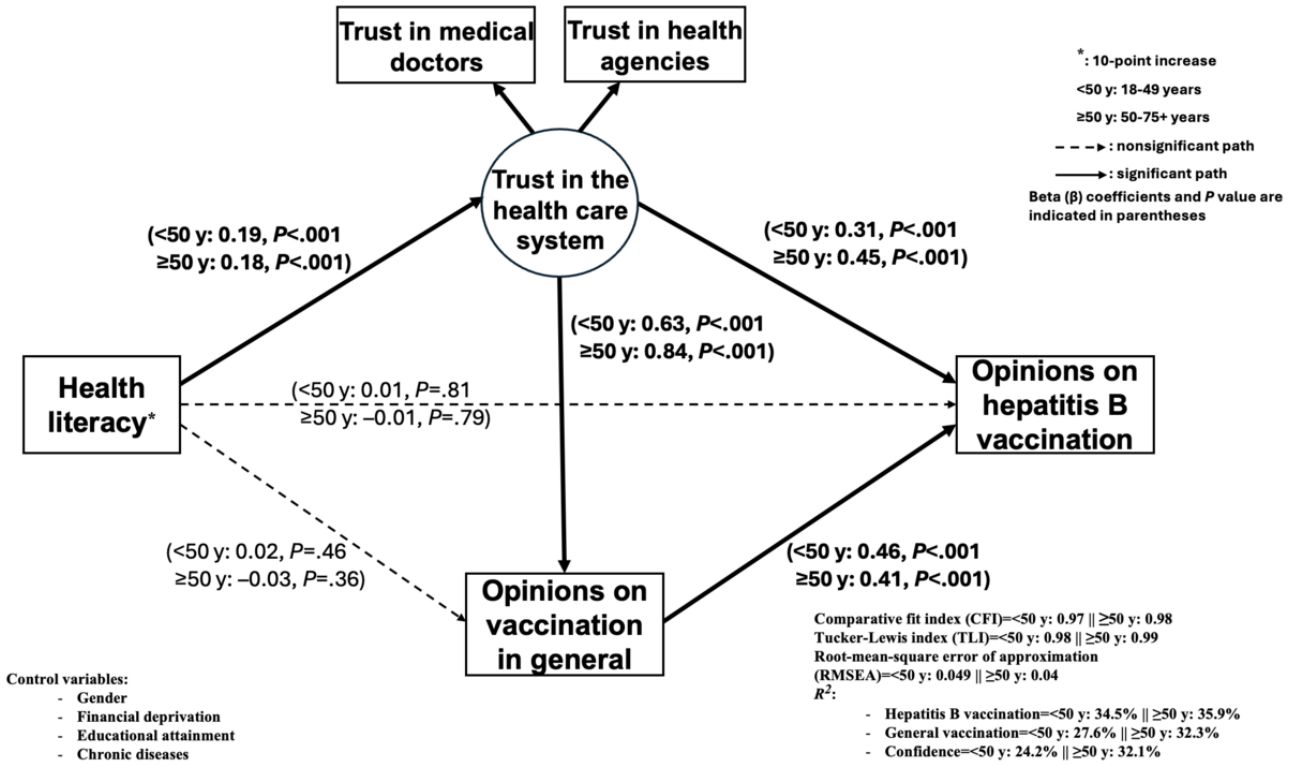
Figure 3. Mediation model of health literacy on opinions about hepatitis B vaccination according to gender (number of women=1016; number of men=916).



XSL•FO
RenderX

In terms of age (Figure 4), similar associations were observed in the <50 years and ≥50 years age groups. A higher level of trust was positively associated with a favorable opinion on hepatitis B vaccination, with a stronger association observed in adults ≥50 years (β=0.45; *P*<.001) than in younger adults (β=0.31; *P*<.001). The association between trust in the health care system and a favorable opinion on vaccination in general was strong for both age groups (<50 years: β=0.63; ≥50 years: β=0.84; *P*<.001). A higher HL level had a positive influence on trust in the health care system in both age groups (<50 years: β=0.19; ≥50 years: β=0.18; *P*<.001) but had no significant direct effect on a favorable opinion about vaccination (general or hepatitis B).

Figure 4. Mediation model of health literacy on opinions about hepatitis B vaccination according to age (number of participants aged <50 years=943; number of participants aged ≥50 years=989).



Lastly, after stratifying by financial deprivation (Figure 5), the association between trust in the health care system and a favorable hepatitis B vaccination opinion was stronger among individuals who did not have financial deprivation (β=0.41; *P*<.001) than among those who did (β=0.29; *P*=.006). In contrast, the association between trust in the health care system and a favorable opinion on vaccination in general was more pronounced among individuals who had financial difficulties (β=0.82; *P*<.001) than among those who did not (β=0.69; *P*<.001). A higher HL level remained positively associated with trust in the health care system in both subgroups (no financial deprivation: β=0.19; financial deprivation: β=0.16; *P*<.001), although there was no direct effect on favorable vaccination opinions (in general or hepatitis B).

XSL•FO
RenderX

Figure 5. Mediation model of health literacy on opinions about hepatitis B vaccination according to financial deprivation (number of participants with no financial deprivation=1210; number of participants with financial deprivation=722).

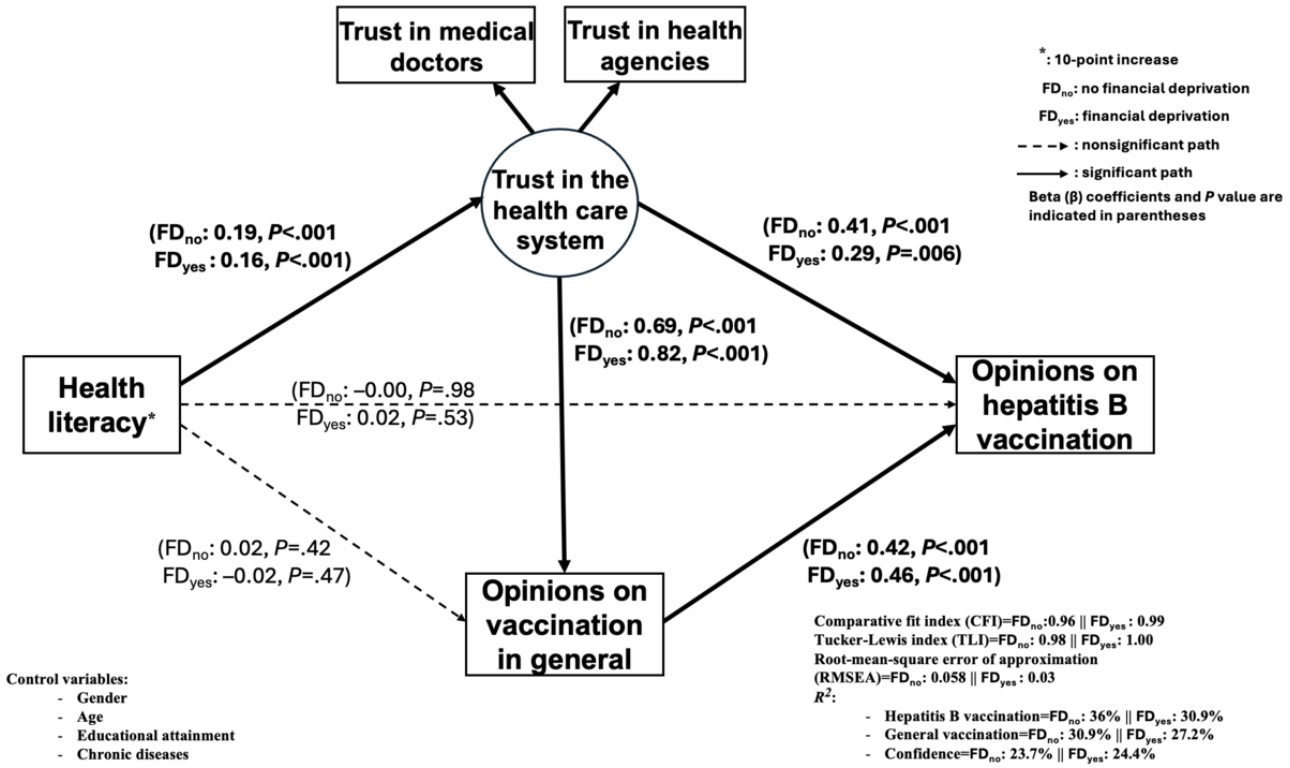


Table 3 shows that the effect of HL on hepatitis B vaccination opinions among all participants was fully mediated by trust in the health care system, after adjusting for educational attainment and the presence of chronic diseases. When trust was the only mediator, the indirect effect of HL was estimated at 0.068 (95% CI 0.042-0.093), accounting for 52.4% (0.068/0.1297) of the total effect. When trust and opinion on vaccination in general were combined as parallel mediators, the indirect effect attributable to HL was 0.059 (95% CI 0.048-0.07), representing 45.5% (0.059/0.1297) of the total effect. Mediation analyses stratified by gender, age, and financial deprivation status also confirmed the complete mediation of the relationship between HL and hepatitis B vaccination opinion through trust in the health care system, even after adjustment for different sets of control variables depending on the stratum (gender, age, educational attainment, financial deprivation, and chronic diseases). The total effect of HL on hepatitis B vaccination opinion was more pronounced among women and those without financial deprivation.

XSL•FO
RenderX

Table 3. Direct, indirect, and total effects for the mediation model of health literacy on opinions about hepatitis B vaccination (nonstratified model and models stratified by gender, age, and financial deprivation).

Model pathways	Estimated effects	95% CIs	*P* value
Direct effect (H1; health literacy[a]>hepatitis B vaccination)			
All participants[b] (N=1932)	0.003	–0.035 to 0.04	.89
Women[c] (n=1016)	0.01	–0.04 to 0.06	.69
Men[c] (n=916)	–0.006	–0.062 to 0.05	.82
18-49 years[d] (n=943)	0.005	–0.036 to 0.046	.81
50-75+ years[d] (n=989)	–0.006	–0.051 to 0.039	.79
No financial deprivation[e] (n=1210)	–0.001	–0.05 to 0.048	.98
Financial deprivation[e] (n=722)	0.015	–0.033 to 0.063	.53
Indirect effects (H2; health literacy[a]>vaccination in general>hepatitis B vaccination)			
All participants[b] (N=1932)	–0	–0.017 to 0.017	.98
Women[c] (n=1016)	0.003	–0.020 to 0.027	.79
Men[c] (n=916)	–0.003	–0.028 to 0.022	.81
18-49 years[d] (n=943)	0.009	–0.015 to 0.032	.47
50-75+ years[d] (n=989)	–0.01	–0.032 to 0.011	.34
No financial deprivation[e] (n=1210)	0.009	–0.013 to 0.031	.43
Financial deprivation[e] (n=722)	–0.01	–0.035 to 0.016	.46
Indirect effects (H3; health literacy[a]> trust>hepatitis B vaccination)			
All participants[b] (N=1932)	0.068	0.042 to 0.093	<.001
Women[c] (n=1016)	0.066	0.033 to 0.100	<.001
Men[c] (n=916)	0.071	0.033 to 0.11	<.001
18-49 years[d] (n=943)	0.059	0.031 to 0.087	<.001
50-75+ years[d] (n=989)	0.084	0.042 to 0.126	<.001
No financial deprivation[e] (n=1210)	0.079	0.046 to 0.112	<.001
Financial deprivation[e] (n=722)	0.048	0.012 to 0.083	.008
Indirect effects (H4; health literacy[a]>trust>vaccination in general>hepatitis B vaccination)			
All participants[b] (N=1932)	0.059	0.048 to 0.070	<.001
Women[c] (n=1016)	0.058	0.044 to 0.073	<.001
Men[c] (n=916)	0.06	0.044 to 0.077	<.001
18-49 years[d] (n=943)	0.056	0.041 to 0.07	<.001
50-75+ years[d] (n=989)	0.063	0.046 to 0.08	<.001
No financial deprivation[e] (n=1210)	0.056	0.043 to 0.069	<.001
Financial deprivation[e] (n=722)	0.062	0.044 to 0.083	<.001
Total effect			
All participants[b] (N=1932)	0.129	0.101 to 0.157	<.001
Women[c] (n=1016)	0.138	0.101 to 0.177	<.001

XSL•FO
RenderX

Model pathways	Estimated effects	95% CIs	*P* value
Men[c] (n=916)	0.123	0.08 to 0.165	<.001
18-49 years[d] (n=943)	0.128	0.087 to 0.169	<.001
50-75+ years[d] (n=989)	0.13	0.092 to 0.168	<.001
No financial deprivation[e] (n=1210)	0.143	0.106 to 0.181	<.001
Financial deprivation[e] (n=722)	0.115	0.073 to 0.158	<.001

[a]10-point increase.

[b]Estimated effects are controlled for gender, age, educational attainment, financial deprivation, and chronic diseases.

[c]Estimated effects are controlled for age, educational attainment, financial deprivation, and chronic diseases.

[d]Estimated effects are controlled for gender, educational attainment, financial deprivation, and chronic diseases.

[e]Estimated effects are controlled for gender, age, educational attainment, and chronic diseases.

For the control variables (Multimedia Appendix 1), the presence of chronic diseases was significantly associated with a higher level of trust in the health care system (β=0.11; *P*=.004), particularly among women (β=0.15; *P*=.004) and individuals without financial deprivation (β=0.14; *P*=.006). About educational attainment, having only a lower or upper secondary school certificate was negatively associated with the level of trust in the health care system among women (lower secondary school certificate β=–0.29; *P*=.004; upper secondary school certificate β=–0.23; *P*=.03) and among financially deprived individuals (β=–0.27; *P*=.01 and β=–0.26; *P*=.03, respectively).

All models showed an excellent goodness-of-fit (Tucker-Lewis Index >0.95, comparative fit index >0.95, and RMSEA <0.05), except for the men stratum (RMSEA=0.057) and for individuals without financial deprivation (RMSEA=0.058).

Discussion

Principal Findings

France has one of the highest rates of vaccine hesitancy in the world. According to an international study conducted in 2015, a total of 41% of French people considered that vaccines might be dangerous, which was the highest rate among the 67 countries studied [40]. The significant proportion of individuals expressing an unfavorable opinion on hepatitis B vaccination (495/1932, 25.6%) in our present study highlights the persisting skepticism concerning vaccines in the French population. Our findings highlight that this skepticism was more pronounced than for vaccination in general. The moderate correlation between opinions on vaccination in general and those on hepatitis B vaccination (*r*=0.44) indicates that these 2 dimensions only partially overlap; this justifies the detailed focus on this specific vaccine.

The particularly high level of skepticism toward hepatitis B vaccination, which we identified, could be the result of a historical context marked by health controversies, particularly in the 1990s, when cases of central nervous system demyelination after hepatitis B vaccination raised fears in the general public of a potential link [41]. Although subsequent investigations did not establish a causal link [24,42] and even though French authorities officially declared that there was no risk of developing a demyelinating disease from hepatitis B vaccination [43], the impact of these events on the general perception of hepatitis B vaccination persists today.

Our study results showed that women, people with a lower level of educational attainment, and individuals with no chronic disease were all significantly more likely to have an unfavorable opinion about hepatitis B vaccination. This greater vaccine hesitancy in women might reflect their greater engagement with health information than men. This engagement may, in turn, leave them more exposed to negative discourses, especially if they echo personal or shared experiences linked to past health controversies [44]. Although men in our study were more likely to have a favorable opinion of hepatitis B vaccination, some studies have suggested that women are more invested in health care and child-rearing than men, and so are more likely to get this vaccine [45,46].

The significant association that we found between the presence of chronic diseases and a higher level of trust in the health care system might partly explain the more favorable opinions toward hepatitis B vaccination, which we observed among people with chronic diseases. In addition to a lack of trust in the health care system, the greater reluctance to hepatitis B vaccination we observed in individuals with no chronic disease could be explained by greater complacency, identified as one of the key determinants of vaccine hesitancy. Complacency occurs when the perception of risk from vaccine-preventable diseases is low and when vaccination is not perceived as necessary [47]. Without a perceived risk to their health and without regular contact with health care services, these individuals might develop a less favorable opinion about vaccination.

The association we observed between intermediate educational attainment (ie, having only a lower or upper secondary school certificate) and less trust in the health care system among women and people in financial difficulty suggests that this level of education may sometimes be associated with a more skeptical attitude toward medical recommendations. This association is also reflected in the relatively high proportion of individuals in our study who expressed an unfavorable opinion about hepatitis B vaccination within these 2 subgroups. About the influence of educational attainment on vaccination opinions, findings in the literature are mixed; some studies emphasize its central role in shaping provaccination opinions [40,48], while others show that individuals with higher educational attainment might also

XSL•FO
RenderX

be more likely to adopt skeptical attitudes [49]. These contrasting views highlight the need to develop other, more actionable and multidimensional measures, such as HL.

In our study, a favorable hepatitis B vaccination opinion was significantly positively associated with a higher HL level before mediation was taken into account. In the mediation model, this association was fully mediated by trust in the health care system, even after stratifying by gender, age group, and financial difficulties. There was no significant direct relationship between HL and opinions on vaccination (ie, general or hepatitis B) when trust was considered in all strata. Although no study to date has directly investigated the mediating role of trust in the health care system on the relationship between HL and hepatitis B vaccination opinion, many studies have highlighted the positive influence of HL and trust in the health care system on hepatitis B prevention behaviors. For example, a group study of sociocultural barriers to hepatitis B prevention among Korean Americans suggested that there were specific HL barriers that governed individuals' behavior in terms of hepatitis B prevention and care access [50]. Similarly, in a study in the Democratic Republic of Congo [51], trust in the health care staff increased the public's acceptance and the effectiveness of community initiatives to prevent mother-to-child transmission of the disease, especially initiatives focusing on adherence to vaccination and antiviral uptake. HL could become an essential tool to overcome the challenge of improving trust in the health care system to reduce vaccine hesitancy [52]. By strengthening individuals' ability to understand how the health care system functions, to evaluate the reliability of health information, and to communicate effectively with professionals in order to both avoid negative interactions and engage proactively in their own care [53], HL could contribute to creating sustainable trust in the health care system.

Our key finding—the absence of a direct relationship between HL and a favorable opinion on hepatitis B vaccination—might be specific to this vaccination. This is because, in order to have a favorable opinion, a degree of trust in the expected long-term benefits (which are not visible) is needed, especially given the controversy surrounding this vaccine in France. This finding may also reflect the specific nature of vaccination as a public health issue; vaccination is particularly exposed to conflicting information at the heart of political and social debate, and is increasingly shaped by dynamics of trust or the lack thereof, which in turn are often associated with collective reasoning. Accordingly, the absence of a direct relationship between HL and opinions on vaccination in general, which we observed in the mediation model, might indicate the limitations of using an approach based solely on increasing general HL to combat vaccine hesitancy. This is a very important point, because over the last decade, the vaccine debate has become a way for the French public to express social tensions and political mistrust [54]. As our results suggest, interventions to improve the level of HL, such as cross-cutting strategies, might contribute to strengthening trust in the health care system.

Limitations

This study has limitations. First, its cross-sectional design prevented us from establishing causal relationships between the variables analyzed. Second, the quota sampling method we used does not guarantee the geographical and socioeconomic representativeness of the whole population of French adults. Our findings must therefore be interpreted with caution. Third, the SLAVACO survey was conducted during the COVID-19 pandemic, when vaccination was the subject of intense media attention, frequently affected by misinformation and disinformation. This may have strongly influenced participants' opinions on vaccination and amplified the mediating effect of trust in the health care system, which we observed. Lastly, following the framework by Parker [55], HL is influenced by both internal (cognitive, educational, and socioeconomic) and external (social, health care, and macroeconomic) factors. Our study did not systematically assess these determinants, highlighting the need for tools that capture the full spectrum of HL in future research.

Conclusions

Enhancing the level of HL may lead to greater trust in government health agencies and medical doctors, which are 2 essential dimensions for positively influencing hepatitis B vaccination opinions. To achieve the WHO's goal of hepatitis elimination by 2030, future public health strategies in France should take HL and these 2 dimensions into account to effectively reduce the country's currently high rate of hepatitis B vaccine hesitancy. From a public health perspective, these findings suggest that strategies should not only focus on providing clear information about vaccines and ensuring access, but also on actively strengthening trust in health care institutions and professionals. National communication campaigns aimed at correcting misconceptions about hepatitis B vaccination could be complemented by targeted interventions for groups most likely to hold negative opinions, such as women, individuals with intermediate education levels, those without chronic diseases, and higher-income populations. In parallel, incorporating HL modules into school curricula and community programs could provide a sustainable means of fostering trust and supporting informed decision-making.

For future research, it will be important to develop and apply vaccine-specific HL tools, to longitudinally assess the causal relationships between HL and the key determinants of vaccine hesitancy (7C-model), and to compare these dynamics across different vaccines. A broader assessment of trust encompassing all 6 components of the health care system—governance, financing, service delivery, human resources, health products and interventions, and health information—as defined by the WHO [56], would also provide a more comprehensive understanding. Conducting a similar survey in the post–COVID-19 context could also reveal different trends, reflecting the evolving public perceptions of vaccines and health authorities.

XSL•FO
RenderX

Acknowledgments

We would like to thank all the persons who contributed to the SLAVACO (Suivi Longitudinal des Attitudes à l'Égard d'un Vaccin Contre la COVID-19) project. Our thanks to Jude Sweeney, MSc, for the English revision and copyediting of this paper.

Funding

This research was supported by grants from the Agence Nationale de la Recherche (ANR-20-COVI-0035-01, ANR-20-COVI-0102, and ANR-22-CE36-0015-01) and the Agence Nationale de la Recherche Scientifique - Maladie Infectieuses Emergentes (ICOVAC). However, the funding sources had no part in the design of this study, the data analysis, or the writing of this paper. Additionally, BC is personally supported by the Aix-Marseille University PhD Programme.

Data Availability

The datasets used and/or analyzed during this study are available from the corresponding author on reasonable request. Information and agreement to use HLS_{19}-Q12 (12-item general health literacy scale) can be found on the M-POHL (Action Network on Measuring Population and Organizational Health Literacy) web page.

Authors' Contributions

BC, JM, and SB conceptualized this study. BC formally analyzed the data. BC wrote the original draft in preparation. BC, JM, and JKW wrote, reviewed, and edited this paper. All authors have read and agreed to the published version of this paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Table showing estimates of control variables for different mediation models.
[DOCX File , 59 KB - publichealth_v12i1e82496_app1.docx]

References

1. Gerlich WH. Medical virology of hepatitis B: how it began and where we are now. Virol J 2013;10:239 [FREE Full text] [doi: 10.1186/1743-422X-10-239] [Medline: 23870415]
2. Hepatitis B. World Health Organization. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/hepatitis-b [accessed 2026-01-29]
3. Hsu Y, Huang DQ, Nguyen MH. Global burden of hepatitis B virus: current status, missed opportunities and a call for action. Nat Rev Gastroenterol Hepatol 2023;20(8):524-537. [doi: 10.1038/s41575-023-00760-9] [Medline: 37024566]
4. Hepatitis B - Annual Epidemiological Report for 2022. European Centre for Disease Prevention and Control. 2025. URL: https://www.ecdc.europa.eu/en/publications-data/hepatitis-b-annual-epidemiological-report-2022 [accessed 2026-01-29]
5. Global health sector strategy on viral hepatitis 2016-2021, towards ending viral hepatitis. World Health Organization. 2016. URL: https://iris.who.int/bitstream/handle/10665/246177/WHO-HIV-2016.06-eng.pdf?sequence=1 [accessed 2026-03-01]
6. Guidelines for the prevention, diagnosis, care and treatment for people with chronic hepatitis B infection. World Health Organization. 2024. URL: https://www.who.int/publications/i/item/9789240090903 [accessed 2026-01-30]
7. Immunization coverage. World Health Organization. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/immunization-coverage [accessed 2026-01-30]
8. Brouard C, Saboni L, Gautier A, Chevaliez S, Rahib D, Richard J, 2016 Health Barometer Group. HCV and HBV prevalence based on home blood self-sampling and screening history in the general population in 2016: contribution to the new French screening strategy. BMC Infect Dis 2019;19(1):896 [FREE Full text] [doi: 10.1186/s12879-019-4493-2] [Medline: 31660879]
9. Prevention of hepatitis B and C in the EU/EEA. European Centre for Disease Prevention and Control. 2024. URL: https://data.europa.eu/doi/10.2900/703244 [accessed 2026-01-30]
10. Saboni L, Brouard C, Chevaliez S, Rahib D, Richard J. Prévalence des hépatites chroniques C et B, et antécédents de dépistage en population générale en 2016: contribution à une nouvelle stratégie de dépistage, Baromètre de Santé publique France-BaroTest. Bull Épidémiologique Hebd 2019 Sep 24(24-25):469-477 [FREE Full text]
11. Laporal S, Bruyand M, Lot F, Pillonel J, Brouard C. Fréquence des hépatites chroniques B et C, morbidité et mortalité parmi la population hospitalisée en France, 2005-2020. Bull Épidémiologique Hebd 2022 Feb 8(3-4):69-85 [FREE Full text]
12. Bordbar S, Hadian M, Mazaheri E, Shoara Z, Jafari A. Economic burden of hepatitis b at different stages of the disease: a systematic review study. Int J Prev Med 2024;15:50. [doi: 10.4103/ijpvm.ijpvm_3_24] [Medline: 39539576]

XSL•FO
RenderX

13. Ramière C, Roche L, Scholtès C, Iwaz J, Saison J, Ecochard R, et al. Evolution of the incidence of hepatitis B virus infection and immunization rates in a large French cohort born between 1960 and 1994. Clin Microbiol Infect 2016;22(10):889.e1-889.e7 [FREE Full text] [doi: 10.1016/j.cmi.2016.07.014] [Medline: 27451939]
14. Article L3111-2 - Code de la santé publique. Légifrance. 2018. URL: https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000036393260/2018-01-20 [accessed 2026-01-30]
15. Hépatite B | vaccination info service. Santé Publique France. URL: https://vaccination-info-service.fr/Les-maladies-et-leurs-vaccins/Hepatite-B [accessed 2025-10-01]
16. Laporal S, Tamandjou C, Lot F, Brouard C. Surveillance de l'hépatite B aiguë en France : données issues de la déclaration obligatoire, 2011-2022. Surveillance de l'hépatite B aiguë en France?: Données issues de la déclaration obligatoire 2024 Sep 24(16-17):369-377 [FREE Full text]
17. Brouard C, Laporal S, Cazein F, Saboni L, Bruyand M, Lot F. Dépistage des hépatites B et C en 2021 en France, enquête LaboHEP. Bull Épidémiologique Hebd 2023 Jul 25(15-16):276-286 [FREE Full text]
18. Tamandjou C, Laporal S, Lot F, Brouard C. Données épidémiologiques récentes sur les hépatites C, B et delta. Bull Épidémiologique Hebd 2023 Jul 25(311-317):15-16 [FREE Full text]
19. Stahl JP, Denis F, Gaudelus J, Cohen R, Lepetit H, Martinot A. Hepatitis B vaccination and adolescents: a lost generation. Med Mal Infect 2016;46(1):1-3. [doi: 10.1016/j.medmal.2015.11.002] [Medline: 26746325]
20. Cazals N. Couverture vaccinale de l'hépatite B chez l'adulte en cabinet de médecine générale: déterminants de la non vaccination et connaissance des modes de transmission. Méd Hum Pathol 2015:dumas-01513835 [FREE Full text] [doi: 10.1016/j.medmal.2018.04.345]
21. Vaux S, Gautier A, Nassany O, Bonmarin I, Randriamampianina S. Évolution de l'adhésion vaccinale en France et facteurs associés de 2000 à 2023. Méd Mal Infectieuses Form 2024 Jun;3(2):S127 [FREE Full text] [doi: 10.1016/j.mmifmc.2024.04.358]
22. Sørensen K, Van den Broucke S, Fullam J, Doyle G, Pelikan J, Slonska Z, (HLS-EU) Consortium Health Literacy Project European. Health literacy and public health: a systematic review and integration of definitions and models. BMC Public Health 2012;12:80 [FREE Full text] [doi: 10.1186/1471-2458-12-80] [Medline: 22276600]
23. Lorini C, Santomauro F, Donzellini M, Capecchi L, Bechini A, Boccalini S, et al. Health literacy and vaccination: a systematic review. Hum Vaccines Immunother 2018;14(2):478-488 [FREE Full text] [doi: 10.1080/21645515.2017.1392423] [Medline: 29048987]
24. Confavreux C, Suissa S, Saddier P, Bourdès V, Vukusic S, Vaccines in Multiple Sclerosis Study Group. Vaccinations and the risk of relapse in multiple sclerosis. N Engl J Med 2001;344(5):319-326. [doi: 10.1056/NEJM200102013440501] [Medline: 11172162]
25. Dubé E, Laberge C, Guay M, Bramadat P, Roy R, Bettinger JA. Vaccine hesitancy: an overview. Hum Vaccines Immunother 2013;9(8):1763-1773 [FREE Full text] [doi: 10.4161/hv.24657] [Medline: 23584253]
26. MacDonald NE, SAGE Working Group on Vaccine Hesitancy. Vaccine hesitancy: definition, scope and determinants. Vaccine 2015;33(34):4161-4164 [FREE Full text] [doi: 10.1016/j.vaccine.2015.04.036] [Medline: 25896383]
27. Galagali PM, Kinikar AA, Kumar VS. Vaccine hesitancy: obstacles and challenges. Curr Pediatr Rep 2022;10(4):241-248 [FREE Full text] [doi: 10.1007/s40124-022-00278-9] [Medline: 36245801]
28. Kisa A, Kisa S. Health conspiracy theories: a scoping review of drivers, impacts, and countermeasures. Int J Equity Health 2025;24(1):93 [FREE Full text] [doi: 10.1186/s12939-025-02451-0] [Medline: 40181408]
29. Ward JK, Peretti-Watel P, Bocquier A, Seror V, Verger P. Vaccine hesitancy and coercion: all eyes on France. Nat Immunol 2019;20(10):1257-1259. [doi: 10.1038/s41590-019-0488-9] [Medline: 31477920]
30. Tsai TI, Yu WR, Lee SYD. Is health literacy associated with greater medical care trust? Int J Qual Health Care 2018;30(7):514-519. [doi: 10.1093/intqhc/mzy043] [Medline: 29608676]
31. Larson HJ, Jarrett C, Eckersberger E, Smith DM, Paterson P. Understanding vaccine hesitancy around vaccines and vaccination from a global perspective: a systematic review of published literature, 2007-2012. Vaccine 2014;32(19):2150-2159. [doi: 10.1016/j.vaccine.2014.01.081] [Medline: 24598724]
32. Hair JF, Hult GTM, Ringle CM, Sarstedt M, Danks NP, Ray S. An introduction to structural equation modeling. In: Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R: A Workbook. Cham: Springer International Publishing; 2025:1-29.
33. The weight sharing method. Institut National de la Statistique et des Études Économiques. 2017. URL: https://www.insee.fr/en/statistiques/fichier/5398341/7-weight-sharing.pdf [accessed 2026-01-30]
34. Calibration techniques. Institut National de la Statistique et des Études Économiques. 2018. URL: https://www.insee.fr/en/statistiques/fichier/5398341/5-calibration-methods.pdf [accessed 2026-01-30]
35. Baromètre de Santé publique France 2021. Questionnaire / Volet métropole. Santé Publique France. 2022. URL: https://www.santepubliquefrance.fr/import/barometre-de-sante-publique-france-2021.-questionnaire-volet-metropole [accessed 2025-05-08]
36. Pelikan JM, Link T, Straßmayr C, Waldherr K, Alfers T, Bøggild H, HLS19 Consortium of the WHO Action Network M-POHL. Measuring comprehensive, general health literacy in the general adult population: the development and validation

XSL•FO
RenderX

of the HLS-Q12 instrument in seventeen countries. Int J Environ Res Public Health 2022;19(21):14129 [FREE Full text] [doi: 10.3390/ijerph192114129] [Medline: 36361025]
37. Touzani R, Rouquette A, Schultz E, Allaire C, Carrieri P, Mancini J, et al. Psychometric validation of the French version of two scales measuring general (HLS-Q12) and navigational (HLS-NAV) health literacy using the Rasch model. BMC Public Health 2024;24(1):3079 [FREE Full text] [doi: 10.1186/s12889-024-20504-x] [Medline: 39511498]
38. Tournay V, Cautrès B. Les déterminants politiques de la rationalité vaccinale: analyse d'une passion Française (II). Le Baromètre de la Confiance Politique, Science Po CEVIPOF, vague 12bis 2021 Sep:1-12 [FREE Full text]
39. Bach A. La bibliothèque de l'Institut national de la statistique et des études économiques (INSEE). In: histoirepolitique. Paris: INSEE; Jun 01, 2017:1-90.
40. Larson HJ, de Figueiredo A, Xiahong Z, Schulz WS, Verger P, Johnston IG, et al. The state of vaccine confidence 2016: global insights through a 67-country survey. eBioMedicine 2016;12:295-301 [FREE Full text] [doi: 10.1016/j.ebiom.2016.08.042] [Medline: 27658738]
41. Herroelen L, de Keyser J, Ebinger G. Central-nervous-system demyelination after immunisation with recombinant hepatitis B vaccine. Lancet 1991;338(8776):1174-1175. [doi: 10.1016/0140-6736(91)92034-y] [Medline: 1682594]
42. Mikaeloff Y, Caridade G, Rossier M, Suissa S, Tardieu M. Hepatitis B vaccination and the risk of childhood-onset multiple sclerosis. Arch Pediatr Adolesc Med 2007;161(12):1176-1182. [doi: 10.1001/archpedi.161.12.1176] [Medline: 18056563]
43. Bilan de pharmacovigilance et profil de sécurité d'emploi des vaccins contre l'hépatite B. Agence Française de Sécurité Sanitaire des Produits de Santé. 2012. URL: https://www.sentinelles971.com/wp-content/uploads/2012/12/HASBilan-de-pharmacovigilance-vaccins-contre-lhépatite-B.pdf [accessed 2026-01-30]
44. Berezin M, Eads A. Risk is for the rich? Childhood vaccination resistance and a culture of health. Soc Sci Med 2016;165:233-245. [doi: 10.1016/j.socscimed.2016.07.009] [Medline: 27499069]
45. Akosionu O, Virnig B, Call KT, Yuan J, Chanthanouvong S, Nguyen RHN. The impact of gender differences in attitudes and beliefs concerning HBV vaccination and screening in the Lao community. J Immigr Minor Health 2016;18(1):277-281. [doi: 10.1007/s10903-015-0160-x] [Medline: 25612922]
46. de Figueiredo A, Simas C, Karafillakis E, Paterson P, Larson HJ. Mapping global trends in vaccine confidence and investigating barriers to vaccine uptake: a large-scale retrospective temporal modelling study. Lancet 2020;396(10255):898-908 [FREE Full text] [doi: 10.1016/S0140-6736(20)31558-0] [Medline: 32919524]
47. European Union's Health Programme under the SANTE. 2019. URL: https://health.ec.europa.eu/system/files/2023-10/vac-pact_uptake_quick-guide_en.pdf [accessed 2026-01-30]
48. Czajka H, Czajka S, Biłas P, Pałka P, Jędrusik S, Czapkiewicz A. Who or what influences the individuals' decision-making process regarding vaccinations? Int J Environ Res Public Health 2020;17(12):4461 [FREE Full text] [doi: 10.3390/ijerph17124461] [Medline: 32575879]
49. Tran BX, Boggiano V, Nguyen LH, Latkin C, Nguyen H, Tran TT, et al. Media representation of vaccine side effects and its impact on utilization of vaccination services in Vietnam. Patient Prefer Adherence 2018;12:1717-1728 [FREE Full text] [doi: 10.2147/PPA.S171362] [Medline: 30233151]
50. Hyun S, Ko O, Kim S, Ventura WR. Sociocultural barriers to hepatitis B health literacy in an immigrant population: a focus group study in Korean Americans. BMC Public Health 2021;21(1):404 [FREE Full text] [doi: 10.1186/s12889-021-10441-4] [Medline: 33632203]
51. Thahir S, Morgan C, Ngimbi P, Kashamuka M, Ntambua S, Matondo J, et al. Promoting antiviral access and adherence in the expansion of hepatitis B prevention programs: insights from the Democratic Republic of Congo. J Pediatric Infect Dis Soc 2024;13(Supplement_5):S148-S152. [doi: 10.1093/jpids/piae081] [Medline: 39570895]
52. Lo Moro G, Ferrara M, Langiano E, Accortanzo D, Cappelletti T, De Angelis A, et al. Countering vaccine hesitancy: a systematic review of interventions to strengthen healthcare professionals' action. Eur J Public Health 2023;33(5):905-915 [FREE Full text] [doi: 10.1093/eurpub/ckad134] [Medline: 37581903]
53. Brosset E, Anderson E, Garrison A, Holford D, Fisher H, Peretti-Watel P, et al. Healthcare professionals' emotions of distance and connection when dealing with patients' vaccine hesitancy: interaction styles, values, and implications. Qual Health Res 2025:10497323251320921 [FREE Full text] [doi: 10.1177/10497323251320921] [Medline: 40211704]
54. Ward JK, Cortaredona S, Touzet H, Gauna F, Peretti-Watel P. Explaining political differences in attitudes to vaccines in France: partisan cues, disenchantment with politics, and political sophistication. J Health Polit Policy Law 2024;49(6):961-988. [doi: 10.1215/03616878-11373758] [Medline: 38836412]
55. Parker R. Measuring health literacy: What? So what? Now what? In: Measures of Health Literacy: Workshop Summary. Washington (DC): National Academies Press (US); 2009.
56. Monitoring the building blocks of health systems: a handbook of indicators and their measurement strategies. World Health Organization. 2010. URL: https://iris.who.int/bitstream/handle/10665/258734/9789241564052-eng.pdf [accessed 2026-01-30]

Abbreviations

HBV: hepatitis B virus
HL: health literacy

XSL•FO
RenderX

HLS19-Q12: 12-item general health literacy questionnaire used in Health Literacy Survey 2019-2021
INSEE: National Institute for Statistics and Economic Studies
RMSEA: root-mean-square error of approximation
SLAVACO: Suivi Longitudinal des Attitudes à l'Égard d'un Vaccin Contre la COVID-19
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 16.Aug.2025; peer-reviewed by K Achstetter, J Sun; comments to author 24.Sep.2025; revised version received 13.Oct.2025; accepted 27.Oct.2025; published 13.Feb.2026.

Please cite as:
Cissé B, Boyer S, Ward JK, Mancini J
Mediation Analysis of the Relationship Between Health Literacy and the French General Population's Opinions on Hepatitis B Vaccination: Representative Cross-Sectional Survey of the SLAVACO Project
JMIR Public Health Surveill 2026;12:e82496
URL: https://publichealth.jmir.org/2026/1/e82496
doi:10.2196/82496
PMID:41687090

©Bakary Cissé, Sylvie Boyer, Jeremy K Ward, Julien Mancini. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 13.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Association Between Gestational Blood Pressure Trajectories and Postpartum Normotension Recovery in Hypertensive Disorders: Retrospective Cohort Study

Zhijiang Liang[1*], MD, PhD; Huili Wei[2*], MA; Xiaojun Xu[3*], MA; Rong Xu[2], MA; Xinyue Yang[4], MA; Jinlian Dong[5], BMed; Xiaoyan Fan[5], MA; Yukun Chen[5], BMed; Meiling Feng[5], BMed; Xin Zhou[3], MD, PhD; Yijun Cai[5], BMed; Lijuan Lv[4], MD, PhD

[1]Department of Medical Information and Statistics, Guangdong Women and Children Hospital, Guangzhou, Guangdong, China

[2]Department of Public Health, School of Medicine, Jinan University, Guangzhou, Guangdong, China

[3]Department of Cardiology, Tianjin Medical University General Hospital, Tianjin, China

[4]Medical Genetic Center and Department of Obstetrics, Guangdong Women and Children Hospital, Guangzhou, Guangdong, China

[5]Medical Genetic Center and Department of Obstetrics, Gaozhou Hospital, Guangdong Women and Children Hospital, Gaozhou, Guangdong, China

[*]these authors contributed equally

Corresponding Author:
Lijuan Lv, MD, PhD
Medical Genetic Center and Department of Obstetrics
Guangdong Women and Children Hospital
521, Xingnan Avenue
Panyu District
Guangzhou, Guangdong, 511422
China
Phone: 86 02039151548
Email: bangbang10@163.com

Abstract

Background: Hypertensive disorders of pregnancy (HDP) may cause lasting vascular, cardiac, and renal damage, potentially increasing the risk of postpartum cardiovascular disease.

Objective: This study aimed to examine the association between gestational blood pressure (BP) trajectories in HDP and the risk of unrecovered BP at 6 weeks post partum.

Methods: A total of 3162 women with HDP were obtained from the antenatal care and the postpartum follow-up information system, between January 1, 2018, and December 31, 2024. Of the 3162 women included, 1674 had gestational hypertension, 607 had preeclampsia, 246 had chronic hypertension with superimposed preeclampsia, and 635 had chronic hypertension. Group-based trajectory modeling was used to fit systolic BP (SBP), diastolic BP (DBP), and mean arterial pressure (MAP) trajectories during pregnancy. Modified Poisson regression was used to assess the association between gestational BP trajectories and the risk of unrecovered BP at 6 weeks post partum.

Results: Trajectories of SBP, DBP, and MAP during pregnancy were significantly associated with unrecovered BP at 6 weeks post partum. For gestational hypertension, those with the high-consistent rise (adjusted relative risk [aRR] 2.493, 95% CI 1.093-5.689) and high-late surge SBP trajectories (aRR 4.535, 95% CI 1.884-10.917) were associated with a significantly increased risk of BP nonrecovery at 6 weeks post partum. Similar associations were observed for DBP and MAP. For chronic hypertension with superimposed preeclampsia, women with high-late surge in SBP (aRR 2.792, 95% CI 1.081-7.214), DBP (aRR 4.043, 95% CI 1.327-12.324), or MAP (aRR 4.018, 95% CI 1.462-11.045) had a significantly increased risk of BP nonrecovery at 6 weeks post partum. Among women with chronic hypertension, those with the high-consistent rise trajectories of SBP (aRR 2.557, 95% CI 1.256-5.207), DBP (aRR 3.862, 95% CI 1.673-8.913), and MAP (aRR 3.714, 95% CI 1.682-8.201) had a significantly increased risk of BP nonrecovery at 6 weeks post partum. Among women with preeclampsia, only high-consistent rise SBP trajectory remained significantly associated with unrecovered BP post partum (aRR 3.355, 95% CI 1.140-9.873). The high-consistent rise and high-late surge trajectories of SBP, DBP, and MAP in gestational hypertension started at similar initial levels and crossed at approximately 22 weeks of gestation.

XSL•FO
RenderX

Conclusions: The gestational BP trajectories in women with HDP are positively associated with the risk of unrecovered BP at 6 weeks post partum. Early identification of women at high risk for poor postpartum BP recovery through BP trajectory analysis may have important clinical implications for improving long-term cardiovascular outcomes in this population.

(JMIR Public Health Surveill 2026;12:e89295) doi:10.2196/89295

KEYWORDS

blood pressure trajectory; group-based trajectory modeling; hypertensive disorders of pregnancy; postpartum BP recovery status

Introduction

Hypertensive disorders of pregnancy (HDP) encompass a spectrum of conditions characterized by elevated blood pressure (BP) during gestation, affecting approximately 5%-12% of pregnancies. Based on the type of disease, HDP include gestational hypertension (GH), preeclampsia-eclampsia, chronic hypertension (CHTN) with superimposed preeclampsia, and CHTN [1,2]. HDP pose a serious threat to maternal health and are among the leading causes of maternal morbidity and mortality worldwide [3]. Women with a history of HDP are at increased risk of developing cardiovascular disease (CVD), the leading cause of death among women [4,5].

Although BP in most patients with HDP gradually normalizes after delivery, the exact time to recovery remains inconsistent across studies [6-8]. The vascular, cardiac, and renal damage caused by HDP may persist beyond pregnancy [9], potentially resulting in persistent postpartum hypertension [10] and an increased long-term risk of adverse cardiovascular events [9]. Postpartum hypertension after preeclampsia may take up to 2 years to resolve [11]. A study found that 39% of women still had hypertension three months after delivery, while this figure dropped to 18% after 2 years [11]. Therefore, when studying remaining hypertension after pregnancies complicated by hypertensive disorders, it is important to take this time frame into consideration [12].

Evidence suggests that the first 6 weeks post partum represent a critical window for cardiovascular risk intervention [8,13]. This aligns with Chinese clinical guidelines, which recommend a postpartum follow-up visit at 6 weeks for women with HDP [14,15]. The Hypertension Canada guidelines recommend continued hypertension management and monitoring until at least 6 weeks post partum in women with cardiovascular risk factors [16]. From 2014 to 2017, HDP, including preeclampsia, accounted for 6.8% of pregnancy-related deaths in the United States, with 65% occurring within 6 weeks post partum [17]. Thus, assessment of BP recovery at this time point is of significant clinical relevance.

While clinical risk factors for impaired postpartum BP recovery are generally considered to include HDP subtype, BP at onset, and maximum antepartum systolic BP (SBP) [18-20], these single-time measurements are susceptible to variability. In contrast, BP trajectories derived from repeated measurements better capture cardiovascular adaptation during pregnancy and are stronger predictors of adverse outcome. For instance, first-trimester BP trajectories correlate strongly with later hypertension [21]. Thus, analyzing BP trajectories rather than isolated readings may offer superior insight into postpartum BP recovery in HDP.

In this study, we classified patients with HDP into 4 subtypes: GH, preeclampsia, CHTN with superimposed preeclampsia, and CHTN [14]. We modeled trajectories of SBP, diastolic BP (DBP), and mean arterial pressure (MAP) across multiple time points during pregnancy within each subtype to evaluate the association between BP trajectories during pregnancy and the risk of unrecovered BP at 6 weeks post partum. By elucidating the heterogeneity of BP changes across HDP subtypes, we aim to improve risk stratification and management of HDP. Furthermore, identifying high-risk trajectory may enable early detection of women at increased risk of poor BP recovery, thereby providing a scientific basis for individualized postpartum follow-up and early intervention of future CVD.

Methods

Study Participants

This study was mainly conducted in the western region of Guangdong Province, China, including Zhanjiang, Maoming, and Yangjiang. Data were obtained from the antenatal care and the postpartum follow-up information system. Routine clinical data from local maternity care, generated through standard practice, were used in this study without active recruitment. Diagnostic criteria for HDP were based on the Chinese Society of Obstetrics and Gynecology guidelines [14], consistent with international standards [2,22]. GH was defined as new-onset hypertension (SBP ≥140 mm Hg and/or DBP ≥90 mm Hg) after 20 weeks of gestation without proteinuria or other features of preeclampsia. Preeclampsia was diagnosed as new-onset hypertension after 20 weeks with either significant proteinuria or, in its absence, new-onset maternal organ dysfunction or placental-fetal compromise (eg, fetal growth restriction). CHTN referred to hypertension present before 20 weeks of gestation. CHTN with superimposed preeclampsia was diagnosed in women with CHTN who developed new features of preeclampsia after 20 weeks. Baseline information was collected for pregnant women who initiated antenatal records between January 1, 2018, and December 31, 2024. Inclusion criteria were (1) diagnosis of 1 of the 4 HDP subtypes (GH, preeclampsia, CHTN with superimposed preeclampsia, or CHTN), (2) at least 1 BP measurement in each trimester, with a total of ≥3 measurements during pregnancy, and (3) available BP data at 6-week postpartum follow-up. Exclusion criteria were (1) age <18 years, (2) history of smoking or alcohol consumption, (3) multifetal pregnancy, (4) women with severe CVDs (such as valvular heart disease, myocarditis sequelae and arrhythmias requiring medication), hematologic disorders (including anemia, thrombophilia, or coagulation dysfunction),

XSL•FO
RenderX

endocrine diseases (such as medication-requiring diabetes mellitus, thyroid diseases, and prolactinoma), or immune system diseases requiring hormone therapy, and (5) women with missing or incomplete BP data from antenatal or follow-up visits (only SBP or DBP available). After applying the inclusion and exclusion criteria, 3162 women were included in the final analysis. Specifically, 1674, 607, 246, and 635 pregnant women were classified as having GH, preeclampsia, CHTN with superimposed preeclampsia, and CHTN, respectively. A flowchart for the selection process of the study is provided in Figure S1 in Multimedia Appendix 1.

BP Measurement During Pregnancy

According to the Chinese Guideline of Preconception and Prenatal Care [23], pregnant women are recommended to have regular checkups, including BP measurement, at 7 specific gestational windows. BP is measured at every antenatal visit. However, as this is a retrospective study relying on routine health information systems, not all measured BP values were consistently recorded or available in the databases. Therefore, to ensure data completeness and consistency across all included women, we applied the criterion of at least 1 recorded BP measurement per trimester (first, second, and third). The description of BP measurement frequency during pregnancy is provided in Tables S10 and S11 in Multimedia Appendix 1. SBP and DBP measurements during pregnancy were obtained from the electronic medical record system. MAP was calculated as (SBP + 2 × DBP)/3. According to the Guidelines for the Management of Hypertensive Disorders of Pregnancy (2015), participants were required to rest for at least 5 minutes prior to each BP measurement. HDP are defined as SBP ≥140 mm Hg or DBP ≥90 mm Hg on at least 2 measurements in the same arm. In women with newly detected elevated BP, a repeat measurement after ≥4 hours is required, and hypertension is diagnosed if the initial and repeat measurements both show SBP ≥140 mm Hg or DBP ≥90 mm Hg. In cases of severe hypertension, defined as SBP ≥160 mm Hg or DBP ≥110 mm Hg, the diagnosis can be confirmed by a repeat measurement after several minutes.

Assessment of Covariates

Maternal baseline information during the study period was extracted from the electronic medical record system. Data included demographic characteristics (maternal age and education level), obstetric history (parity status, vaginal deliveries, cesarean deliveries, gravidity, abortion history, and gynecological surgery history), delivery-related factors (mode of delivery, delivery gestational age, and intrapartum BMI), and neonatal characteristics (sex, birth weight, and Apgar scores at 1, 5, and 10 minutes). Intrapartum BMI was calculated as weight at delivery (kg)/height (m^2). Missing covariates were imputed using a random forest–based method to ensure the reliability of the analyses. Covariates adjusted in modified Poisson regression models were determined based on previous literature [21,24] and results from univariate analyses, including maternal age, education level, gravidity, delivery gestational age, parity status, mode of delivery, neonatal birth weight, and sex. As data on medication use during pregnancy and the postpartum period were not available, this variable could not be adjusted for in this analysis.

Exposure and Outcome

In this study, the exposure was defined as BP trajectory during pregnancy. Group-based trajectory modeling (GBTM) was conducted using the TRAJ procedure (the Stata command for group-based trajectory modeling); in Stata (version 18.0; Stata Corp) to identify SBP, DBP, and MAP trajectories. GBTM classifies individuals with similar developmental patterns within a heterogeneous population, allowing for better subgroup characterization [25,26]. As BP measurements are continuous and normally distributed, trajectory fitting used a censored normal model [27]. Since no universally accepted criteria exist for optimal model selection [28], objective indices, trajectory plot simplicity, clinical interpretability, and practical use were considered to select the final model in this study [29]. Starting from the third-order polynomial, models with various polynomial orders were repeatedly constructed and compared. Model fitting parameters for GBTM are presented in Tables S1-S8 in Multimedia Appendix 1. The objective criteria for selecting the optimal GBTM model included (1) lower absolute values of Bayesian information criterion and Akaike information criterion indicating better model fit, (2) average posterior probability of assignment >0.7 for each trajectory group, (3) odds of correct classification >5.0 for each trajectory group, (4) each trajectory group accounting for at least 5% of the sample, (5) good concordance between the posterior probability of group membership (Pj) and the actual proportion of group membership (πj), (6) minimum proportion of each trajectory group [30], and (7) statistical significance of the highest-order polynomial term in each trajectory.

The outcome was the BP recovery status at 6 weeks post partum, defined as a binary variable. BP recovery to normotension at 6 weeks post partum was defined as SBP <140 mm Hg and DBP <90 mm Hg, while BP nonrecovery to normotension was defined as SBP ≥140 mm Hg or DBP ≥90 mm Hg (1 mm Hg=0.133 kPa). BP measurements at 6 weeks post partum were measured in patients with resting HDP using an electronic sphygmomanometer. BP measured at 6 weeks post partum was obtained by a qualified physician. If the reading was abnormal (≥140/90 mm Hg), the measurement was repeated.

Statistical Analysis

Statistical analyses were conducted using R software (version 4.4.2; R Foundation for Statistical Computing) and Stata 18.0. Statistical tests were 2-sided, and $P \leq .05$ was considered statistically significant. The normality of continuous variables was assessed by the Shapiro-Wilk test, and homogeneity of variance by the Bartlett test. Variables that met normality and homogeneity of variance assumptions were presented as mean (SD), and for multiple group comparisons, one-way analysis of variance (ANOVA) was performed. Variables not meeting normality assumptions were presented as median (IQR), and the Kruskal-Wallis test was used for multiple group comparisons. Categorical variables were expressed as n (%), with group comparisons performed using the chi-square test or Fisher exact test. Postpartum BP recovery status at 6 weeks (nonrecovery=1 and recovery=0) was used as the dependent

XSL•FO
RenderX

variable, and SBP, DBP, and MAP trajectory groups as independent variables. Modified Poisson regression was used to estimate risk ratios and 95% CIs for the association between gestational BP trajectories and postpartum BP nonrecovery. Sensitivity analysis was conducted among individuals with at least 4 BP measurements during pregnancy, including at least 1 per trimester, by repeating the trajectory fitting and regression analyses to assess the robustness of the main findings (Figure S2 and Table S9 in Multimedia Appendix 1).

Ethical Considerations

This study complied with the Declaration of Helsinki and was approved by the ethics committee of Guangdong Women and Children Hospital (number 20251125). Informed consent was waived due to the retrospective study design. All data were anonymized to protect participant confidentiality. No compensation was provided to participants.

Results

Baseline Characteristics of Study Participants

A total of 3162 pregnant women were included in the study. Of these, 1674 were diagnosed with GH, 607 with preeclampsia, 246 with CHTN with superimposed preeclampsia, and 635 with CHTN. The proportions of women with gravidity ≥3 were 38.11% (638/1674), 42.17% (256/607), 41.87% (103/246), and 50.08% (318/635), respectively. The proportions of multiparous women were 53.11% (889/1674), 60.13% (365/607), 55.69% (137/246), and 68.35% (434/635), respectively. The median maternal age was 31.00 (27.00–35.00), 30.00 (27.00–34.00), 32.00 (28.00–36.00), and 33.00 (30.00–37.00) years, respectively. The median gestational age at delivery was 38.71 (37.57–39.71), 38.71 (37.43–39.57), 38.00 (36.71–39.11), and 38.29 (37.29–39.29) weeks, respectively (Table 1).

XSL•FO
RenderX

Table 1. Baseline characteristics of 3162 women with different types of hypertensive disorders of pregnancy in Western Guangdong, China, 2018-2024.

Characteristics	GH[a] (n=1674)	Preeclampsia (n=607)	CHTN[b] with superimposed preeclampsia (n=246)	CHTN (n=635)	*P* value
Maternal age, median (IQR)	31.00 (27.00-35.00)	30.00 (27.00-34.00)	32.00 (28.00-36.00)	33.00 (30.00-37.00)	<.001
Education level, n (%)					.06
Above college	404 (24.13)	130 (21.42)	43 (17.48)	157 (24.72)	
Below high school	1270 (75.87)	477 (78.58)	203 (82.52)	478 (75.28)	
Gravidity, n (%)					<.001
1	633 (37.81)	187 (30.81)	81 (32.93)	151 (23.78)	
2	403 (24.07)	164 (27.02)	62 (25.20)	166 (26.14)	
≥3	638 (38.11)	256 (42.17)	103 (41.87)	318 (50.08)	
Parity status, n (%)					<.001
Primipara	785 (46.89)	242 (39.87)	109 (44.31)	201 (31.65)	
Multipara	889 (53.11)	365 (60.13)	137 (55.69)	434 (68.35)	
Number of cesarean sections, n (%)					<.001
≥1	234 (13.98)	85 (14.00)	54 (21.95)	168 (26.46)	
0	1440 (86.02)	522 (86.00)	192 (78.05)	467 (73.54)	
Number of vaginal deliveries, n (%)					.003
≥1	698 (41.70)	297 (48.93)	98 (39.84)	299 (47.09)	
0	976 (58.30)	310 (51.07)	148 (60.16)	336 (52.91)	
Gynecological surgery history, n (%)					<.001
No	1473 (87.99)	538 (88.63)	201 (81.71)	512 (80.63)	
Yes	201 (12.01)	69 (11.37)	45 (18.29)	123 (19.37)	
Abortion history, n (%)					<.001
No	1288 (76.94)	447 (73.64)	175 (71.14)	438 (68.98)	
Yes	386 (23.06)	160 (26.36)	71 (28.86)	197 (31.02)	
Delivery gestational age, median (IQR)	38.71 (37.57-39.71)	38.71 (37.43-39.57)	38.00 (36.71-39.11)	38.29 (37.29-39.29)	<.001
Mode of delivery, n (%)					<.001
Cesarean section	926 (55.32)	307 (50.58)	176 (71.54)	390 (61.42)	
Vaginal delivery	748 (44.68)	300 (49.42)	70 (28.46)	245 (38.58)	
Intrapartum BMI, median (IQR)	28.08 (26.95-29.21)	28.00 (26.79-29.30)	28.38 (27.18-29.61)	28.31 (27.28-29.50)	.002
Neonatal sex, n (%)					.60
Male	900 (53.76)	318 (52.39)	137 (55.69)	355 (55.91)	
Female	774 (46.24)	289 (47.61)	109 (44.31)	280 (44.09)	
Neonatal birth weight, median (IQR)	3000.00 (2650.00-3387.50)	3000.00 (2580.00-3400.00)	2850.00 (2270.00-3300.00)	3000.00 (2650.00-3300.00)	.001
Apgar score at 1 minute, median (IQR)	10.00 (10.00-10.00)	10.00 (10.00-10.00)	10.00 (9.00-10.00)	10.00 (10.00-10.00)	.003
Apgar score at 5 minutes, median (IQR)	10.00 (10.00-10.00)	10.00 (10.00-10.00)	10.00 (10.00-10.00)	10.00 (10.00-10.00)	.10
Apgar score at 10 minutes, median (IQR)	10.00 (10.00-10.00)	10.00 (10.00-10.00)	10.00 (10.00-10.00)	10.00 (10.00-10.00)	.01

[a]GH: gestational hypertension.

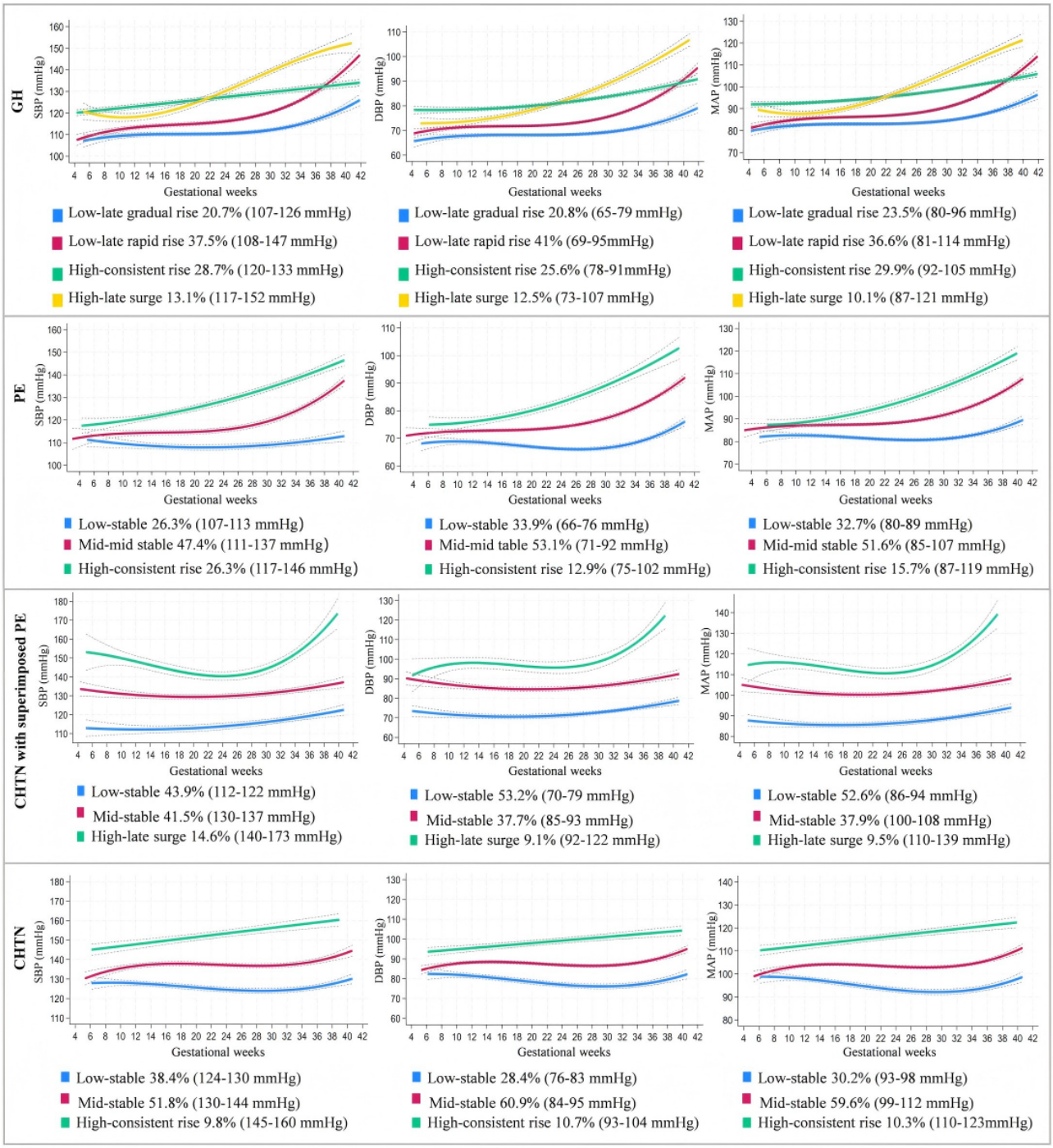
[b]CHTN: chronic hypertension.

XSL•FO
RenderX

Trajectories of BP During Pregnancy

A total of 4 trajectories of SBP, DBP, and MAP were identified for GH, and 3 for preeclampsia, CHTN with superimposed preeclampsia, and CHTN. Trajectories were named according to overall relative position and slope and were consistent across SBP, DBP, and MAP within each subtype. For GH, the identified trajectories were (1) low-late gradual rise trajectory, (2) low-late rapid rise trajectory, (3) high-consistent rise trajectory, and (4) high-late surge trajectory (Figure 1). For preeclampsia, trajectories were (1) low-stable trajectory, (2) mid-mid stable trajectory, and (3) high-consistent rise trajectory (Figure 1). For CHTN with superimposed preeclampsia, trajectories were (1) low-stable trajectory, (2) mid-stable trajectory, and (3) high-late surge trajectory (Figure 1). For CHTN, the trajectories were (1) low-stable trajectory, (2) mid-stable trajectory, and (3) high-consistent rise trajectory (Figure 1).

Figure 1. Trajectories of systolic blood pressure (SBP), diastolic blood pressure (DBP), and mean arterial pressure (MAP) during pregnancy in different types of hypertensive disorders of pregnancy in Western Guangdong, China, 2018-2024. CHTN: chronic hypertension; GH: gestational hypertension; PE: preeclampsia.

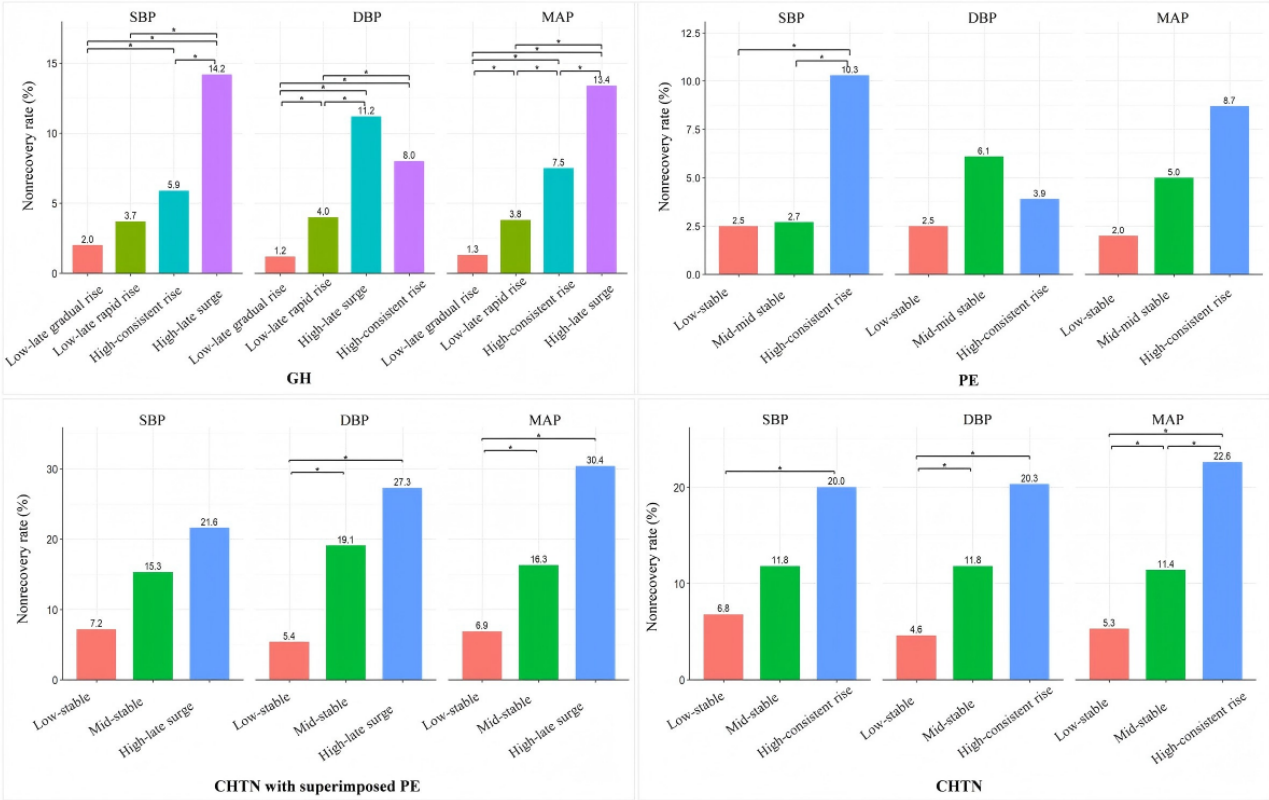


XSL•FO
RenderX

Incidence of Unrecovered BP at 6 Weeks Post Partum by BP Trajectories During Pregnancy

For GH, the high-late surge SBP trajectory was associated with the highest rate of BP nonrecovery at 6 weeks post partum (29/204, 14.2%) compared to the other 3 SBP trajectories. For DBP, there was no statistically significant difference in nonrecovery rates between the high-late surge and high-consistent rise trajectories, while significant differences were observed among the remaining DBP trajectories. For MAP, significant differences were observed among all trajectories, with the high-late surge trajectory showing the highest nonrecovery rate (22/164, 13.4%; Figure 2). For preeclampsia, only the high-consistent rise SBP trajectory had a higher rate of BP nonrecovery at 6 weeks post partum (16/155, 10.3%) than the mid-mid stable (8/293, 2.7%) and low-stable (4/159, 2.5%) SBP trajectories (Figure 2). For CHTN with superimposed preeclampsia, the high-late surge DBP trajectory had a higher rate of BP nonrecovery (6/22, 27.3%), compared to the low-stable (7/130, 5.4%) trajectory. MAP trajectories showed a similar distribution (Figure 2). For CHTN, the high-consistent rise SBP trajectory was associated with a higher rate of BP nonrecovery at 6 weeks post partum (12/60, 20%) than the low-stable trajectory (16/235, 6.8%). DBP and MAP trajectories showed a similar distribution of nonrecovery rates as the corresponding SBP trajectories (Figure 2).

Figure 2. Incidence of unrecovered blood pressure (BP) at 6 weeks postpartum by BP trajectories during pregnancy in different types of hypertensive disorders of pregnancy in Western Guangdong, China, 2018-2024. CHTN: chronic hypertension; DBP: diastolic blood pressure; GH: gestational hypertension; MAP: mean arterial pressure; PE: preeclampsia; SBP: systolic blood pressure.



Association Between BP Trajectories During Pregnancy and the Risk of Nonrecovery to Normotension at 6 Weeks Post Partum

In the adjusted model for GH, compared to pregnant women with the low-late gradual rise SBP trajectory, those with the high-consistent rise (adjusted relative risk [aRR] 2.493, 95% CI 1.093-5.689) and high-late surge SBP trajectories (aRR 4.535, 95% CI 1.884-10.917) were associated with a significantly increased risk of BP nonrecovery at 6 weeks post partum. Similar associations were observed for MAP, with the high-consistent rise (aRR 4.838, 95% CI 1.890-12.382) and high-late surge (aRR 6.717, 95% CI 2.414-18.691) trajectories showing increased risks. For DBP trajectories, the low-late rapid rise (aRR 3.064, 95% CI 1.075-8.736), high-late surge (aRR 6.452, 95% CI 2.103-19.798), and high-consistent rise (aRR 5.603, 95% CI 1.941-16.177) trajectories were all significantly associated with increased risk of BP nonrecovery at 6 weeks post partum, compared to the low-late gradual rise trajectory (Table 2).

In the adjusted model for preeclampsia, only high-consistent rise SBP trajectory remained significantly associated with BP nonrecovery (aRR 3.355, 95% CI 1.140-9.873; Table 2).

In the adjusted model for CHTN with superimposed preeclampsia, compared to the low-stable trajectory, women with high-late surge in SBP (aRR 2.792, 95% CI 1.081-7.214), DBP (aRR 4.043, 95% CI 1.327-12.324), or MAP (aRR 4.018, 95% CI 1.462-11.045) had a significantly increased risk of BP nonrecovery at 6 weeks post partum. Additionally, women with mid-stable DBP trajectory (aRR 3.412, 95% CI 1.467-7.937) and MAP trajectory (aRR 2.280, 95% CI 1.030-5.049) were also associated with higher risk (Table 2).

XSL•FO
RenderX

Table 2. The association between blood pressure (BP) trajectories during pregnancy and the risk of unrecovered BP at 6 weeks among women with different types of hypertensive disorders of pregnancy (HDP) in Western Guangdong, China, 2018-2024.

HDP subtype, BP index, and trajectory group	aRR[a] (95% CI)
GH[b]	
SBP[c]	
Low-late rapid rise	1.549 (0.662-3.623)
High-consistent rise	2.493 (1.093-5.689)
High-late surge	4.535 (1.884-10.917)
DBP[d]	
Low-late rapid rise	3.064 (1.075-8.736)
High-consistent rise	5.603 (1.941-16.177)
High-late surge	6.452 (2.103-19.798)
MAP[e]	
Low-late rapid rise	2.558 (0.983-6.656)
High-consistent rise	4.838 (1.890-12.382)
High-late surge	6.717 (2.414-18.691)
Preeclampsia	
SBP	
Mid-mid stable	0.976 (0.299-3.187)
High-consistent rise	3.355 (1.140-9.873)
DBP	
Mid-mid stable	2.018 (0.758-5.373)
High-consistent rise	0.973 (0.227-4.180)
MAP	
Mid-mid stable	2.144 (0.692-6.643)
High-consistent rise	3.445 (0.919-12.919)
CHTN[f] with superimposed preeclampsia	
SBP	
Mid-stable	1.937 (0.826-4.543)
High-late surge	2.792 (1.081-7.214)
DBP	
Mid-stable	3.412 (1.467-7.937)
High-late surge	4.043 (1.327-12.324)
MAP	
Mid-stable	2.280 (1.030-5.049)
High-late surge	4.018 (1.462-11.045)
CHTN	
SBP	
Mid-stable	1.638 (0.932-2.877)
High-consistent rise	2.557 (1.256-5.207)
DBP	
Mid-stable	2.536 (1.224-5.256)
High-consistent rise	3.862 (1.673-8.913)

XSL•FO
RenderX

HDP subtype, BP index, and trajectory group	aRR[a] (95% CI)
MAP	
Mid-stable	2.095 (1.071-4.098)
High-consistent rise	3.714 (1.682-8.201)

[a]aRR: adjusted relative risk.
[b]GH: gestational hypertension.
[c]SBP: systolic blood pressure.
[d]DBP: diastolic blood pressure.
[e]MAP: mean arterial pressure.
[f]CHTN: chronic hypertension.

In the adjusted model for CHTN, compared to pregnant women with the low-stable trajectory, those with the high-consistent rise trajectories of SBP (aRR 2.557, 95% CI 1.256-5.207), DBP (aRR 3.862, 95% CI 1.673-8.913), and MAP (aRR 3.714, 95% CI 1.682-8.201), as well as mid-stable trajectories of DBP (aRR 2.536, 95% CI 1.224-5.256) and MAP (aRR 2.095, 95% CI 1.071-4.098) had a significantly increased risk of BP nonrecovery at 6 weeks post partum (Table 2).

Sensitivity Analysis

In the sensitivity analysis, 1656 individuals were diagnosed with GH, 603 with preeclampsia, 245 with CHTN with superimposed preeclampsia, and 631 with CHTN. For GH, 4 trajectories were identified for each BP type, while 3 trajectories were identified for each of the other subtypes. The changing trends of each BP trajectory were highly similar to those in the main analysis: most trajectories remained stable in early pregnancy, followed by a marked increase in the third trimester. Sustained-rising and late-acceleration BP patterns were still observed and remained statistically associated with unresolved postpartum BP ($P<.05$). For instance, regarding SBP, pregnant women who exhibited the high-consistent rise trajectories were associated with a significantly increased risk of BP nonrecovery at 6 weeks post partum (GH: aRR 2.529, 95% CI 1.110-5.765; preeclampsia: aRR 3.230, 95% CI 1.101-9.477; CHTN: aRR 2.575, 95% CI 1.262-5.250). For CHTN with superimposed preeclampsia, women with high-late surge in SBP (aRR 2.759, 95% CI 1.068-7.133) had a significantly increased risk of BP nonrecovery at 6 weeks post partum. This is similar to the results of our main analysis (Figure S2 and Table S9 in Multimedia Appendix 1).

Discussion

Principal Findings

In this study, GBTM were used to fit BP trajectories in pregnancy for different HDP subtypes of pregnant women. The results revealed variations in the number and shape of trajectories across HDP subtypes, which is consistent with the findings by Sinkey et al [31], who also reported divergent BP trajectories among different HDP categories [31]. Another study investigating dynamic BP changes during pregnancy similarly indicated that BP trajectories reflect, to some extent, the type of HDP, with SBP, DBP, and MAP following generally consistent overall trends within the same subtype, differing mainly in magnitude and timing [32]. In our analysis, we also found that SBP and MAP trajectories within the same subtype were highly comparable. Although SBP and DBP trends were not entirely consistent within subtypes, this finding supports the view that these 2 parameters may follow a distinct trajectory pattern [33]. Furthermore, our results demonstrated that women with HDP who exhibited higher overall BP trajectory levels during pregnancy had a greater risk of being nonrecovery to normotension at 6 weeks post partum. A prospective cohort study examining the association between early-pregnancy BP trajectories and postpartum hypertension also found a graded increase in risk for postpartum hypertension, ranging from moderate-rapid-decline to sustained-high trajectories [21]. Other studies have similarly identified risk gradients from low to high for DBP and MAP trajectories [34,35]. Within the GH subgroup, we observed that both the high-consistent rise trajectory and high-late surge trajectory for SBP, DBP, and MAP started at similar baseline levels but crossed around 22 weeks of gestation. Moreover, women with high-late surge trajectory had a higher risk of unrecovered BP at 6 weeks post partum compared to those with high-consistent rise trajectory. These findings suggest that the period around 22 weeks of gestation may represent a critical window for risk stratification and intervention. This timing aligns with the characteristic hemodynamic shift in pregnancy, during which BP normally reaches its lowest point before commencing a subsequent rise [36-38]. The divergence of pathological trajectories we identified at this juncture indicates that women who progress to GH begin to exhibit distinct deviations from the expected physiological pattern. The crossing of trajectories likely captures differing degrees of maladaptation in vascular function. Consequently, assessing BP trends across this window thus offers a dynamic metric for early risk stratification regarding postpartum BP recovery [39,40]. For women with elevated BP in the first and second trimester, intensified monitoring during the second and third trimester is recommended to facilitate timely detection of accelerating BP trends, accurate trajectory classification, and informed initiation of antihypertensive therapy. However, further research and clinical trials are warranted to validate these observations.

A systematic review showed that in women with HDP, SBP did not change significantly in the first 6 months of pregnancy and began to rise by 31 mm Hg in the third trimester [33]. Our study also identified similar patterns. Specifically, the low-late gradual rise and low-late rapid rise SBP trajectories for GH, the low-stable and mid-mid stable SBP trajectories for preeclampsia,

XSL•FO
RenderX

the low-stable and mid-stable SBP trajectories for CHTN with superimposed preeclampsia, and CHTN were all relatively stable during the first and second trimesters but increased in the third trimester, with the degree of increase varying by subtype. Additionally, across all 4 HDP subtypes, we identified a subset of women with specific SBP trajectories characterized by high baseline levels and SBP values exceeding 140 mm Hg in the third trimester. Although the high-consistent rise trajectory in GH did not exceed this threshold in the third trimester, it had the highest baseline level and showed a linear rising trend. Our results indicate that these women had greater difficulty achieving BP recovery at 6 weeks post partum, consistent with previous studies reporting that more severe and longer duration of HDP are associated with greater vascular, renal, and multisystem injury, leading to poorer postpartum BP recovery [8,18]. In this study, SBP trajectories across all 4 HDP types were significantly associated with the risk of unrecovered BP at 6 weeks post partum. This may be attributable to SBP being a better predictor of CVD [41], providing biological plausibility for our findings.

Regarding DBP trends, our study showed different trends, mainly in the second trimester. Specifically, women with GH, preeclampsia, and CHTN who exhibited high-consistent rise trajectories, as well as those with GH showing high-late surge trajectory, all demonstrated rising DBP during the second trimester. This is similar to previous research reports stating that in women with HDP, DBP shows an upward trend, increasing by approximately 5 mm Hg in the second trimester compared with the first trimester [42,43]. Additionally, some trajectories in our study—such as the low-late gradual rise and low-late rapid rise trajectories in GH, and the low-stable and mid-mid stable trajectories in preeclampsia, as well as the low-stable and mid-stable trajectories in CHTN—did not show significant rising or falling trends during the second trimester but began to increase toward the end of the second trimester. This trend differs from most previous reports. Similar trends were observed in CHTN with superimposed preeclampsia. Notably, previous studies have suggested that women with GH may experience a distinct time point—occurring at various stages of pregnancy—at which DBP rises abruptly. Before this point, DBP typically remains stable and within the normal range, but increases sharply thereafter [44], further supporting the concept of heterogeneous dynamic BP changes during pregnancy. Finally, across all subtypes, DBP trajectories significantly associated with unrecovered BP at 6 weeks post partum ($P<.05$) were characterized by higher overall levels. Specifically, in women with CHTN with superimposed preeclampsia and those with CHTN, the relevant trajectories reached DBP ≥90 mm Hg before 20 weeks of gestation, whereas in women with GH, this threshold was attained after 20 weeks.

In this study, baseline levels and trends of MAP were similar to those of SBP. Similar to the SBP results, pregnant women with GH or CHTN who exhibited high-consistent rise MAP trajectory, as well as those with GH and CHTN with superimposed preeclampsia who exhibited high-late surge MAP trajectory, were all at high risk for unrecovered BP at 6 weeks post partum. Unlike the SBP trajectories, the mid-stable MAP trajectory in the CHTN with superimposed preeclampsia and the CHTN remained statistically associated with the risk of unrecovered BP at 6 weeks post partum. Previous studies have shown that MAP has predictive value for adverse pregnancy outcomes [45] and may be superior to both SBP and DBP in clinical application [46], possibly because elevated MAP induces abnormalities in the nitric oxide (NO) synthesis process, thereby triggering a series of vascular dysfunctions [47]. However, research on the application of MAP in pregnancy is still limited [48], and further studies are needed.

Giorgione et al [49] demonstrated that persistent postpartum hypertension following HDP is predicted by a confluence of demographic factors (age or Afro-Caribbean ethnicity), clinical metrics (early pregnancy MAP and BMI), and, most significantly, peripartum echocardiographic evidence of pathological cardiac remodeling (elevated left ventricular mass index or relative wall thickness) and subclinical myocardial dysfunction (impaired global longitudinal strain and diastolic indices). Their findings posit that persistent hypertension reflects underlying cardiovascular maladaptations unmasked by pregnancy, which can be effectively identified through peripartum screening [49]. Little is known about the potential mechanisms linking dynamic changes in BP during pregnancy to postpartum BP recovery status. A study using a new 3D miniheart model revealed that GH may activate cell death signaling pathways and promote cardiac fibrosis, leading to cardiovascular damage; preeclampsia may cause vascular dysfunction-related inflammation, protein disorders, and endothelial dysfunction, thereby adversely affecting cardiovascular health [50]. A retrospective analysis of patients with preeclampsia and CHTN with superimposed preeclampsia demonstrated that autonomic dysregulation may occur within the first week post partum [51]. An imbalance between sympathetic and parasympathetic regulation can promote elevated BP and peripheral vasoconstriction, thereby increasing cardiac afterload and myocardial oxygen consumption [51]. Moreover, women with CHTN, who typically have hypertension prior to pregnancy, often present with underlying endothelial dysfunction. This dysfunction may be further aggravated by systemic inflammation during pregnancy, contributing to persistent postpartum hypertension [52].

Endothelial dysfunction is an early pathological change in hypertension [53,54]. Soluble fms-like tyrosine kinase-1 (sFlt1) and placental growth factor (PLGF) are important factors affecting maternal endothelial function, and their abnormal expression is closely related to endothelial dysfunction. sFlt1 is an antiangiogenic soluble protein, while PLGF promotes placental angiogenesis and repair [55,56]. Studies have shown that increased circulating sFlt1 and decreased proangiogenic factors such as PLGF contribute to impaired placental angiogenesis [57]. The imbalance of angiogenic factors can lead to maternal systemic endothelial dysfunction and vascular hyperreactivity, and clinical symptoms such as hypertension, proteinuria, and organ dysfunction [58]. One study showed that serum PLGF levels in the HDP group were significantly lower than those in the normal pregnancy group and declined further with increasing disease severity. Specifically, PLGF levels showed a significant stepwise decline from GH to mild preeclampsia and severe preeclampsia [59]. NO is a key endogenous molecule for regulating vascular tone and

XSL•FO
RenderX

maintaining BP stability [60]. Studies indicate that plasma endothelial NO synthase levels are significantly lower in women with GH and preeclampsia compared to normal pregnant women, with the lowest levels in the preeclampsia group. Reduced endothelial NO synthase leads to impaired NO synthesis [61]. Decreased NO bioavailability results in vascular relaxation disorders, increased systemic vascular resistance, elevated BP, and impaired placental and fetal development [62]. In addition, elevated asymmetric dimethylarginine, reduced dimethylarginine dimethylaminohydrolases, and insufficient L-arginine transport in patients with preeclampsia also inhibit NO production [62].

Based on these mechanisms and the trajectory disparities identified in our study, we recommend stratified management for women with HDP. For those with sustained-rise or late-acceleration trajectories, postpartum BP often remains difficult to normalize. These high-risk patients can be classified as key populations for monitoring during pregnancy, with daily BP measurements at fixed times, the plotting of BP trend charts, clarification of the type of hypertension, and early implementation of pharmacological interventions. These patients should also receive prioritized and more frequent postpartum follow-up visits, with prolongation of the follow-up duration, in order to reduce their long-term risk of CVD [63-65]. For pregnant women with low-risk BP trajectories, nonpharmacological interventions should also be emphasized [48]. A meta-analysis demonstrated that moderate walking, upper and lower limb muscles training, and stretching exercises are more effective in reducing gestational SBP and DBP in women with HDP than healthy pregnant women [66]. Deep breathing exercise and progressive muscular relaxation can significantly reduce SBP, DBP, and MAP in pregnant women with mild preeclampsia [67,68]. Therefore, a combination of pharmacological and nonpharmacological approaches tailored to BP trajectory type during pregnancy may help manage BP and reduce the risk of postpartum BP not returning to normal.

Postpartum follow-up guidelines for patients with HDP are lacking or insufficient in most countries [69]. Early diagnosis, risk stratification, standardized postpartum care, and strict BP management can effectively reduce postpartum cardiovascular risk and mortality [64]. Our findings indicate that BP trajectories during pregnancy are valuable for HDP risk stratification and for identifying women at high risk of poor postpartum BP recovery. Incorporating BP trajectories into postpartum follow-up may optimize management and improve both short- and long-term cardiovascular outcomes. Currently, predictive models based on BP trajectories during pregnancy are in the stages of algorithm validation and prospective testing [70]. Trajectories are constructed from BP data collected under routine clinical conditions, demonstrating practicality and feasibility. Emerging evidence suggests that the early postpartum period represents a critical therapeutic window for mitigating the long-term cardiovascular sequelae of HDP. For instance, a self-management intervention optimizing BP control within the first few weeks post partum led to favorable left ventricular and left atrial remodeling that was evident nearly a year later [13]. Developing a trajectory-specific tool to predict postpartum BP nonrecovery could enhance risk assessment and facilitate personalized intervention and follow-up strategies. It should be noted that postpartum BP recovery status is also influenced by multiple factors, such as maternal serum albumin levels, family history of hypertension, and prepregnancy body weight [18]. Further studies should explore the interactions between these risk factors and BP trajectories to better elucidate the relationship between BP dynamics during pregnancy and postpartum BP recovery status.

This study used real-world data from women with HDP, all of whom had at least 3 BP measurements during pregnancy spanning the first, second, and third trimesters. This comprehensive sampling strengthens the reliability of the findings. Sensitivity analyses further supported the robustness of our results: the number, trends, and relative proportion of trajectory groups in each subtype were generally consistent with those in the main analysis, with only minor changes in statistical associations. To enhance internal validity, we excluded multiple pregnancies, as this condition represents an important cardiovascular risk factor that may confound the association between BP trajectories and postpartum outcomes.

Nevertheless, this study has several limitations. First, we lacked data on antihypertensive use, so we could not adjust for potential confounding by treatment. While this may have led to some underestimation of true BP levels, the associations held across HDP subtypes—including in CHTN, where treatment is most likely—and remained consistent in direction and magnitude. Because antihypertensive therapy is part of routine care, the trajectories observed here reflect real-world BP patterns under standard clinical management. Second, owing to follow-up challenges, only BP recovery at 6 weeks post partum was assessed; longer-term BP recovery status remains unexplored. Future large-scale, multicenter cohort studies with extended follow-up are therefore warranted to further validate these findings. This study suggests a strong association between antenatal BP trajectories during pregnancy and failure to achieve normotension by 6 weeks post partum. As an observational study, however, our results indicate correlation rather than causation. These findings contribute to the understanding of postpartum BP changes in women with HDP, but further prospective and mechanistic studies are needed for validation.

Conclusion

This study identified distinct antepartum BP trajectories across HDP subtypes, with MAP and SBP showing similar trends and DBP showing more variability. Sustained-rising and late-acceleration patterns, particularly in GH, CHTN with superimposed preeclampsia, and CHTN, were associated with a higher risk of nonrecovery to normotension at 6 weeks post partum. In preeclampsia, only the high-consistent rise SBP trajectory predicted nonrecovery of normotension postpartum. Trajectory analysis, therefore, could support early risk identification and individualized postpartum care to improve long-term cardiovascular outcomes in patients with HDP.

XSL•FO
RenderX

Acknowledgments

The authors sincerely acknowledge the participants for their involvement in the study.

Funding

This study was funded by the National Natural Science Foundation of China (72274133) and the Guangdong Basic and Applied Basic Research Foundation (2025A1515011861).

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

LL, ZL, and YC participated in the study concept and design. The initial manuscript draft was prepared by ZL and HW, with ZL, HW, XX, and RX performing the statistical analysis. Data collection and management efforts were led by LL, ZL, XY, JD, XF, YC, and MF. The interpretation of results and critical revision of the manuscript for important intellectual content involved LL, ZL, and XZ. LL, YC, and XZ jointly supervised this work and are the guarantors of this study, taking responsibility for the integrity of the data and the accuracy of the data analysis.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary figures and tables.
[DOCX File , 749 KB - publichealth_v12i1e89295_app1.docx]

References

1. Henderson JT, Webber EM, Vesco KK. Screening for hypertensive disorders of pregnancy: an evidence update for the U.S. preventive services task force. JAMA 2023;330(11):1083-1091. [Medline: 37769050]
2. N/A. Gestational hypertension and preeclampsia: ACOG practice bulletin summary, number 222. Obstet Gynecol 2020;135(6):1492-1495. [doi: 10.1097/AOG.0000000000003892] [Medline: 32443077]
3. Say L, Chou D, Gemmill A, Tunçalp Ö, Moller A, Daniels J, et al. Global causes of maternal death: a WHO systematic analysis. Lancet Glob Health 2014;2(6):e323-e333 [FREE Full text] [doi: 10.1016/S2214-109X(14)70227-X] [Medline: 25103301]
4. Chappell LC, Cluver CA, Kingdom J, Tong S. Pre-eclampsia. Lancet 2021;398(10297):341-354. [doi: 10.1016/s0140-6736(20)32335-7]
5. Vogel B, Acevedo M, Appelman Y, Bairey Merz CN, Chieffo A, Figtree GA, et al. The Lancet Women and Cardiovascular Disease Commission: reducing the global burden by 2030. Lancet 2021;397(10292):2385-2438. [doi: 10.1016/S0140-6736(21)00684-X] [Medline: 34010613]
6. Ackerman‐Banks CM, Grechukhina O, Spatz E, Lundsberg L, Chou J, Smith G, et al. Seizing the window of opportunity within 1 year postpartum: early cardiovascular screening. J Am Heart Assoc 2022;11(8):e024443. [doi: 10.1161/jaha.121.024443]
7. Hauspurg A, Bryan S, Jeyabalan A, Davis EM, Hart R, Shirriel J, et al. Blood pressure trajectories through the first year postpartum in overweight or obese individuals following a hypertensive disorder of pregnancy. Hypertension 2024;81(2):302-310. [doi: 10.1161/hypertensionaha.123.22231]
8. Palatnik A, Mukhtarova N, Hetzel SJ, Hoppe KK. Blood pressure changes in gestational hypertension, preeclampsia, and chronic hypertension from preconception to 42-day postpartum. Pregnancy Hypertens 2022;31:25-31. [doi: 10.1016/j.preghy.2022.11.009] [Medline: 36512857]
9. Lailler G, Grave C, Gabet A, Joly P, Regnault N, Deneux‐Tharaux C, et al. Early mortality, cardiovascular, and renal diseases in women's lives following hypertensive disorders of pregnancy: the prospective nationwide study conception. J Am Heart Assoc 2024;13(8):e033252. [doi: 10.1161/jaha.123.033252]
10. Ahmed AA, Willkens MA, Matovelo D, Kiritta R, Kaizilege G, Mathad J, et al. Persistent postpartum hypertension 3 months post-delivery among women with hypertensive disorders of pregnancy in Tanzania: a multicenter cohort study. Pregnancy Hypertens 2024;35:51-54. [doi: 10.1016/j.preghy.2023.12.006] [Medline: 38176194]
11. Berks D, Steegers EAP, Molas M, Visser W. Resolution of hypertension and proteinuria after preeclampsia. Obstet Gynecol 2009;114(6):1307-1314. [doi: 10.1097/AOG.0b013e3181c14e3e] [Medline: 19935034]

XSL•FO
RenderX

12. Duvekot JJ. Do all women after hypertensive diseases of pregnancy have the same long-term risk of cardiovascular disease in later life? Best Pract Res Clin Obstet Gynaecol 2025;100:102597 [FREE Full text] [doi: 10.1016/j.bpobgyn.2025.102597] [Medline: 40138925]
13. Kitt J, Krasner S, Barr L, Frost A, Tucker K, Bateman PA, et al. Cardiac remodeling after hypertensive pregnancy following physician-optimized blood pressure self-management: the POP-HT randomized clinical trial imaging substudy. Circulation 2024;149(7):529-541. [doi: 10.1161/circulationaha.123.067597]
14. Hypertensive Disorders in Pregnancy Subgroup, Chinese Society of ObstetricsGynecology, Chinese Medical Association. [Diagnosis and treatment of hypertension and pre-eclampsia in pregnancy: a clinical practice guideline in China（2020）]. Zhonghua Fu Chan Ke Za Zhi 2020;55(4):227-238. [doi: 10.3760/cma.j.cn112141-20200114-00039] [Medline: 32375429]
15. Chinese Society of Cardiology, Chinese Medical Association, Editorial Board of Chinese Journal of Cardiology. [Expert consensus on the comprehensive cardiovascular risk management of hypertensive disorders in pregnancy]. Zhonghua Xin Xue Guan Bing Za Zhi 2025;53(3):235-246. [doi: 10.3760/cma.j.cn112148-20241024-00643] [Medline: 40091243]
16. Kaminker JD, MacMaster A, Pudwell J, King K, Smith GN. Prevalence of newly diagnosed essential hypertension within one year postpartum: a systematic review and meta-analysis. Am J Obstet Gynecol Glob Rep 2025;5(3):100519 [FREE Full text] [doi: 10.1016/j.xagr.2025.100519] [Medline: 40641834]
17. Ananth CV, Vintzileos AM. Maternal-fetal conditions necessitating a medical intervention resulting in preterm birth. Am J Obstet Gynecol 2006;195(6):1557-1563. [doi: 10.1016/j.ajog.2006.05.021] [Medline: 17014813]
18. Wei J, Wang Y, Xu J, Zhang C, Zhou H, Liu G. Clinical study on the factors affecting the post-partum recovery of patients with hypertensive pregnancy disorders at a Chinese hospital. J Obstet Gynaecol Res 2017;43(3):456-461. [doi: 10.1111/jog.13230] [Medline: 27935226]
19. Bruce KH, Anderson M, Stark JD. Factors associated with postpartum readmission for hypertensive disorders of pregnancy. Am J Obstet Gynecol MFM 2021;3(5):100397 [FREE Full text] [doi: 10.1016/j.ajogmf.2021.100397] [Medline: 33991709]
20. Advani R, Shree R, Albright CM, Chandrasekaran S. Smaller decrease in late second trimester blood pressure is associated with gestational hypertensive disease development. Pregnancy Hypertens 2025;39:101189. [doi: 10.1016/j.preghy.2025.101189] [Medline: 39793458]
21. Roberts JM, Alexeeff SE, Sun B, Greenberg M, King A, Nguyen-Huynh MN, et al. Early pregnancy blood pressure trajectories and hypertension years after pregnancy. Hypertension 2025;82(5):e75-e87. [doi: 10.1161/hypertensionaha.125.24649]
22. N/A. ACOG practice bulletin No. 203 summary: chronic hypertension in pregnancy. Obstet Gynecol 2019;133(1):215-219. [doi: 10.1097/AOG.0000000000003021] [Medline: 30575669]
23. Obstetrics Subgroup, Chinese Society of Obstetrics Gynecology, Chinese Medical Association. [Guideline of preconception and prenatal care (2018)]. Zhonghua Fu Chan Ke Za Zhi 2018;53(1):7-13. [doi: 10.3760/cma.j.issn.0529-567X.2018.01.003] [Medline: 29374879]
24. Niu Z, Anderson EC, Yang T, Chen X, Vigil M, Guo F, et al. Gestational blood pressure trajectories and 5-Year postpartum hypertension risk in the MADRES study. JACC Adv 2025;4(4):101660 [FREE Full text] [doi: 10.1016/j.jacadv.2025.101660] [Medline: 40101499]
25. Peng J, Yang J, Li F. Association of bioimpedance analysis parameters trajectories with clinical outcomes in neurocritical patients. Heliyon 2024;10(12):e32948. [doi: doi: 10.1016/j.heliyon.2024.e32948] [Medline: 38994111]
26. Nagin DS, Jones BL, Elmer J. Recent advances in group-based trajectory modeling for clinical research. Annu Rev Clin Psychol 2024;20(1):285-305 [FREE Full text] [doi: 10.1146/annurev-clinpsy-081122-012416] [Medline: 38382118]
27. Chen Q, Du J, Hong X. Association between blood pressure multi-trajectory and cardiovascular disease among a Chinese elderly medical examination population. Front Cardiovasc Med 2024;11:1363266 [FREE Full text] [doi: 10.3389/fcvm.2024.1363266] [Medline: 39114559]
28. Serra L, Farrants K, Alexanderson K, Ubalde M, Lallukka T. Trajectory analyses in insurance medicine studies: examples and key methodological aspects and pitfalls. PLoS One 2022;17(2):e0263810 [FREE Full text] [doi: 10.1371/journal.pone.0263810] [Medline: 35148351]
29. Hickson RP, Annis IE, Killeya-Jones LA, Fang G. Comparing continuous and binary group-based trajectory modeling using statin medication adherence data. Med Care 2021;59(11):997-1005 [FREE Full text] [doi: 10.1097/MLR.0000000000001625] [Medline: 34644285]
30. Park B, Chang Y, Ryu S, Tran TXM. Trajectories of breast density change over time and subsequent breast cancer risk: longitudinal study. BMJ 2024;387:e079575 [FREE Full text] [doi: 10.1136/bmj-2024-079575] [Medline: 39797631]
31. Sinkey RG, Blanchard CT, Sanusi A, Elkins C, Szychowski JM, Harper LM, et al. Physiologic blood pressure patterns in pregnancies with mild chronic hypertension. Pregnancy Hypertens 2024;36:101118. [doi: 10.1016/j.preghy.2024.101118] [Medline: 38460322]
32. Fang Y, Zuo L, Li J, Shi H, Zhang R, Han C, et al. Optimization of ambulatory blood pressure monitoring during pregnancy: a path toward risk stratification improvement and management of hypertensive disorders. Rev Cardiovasc Med 2025;26(5):27235. [doi: 10.31083/rcm27235] [Medline: 40475728]

XSL•FO
RenderX

33. de Haas S, Mulder E, Schartmann N, Mohseni Z, Abo Hasson F, Alsadah F, et al. Blood pressure adjustments throughout healthy and hypertensive pregnancy: a systematic review and meta-analysis. Pregnancy Hypertens 2022;27:51-58 [FREE Full text] [doi: 10.1016/j.preghy.2021.12.004] [Medline: 34929556]
34. Zhang X, Ouyang Y, Huang F, Zhang J, Su C, Jia X, et al. Modifiable factors of 20-year blood pressure trajectories among normotensives and their associations with hypertension: a prospective study. Br J Nutr 2021;128(2):252-262. [doi: 10.1017/s0007114521003378]
35. Song Z, Zhao H, Wei Z, Zhao W, Tan Y, Yang P, et al. Mean arterial pressure trajectory with premature cardiovascular disease and all-cause mortality in young adults: the Kailuan prospective cohort study. Front Cardiovasc Med 2023;10:1222995 [FREE Full text] [doi: 10.3389/fcvm.2023.1222995] [Medline: 37771669]
36. Tan EK, Tan EL. Alterations in physiology and anatomy during pregnancy. Best Pract Res Clin Obstet Gynaecol 2013;27(6):791-802. [doi: 10.1016/j.bpobgyn.2013.08.001] [Medline: 24012425]
37. Macdonald-Wallis C, Silverwood RJ, Fraser A, Nelson SM, Tilling K, Lawlor DA, et al. Gestational-age-specific reference ranges for blood pressure in pregnancy: findings from a prospective cohort. J Hypertens 2015;33(1):96-105 [FREE Full text] [doi: 10.1097/HJH.0000000000000368] [Medline: 25255393]
38. Lei Q, Zhou X, Duan DM, Lv LJ, Lin XH, Ji WJ, et al. Trimester-specific weight gain and midpregnancy diastolic blood pressure rebound during normotensive pregnancy. Hypertension 2017;70(4):804-812. [doi: 10.1161/HYPERTENSIONAHA.117.09760] [Medline: 28827477]
39. Nobles CJ, Mendola P, Mumford SL, Silver RM, Kim K, Andriessen VC, et al. Preconception blood pressure and its change into early pregnancy: early risk factors for preeclampsia and gestational hypertension. Hypertension 2020;76(3):922-929 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.120.14875] [Medline: 32755413]
40. Macdonald-Wallis C, Lawlor DA, Fraser A, May M, Nelson SM, Tilling K. Blood pressure change in normotensive, gestational hypertensive, preeclamptic, and essential hypertensive pregnancies. Hypertension 2012;59(6):1241-1248 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.111.187039] [Medline: 22526257]
41. Anderson EC, Muse ME, Niu Z, Foley HB, Levy JJ, Romano ME, et al. Blood pressure and late pregnancy circulating miRNAs in the MADRES study. J Am Heart Assoc 2025;14(12):e040416. [doi: 10.1161/jaha.124.040416]
42. Gunnarsdottir J, Akhter T, Högberg U, Cnattingius S, Wikström AK. Elevated diastolic blood pressure until mid-gestation is associated with preeclampsia and small-for-gestational-age birth: a population-based register study. BMC Pregnancy Childbirth 2019;19(1):186 [FREE Full text] [doi: 10.1186/s12884-019-2319-2] [Medline: 31138157]
43. Hauspurg A, Parry S, Mercer BM, Grobman W, Hatfield T, Silver RM, et al. Blood pressure trajectory and category and risk of hypertensive disorders of pregnancy in nulliparous women. Am J Obstet Gynecol 2019;221(3):277.e1-277.e8 [FREE Full text] [doi: 10.1016/j.ajog.2019.06.031] [Medline: 31255629]
44. Mo M, Shen Y, Si S, Xin X, Shao B, Wang S, et al. Feature of trajectory of blood pressure among pregnant women with gestational hypertension. J Hypertens 2020;38(1):127-132. [doi: 10.1097/HJH.0000000000002197] [Medline: 31568054]
45. Cnossen JS, Vollebregt KC, Vrieze ND, Riet GT, Mol BWJ, Franx A, et al. Accuracy of mean arterial pressure and blood pressure measurements in predicting pre-eclampsia: systematic review and meta-analysis. BMJ 2008;336(7653):1117-1120. [doi: 10.1136/bmj.39540.522049.be]
46. Mayrink J, Souza RT, Feitosa FE, Rocha Filho EA, Leite DF, Vettorazzi J, Preterm SAMBA Study Group. Mean arterial blood pressure: potential predictive tool for preeclampsia in a cohort of healthy nulliparous pregnant women. BMC Pregnancy Childbirth 2019;19(1):460 [FREE Full text] [doi: 10.1186/s12884-019-2580-4] [Medline: 31795971]
47. Moore MD, Kuo HC, Sinkey RG, Boggess K, Dugoff L, Sibai B, et al. Mean arterial pressure and neonatal outcomes in pregnancies complicated by mild chronic hypertension. Obstet Gynecol 2024;144(1):101-108. [doi: 10.1097/AOG.0000000000005611] [Medline: 38781591]
48. Ma S, Wu L, Yu Q. Associations between trajectory of different blood pressure components in pregnancy and risk of adverse birth outcomes - a real world study. Risk Manag Healthc Policy 2021;14:3255-3263. [doi: 10.2147/RMHP.S318956]
49. Giorgione V, Khalil A, O'Driscoll J, Thilaganathan B. Peripartum screening for postpartum hypertension in women with hypertensive disorders of pregnancy. J Am Coll Cardiol 2022;80(15):1465-1476 [FREE Full text] [doi: 10.1016/j.jacc.2022.07.028] [Medline: 36202536]
50. Chung Ming CL, Pienaar D, Ghorbanpour S, Chen H, Roberts LM, Cole L, et al. 3D in vitro modelling of post-partum cardiovascular health reveals unique characteristics and signatures following hypertensive disorders in pregnancy. Biol Sex Differ 2024;15(1):94 [FREE Full text] [doi: 10.1186/s13293-024-00672-6] [Medline: 39587642]
51. Van Wingerden AS, Katsidoniotaki M, Haghighi N, Almonte C, Woolcock Martinez H, Valdes E, et al. Postpartum blood pressure variability and heart rate variability in preeclampsia. Hypertension 2024;81(12):2510-2519. [doi: 10.1161/hypertensionaha.124.23321]
52. Kametas NA, Nzelu D, Nicolaides KH. Chronic hypertension and superimposed preeclampsia: screening and diagnosis. Am J Obstet Gynecol 2022;226(2S):S1182-S1195. [doi: 10.1016/j.ajog.2020.11.029] [Medline: 35177217]
53. Bryan NS. Nitric oxide deficiency is a primary driver of hypertension. Biochem Pharmacol 2022;206:115325. [doi: 10.1016/j.bcp.2022.115325] [Medline: 36349641]

XSL•FO
RenderX

54. Andreeva VD, Ehlers H, Presselt M, J van den Broek L, Bonnet S. Combining nitric oxide and calcium sensing for the detection of endothelial dysfunction. Commun Chem 2023;6(1):179 [FREE Full text] [doi: 10.1038/s42004-023-00973-8] [Medline: 37644120]
55. Vincent KMM, Garner T, Stevens A, Cottrell EC, Myers JE, Higgins LE. Assessing placental dysfunction subtypes in pregnancies with a low PlGF centile. Hypertension 2025;82(7):1249-1260. [doi: 10.1161/hypertensionaha.124.24440]
56. Sun F, Peers de Nieuwburgh M, Hubinont C, Debiève F, Colson A. Gene therapy in preeclampsia: the dawn of a new era. Hypertens Pregnancy 2024;43(1):2358761 [FREE Full text] [doi: 10.1080/10641955.2024.2358761] [Medline: 38817101]
57. Famá EAB, Pinhal MAS. Extracellular matrix components in preeclampsia. Clin Chim Acta 2025;568:120132. [doi: 10.1016/j.cca.2025.120132] [Medline: 39798685]
58. Kremer V, Oppelaar JJ, Gimbel T, Koziarek S, Ganzevoort W, van Pampus MG, et al. Neuro-oncological ventral antigen 2 regulates splicing of vascular endothelial growth factor receptor 1 and is required for endothelial function. Reprod Sci 2023;30(2):678-689 [FREE Full text] [doi: 10.1007/s43032-022-01044-4] [Medline: 35927413]
59. Li W, Liu S, Li Y, Wang W, Dong Y, Qi L, et al. Association between serum placental growth factor and vascular endothelial function in hypertensive disorders complicating pregnancy. Exp Biol Med (Maywood) 2022;247(6):480-487 [FREE Full text] [doi: 10.1177/15353702211063833] [Medline: 34904900]
60. Kurhaluk N, Tkaczenko H. L-Arginine and nitric oxide in vascular regulation-experimental findings in the context of blood donation. Nutrients 2025;17(4):665 [FREE Full text] [doi: 10.3390/nu17040665] [Medline: 40004994]
61. Kaihara JNS, Minami CK, Peraçoli MTS, Romão-Veiga M, Ribeiro-Vasques VR, Peraçoli JC, et al. Plasma eNOS concentration in healthy pregnancy and in hypertensive disorders of pregnancy: evidence of reduced concentrations in pre-eclampsia from two independent studies. Diseases 2023;11(4):155 [FREE Full text] [doi: 10.3390/diseases11040155] [Medline: 37987266]
62. Dela Justina V, Dos Passos Júnior RR, Lima VV, Giachini FR. Evidence of nitric oxide impairment during hypertensive pregnancies. Adv Exp Med Biol 2023;1428:99-125. [doi: 10.1007/978-3-031-32554-0_5] [Medline: 37466771]
63. Ackerman-Banks CM, Lipkind HS, Palmsten K, Ahrens KA. Association between hypertensive disorders of pregnancy and cardiovascular diseases within 24 months after delivery. Am J Obstet Gynecol 2023;229(1):65.e1-65.e15 [FREE Full text] [doi: 10.1016/j.ajog.2023.04.006] [Medline: 37031763]
64. Khosla K, Heimberger S, Nieman KM, Tung A, Shahul S, Staff AC, et al. Long-term cardiovascular disease risk in women after hypertensive disorders of pregnancy: recent advances in hypertension. Hypertension 2021;78(4):927-935. [doi: 10.1161/hypertensionaha.121.16506]
65. Ryan K, McGrath L, Brookfield K. Hypertension management in pregnancy. Annu Rev Med 2025;76(1):315-326 [FREE Full text] [doi: 10.1146/annurev-med-050423-085626] [Medline: 39586030]
66. Zhu Z, Xie H, Liu S, Yang R, Yu J, Yan Y, et al. Effects of physical exercise on blood pressure during pregnancy. BMC Public Health 2022;22(1):1733 [FREE Full text] [doi: 10.1186/s12889-022-14074-z] [Medline: 36096756]
67. Ibrahim AA, Gaballah S, Abu Bakr Elsaid NM, Mohamed HA. Progressive muscular relaxation versus breathing excercise techniques to control blood pressure among mild preeclamptic pregnant women. Iran J Nurs Midwifery Res 2024;29(4):411-416. [doi: 10.4103/ijnmr.ijnmr_33_22] [Medline: 39205847]
68. Ertürk Ç, Özden AV. Comparison of the acute effects of auricular vagus nerve stimulation and deep breathing exercise on the autonomic nervous system activity and biomechanical properties of the muscle in healthy people. J Clin Med 2025;14(4):1046 [FREE Full text] [doi: 10.3390/jcm14041046] [Medline: 40004576]
69. Staff AC, Costa ML, Dechend R, Jacobsen DP, Sugulle M. Hypertensive disorders of pregnancy and long-term maternal cardiovascular risk: bridging epidemiological knowledge into personalized postpartum care and follow-up. Pregnancy Hypertens 2024;36:101127 [FREE Full text] [doi: 10.1016/j.preghy.2024.101127] [Medline: 38643570]
70. Gunderson EP, Greenberg M, Sun B, Goler N, Go AS, Roberts JM, et al. Early pregnancy systolic blood pressure patterns predict early - and later - onset preeclampsia and gestational hypertension among ostensibly low - to - moderate risk groups. J Am Heart Assoc 2023;12(15):e029617. [doi: 10.1161/jaha.123.029617]

Abbreviations

aRR: adjusted relative risk
BP: blood pressure
CHTN: chronic hypertension
CVD: cardiovascular disease
DBP: diastolic blood pressure
GBTM: group-based trajectory modeling
GH: gestational hypertension
HDP: hypertensive disorders of pregnancy
MAP: mean arterial pressure
NO: nitric oxide
PLGF: placental growth factor

XSL•FO
RenderX

SBP: systolic blood pressure
sFlt1: soluble fms-like tyrosine kinase-1

Edited by A Mavragani, T Sanchez; submitted 10.Dec.2025; peer-reviewed by H Bao, L Voskamp; comments to author 22.Jan.2026; revised version received 02.Mar.2026; accepted 20.Mar.2026; published 30.Apr.2026.

Please cite as:
Liang Z, Wei H, Xu X, Xu R, Yang X, Dong J, Fan X, Chen Y, Feng M, Zhou X, Cai Y, Lv L
Association Between Gestational Blood Pressure Trajectories and Postpartum Normotension Recovery in Hypertensive Disorders: Retrospective Cohort Study
JMIR Public Health Surveill 2026;12:e89295
URL: https://publichealth.jmir.org/2026/1/e89295
doi:10.2196/89295
PMID:42060921

©Zhijiang Liang, Huili Wei, Xiaojun Xu, Rong Xu, Xinyue Yang, Jinlian Dong, Xiaoyan Fan, Yukun Chen, Meiling Feng, Xin Zhou, Yijun Cai, Lijuan Lv. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 30.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Assessment of Knowledge, Attitudes, and Behaviors Toward Others' Mental Disorders to Improve Mental Health First Aid: Cross-Sectional Study in a French Workplace

Matthieu Lebrat[1,2], MSc, PharmD; Florence Carrouel[1], Prof Dr; Benjamin du Sartz de Vigneulles[1], PhD; Michel Lamure[1], PhD; Corélie Salque[1], PhD; Roger Salamon[3], Prof Dr Med; Catherine Massoubre[4,5], Prof Dr Med; Nicolas Franck[6,7], Prof Dr Med; Claude Dussart[1,2], Prof Dr

[1]P2S UR4129, Université Lyon 1, Lyon Cedex 08, France
[2]Pharmacie et Stérilisation Centrales, Hospices Civils de Lyon, Lyon, Rhône-Alpes, France
[3]Institute of Public Health, Epidemiology and Development (ISPED), Inserm U1219, Université de Bordeaux, Bordeaux, Nouvelle-Aquitaine, France
[4]Research Unit EA7423, Université Jean Monnet, Saint-Étienne, Auvergne-Rhône-Alpes, France
[5]Department of Psychiatry, Université Jean Monnet, Saint-Etienne, Auvergne-Rhône-Alpes, France
[6]Centre Ressource de Réhabilitation psychosociale, Pôle Centre Rive Gauche, Centre Hospitalier Le Vinatier, Bron, Auvergne-Rhône-Alpes, France
[7]UMR 5229 CNRS, Centre Hospitalier Le Vinatier, Bron, Auvergne-Rhône-Alpes, France

Corresponding Author:
Florence Carrouel, Prof Dr
P2S UR4129, Université Lyon 1
7 rue Guillaume PAradin
Lyon Cedex 08, 69372
France
Phone: 33 4 78 78 57 44
Email: florence.carrouel@univ-lyon1.fr

Abstract

Background: Mental distress represents a major public health challenge worldwide, affecting work ability, productivity, and social functioning. Public-sector employees face high psychosocial demands, but data on their mental health literacy and supportive behaviors remain scarce. Mental Health First Aid (MHFA) aims to improve knowledge, attitudes, and behaviors (KAB) toward psychological distress among nonspecialists, but evidence among civil servants and on skill retention over time is limited.

Objective: This study aimed (1) to determine KAB related to psychological distress among French civil servants, (2) to examine associations between KAB scores and MHFA training exposure, and (3) to explore gradients according to time since training as an indirect indicator of potential retention of MHFA-related competencies.

Methods: A nationwide cross-sectional survey was conducted in November 2025 among French civil servants. Participants completed an online questionnaire adapted from a published instrument. The 3 dimension scores were normalized to a 0-40 scale, yielding a total KAB score of 0-120, categorized as inadequate (0-70), marginal (71-88), or adequate (89-120). Multivariable linear regression models identified factors associated with KAB outcomes, adjusting for sociodemographic and professional variables.

Results: Among 338,560 eligible civil servants invited by email, 6526 (1.93%) initiated the questionnaire, and 4679 (1.38% of those invited) fully completed it and were included in the analysis. Participants were predominantly women (3022/4679, 64.59%), aged 40-69 years (4317/4679, 92.26%), and belonged to professional category A (2207/4679, 47.17%). The mean total KAB score was 82.88 (SD 19.14; 95% CI 82.33-83.43), corresponding to a marginal level. Mean subscores were 31.31 (SD 5.26; 95% CI 31.16-31.46) for knowledge, 29.72 (SD 11.52; 95% CI 29.39-30.05) for attitudes, and 21.86 (SD 7.96; 95% CI 21.63-22.09) for behaviors. MHFA-trained participants (545/4679, 11.65%) obtained significantly higher total KAB scores than untrained participants (mean 95.56, SD 14.17; 95% CI 94.37-96.75 vs mean 81.21, SD 19.09; 95% CI 80.63-81.79; $P<.001$), with higher scores for knowledge (mean 33.31, SD 5.26 vs mean 31.04, SD 5.35), attitudes (mean 34.35, SD 8.81 vs mean 29.10, SD 11.70), and behaviors (mean 27.87, SD 6.13 vs mean 21.06, SD 7.84; all $P<.001$). Among MHFA-trained participants, scores decreased over time (<6 months, 6-12 months, >12 months; $P<.001$), but remained higher than in untrained individuals even beyond 12 months after training. In adjusted models, MHFA training was strongly associated with higher KAB scores (β=14.63 for <6

XSL•FO
RenderX

months; β=13.61 for 6-12 months; β=10.33 for >12 months; all *P*<.001). Lower scores were observed in professional categories B and C and among men.

Conclusions: Mental health literacy and supportive behaviors among French civil servants remain heterogeneous and, on average, marginal. MHFA training is associated with substantially higher KAB scores, even beyond 12 months, supporting its role as a public health literacy intervention in occupational settings. Declining scores over time supports reinforcement strategies to sustain competencies.

(JMIR Public Health Surveill 2026;12:e93207) doi:10.2196/93207

KEYWORDS

public health; health promotion; mental health; Mental Health First Aid; knowledge, attitudes, and behaviors; health literacy; civil servants; workplace prevention; stigma

Introduction

Mental health disorders—defined in the *International Classification of Diseases, 11th Revision* as clinically significant disturbances in cognition, emotional regulation, or behavior associated with distress or impairment in personal, social, educational, or occupational functioning—represent one of the leading public health challenges worldwide [1]. Nearly 1 in 7 people globally, more than a billion individuals, are living with a mental disorder, with depressive and anxiety disorders being the most prevalent. Mental disorders are the leading cause of suicide [2], and also the leading cause of years lived with disability globally, accounting for 1 in 6 years lived with disability [3,4]. The COVID-19 pandemic further intensified this burden, leading to an estimated 27.6% increase in major depressive disorders and a 25.6% increase in anxiety disorders worldwide [5].

Against this global background, similar trends have been observed in France, where data from 2017 to 2021 indicate a clear increase in major depressive episodes, by roughly 3 to 4 percentage points, as well as a high prevalence of anxiety disorders, which affects nearly one-fifth of adults depending on the subgroup considered (men or women) [6]. These findings are consistent with international observations [7] and point to a significant deterioration in the mental health of the French population.

Mental health disorders contribute substantially to disability, reduced social participation, and impaired occupational functioning. Psychosocial distress also plays a significant role in long-term sickness absence, with substantial indirect costs (particularly productivity losses linked to difficulties in maintaining employment), which are often more costly than direct health care expenditures [3]. Thus, the burden of mental disorders and their societal impacts has led to the designation of mental health as France's "National Cause" for 2025 and 2026, reflecting its strategic importance for societal sustainability [8].

More specifically, public sector employees are affected. In 2023, 19% of public-sector employees presented a risk of depressive symptoms [9], which is concerning, given their essential societal role and frequent exposure to emotionally demanding or high-responsibility situations [10,11]. Epidemiological studies have shown that workload, low decision latitude, and imbalance between effort and reward significantly contribute to the onset of common mental disorders [12]. Recent international guidelines from the World Health Organization (WHO) and the International Labour Organization emphasize the responsibility of governments, both as public employers and regulators, to prevent work-related mental disorders and support employees experiencing psychological difficulties [13,14]. Explicit consideration of the mental health of public-sector workers is thus essential for service quality, sustainability of public systems, and collective resilience. In this perspective, strengthening the capacity of civil servants to identify, understand, and respond appropriately to psychological distress emerges as a strategic priority.

Mental Health First Aid (MHFA) first emerged in Australia as a relevant resource. It aims to improve mental health literacy and reduce stigma among the general public. It also equips people with the skills to encourage those with mental issues to seek professional help. In France, the MHFA program, named "Premiers Secours en Santé Mentale," has expanded rapidly since 2018, with more than 120,000 trained "first aiders" and over 1400 certified instructors as of May 2024 [15]. This rapid dissemination reflects strong institutional interest in equipping nonclinicians with the skills needed to recognize early signs of psychological distress and provide timely support. In 2022, a ministerial directive requested the establishment of MHFA within public employers [16]. Although several French studies [17,18] have highlighted the promising effects of MHFA training on trainees, these findings emphasize the need for additional analysis to further optimize the program [19]. Indeed, existing evidence is primarily based on student populations and has not specifically examined long-term effects. Consequently, it remains unclear whether similar benefits can be observed beyond the short term and in other populations such as civil servants. This group differs substantially in age and professional constraints, while their predefined socioprofessional classification offers a methodological asset for analyzing socioprofessional gradients.

Within the framework of improving mental health support in the workplace, particularly in the civil service, it is therefore essential to assess whether employees are able to recognize signs of psychological distress in their colleagues and respond appropriately. This requires a structured approach based on the knowledge, attitudes, and behaviors (KAB) model, which captures how individuals perceive, evaluate, and act upon a given situation. This can shed light on levels of mental health literacy. Indeed, KAB allows measurement of (1) knowledge

XSL•FO
RenderX

of symptoms, risk factors, and help-seeking pathways; (2) attitudes regarding stigma and self-efficacy; and (3) supportive behaviors, including intent to help and referral practices. The KAB framework originates from health behavior and public health research and assumes that knowledge influences attitudes, which in turn may shape behaviors. Although these relationships are not necessarily linear, KAB models provide a useful framework for assessing health literacy and informing prevention strategies [20-22]. To date, despite the expansion of MHFA, no study has evaluated mental health KAB among civil servants in France, compared MHFA trained and untrained agents, or examined the persistence of skills over time in this context. Addressing this gap is crucial for informing national strategies aimed at strengthening mental health–related competencies within the civil service.

This study aimed (1) to determine KAB related to psychological distress among French civil servants, (2) to examine associations between KAB scores and MHFA training exposure, and (3) to explore gradients according to time since training as an indirect indicator of potential retention of MHFA-related competencies.

Methods

Study Design

This study adopted a descriptive, cross-sectional, quantitative design to examine KAB regarding mental disorders in others. Such a design is well suited to KAB surveys, particularly when data are scarce and baseline information is needed before developing interventions [21,23]. This study was performed in accordance with the CROSS (Checklist for Reporting Of Survey Studies; Multimedia Appendix 1) [24].

Study Setting

This cross-sectional KAB survey was conducted in France from November 13-27, 2025, among active French civil servants who had a digital personal account on the national health insurance website and had agreed to receive information by email. Within the French civil service, positions are organized into 3 main professional categories: A, B, and C, which represent different corps of professions characterized by varying levels of responsibility, qualification, remuneration, and recruitment criteria. These categories structure the civil service hierarchy, with specific access conditions and baseline remuneration associated with each [9].

Participants

Eligibility Criteria

To be included, participants had to (1) be at least 18 years of age, (2) be active public service employees, (3) have agreed to receive information at their email address, and (4) consent to complete the online questionnaire.

Retired individuals and nonpermanent civil servants (ie, public sector contract employees not belonging to public service category A, B, or C) were excluded from the study.

Recruitment

The recruitment frame consisted of all French civil servants who had a personal digital account within the Union Régime Obligatoire Prévention Santé system and had previously agreed to receive email communications. A total of 338,560 individuals met these criteria and received the study invitation.

No sampling, stratification, or additional selection procedures were applied because the invitation was sent to the entire eligible recruitment frame. The invitation email and study information letter were distributed to the entire eligible recruitment frame, followed by 1 reminder sent 1 week later to the same distribution list.

Participation was entirely voluntary. Individuals who provided informed consent, met the eligibility criteria, and initiated the online questionnaire were included in the study. Participants who completed all items required for KAB score calculation were included in the final analyses.

Questionnaire

The French questionnaire developed and used by Llopis et al [17] to assess mental health–related knowledge, attitudes, and practices (KAP) among French university students who received MHFA training was adapted for the study population. The questionnaire was not originally developed as a psychometric instrument, and no formal psychometric validation was performed in this study. Specifically, age and date of MHFA training were collected using predefined categories, job position and education level items were adapted for relevance to French public service classifications, and some items were removed to reduce questionnaire length. In total, 5 questions were adapted for this study population, 1 question was shortened, and the other 32 items were retained in their original form. No new items were created. The wording of the questions was kept as original.

It comprised 38 closed-ended questions and covered the following domains: (1) sociodemographic data (age collected in ranges and job position collected according to French public service categories A, B, and C) and information on MHFA training (including motivations and time since training: <6, 6-12, and >12 months)—7 questions, (2) mental health knowledge—10 questions, (3) attitudes toward people with mental health disorders—4 questions, and (4) behaviors toward individuals experiencing psychological difficulties—17 questions. Among these questions, 5 were specifically reserved for MHFA trainees. The survey required approximately 10 minutes to complete. The original French version and an English translation of the questionnaire are provided in Multimedia Appendices 2 and 3.

Scoring System

From the collected responses, new variables were created: knowledge score related to mental health: sum of items 8 to 17 (1 point per correct response and 0 points per incorrect response or “do not know”). Attitude score toward people with psychological distress: sum of items 18 to 21 (1 point per correct response and 0 points per incorrect response or “do not know”). Behavior score toward individuals experiencing psychological distress, to assess the level of consistency with the recommendations made during MHFA training: sum of items 22 to 26 and 31 to 38 (1 point per correct response and 0 points per incorrect response or “do not know.” Likert items scored

XSL•FO
RenderX

0-4). This scoring system was developed in accordance with the WHO KAP survey methodology [20], which recommends simplified and direction-oriented coding approaches for population-based assessments of health-related KAB. The normalization and categorization procedures were adopted to facilitate interpretation and comparison of KAB dimensions at the population level. Equal weighting of items was intentionally adopted because the objective of the study was to generate descriptive population-level indicators rather than to develop a diagnostic or psychometric scale. No empirical or theoretical framework was available to justify differential weighting of individual items. Consequently, each item contributed equally to its corresponding KAB dimension, consistent with WHO KAP survey methodology and previous population-based KAP or KAB surveys.

Scores for each KAB dimension were normalized on a scale from 0 to 40, with total KAB score ranging from 0 to 120. Because the 3 KAB dimensions were based on different numbers of items and score ranges, raw scores were transformed to a common 0-40 scale using a linear normalization procedure according to the following formula:

$$\text{Normalized score}=(\text{observed score}/\text{maximum possible score})\times 40 \quad \textbf{(1)}$$

This transformation allowed direct comparison of the 3 dimensions while preserving the relative position of participants within each dimension. The total KAB score corresponded to the sum of the 3 normalized dimension scores and therefore ranged from 0 to 120. Details of the raw and normalized score ranges are provided in Multimedia Appendix 4.

Categories were set as follows: inadequate (0-70), marginal (71-88), and adequate (89-120) based on previous questionnaires [23]. These categories were intended to facilitate interpretation of population-level KAB results and should not be considered clinically validated thresholds.

Survey Platform

The questionnaire was hosted online using LimeSurvey (version 6.16.5; LimeSurvey GmbH) on a dedicated server with 2048 bitset Secure Sockets Layer/Transport Layer Security encryption and without collecting IP addresses. Measures were taken to ensure secure, anonymous participation, cookies to block duplicate submissions, and unique tokens for log-in. Respondents could pause and resume the survey as needed.

Study Outcomes

The primary outcome was to determine KAB of civil servants when addressing psychological distress in others.

The secondary outcomes were to examine associations between KAB scores and MHFA training exposure and to explore gradients according to time since training as an indirect indicator of potential retention of MHFA-related competencies. Accordingly, analyses comparing MHFA-trained and untrained participants, as well as analyses according to time since training (<6, 6-12, and >12 months), were predefined secondary exploratory analyses.

Statistical Analysis

Sample Size Calculation

Based on an eligible population of 338,560 civil servants who received the email invitation, corresponding to the entire accessible Union Régime Obligatoire Prévention Santé recruitment frame, a minimum of 1064 completed questionnaires was calculated to ensure a 95% confidence level and a ±3 percentage-point margin of error for proportion estimates (finite population correction applied). This calculation was derived from the standard formula for estimating a proportion, assuming a conservative prevalence of 50% to maximize variance:

$$n = Z^2 \times p(1-p)/e^2 \quad \textbf{(2)}$$

where Z=1.96 (95% confidence level), p=0.50 (maximum variance assumption), and e=0.03 (margin of error). The resulting sample size was then adjusted using the finite population correction:

$$n = n/[1+(n-1)/N] \quad \textbf{(3)}$$

where N=338,560. The final minimum sample size was estimated at 1064 completed questionnaires.

Data Analysis

Data were analyzed using R (R Foundation for Statistical Computing) via RStudio. Descriptive statistics were computed using standard R packages, and forest plots were generated using the *ggplot2* package. Descriptive statistics (counts and percentages) were used for categorical sociodemographic variables and general information related to MHFA training. Means and SDs were reported for continuous KAB score variables.

The sample was divided into 2 groups according to MHFA training status (yes vs no), and KAB scores were compared using 1-way ANOVA. Among participants who received MHFA training, 3 categories were defined based on time since training, and KAB scores were compared across these groups using 1-way ANOVA. Prior to conducting ANOVA analyses, normality was assessed using the Shapiro-Wilk test and visual inspection of quantile-quantile plots. Although the formal Shapiro-Wilk test indicated departures from normality, visual inspection suggested only minor deviations, and parametric analyses were considered appropriate, given the large sample size and the robustness of ANOVA to modest departures from normality. Homogeneity of variances was assessed using the Levene test. When the Levene test indicated heterogeneity of variances, corresponding nonparametric tests (Wilcoxon rank sum test with continuity correction or Kruskal-Wallis rank sum test, as appropriate) were performed as sensitivity analyses and yielded the same conclusions regarding statistical significance. Moreover, because comparisons according to MHFA training status and time since training were predefined secondary analyses addressing specific study objectives rather than exploratory post hoc analyses, no formal adjustment for multiple comparisons was applied. Achieved statistical power was estimated for the predefined subgroup analyses according to time since MHFA training.

All complete questionnaires were included in descriptive and inferential analyses, as complete KAB score calculation required

XSL•FO
RenderX

complete responses for the corresponding items. To evaluate the potential for selection bias related to missing data, respondents with complete questionnaires were compared with participants who did not complete the questionnaire but who had provided at least the sociodemographic information required for comparison analyses (Multimedia Appendix 5). Because KAB scores were directly derived from questionnaire responses, imputation of missing outcome items would have required strong assumptions regarding unobserved responses. Consequently, analyses were restricted to complete questionnaires, and potential selection bias was evaluated through comparison of complete and incomplete respondents. Inverse probability weighting was not performed because comparisons between complete and incomplete respondents showed only negligible-to-small differences across the available baseline characteristics and no difference in MHFA training status, the primary exposure variable. In addition, only a limited set of baseline variables was available for incomplete respondents, preventing construction of a robust response-propensity model. Determinants of KAB scores were examined using multivariable linear regression models. Covariates included job position, level of education, age, gender, MHFA training status, and time since training. All covariates were modeled as categorical variables according to the categories in which they were collected. For each variable, one category was used as the reference category in the regression models. Models were estimated by using generalized linear modeling procedures in R.

Regression coefficients (β) were reported with 95% CIs and associated *P* values. A 2-sided *P* value <.05 was considered statistically significant.

Ethical Considerations

The study was conducted in accordance with the Declaration of Helsinki and approved by the research ethics committee of the Vinatier Hospital, Lyon, France (CEREVI/2025/45; July 7, 2025). Before accessing the questionnaire, all participants received study information electronically and provided informed consent to participate. Participation was entirely voluntary, and participants could discontinue the questionnaire at any time without consequence. To protect privacy and confidentiality, the questionnaire was administered anonymously through a secure online platform. No IP addresses were collected, and all data were analyzed in aggregated form only. Participants did not receive any financial compensation or incentive for participation.

Results

Characteristics of the Participants

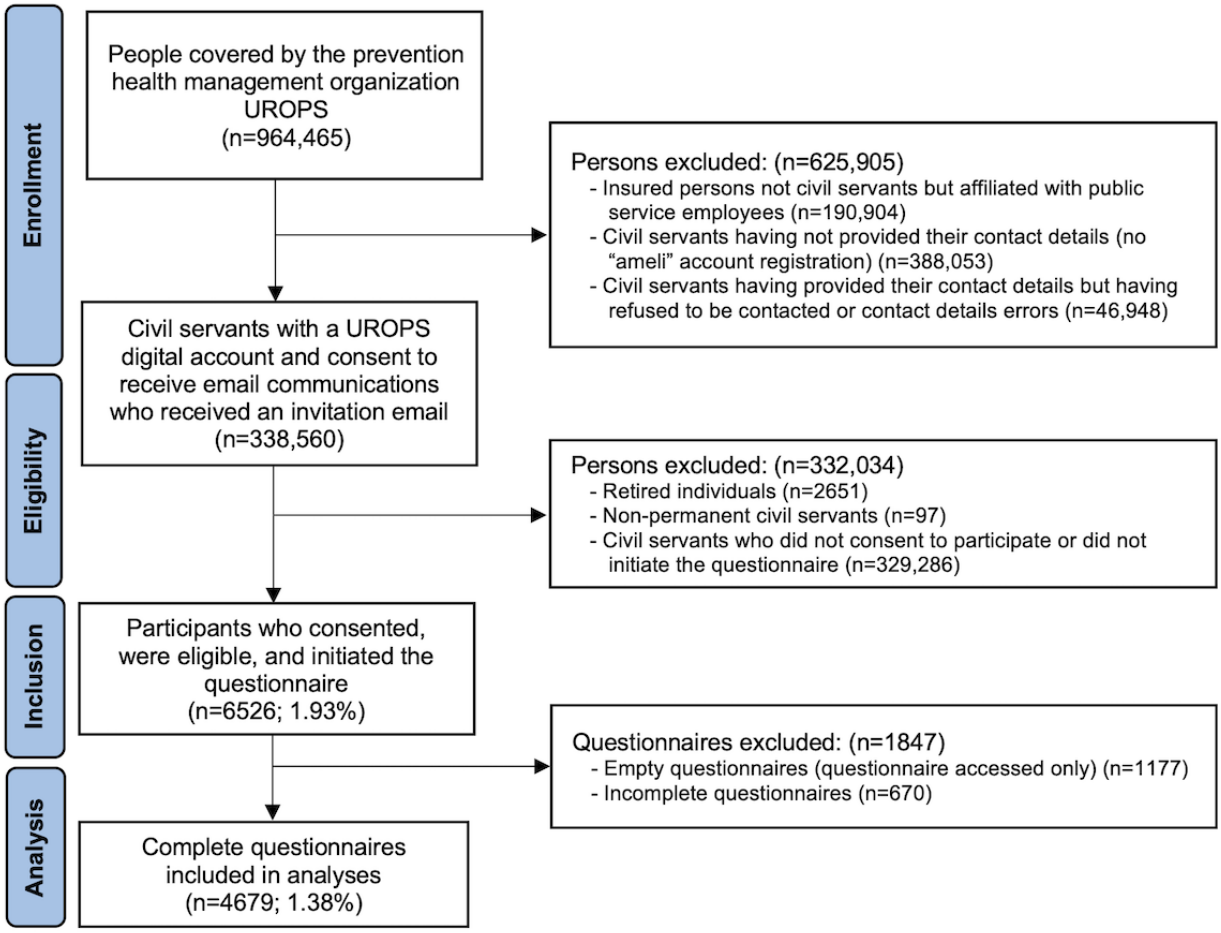
Figure 1 shows the flowchart of the study. A total of 338,560 civil servants received the study invitation. Among them, 6526 (1.93%) individuals provided consent, were eligible, and initiated the questionnaire. Of these participants, 4679 (1.38% of invited individuals) completed all items required for KAB score calculation and were included in the final analyses. This sample size, well above the calculated minimum, ensured high statistical precision, whereas the voluntary recruitment process may limit full population representativeness.

Table 1 shows the sociodemographic characteristics and information related to MHFA training of the participants. Most participants were women (3022/4679, 64.59%), aged between 40 and 69 years (4317/4679, 92.26%), and belonged to the A professional category (2207/4679, 47.17%). Participants with an undergraduate degree (first-cycle university degree and second-cycle university degree) were predominant (2220/4679, 47.45%). Most participants had never received training in mental health (3652/4679, 78.05%) and had not had the opportunity to follow an MHFA training (4134/4679, 88.35%).

The distribution of questionnaire breakoff according to the last completed questionnaire item is presented in Multimedia Appendix 6. Characteristics of complete and incomplete respondents are presented in Multimedia Appendix 5. Compared with respondents who completed the questionnaire, those with incomplete questionnaires differed slightly in age, gender, professional category, and educational level (all *P*<.001). However, effect sizes were negligible to small (Cramer *V*=0.06-0.11), suggesting limited practical differences between groups. No significant differences were observed regarding previous mental health training (*P*=.20) or MHFA training status (*P*=.75).

XSL•FO
RenderX

Figure 1. Flowchart of the study. UROPS: Union Régime Obligatoire Prévention Santé.



XSL•FO
RenderX

Table 1. Sociodemographic characteristics and information related to Mental Health First Aid (MHFA) training of the study population (N=4679).

Variables	Participants, n (%)
Age range (years)	
18-29	59 (1.26)
30-39	303 (6.48)
40-49	1075 (22.97)
50-59	1924 (41.12)
60-69	1318 (28.17)
Gender	
Woman	3022 (64.59)
Man	1650 (35.26)
Other	7 (0.15)
Professional category	
A	2207 (47.17)
B	1777 (37.98)
C	695 (14.85)
Level of education	
No diploma	39 (0.83)
Below high school diploma	433 (9.25)
High school diploma[a]	959 (20.50)
First-cycle university degree[b]	878 (18.76)
Second-cycle university degree[c]	1342 (26.68)
Third-cycle university degree[d]	1028 (21.97)
Previous mental health training	
Yes	1027 (21.95)
No or do not know	3652 (78.05)
Previous MHFA training	
Yes, <6 months ago	216 (4.62)
Yes, 6-12 months ago	117 (2.50)
Yes, >12 months ago	212 (4.53)
No	4134 (88.35)

[a]French baccalaureate.
[b]Baccalaureate+2 years.
[c]Baccalaureate+3 years.
[d]Baccalaureate+5 years and more.

KAB Scores in the Civil Servant Population in France

Table 2 presents descriptive statistics of KAB scores. In the overall civil servant population (N=4679), the mean total KAB score was 82.88 (SD 19.14; 95% CI 82.33-83.43), based on a normalized scale theoretically ranging from 0 to 120 points, corresponding to a “marginal” level according to the predefined scoring thresholds.

The mean scores for the 3 subdimensions—knowledge, attitudes, and behaviors—were 31.31 (SD 5.26; 95% CI 31.16-31.46), 29.72 (SD 11.52; 95% CI 29.39-30.05), and 21.86 (SD 7.96; 95% CI 21.63-22.09), respectively, each based on a normalized scale theoretically ranging from 0 to 40 points.

KAB scores differed according to MHFA training status and, among trained participants, by time since training (Table 2). Participants who had received MHFA training had a significantly higher mean KAB score than those without training (mean 95.56, SD 14.17; 95% CI 94.37-96.75 vs mean 81.21, SD 19.09; 95% CI 80.63-81.79; $P<.001$). Similar patterns were observed for each KAB dimension. Trained participants showed higher mean scores for knowledge (mean 33.31, SD 3.96; 95% CI 32.98-33.64 vs mean 31.04, SD 5.35; 95% CI 30.88-31.20;

XSL•FO
RenderX

P<.001), attitudes (mean 34.35, SD 8.81; 95% CI 33.61-35.09 vs mean 29.10, SD 11.70; 95% CI 28.74-29.46; *P*<.001), and behaviors (mean 27.87, SD 6.13; 95% CI 27.36-28.38 vs mean 21.06, SD 7.84; 95% CI 20.82-21.30; *P*<.001).

In addition to statistically significant differences, the comparison between MHFA-trained and untrained participants showed small-to-moderate effect sizes for knowledge (Cohen *d*=0.44) and attitudes (*d*=0.46), a large effect size for behaviors (*d*=0.89), and a moderate-to-large effect size for the overall KAB score (*d*=0.77; Multimedia Appendix 7).

Among participants who had completed MHFA training, a decreasing gradient of mean scores was observed with increasing time since training (<6, 6-12, and >12 months) for the total KAB score (*P*<.001) as well as for knowledge, attitudes, and behaviors scores over time (*P*<.001, *P*=.04, and *P*<.001, respectively). Results of the post-hoc power calculations for these subgroup analyses are presented in Multimedia Appendix 8.

Table 2. Descriptive statistics and univariable comparisons of knowledge, attitudes, and behaviors (KAB) scores by Mental Health First Aid (MHFA) training status and time since training.

MHFA training status	All (N=4679), mean (SD; 95% CI)	No (n=4134), mean (SD; 95% CI)	Yes (n=545), mean (SD; 95% CI)	Yes (<6 months; n=216), mean (SD; 95% CI)	Yes (6-12 months; n=117), mean (SD; 95% CI)	Yes (>12 months; n=212), mean (SD; 95% CI)	*P* value (yes or no)[a]	Cohen *d*	*P* value (<6 or 6-12 or >12 months)[b]
Knowledge score[c]	31.31 (5.26; 31.16-31.46)	31.04 (5.35; 30.88-31.20)	33.31 (3.96; 32.98-33.64)	34.07 (3.58; 33.59-34.55)	33.50 (3.76; 32.82-34.18)	32.42 (4.27; 31.85-32.99)	<.001	0.44	<.001
Attitude score[c]	29.72 (11.52; 29.39-30.05)	29.10 (11.70; 28.74-29.46)	34.35 (8.81; 33.61-35.09)	35.32 (7.53; 34.32-36.32)	34.10 (9.48; 32.38-35.82)	33.58 (9.56; 32.29-34.87)	<.001	0.46	.04
Behavior score[c]	21.86 (7.96; 21.63-22.09)	21.06 (7.84; 20.82-21.30)	27.87 (6.13; 27.36-28.38)	28.7 (5.47; 27.97-29.43)	28.49 (5.38; 27.52-29.46)	26.66 (6.93; 25.73-27.59)	<.001	0.89	<.001
KAB score[d]	82.88 (19.14; 82.33-83.43)	81.21 (19.09; 80.63-81.79)	95.56 (14.17; 94.37-96.75)	98.12 (11.74; 96.55-99.69)	96.09 (14.33; 93.49-98.69)	92.67 (15.79; 90.54-94.80)	<.001	0.77	<.001

[a]One-way ANOVA comparing MHFA-trained and untrained participants. When the Levene test indicated heterogeneity of variances, results were confirmed using the Wilcoxon rank sum test with continuity correction.

[b]One-way ANOVA comparing KAB scores according to time since MHFA training. When the Levene test indicated heterogeneity of variances, results were confirmed using the Kruskal-Wallis rank sum test.

[c]Score between 0 and 40.

[d]Score between 0 and 120.

Identification of Determinants Associated With KAB Scores

To examine factors associated with KAB scores while accounting for potential confounders, multivariable linear regression models were fitted. These analyses aimed to identify variables statistically associated with KAB outcomes, rather than to infer causal effects.

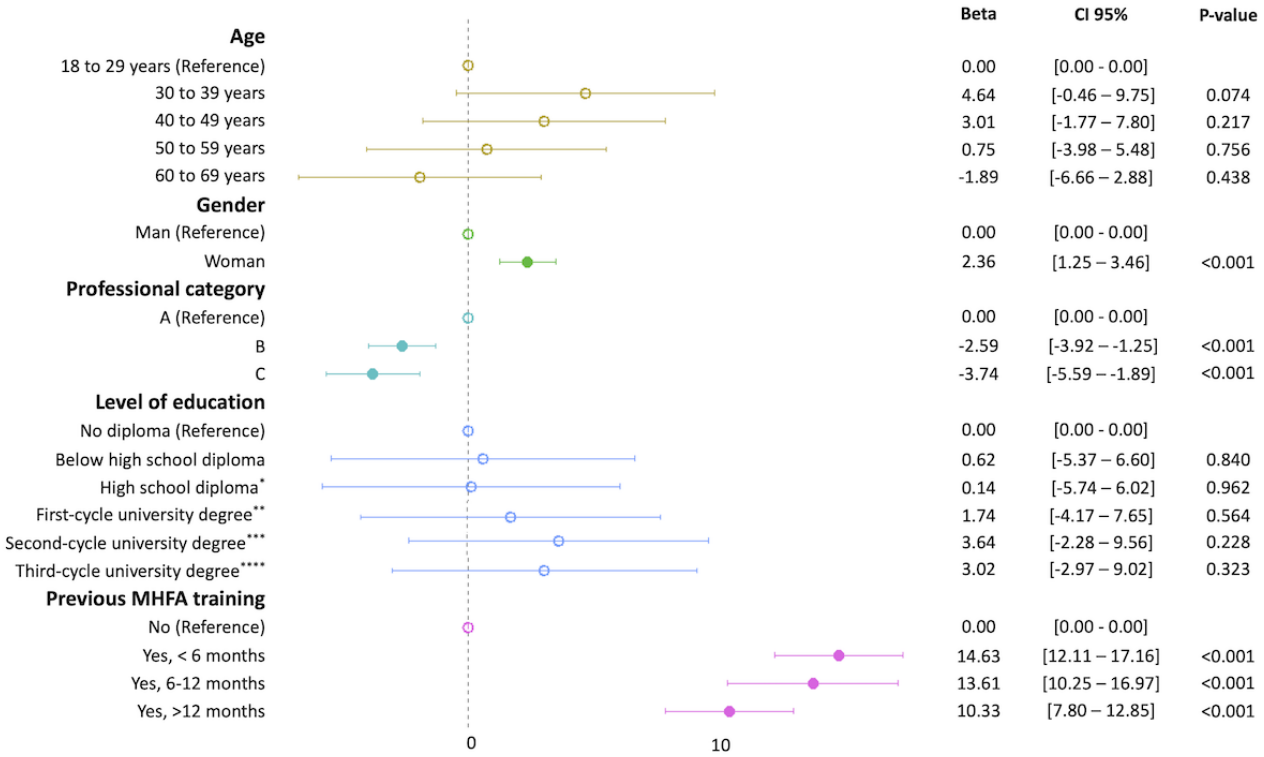
Total KAB Score

As shown in Figure 2, previous MHFA training was strongly associated with higher total KAB scores. Compared with untrained participants, mean KAB scores were higher among those trained: less than 6 months ago (β=14.63, 95% CI 12.11-17.16; *P*<.001), 6-12 months ago (β=13.61, 95% CI 10.25-16.97; *P*<.001), and more than 12 months ago (β=10.33, 95% CI 7.80-12.85; *P*<.001).

Women had higher total KAB scores than men (β=2.36, 95% CI 1.25-3.46; *P*<.001). In contrast, participants in professional categories B and C had lower KAB scores compared with those in category A, with mean decreases of 2.59 (95% CI –3.92 to –1.25) points and 3.74 (95% CI –5.59 to –1.89) points, respectively (all *P*<.001).

XSL•FO
RenderX

Figure 2. Forest plot representing the determinants of the KAB score regarding psychological distress in others. Filled circles indicate statistical significance (P≤.05), while empty circles indicate nonsignificance (*P*>.05). Reference categories are specified in parentheses next to the variable labels. KAB: knowledge, attitudes, and behaviors; MHFA: Mental Health First Aid. *French baccalaureate, **Baccalaureate+2 years, ***Baccalaureate+3 years, ****Baccalaureate+5 years and more.

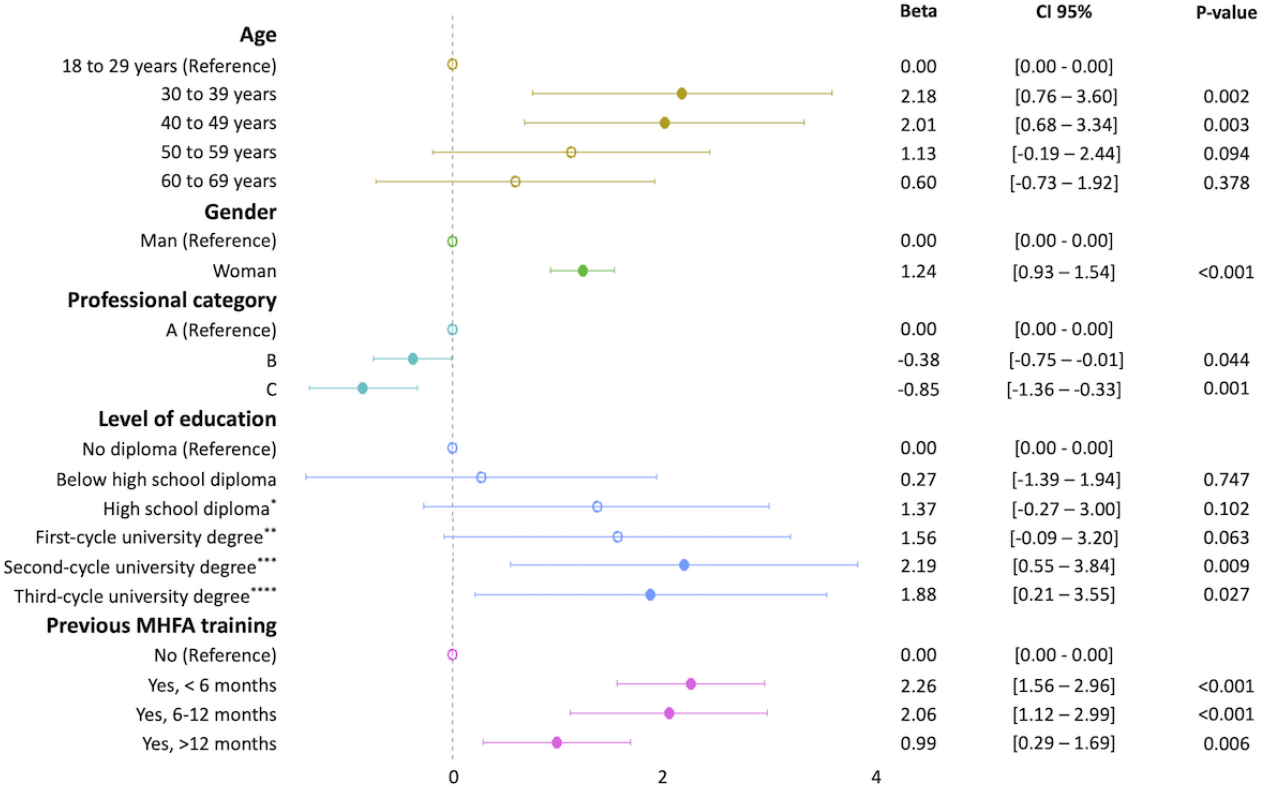


Knowledge Score

As shown in Figure 3, higher knowledge scores were associated with MHFA training status, women gender, and higher professional category. Participants aged 30-39 and 40-49 years had higher knowledge scores than those aged 18-29 years (β=2.18, 95% CI 0.76-3.60; *P*=.002 and β=2.01, 95% CI 0.68-3.34; *P*=.003, respectively). Higher knowledge scores were also observed among participants with an undergraduate degree or a postgraduate or graduate degree compared with those without a diploma (β=2.19, 95% CI 0.55-3.84; *P*=.009 and β=1.88, 95% CI 0.21-3.55; *P*=.03, respectively).

XSL•FO
RenderX

Figure 3. Forest plot representing the determinants of the knowledge score regarding psychological distress in others. Filled circles indicate statistical significance (P≤.05), while empty circles indicate nonsignificance (*P*>.05). Reference categories are specified in parentheses next to the variable labels. MHFA: Mental Health First Aid. *French baccalaureate, **Baccalaureate+2 years, ***Baccalaureate+3 years, ****Baccalaureate+5 years and more.



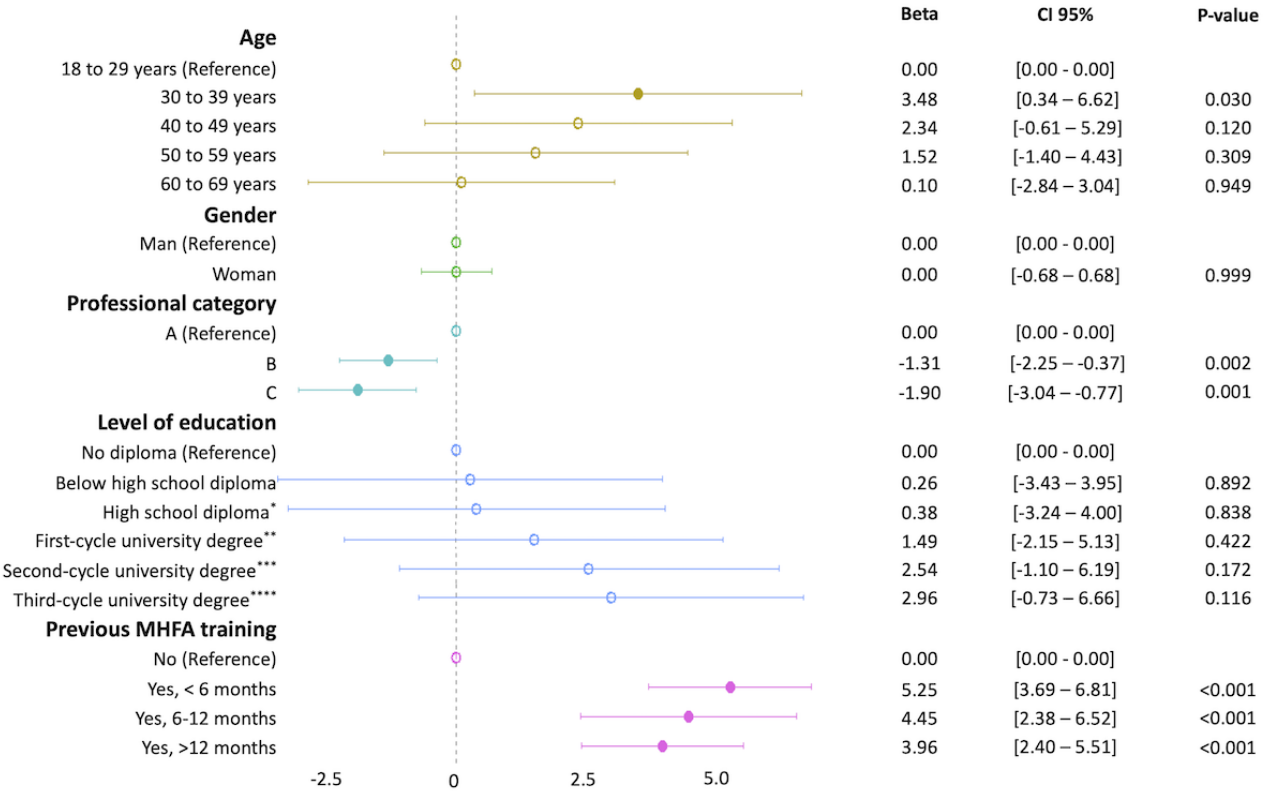
Attitude Score

As shown in Figure 4, attitude scores were positively associated with MHFA training across all time categories. Participants trained less than 6 months ago had the highest adjusted attitude scores, followed by those trained 6-12 months and more than 12 months ago.

Professional categories B and C were negatively associated with attitude scores compared with category A. Age, gender, and education level were not significantly associated with attitude scores in adjusted models, except participants aged 30-39 years who had significantly higher attitude scores than those aged 18-29 years (β=3.48, 95% CI 0.34-6.62; *P*=.03).

XSL•FO
RenderX

Figure 4. Forest plot representing the determinants of the attitude score regarding psychological distress in others. Filled circles indicate statistical significance (P≤.05), while empty circles indicate nonsignificance (*P*>.05). Reference categories are specified in parentheses next to the variable labels. MHFA: Mental Health First Aid. *French baccalaureate, **Baccalaureate+2 years, ***Baccalaureate+3 years, ****Baccalaureate+5 years and more.

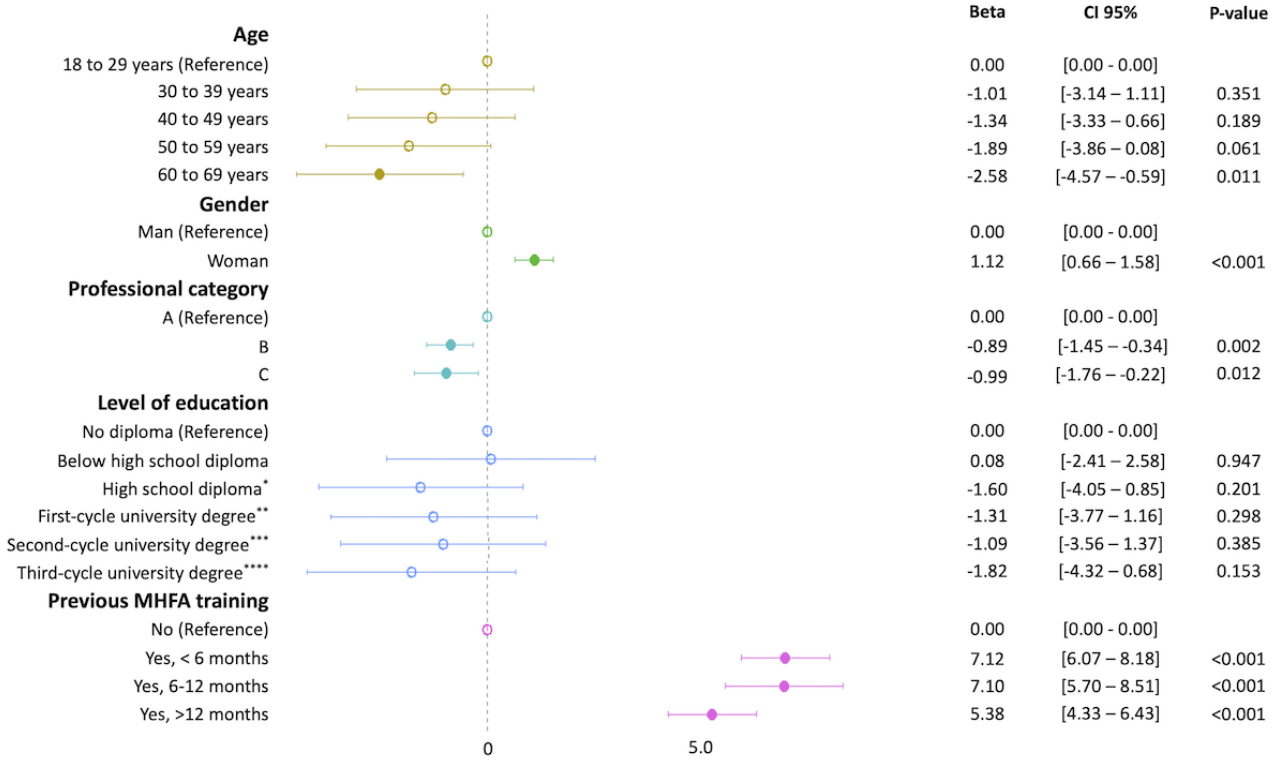


Behavior Score

As shown in Figure 5, MHFA training status was strongly associated with higher behavior scores across all time categories. Woman gender was also associated with higher behavior scores, whereas increasing age and lower professional categories were associated with lower scores.

XSL•FO
RenderX

Figure 5. Forest plot representing the determinants of the behavior score regarding psychological distress in others. Filled circles indicate statistical significance (P≤.05), while empty circles indicate nonsignificance (*P*>.05). Reference categories are specified in parentheses next to the variable labels. MHFA: Mental Health First Aid. *French Baccalaureate, **Baccalaureate+2 years, ***Baccalaureate+3 years, ****Baccalaureate+5 years and more.



Motivations, Knowledge Retention, and Application of MHFA Skills Among Trained Civil Servants

Table 3 describes the motivations for MHFA training, knowledge of ALGEE (Approach, Listen, Give support, Encourage professional help, Encourage other supports), and application in practice. Among the 545 participants who had completed MHFA training, the majority reported enrolling for reasons related to their professional or academic environment (384/545, 70.46%), while over half cited the desire to better support family members or friends (290/545, 53.21%). Additional motivations included a personal interest in mental health (228/545, 41.83%), the wish to better understand their own mental health (188/545, 34.50%), and civic engagement (193/545, 35.41%). Fewer participants indicated motivations related to the COVID-19 crisis (26/545, 4.77%) or the free nature of the training (33/545, 6.06%).

Assessment of core MHFA concepts showed high levels of retention of key components of the ALGEE action plan (a major component of MHFA training), with over 97% correctly identifying the importance of listening without judgment (531/545, 97.43%) and providing comfort and information (528/545, 96.88%). However, some participants still endorsed less appropriate responses, such as “encouraging the person to pull themselves together” (206/545, 37.80%) or “trying to solve their problems” (137/545, 25.14%).

Approximately one-third of trained participants (194/545, 35.60%) reported having used MHFA skills to help someone experiencing psychological distress. Among these, most had supported 1 to 4 individuals (145/545, 26.61%), while a smaller number reported helping 5 or more (49/545, 8.99%). Among those who had used their training, the majority (170/194, 87.63%) perceived it as helpful in supporting the person in distress.

XSL•FO
RenderX

Table 3. Motivations for Mental Health First Aid (MHFA) training, knowledge of ALGEE (Approach, Listen, Give support, Encourage professional help, Encourage other supports), and application in practice (n=545).

Variables	Participants, n (%)
Motivation for participation to MHFA training	
To better support family members or friends	290 (53.21)
Related to and useful for my studies and/or profession	384 (70.46)
To improve my understanding of my own mental health	188 (34.50)
Due to my interest in the field of mental health	228 (41.83)
To engage in a civic action and act in the public interest	193 (35.41)
To address a mental health need arising from the COVID-19 crisis	26 (4.77)
Because this training was free	33 (6.06)
Other	22 (4.04)
Level of knowledge of the ALGEE action plan among participants who had completed MHFA training	
Trying to treat the person or find solutions to their problems[a]	408 (74.86)
Providing support and assessment in case of a crisis[b]	476 (87.34)
Providing comfort and information[b]	528 (96.88)
Listening actively and without judgment[b]	531 (97.43)
Encouraging the person to pull themselves together[a]	339 (62.20)
Applied MHFA skills to help someone with a mental health problem	
Yes	194 (35.60)
No or prefer not to say	351 (64.40)
Number of people assisted (approximate)	
0	351 (64.40)
1	58 (10.64)
2-4	87 (15.96)
5-7	22 (4.04)
8-10	8 (1.47)
>10	19 (3.49)
Perceived usefulness of MHFA knowledge in helping the persons	
Yes	170 (87.63)
No or do not know	24 (12.37)

[a]Correct answer “No.”

[b]Correct answer “Yes.”

Discussion

Principal Findings

This large cross-sectional study provides the first comprehensive assessment of KAB related to psychological distress among French civil servants. The results reveal substantial heterogeneity within this population, with an overall mean KAB score of 82.88 of 120 (SD 19.14), classified as marginal (95% CI 82.33-83.43). From a public health perspective, it highlights that mental health competencies remain insufficiently consolidated in a workforce central to public service delivery and social functioning. While these observations provide valuable insight into mental health literacy within the French civil service, they are based primarily on self-reported measures and should therefore be interpreted with appropriate caution. In particular, reported use of MHFA-related skills and perceived usefulness of training reflect participants’ own perceptions and were not independently validated through recipient-level outcomes.

Mental health disorders are now among the leading causes of disability worldwide, with well-documented consequences for work ability, productivity, and long-term labor participation [3,25,26]. In this context, the marginal level observed highlights a gap between the growing burden of mental health needs and the collective capacity of institutions to recognize and respond to psychological distress. Addressing mental health in the

XSL•FO
RenderX

workplace must thus go beyond clinical care and be integrated into public health strategies.

The strongest finding of this study is that MHFA-trained participants displayed substantially higher KAB scores than nontrained participants across all dimensions: knowledge (+2.3 points), attitudes (+5.3 points), and behaviors (+6.8 points). Trained participants had a total KAB score of 95.56 versus 81.21 for untrained individuals ($P<.001$). This interpretation is further supported by effect size estimates, which were moderate for the overall KAB score (Cohen $d=0.77$) and particularly large for the behavior dimension ($d=0.89$). However, given the cross-sectional design and the voluntary nature of participation in MHFA training, these differences should be interpreted as associations rather than evidence of a causal training effect. Importantly, participants who had completed MHFA training more than 12 months previously continued to display higher KAB scores than untrained individuals. Although lower scores were observed among participants reporting longer intervals since training, this pattern should be interpreted cautiously because time since training represents only an indirect indicator of retention and may also reflect differences in opportunities for skill application, workplace environment, refresher activities, or other unmeasured factors. Beyond these quantitative findings, responses from trained participants also provide insight into the motivations for attending MHFA training and the practical application of acquired skills. Most cited professional or academic reasons for attending MHFA (384/545, 70.46%), followed by the desire to support relatives or friends (290/545, 53.21%). These findings highlight the strategic relevance of MHFA training within the civil service and underscore participants' recognition of the importance of mental health considerations in the workplace.

Core ALGEE principles were well retained—for example, 97.43% (531/545) recalled the importance of nonjudgmental listening—though some suboptimal reflexes persisted (eg, 206/545, 37.80% agreed with "pulling oneself together"). About 35.60% (194/545) had applied MHFA skills to help someone, most often 1 to 4 individuals, and 87.63% (170/194) found the training useful in doing so. These findings suggest that MHFA can be meaningfully translated into supportive behaviors.

Finally, the sample was predominantly women (3022/4679, 64.59%) and aged 50 years and older (3242/4679, 69.29%), aligning with broader public service demographics [9]. This should be considered when tailoring future training strategies to subgroups with differing needs or exposure.

Comparison With Prior Work

The positive impact of MHFA training found in this study aligns with existing literature [27-31]. A 2018 meta-analysis by Morgan et al [27] showed consistent improvements in mental health knowledge, stigma reduction, and helping behaviors across diverse international contexts. Our findings extend these results to a civil service context—a group not yet well represented in MHFA research. Nevertheless, unlike the intervention studies cited earlier, this study was not designed to evaluate MHFA effectiveness. Participation in MHFA training was voluntary, and self-selection mechanisms may partly explain the higher KAB scores observed among trained participants. Consequently, our findings should be interpreted as evidence of an association between MHFA training exposure and KAB outcomes rather than proof of a causal effect.

French studies conducted among university students [17,18] as well as a randomized controlled trial in Switzerland [32] have shown significant posttraining improvements in knowledge, attitudes, and peer support capacities. Effects were particularly strong for proximal outcomes such as confidence and willingness to help. Our study confirms that these improvements are also observable in older, more professionally diverse cohorts, offering useful insights for public policy and workforce-based mental health strategies.

Beyond perceived improvements, this study documents the self-reported application of MHFA skills. Over one-third of trained participants reported supporting someone in psychological distress, with the vast majority considering their training helpful in doing so. These findings suggest that MHFA-related competencies may be actively mobilized in professional or social contexts. However, as highlighted in a systematic review by Forthal et al [33], the evidence on real-world outcomes remains mixed, and few studies have rigorously assessed whether actions taken by trainees are truly helpful or improve recipient outcomes. Consequently, this study cannot determine whether actions taken by trainees actually improved recipient outcomes, increased professional help-seeking, reduced psychological distress, or improved other recipient-level outcomes. Our findings therefore extend the MHFA literature by documenting self-reported application of MHFA skills, but independent validation of recipient outcomes remains an important gap in the literature. Further research using validated recipient-level outcomes is needed to establish the real-world impact of MHFA training.

Among MHFA-trained civil servants, a high level of retention of key elements of the ALGEE action plan was observed, particularly nonjudgmental listening (531/545, 97.43%) and providing comfort and information (528/545, 96.88%). These rates are in line with previous findings from student populations, where essential MHFA principles remained well retained months after training [17,32]. Nonetheless, some persistent misconceptions were noted—nearly 37.80% (206/545) of trained participants agreed with the idea of "encouraging the person to pull themselves together" [17], a response inconsistent with MHFA principles. This enduring influence of cultural or intuitive norms highlights the limits of single-session training and reinforces the need for ongoing reinforcement.

Our study provides exploratory observations regarding differences in KAB scores according to time since training. While lower scores were observed among participants reporting longer intervals since MHFA training, the cross-sectional design does not allow these differences to be interpreted as direct evidence of skill decay. Alternative explanations, including cohort effects, differences in training implementation, opportunities for skill application, workplace support, refresher training, or other unmeasured factors, cannot be excluded. This pattern is consistent with findings in health education and learning science, particularly the "spacing effect," which shows that spaced learning supports long-term retention more

XSL•FO
RenderX

effectively than one-time exposure [34,35]. These results argue for the inclusion of refresher sessions as part of long-term MHFA implementation strategies, especially in large and diverse workforces.

Interestingly, 11.65% (545/4679) of our sample reported having received MHFA training—substantially higher than national averages (estimated at ~0.2% of the general population) [15]. This suggests a possible selection bias, but may also reflect higher interest, access, or perceived relevance of MHFA among public sector employees. Notably, 70.46% (384/545) of trained participants cited professional or academic motivations, reinforcing the relevance of MHFA for emotionally demanding roles and public service functions.

Sociodemographic gradients in KAB scores were consistent with prior findings. Women scored higher across all dimensions, in line with studies showing greater mental health literacy and empathy among women respondents [36,37]. Participants in higher professional categories (category A) and with higher education levels also showed stronger outcomes, likely reflecting greater exposure to psychosocial issues or prior training. Similar patterns were reported in KAB-focused studies among civil servants, including in the context of long COVID [38], or oral health literacy [39], underlining the structural dimensions that shape prevention capacities.

Taken together, these findings confirm the relevance of MHFA in occupational settings and provide strong support for its application in workforces. However, to ensure sustainable impact, training should be combined with targeted reinforcement strategies and equitable access across all staff categories.

Limitations

Several limitations should be acknowledged. First, the cross-sectional nature of the study precludes causal inference regarding MHFA training effectiveness. Because participation in MHFA training was voluntary, self-selection cannot be excluded, and trained participants may have differed from untrained participants in baseline mental health literacy, motivation, or interest in mental health. Consequently, the observed differences in KAB scores may partly reflect preexisting characteristics rather than training effects. Although multivariable analyses adjusted for several sociodemographic variables, residual confounding and reverse causation remain possible. Longitudinal or pre- and poststudies would be required to better assess the causal impact of MHFA training.

Second, although the invitation was distributed to the entire recruitment frame of French civil servants who had agreed to receive email communications, the low participation rate (1.93%) raises concerns about selection bias, increasing the risk of self-selection and limiting population representativeness. This low participation rate is consistent with voluntary nationwide web-based surveys conducted without incentives and may also reflect the sensitive nature of mental health as a survey topic. Individuals who chose to participate may have differed systematically from nonparticipants with respect to mental health interest, literacy, previous training experiences, or personal exposure to psychological distress. Consequently, the sample may not fully represent the overall French civil servant population.

In addition, participation in MHFA training itself is voluntary. Individuals who attend such training are likely to differ from nontrained individuals in motivation, professional interest, empathy, or preexisting mental health knowledge. Therefore, the higher KAB scores observed among MHFA-trained participants cannot be interpreted as evidence of a causal training effect and may partly reflect self-selection. Although multivariable analyses adjusted for age, gender, educational level, and professional category, residual confounding due to unmeasured characteristics remains possible. Regarding incomplete questionnaires, although some statistically significant differences were observed between complete and incomplete respondents, the corresponding effect sizes were negligible to small (Multimedia Appendix 5), suggesting limited selection based on the observed baseline characteristics. Only a limited set of baseline variables was available for incomplete respondents, preventing the development of a robust response-propensity model. Consequently, inverse probability weighting would have been limited in its ability to account for potentially outcome-related noncompletion and would itself have required substantial assumptions about the missingness mechanism. However, because KAB outcomes were unavailable for participants who did not complete the questionnaire, we cannot exclude the possibility that complete-case estimates remain affected by outcome-related bias. Therefore, the associations should be interpreted as observed within a large volunteer sample rather than estimates that can be directly generalized to all French civil servants.

Third, the questionnaire was adapted from a previously developed instrument but was not formally psychometrically validated in this population. Consequently, the KAB scores should be interpreted as descriptive indicators of mental health–related knowledge, attitudes, and behaviors rather than as measurements derived from a validated psychometric scale.

Fourth, the study population consisted exclusively of French civil servants affiliated with a specific health insurance and prevention system and who had agreed to receive email communications. Consequently, the findings may not be directly generalizable to other occupational groups, private-sector employees, or populations from different health care, organizational, or cultural contexts.

Fifth, the use of self-reported data could introduce social desirability bias, particularly regarding attitudes or prosocial behaviors. The anonymity of the survey and validation of the questionnaire mitigate this risk, but it cannot be excluded.

Sixth, time since training was used as an indirect proxy for retention and did not capture frequency of skill application, refresher training, workplace support, or other factors that may influence the long-term maintenance of MHFA-related competencies.

Despite these limitations, this study provides valuable descriptive data on mental health literacy and supportive behaviors among civil servants and offers a relevant baseline for future longitudinal and intervention-based research.

XSL•FO
RenderX

Conclusions

This nationwide cross-sectional study highlights substantial heterogeneity in KAB related to psychological distress among French civil servants, with overall KAB levels remaining in the marginal range. Exposure to MHFA training was consistently associated with higher KAB scores, supporting its role as a public health literacy intervention in occupational settings. However, the observed attenuation of scores over time underscores the need to consider mental health competencies as a potentially declining dynamic rather than permanently acquired. Taken together, these findings support the relevance of integrating MHFA within broader and sustained prevention strategies, combining initial training with regular reinforcement mechanisms to strengthen mental health literacy. Their deployment in the civil service, and other workplace settings, would improve the collective capacity to respond to psychological distress, leading to greater social solidarity.

Acknowledgments

The authors would like to thank all participants who responded to the questionnaire. The authors would also like to thank the authors of a previous similar study for sharing their questionnaires, especially Ilaria Montagni. The authors would also like to thank Union Régime Obligatoire Prévention Santé, a French health prevention organization that manages the compulsory health insurance scheme for civil servants, for distributing the questionnaire to their insured members, with special thanks to Marina Molin Orru, David Ollivier-Lannuzel, and José de Freitas. The authors did not use generative AI for this manuscript.

Funding

The authors declared no financial support was received for this work.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: M Lebrat, FC, BSV, M Lamure, RS, NF, and CD
Methodology: M Lebrat, FC, BSV, and M Lamure
Project administration: FC, BSV, and CD
Supervision: M Lebrat, FC, BSV, and CD
Visualization: M Lebrat, FC, BSV, M Lamure, and CS
Writing—original draft: M Lebrat, FC, and BSV
Writing—review and editing: M Lebrat, FC, BSV, M Lamure, CS, RS, CM, NF, and CD

Conflicts of Interest

None declared.

Multimedia Appendix 1
CROSS checklist.
[DOCX File , 27 KB - publichealth_v12i1e93207_app1.docx]

Multimedia Appendix 2
Original French questionnaire.
[DOCX File , 28 KB - publichealth_v12i1e93207_app2.docx]

Multimedia Appendix 3
English translation of the original French questionnaire.
[DOCX File , 28 KB - publichealth_v12i1e93207_app3.docx]

Multimedia Appendix 4
Raw score ranges, normalization procedure, and normalized score ranges for the knowledge, attitudes, and behaviors dimensions.
[DOCX File , 20 KB - publichealth_v12i1e93207_app4.docx]

Multimedia Appendix 5
Comparison between respondents who completed the entire questionnaire (N=4679) and those who completed only part of it but who at least provided their sociodemographic data (n=670).

XSL•FO
RenderX

[DOCX File , 21 KB - publichealth_v12i1e93207_app5.docx]

Multimedia Appendix 6
Distribution of incomplete questionnaires according to the last completed questionnaire item.
[DOCX File , 20 KB - publichealth_v12i1e93207_app6.docx]

Multimedia Appendix 7
Cohen *d* effect sizes for the comparison between Mental Health First Aid–trained and untrained participants.
[DOCX File , 17 KB - publichealth_v12i1e93207_app7.docx]

Multimedia Appendix 8
Post-hoc power calculations for subgroup analyses according to time since Mental Health First Aid training.
[DOCX File , 20 KB - publichealth_v12i1e93207_app8.docx]

References

1. Clinical descriptions and diagnostic requirements for ICD-11 mental, behavioural and neurodevelopmental disorders (CDDR). World Health Organization. 2024. URL: https://www.who.int/publications/i/item/9789240077263 [accessed 2026-01-26]
2. Kim H, Jung JH, Han K, Jeon HJ. Risk of suicide and all-cause death in patients with mental disorders: a nationwide cohort study. Mol Psychiatry 2025;30(7):2831-2839. [doi: 10.1038/s41380-025-02887-4] [Medline: 39843548]
3. World mental health today: latest data. World Health Organization. 2025. URL: https://www.who.int/publications/i/item/9789240113817 [accessed 2026-01-26]
4. GBD 2023 Mental Disorder Collaborators. Updated trends in the global prevalence and burden of mental disorders, 1990-2023: a systematic analysis for the Global Burden of Disease Study 2023. Lancet 2026;407(10543):2040-2064. [doi: 10.1016/S0140-6736(26)00519-2] [Medline: 42167272]
5. COVID-19 Mental Disorders Collaborators. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. Lancet 2021;398(10312):1700-1712 [FREE Full text] [doi: 10.1016/S0140-6736(21)02143-7] [Medline: 34634250]
6. Troubles anxio-dépressifs et conduites suicidaires?: Déterminants sociaux des différences de genre. Synthèse élaborée à partir des données produites par Santé publique France sur la période 2010-2023. Santé Publique France. 2010. URL: https://www.santepubliquefrance.fr/import/troubles-anxio-depressifs-et-conduites-suicidaires-determinants-sociaux-des-differences-de-genre.-synthese-elaboree-a-partir-des-donnees-produite [accessed 2026-01-26]
7. World mental health report: transforming mental health for all. World Health Organization. 2022. URL: https://www.who.int/publications/i/item/9789240049338 [accessed 2026-01-27]
8. Le gouvernement prolonge en 2026 la grande cause nationale dédiée à la santé mentale. info.gouv.fr. 2025. URL: https://www.info.gouv.fr/communique/le-gouvernement-prolonge-en-2026-la-grande-cause-nationale-dediee-a-la-sante-mentale [accessed 2026-01-26]
9. Rapport annuel sur l'état de la fonction publique—édition 2025. Le Portail de la Fonction Publique. 2025. URL: https://www.fonction-publique.gouv.fr/toutes-les-publications/rapport-annuel-sur-letat-de-la-fonction-publique-edition-2025 [accessed 2025-11-22]
10. Borst RT, Knies E. Well-being of public servants under pressure: the roles of job demands and personality traits in the health-impairment process. Rev Public Pers Adm 2021;43(1):159-184. [doi: 10.1177/0734371x211052674]
11. Niedhammer I, Bertrais S. Organisational changes and psychosocial work factors: prospective findings from the national French working conditions survey. Saf Health Work 2024;15(4):496-499 [FREE Full text] [doi: 10.1016/j.shaw.2024.10.002] [Medline: 39697314]
12. Theorell T, Hammarström A, Aronsson G, Träskman Bendz L, Grape T, Hogstedt C, et al. A systematic review including meta-analysis of work environment and depressive symptoms. BMC Public Health 2015;15:738 [FREE Full text] [doi: 10.1186/s12889-015-1954-4] [Medline: 26232123]
13. WHO and ILO call for new measures to tackle mental health issues at work. World Health Organization. 2022. URL: https://www.who.int/news/item/28-09-2022-who-and-ilo-call-for-new-measures-to-tackle-mental-health-issues-at-work [accessed 2026-01-26]
14. Mental health at work. World Health Organization. 2024. URL: https://www.who.int/news-room/fact-sheets/detail/mental-health-at-work [accessed 2026-01-26]
15. Suicide: mal-être croissant des jeunes femmes et fin de vie. Statistique publique de la santé et des solidarités. 2025 Feb. URL: https://drees.solidarites-sante.gouv.fr/sites/default/files/2025-04/ONS%202025-6e-RAPPORT%20BAT%20WEB%20V2.pdf [accessed 2026-01-26]
16. Circulaire du 23 février 2022 relative aux actions de sensibilisation et de formation au secourisme en santé mentale dans la fonction publique. Légifrance. URL: https://www.legifrance.gouv.fr/circulaire/id/45298 [accessed 2026-01-26]

XSL•FO
RenderX

17. Llopis N, Antoine E, Grové C, Marinucci A, Touchard F, Montagni I. A mixed-methods project on the impact of the mental health first aid training on French university students' knowledge, attitudes and practices. Early Interv Psychiatry 2024;18(5):366-373. [doi: 10.1111/eip.13471] [Medline: 37749765]
18. Sibeoni J, Ellul P, Bubola T, Debiche Y, Piot M. Mental health first aid training among healthcare French students: a qualitative study. Front Med (Lausanne) 2023;10:1268277 [FREE Full text] [doi: 10.3389/fmed.2023.1268277] [Medline: 37937137]
19. Mei C, McGorry PD. Mental health first aid: strengthening its impact for aid recipients. Evid Based Ment Health 2020;23(4):133-134 [FREE Full text] [doi: 10.1136/ebmental-2020-300154] [Medline: 32727814]
20. Advocacy, communication and social mobilization for TB control: a guide to developing knowledge, attitude and practice surveys. World Health Organization. 2008. URL: https://www.who.int/publications/i/item/9789241596176 [accessed 2026-01-26]
21. Andrade C, Menon V, Ameen S, Kumar Praharaj S. Designing and conducting knowledge, attitude, and practice surveys in psychiatry: practical guidance. Indian J Psychol Med 2020;42(5):478-481 [FREE Full text] [doi: 10.1177/0253717620946111] [Medline: 33414597]
22. Conner M, Norman P. Predicting and Changing Health Behaviour: Research and Practice with Social Cognition Models. Berkshire, United Kingdom: McGraw-Hill Education (UK); 2015.
23. Carrouel F, Nugem R, Fraticelli L, Fayet Y, Basbous L, Salque C, et al. Behavioral profiles and sociodemographic predictors of planetary health diet engagement among health care professionals to inform public health promotion: cross-sectional study. JMIR Public Health Surveill 2025;11:e67633 [FREE Full text] [doi: 10.2196/67633] [Medline: 41223423]
24. Sharma A, Minh Duc NT, Luu Lam Thang T, Nam NH, Ng SJ, Abbas KS, et al. A consensus-based Checklist for Reporting Of Survey Studies (CROSS). J Gen Intern Med 2021;36(10):3179-3187 [FREE Full text] [doi: 10.1007/s11606-021-06737-1] [Medline: 33886027]
25. de Oliveira C, Saka M, Bone L, Jacobs R. The role of mental health on workplace productivity: a critical review of the literature. Appl Health Econ Health Policy 2023;21(2):167-193 [FREE Full text] [doi: 10.1007/s40258-022-00761-w] [Medline: 36376610]
26. OECD. Health at a Glance 2025: OECD Indicators. Paris, France: Health Glance OECD Publishing; 2025.
27. Morgan AJ, Ross A, Reavley NJ. Systematic review and meta-analysis of Mental Health First Aid training: effects on knowledge, stigma, and helping behaviour. PLoS One 2018;13(5):e0197102 [FREE Full text] [doi: 10.1371/journal.pone.0197102] [Medline: 29851974]
28. Burns S, Crawford G, Hallett J, Hunt K, Chih HJ, Tilley PM. What's wrong with John? A randomised controlled trial of Mental Health First Aid (MHFA) training with nursing students. BMC Psychiatry 2017;17(1):111 [FREE Full text] [doi: 10.1186/s12888-017-1278-2] [Medline: 28335758]
29. Hung MSY, Lam SKK, Chow MCM. Nursing students' experiences of mental health first aid training: a qualitative descriptive study. Collegian 2019;26(5):534-540. [doi: 10.1016/j.colegn.2019.02.006]
30. Liang Md M, Chen Md Q, Guo Md J, Mei Ph D, Wang Md J, Zhang Md Y, et al. Mental Health First Aid improves mental health literacy among college students: a meta-analysis. J Am Coll Health 2023;71(4):1196-1205. [doi: 10.1080/07448481.2021.1925286] [Medline: 34242537]
31. Schick A, Henn C, Send TS, Maltese S, Deuschle M, Reininghaus U. Initial effects of Mental Health First Aid training on mental health literacy, social distance, stigma, mental health, and quality of life in the German adult population—findings from a pilot study. Eur Psychiatry 2025;68(1):e164. [doi: 10.1192/j.eurpsy.2025.10114] [Medline: 41140121]
32. Dzemaili S, Pasquier J, Oulevey Bachmann A, Mohler-Kuo M. The effectiveness of Mental Health First Aid training among undergraduate students in Switzerland: a randomized control trial. Int J Environ Res Public Health 2023;20(2):1303 [FREE Full text] [doi: 10.3390/ijerph20021303] [Medline: 36674060]
33. Forthal S, Sadowska K, Pike KM, Balachander M, Jacobsson K, Hermosilla S. Mental Health First Aid: a systematic review of trainee behavior and recipient mental health outcomes. Psychiatr Serv 2022;73(4):439-446 [FREE Full text] [doi: 10.1176/appi.ps.202100027] [Medline: 34346736]
34. Cepeda NJ, Vul E, Rohrer D, Wixted JT, Pashler H. Spacing effects in learning: a temporal ridgeline of optimal retention. Psychol Sci 2008;19(11):1095-1102 [FREE Full text] [doi: 10.1111/j.1467-9280.2008.02209.x] [Medline: 19076480]
35. Maye JA, Hurley F. The effectiveness of spaced repetition in medical education: a systematic review and meta-analysis. Clin Teach 2026;23(2):e70353. [doi: 10.1111/tct.70353] [Medline: 41601436]
36. Fischer A, LaFrance M. What drives the smile and the tear: why women are more emotionally expressive than men. Emot Rev 2014;7(1):22-29. [doi: 10.1177/1754073914544406]
37. Else-Quest NM, Hyde JS, Goldsmith HH, Van Hulle CA. Gender differences in temperament: a meta-analysis. Psychol Bull 2006;132(1):33-72. [doi: 10.1037/0033-2909.132.1.33] [Medline: 16435957]
38. Carrouel F, Lvovschi V, du Sartz de Vigneulles B, Rhanoui M, Salamon R, Lamure M, et al. Prevalence, risk factors, disease-related knowledge, and vaccination attitudes and behaviors for long COVID among French civil servants: cross-sectional survey. JMIR Public Health Surveill 2025;11:e83323 [FREE Full text] [doi: 10.2196/83323] [Medline: 41348962]

XSL•FO
RenderX

39. Carrouel F, du Sartz de Vigneulles B, Clément C, Lvovschi V, Verot E, Tantardini V, et al. Promoting health literacy in the workplace among civil servants: cross-sectional study. JMIR Public Health Surveill 2024;10:e58942 [FREE Full text] [doi: 10.2196/58942] [Medline: 39149854]

Abbreviations

ALGEE: Approach, Listen, Give support, Encourage professional help, Encourage other supports
CROSS: Checklist for Reporting Of Survey Studies
KAB: knowledge, attitudes, and behaviors
KAP: knowledge, attitudes, and practices
MHFA: Mental Health First Aid
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 10.Feb.2026; peer-reviewed by F Denis, M Chakit, X Tao, A Kermis; comments to author 15.Jun.2026; revised version received 20.Jul.2026; accepted 28.Jul.2026; published 21.Aug.2026.

Please cite as:
Lebrat M, Carrouel F, du Sartz de Vigneulles B, Lamure M, Salque C, Salamon R, Massoubre C, Franck N, Dussart C
Assessment of Knowledge, Attitudes, and Behaviors Toward Others' Mental Disorders to Improve Mental Health First Aid: Cross-Sectional Study in a French Workplace
JMIR Public Health Surveill 2026;12:e93207
URL: https://publichealth.jmir.org/2026/1/e93207
doi:10.2196/93207
PMID:

©Matthieu Lebrat, Florence Carrouel, Benjamin du Sartz de Vigneulles, Michel Lamure, Corélie Salque, Roger Salamon, Catherine Massoubre, Nicolas Franck, Claude Dussart. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 21.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Psychosocial Correlates of Sedentary Behavior Reduction Readiness Relevant to Zero-Time Exercise Among Chinese Adults in Sedentary Occupations: Cross-Sectional Study

Ziyan Xiong[1], MSc; Xiangyuan Li[1], MSc; Lu Chai[1,2], MSc; Xiongqiang Duan[1], BSc; Xiang Hu[1], BSc; Yuan Xie[1], BSc; Li Cong[1], PhD

[1]Health Science Center, Hunan Normal University, Changsha, Hunan, China
[2]Kiang Wu Nursing College of Macau, Macau, China

Corresponding Author:
Li Cong, PhD
Health Science Center
Hunan Normal University
371 Tongzipo Road
Changsha, Hunan, 410013
China
Phone: 86 15974294346
Email: congli@hunnu.edu.cn

Abstract

Background: Prolonged sitting is an increasingly important public health concern, particularly in occupations that offer few opportunities for conventional exercise. Zero-time exercise (ZTEx) integrates brief muscle-engaging movements into daily routines without requiring dedicated time or equipment and may offer a practical way to interrupt prolonged sitting. However, the motivational, cognitive, and contextual factors associated with intention relevant to ZTEx adoption remain poorly understood.

Objective: This study examined the association between exercise motivation and general sedentary behavior reduction intention, used as a zero-time exercise–related intention proxy (ZTEx-IP), and assessed model-implied indirect associations involving general self-efficacy and perceived physical activity triggers among Chinese adults in sedentary occupations.

Methods: This cross-sectional survey included 790 adults working full-time in sedentary occupations across 20 provincial-level administrative divisions in China between May and September 2025. Eligible participants reported at least 6 hours of work-related sitting on a typical workday. Standardized questionnaires assessed exercise motivation, general self-efficacy, perceived physical activity triggers, and the ZTEx-IP. Multivariable regression and restricted cubic splines examined graded and nonlinear associations between exercise motivation and the ZTEx-IP. PROCESS Models 4 and 6 estimated covariate-adjusted direct and model-implied indirect associations using 5000 bootstrap resamples.

Results: Higher exercise motivation scores were associated with higher ZTEx-IP scores. Compared with participants in the lowest motivation tertile, those in the highest tertile had an adjusted score difference of 8.49 points (unstandardized regression coefficient, B=8.49, 95% CI 6.43-10.60; $P<.001$), and the association was nonlinear (P for nonlinearity<.001). In the prespecified ordered model, exercise motivation retained a direct association with the ZTEx-IP (B=0.167, 95% CI 0.088-0.246). Model-implied indirect associations were observed through general self-efficacy (B=0.104, 95% CI 0.057-0.155), perceived physical activity triggers (B=0.157, 95% CI 0.110-0.205), and their prespecified ordering (B=0.025, 95% CI 0.013-0.039). Perceived triggers accounted for a larger increment in explained variance than general self-efficacy (ΔR^2=0.054 vs 0.033), and the adjusted outcome model explained 37.6% of the variance in ZTEx-IP scores (R^2=0.376).

Conclusions: Among Chinese adults in sedentary occupations, exercise motivation was positively associated with the ZTEx-IP, with model-implied indirect associations involving general self-efficacy and perceived physical activity triggers. The statistical ordering of these associations does not establish that the constructs develop sequentially over time. The comparatively larger trigger-related estimate identifies contextual cue salience as a candidate mechanism for prospective testing but does not demonstrate intervention effectiveness. These hypothesis-generating findings provide a basis for workplace trials of contextually adapted, multicomponent strategies that evaluate actual ZTEx frequency, objectively measured sitting interruptions, adherence over time, and workplace acceptability.

XSL•FO
RenderX

(JMIR Public Health Surveill 2026;12:e95707) doi:10.2196/95707

KEYWORDS

zero-time exercise; sedentary behavior; occupational health; exercise motivation; self-efficacy; physical activity triggers; cross-sectional study

Introduction

Sedentary behavior is defined as any waking behavior characterized by an energy expenditure of 1.5 metabolic equivalents (METs) or less while sitting, reclining, or lying down [1]. Prolonged sitting has been associated with increased risks of dementia, cardiovascular disease, metabolic disorders, and all-cause mortality [2-4]. In China, approximately one-quarter of adults report sitting for more than 6 hours per day, and prolonged occupational sitting is increasingly common [5]. Given that work organization can limit opportunities for conventional exercise, feasible strategies for interrupting sitting without substantially disrupting routine tasks are needed. Zero-time exercise (ZTEx) refers to brief, simple, muscle-engaging movements performed as part of an existing daily activity, without a separate exercise session, specialized equipment, or additional scheduled time [6]. Examples include seated leg raises during desk work and shoulder retraction while typing. Unlike accumulated exercise, which divides physical activity into several discrete bouts, ZTEx is performed alongside occupational or household tasks. This feature may be particularly relevant to workers whose schedules or job demands limit participation in conventional exercise. Evidence that brief, lifestyle-integrated activity can improve cardiometabolic and functional outcomes further supports the potential value of incorporating movement into sedentary routines [7-9].

Research on ZTEx has largely examined its feasibility and preliminary health outcomes [10], while less is known about the psychosocial factors associated with intention to adopt it. Motivation, self-regulation, and perceived capability are established correlates of physical activity [11,12], but ZTEx may also depend on whether brief movement is perceived as feasible within an ongoing task. Workplace conditions such as time pressure, environmental cues, task structure, and organizational norms may therefore be especially relevant to this form of lifestyle-integrated activity [13]. Yet these factors have generally been studied separately, leaving their joint associations with intention relevant to ZTEx adoption unclear.

General sedentary behavior reduction intention and readiness to adopt ZTEx are conceptually related but not equivalent. Their overlap lies in the evaluation of interrupting sitting, perceived social expectations regarding such interruptions, perceived control over doing so, and willingness to change sedentary routines [14]. However, reducing sedentary behavior is a broad behavioral goal that may be pursued through standing, walking, conventional activity breaks, or ZTEx. ZTEx is a more specific means of pursuing this goal through brief muscle-engaging movements embedded within ongoing activities. The relationship between the constructs can therefore be understood as increasing behavioral specificity, from a general intention to reduce sedentary behavior, through willingness to interrupt a sitting episode and readiness to integrate movement into an ongoing activity, to performance of a defined ZTEx behavior in a particular context [15]. The zero-time exercise–related intention proxy (ZTEx-IP) is positioned at the broader end of this framework because it reflects general willingness and perceived control relevant to changing sedentary routines without specifying ZTEx as the means of change. ZTEx-specific readiness additionally involves beliefs about performing particular movements, knowledge of appropriate techniques, movement-specific confidence and outcome expectations, perceived feasibility and compatibility with ongoing tasks, and workplace appropriateness [16,17].

Because ZTEx is embedded within an ongoing activity, it must be initiated while another goal, such as completing a work task, is already active. This feature guided the selection of exercise motivation, general self-efficacy, and perceived physical activity triggers. Exercise motivation reflects the quality and internalization of the value placed on movement. This motivational quality is relevant when movement competes with the immediate convenience of sitting, habitual behavior, and task demands. Self-determination theory provides a theoretical account of this variation in motivational quality by distinguishing autonomous and controlled forms of regulation [18]. General self-efficacy represents a broader agentic resource for managing obstacles and competing demands. Social cognitive theory supports its inclusion because efficacy beliefs influence goal pursuit, effort, and persistence in the presence of impediments [19]. Perceived physical activity triggers represent the salience of cues and facilitators that make otherwise overlooked movement opportunities noticeable during work. This is particularly relevant when employees remain seated because their attention is absorbed by ongoing tasks. Workplace evidence identifies competing motivations, limited perceived control, task and organizational constraints, and missed or ignored cues as barriers to interrupting sitting. The three variables were therefore selected to represent motivational regulation, generalized perceived agency, and cue salience within the specific demands of lifestyle-integrated movement. They were not intended to capture all determinants of ZTEx adoption.

The capability, opportunity, motivation-behavior (COM-B) model and the Fogg behavior model provide complementary frameworks for organizing these factors at different explanatory levels rather than functioning as competing models. COM-B provides a system-level account in which behavior depends on motivation, capability, and opportunity [20,21]. At a broad conceptual level, exercise motivation relates to motivational resources, general self-efficacy to perceived psychological capability, and perceived physical activity triggers to the salience of cues and facilitators that signal opportunities for action. Nevertheless, the constructs measured in this study are not direct operational equivalents of these framework components. For ZTEx, capability also encompasses technique-specific

XSL•FO
RenderX

knowledge and movement skills, whereas opportunity encompasses task compatibility and workplace norms. The Fogg behavior model provides a more immediate account of behavioral initiation: a prompt is expected to elicit action only when sufficient motivation and ability are present at the same moment [22]. Thus, self-determination theory provides a theoretical account of motivational quality, social cognitive theory supports the role of generalized perceived agency, COM-B situates these factors within broader behavioral conditions, and the Fogg behavior model highlights the conditional role of prompts at the moment of action.

Therefore, this study aimed to examine the association between exercise motivation and the ZTEx-IP among Chinese adults in sedentary occupations and to evaluate model-implied indirect associations involving general self-efficacy and perceived physical activity triggers. We also characterized participants' sedentary behavior and physical activity patterns and examined whether the association between exercise motivation and the ZTEx-IP was graded or nonlinear. We hypothesized that higher exercise motivation would be positively associated with the ZTEx-IP and that the fitted models would yield positive indirect associations involving general self-efficacy and perceived physical activity triggers.

Methods

Design

This cross-sectional online survey examined contemporaneous associations among exercise regulation, general self-efficacy, perceived physical activity triggers, and the ZTEx-IP. Theory-informed parallel and ordered indirect-association analyses were used to quantify model-implied indirect associations. Because all variables were measured at the same time, these analyses were not intended to establish temporal or causal mediation. Reporting followed the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist (Multimedia Appendix 1).

Setting and Sample

Data were collected through an online questionnaire in China from May to September 2025. The study included participants from 20 of China's 34 provincial-level administrative divisions, covering eastern, central, and western regions. Recruitment used a multichannel convenience-sampling strategy. Recruitment information was distributed through workplace WeChat (Tencent Holdings Limited) groups, occupation-related online communities, and workplace or institutional contacts. Individual recruitment source was not recorded; therefore, channel-specific recruitment proportions could not be calculated. Self-reported work-unit information indicated participation from public-service institutions, enterprises, and other workplace settings. Organization size and ownership were not systematically recorded. A standardized recruitment notice described the study purpose, eligibility criteria, voluntary participation, confidentiality, approximate completion time, and the right to withdraw before submission. Contacts only forwarded the notice and questionnaire link and had no access to individual responses. The electronic questionnaire was administered through Wenjuanxing and accessed via a web link or QR code. Participants provided electronic informed consent before entering the survey. A standardized definition of ZTEx was presented before the ZTEx-related items. The inclusion criteria required (1) adults aged ≥18 years; (2) employed full-time in sedentary occupations, operationally defined as jobs predominantly involving seated work for ≥6 hours per working day with low energy expenditure (≤1.5 METs), consistent with the Sedentary Behaviour Research Network definition and commonly used cutoffs in occupational sedentary epidemiology; and (3) able to provide informed consent. Exclusion criteria eliminated individuals on long-term leave, with cognitive impairment affecting comprehension, or participating in concurrent behavioral intervention programs.

Of the 820 submitted questionnaires, 30 were excluded because missing data precluded calculation of at least one core scale score or analysis variable, leaving 790 questionnaires for analysis. Thus, 96.3% (790/820) of submitted questionnaires were included; this percentage represents analytic inclusion rather than a response rate because the number of individuals who received or viewed the recruitment information was unavailable. No item-level imputation was performed. Before data collection, sample-size planning used a pragmatic ratio of 10 to 20 participants per anticipated model parameter. Allowing 20% for unusable questionnaires yielded a recruitment target of 325 to 650 participants. Because sample-size requirements for indirect associations depend on path magnitudes and model characteristics [23,24], this heuristic was not regarded as a model-specific power calculation. A subsequent scenario-based Monte Carlo analysis of sample adequacy is reported in Multimedia Appendix 2.

Data Collection Tools

General Information Questionnaire

The questionnaire was developed based on a literature review and study goals, covering 17 items and related sociodemographic and work characteristics. Variables included sex, age, educational attainment, monthly income, occupation category, job tenure, cumulative sedentary hours per day, musculoskeletal discomfort, recent health examination, chronic disease history, physical activity duration, exercise frequency and duration, use of wearable devices, and familiarity with ZTEx. For descriptive reporting, physical activity was separated based on context and intent. Incidental physical activity was defined as physical activity accumulated during work-related activities, household duties, transportation, and other routine daily activities, other than physical activity undertaken during a planned exercise session. Structured exercise was defined as a planned, intentional, and repetitive form of physical activity that was undertaken specifically with the aim of improving or maintaining health or physical fitness [25]. Frequency and duration were recorded separately for these 2 activity categories. Before answering the familiarity item, participants were shown the standardized definition of ZTEx. Familiarity was assessed with the single self-report item, "How familiar are you with zero-time exercise?" Responses were categorized as "not familiar," "somewhat familiar," or "very familiar" and retained in these categories for descriptive analyses. For the exploratory familiarity-restricted analysis, "somewhat familiar" and "very

XSL•FO
RenderX

familiar" were combined as "at least somewhat familiar," whereas participants who selected "not familiar" were excluded.

Adult Sedentary Behavior Reduction Intention Questionnaire

The Adult Sedentary Behavior Reduction Intention Questionnaire was developed by Ye et al [26] from the theory of planned behavior to measure adults' intention to reduce sedentary behavior. It contains 4 dimensions: behavioral intention, attitude, subjective norms, and perceived behavioral control. The scale has 16 items, each rated on a 5-point Likert scale (1 to 5), with higher values showing greater intention to reduce sedentary behavior. In the absence of a validated ZTEx-specific intention measure, the questionnaire total score was operationalized as the ZTEx-IP. This operationalization was based on the conceptual overlap between reducing sedentary behavior and incorporating brief, routine-integrated movements into otherwise sedentary daily activities. However, the questionnaire assesses general sedentary behavior reduction intention rather than intention to perform ZTEx specifically; therefore, the outcome should be interpreted as a proxy indicator rather than a direct measure of ZTEx-specific intention. The validation study reported excellent internal consistency (Cronbach α=0.967) and strong construct validity, with subscale α ranging from 0.911 to 0.950. In this study, Cronbach α for this scale was 0.953.

Behavioral Regulation in Exercise Questionnaire

The Behavioral Regulation in Exercise Questionnaire (BREQ-2) was developed by Markland and Tobin in 2004 to assess exercise motivation within self-determination theory [27]. It includes 5 dimensions: amotivation, external regulation, introjected regulation, identified regulation, and intrinsic motivation. The Chinese version contains 18 items rated on a 5-point Likert scale from 0 to 4. In this study, amotivation and external regulation items were reverse scored before all items were summed to create a composite score; higher scores represented a more self-determined motivational profile. The BREQ-2 captures relatively stable forms of behavioral regulation and motivational internalization rather than the momentary intensity of motivation immediately before an action, as conceptualized in the Fogg behavior model. Liu et al [28] translated and validated the Chinese version in 2015, reporting Cronbach α values ranging from 0.72 to 0.85 and good construct validity. Cronbach α was 0.882 in this study.

General Self-Efficacy Scale

General Self-Efficacy Scale (GSES) was developed by Schwarzer and Jerusalem in 1995 to assess individuals' confidence in coping with challenges and controlling demanding situations [29]. The unidimensional scale contains 10 items rated from 1 to 4, with higher scores indicating greater general self-efficacy. Wang et al [30] translated and validated the Chinese version in 2001 [30], reporting Cronbach α=0.87, test-retest reliability r=0.83, and split-half reliability r=0.82. Cronbach α was 0.938 in this study. The GSES assesses generalized rather than ZTEx-specific self-efficacy and was therefore not treated as a direct operationalization of behavior-specific capability in the COM-B model. Instead, it was used as a theoretically relevant measure of generalized perceived agency within the proposed psychosocial model.

Physical Activity Triggers Questionnaire

The Physical Activity Triggers Questionnaire (PATQ) was developed by Wang and Kang in 2022 to assess physical activity-related triggers [31]. It consists of 3 dimensions: spark, signal, and facilitator. The PATQ has 14 items measured on a 5-point Likert scale (1=strongly disagree, 5=strongly agree). The PATQ captures perceived exposure or receptivity to physical activity cues in daily life rather than brief prompts occurring immediately before a specific action. It was therefore used to examine cue-related associations with the other study variables. The original validation among Chinese university students reported strong internal consistency (Cronbach α=0.925), split-half reliability of 0.821, test-retest reliability of 0.860, and acceptable construct validity. Cronbach α was 0.934 in this study.

Data Analysis

All statistical analyses were conducted using R (version 4.5.1; R Core Team) and IBM SPSS Statistics version 29.0 with the PROCESS macro. Scale-score distributions were assessed using the Shapiro-Wilk test. Categorical variables were summarized as frequencies and percentages, and scale scores as medians and IQRs (25th-75th percentiles). Internal consistency was evaluated using Cronbach α. Differences in the ZTEx-IP across participant characteristics were examined using Mann-Whitney U tests or Kruskal-Wallis H tests, as appropriate. The Benjamini-Hochberg procedure was applied to the 15 comparisons between-group comparisons; unadjusted and adjusted P values are reported in Multimedia Appendix 3. Spearman rank correlations were used to examine associations among the 4 focal scores. Harman's unrotated single-factor analysis of the 58 core questionnaire items was used as a limited diagnostic for common method variance.

The adjustment set was defined using a combination of substantive considerations and nominal empirical screening. Candidate variables represented 4 relevant domains. Age and monthly income represented sociodemographic context; incidental physical activity duration and structured exercise frequency and duration represented existing movement patterns; wearable-device use represented potential exposure to activity monitoring and digital cues; and ZTEx familiarity represented prior awareness of the target strategy. The socioeconomic and activity-related rationale was informed by previous literature [32], whereas wearable-device use and ZTEx familiarity were included on substantive grounds because cue exposure and prior knowledge could be related to the focal study variables. To obtain a parsimonious adjustment set for this exploratory analysis, candidate variables meeting the nominal significance criterion (α=.05) in the initial unadjusted comparisons were retained. The resulting set comprised age, monthly income, incidental physical activity duration, structured exercise frequency, structured exercise duration, wearable-device use, and ZTEx familiarity. Monthly income met the nominal screening criterion before multiplicity adjustment (unadjusted P=.04), although the Benjamini-Hochberg–adjusted comparison was not statistically significant (adjusted P=.07). The

XSL•FO
RenderX

Benjamini-Hochberg procedure was used only for descriptive multiplicity control and did not redefine the adjustment set. The selected variables were treated as adjustment covariates rather than empirically established confounders. Age, income, incidental physical activity duration, structured exercise frequency and duration, and ZTEx familiarity were entered as ordinal category scores, whereas wearable-device use was entered as a binary indicator. Ordinary least squares regression was used to examine the association between exercise motivation and the ZTEx-IP. Motivation was modeled using sample-specific tertiles, with the lowest tertile as the reference and an ordinal term to test for trend. The continuous association was examined using a restricted cubic spline with knots at motivation scores of 33, 48, 54, and 72 and a reference score of 51; overall and nonlinear associations were tested.

Model-implied indirect associations were estimated using PROCESS Models 4 and 6. Model 4 specified general self-efficacy and perceived physical activity triggers as parallel intermediate variables, whereas Model 6 imposed the prespecified theory-informed ordering of exercise motivation, general self-efficacy, perceived physical activity triggers, and the ZTEx-IP to estimate the corresponding model-implied indirect association. The same covariates were included in every component equation. Direct, specific indirect, total indirect, and total associations were estimated using 5000 nonparametric percentile bootstrap resamples. A 95% bootstrap CI excluding 0 was considered statistical evidence of an indirect association. Unstandardized coefficients were the primary estimates, with standardized coefficients additionally reported for the focal associations. Model R^2 and overall F statistics were reported. The incremental contribution of each intermediate variable was assessed using ΔR^2 and Cohen f^2. Because all variables were measured concurrently, the indirect associations were interpreted as statistical decompositions of cross-sectional associations rather than evidence of temporal or causal mediation. Model assumptions were assessed using variance inflation factors and residual diagnostics, with HC3 heteroscedasticity-consistent SEs examined as a robustness analysis. Three theoretically motivated sensitivity analyses were prioritized: alternative outcome definitions using the behavioral intention subscale and a score excluding perceived behavioral control items, restriction to participants who were at least somewhat familiar with ZTEx, and an alternative ordering of general self-efficacy and perceived physical activity triggers. Additional regression diagnostics and technical robustness analyses are reported in Multimedia Appendices 4 and 5. Bootstrap-based sensitivity analyses used 5000 resamples. All sensitivity analyses were considered exploratory, and no additional multiplicity adjustment was applied.

Ethical Considerations

The study was conducted in accordance with the Declaration of Helsinki, and ethical approval was granted by the Ethics Committee of Hunan Normal University (2025889). Before beginning the online questionnaire, participants received information about the study purpose, voluntary participation, data confidentiality, and their right to decline participation or withdraw before submission without penalty. All participants provided electronic informed consent before proceeding to the survey items. Participants received no financial or material compensation for participation. To protect participant privacy and confidentiality, the analytic dataset was deidentified before analysis by removing information that could identify individual participants or participating institutions. Results were reported in aggregate, data were used solely for the purposes of this study, and access was restricted to the research team.

Results

Descriptive Statistics

Of the 820 online questionnaire submissions received, 790 met the completeness requirements described in the “Methods” section and were included in the analyses, corresponding to an analytic inclusion proportion of 96.3% (790/820). Table 1 summarizes participants’ sociodemographic, occupational, and behavioral characteristics and their median ZTEx-IP scores. The sample was relatively balanced by sex, with female participants accounting for 56.6% (447/790), and 75.4% (596/790) of participants were younger than 40 years. Education and research professionals formed the largest occupational group (297/790, 37.6%). Most participants reported middle-income levels and had completed higher education.

XSL•FO
RenderX

Table 1. Demographic, occupational, and behavioral characteristics and ZTEx-IP[a] scores among participants (n=790)[b].

Variables	n (%)	Median (IQR)	BH[c]-adjusted *P*
Sex			.80
Male	343 (43.4)	58 (48-67)	
Female	447 (56.6)	59 (51-64)	
Age (y)			.04
<30	324 (41)	58 (49-64)	
30-40	272 (34.4)	59 (50-66)	
41-50	128 (16.2)	62 (52-69)	
51-60	52 (6.6)	61 (50-65)	
>60	14 (1.8)	52 (48-63)	
Educational attainment			.38
High school diploma or below	272 (34.4)	58 (49-66)	
Undergraduate degree	400 (50.6)	60 (50-65)	
Master's degree or above	118 (14.9)	60 (53-66)	
Average monthly income (RMB)[d]			.07
<2000	98 (12.4)	55 (48-64)	
2000-5000	384 (48.6)	59 (50-65)	
5000-10,000	247 (31.3)	59 (51-66)	
>10,000	61 (7.7)	61 (52-70)	
Occupation category			.14
Office-based workers	161 (20.4)	57 (48-64)	
Education and research professionals	297 (37.6)	60 (51-66)	
IT and technical professionals	220 (27.8)	60 (53-67)	
Drivers and transport personnel	112 (14.2)	58 (48-65)	
Job tenure (y)			.20
<5	283 (35.8)	58 (49-64)	
5-15	289 (36.6)	59 (51-66)	
15-25	121 (15.3)	58 (48-68)	
>25	97 (12.3)	63 (53-68)	
Cumulative daily sedentary duration (h)			.37
<8	555 (70.3)	60 (51-66)	
8-10	166 (21)	58 (49-64)	
>10	69 (8.7)	56 (48-67)	
Musculoskeletal discomfort			.37
No	137 (17.3)	57 (48-67)	
Mild (occasional soreness)	535 (67.7)	60 (51-65)	
Moderate (affecting work)	93 (11.8)	58 (48-66)	
Severe (requiring treatment)	25 (3.2)	55 (47-65)	
Health examination within the past 6 months			.32
Yes	466 (59)	59 (50-67)	
No	324 (41)	59 (51-64)	
Chronic disease			.25

XSL•FO
RenderX

Variables	n (%)	Median (IQR)	BH[c]-adjusted *P*
Yes	217 (27.5)	60 (52-67)	
No	573 (72.5)	59 (50-65)	
Cumulative duration of daily incidental physical activity (min)			.001
<30	190 (24.1)	55 (48-64)	
30-60	328 (41.5)	59 (52-65)	
60-90	114 (14.4)	62 (51-69)	
>90	158 (20)	60 (52-69)	
Frequency of structured exercise (sessions/wk)			<.001
0	255 (32.3)	53 (48-63)	
1-2	366 (46.3)	60 (53-66)	
3-4	117 (14.8)	62 (53-69)	
≥5	52 (6.6)	68 (60-78)	
Average duration per structured exercise session (min)			<.001
<10	246 (31.1)	53 (48-63)	
10-30	282 (35.7)	59 (52-65)	
30-60	180 (22.8)	62 (56-69)	
>60	82 (10.4)	64 (49-73)	
Use of wearable health-tracking devices			.001
Yes	263 (33.3)	61 (53-67)	
No	527 (66.7)	58 (48-65)	
Familiarity with ZTEx[e]			.04
Not familiar	603 (76.3)	58 (49-65)	
Somewhat familiar	154 (19.5)	61 (53-67)	
Very familiar	33 (4.2)	59 (53-80)	

[a]ZTEx-IP: zero-time exercise–related intention proxy.

[b]Incidental physical activity refers to movement accumulated through occupational duties, household tasks, transportation, and other routine daily activities outside planned exercise sessions. Structured exercise refers to planned, intentional, and repetitive physical activity undertaken to improve or maintain health or physical fitness. The presented *P* values were adjusted using the Benjamini-Hochberg procedure. Detailed test statistics, unadjusted *P* values, and Benjamini-Hochberg–adjusted results are provided in Multimedia Appendix 3.

[c]Benjamini-Hochberg.

[d]The 2025 average exchange rate was US $1=RMB 7.1875.

[e]ZTEx: zero-time exercise.

The median ZTEx-IP score was 59 (IQR 50-65.75). After Benjamini-Hochberg adjustment, differences remained statistically significant across age groups (adjusted *P*=.04), incidental physical activity duration (adjusted *P*=.001), structured exercise frequency (adjusted *P*<.001), structured exercise duration (adjusted *P*<.001), wearable-device use (adjusted *P*=.001), and ZTEx familiarity (adjusted *P*=.04). No statistically significant differences were observed by sex, educational attainment, occupation, job tenure, daily sedentary duration, musculoskeletal discomfort, health examination in the previous 6 months, or chronic disease status (all adjusted *P*>.05).

Correlation Analysis

Spearman rank correlations were used to examine associations among exercise motivation, general self-efficacy, perceived physical activity triggers, and the ZTEx-IP (Table 2). All correlations were positive and statistically significant (all *P*<.001). Exercise motivation was positively correlated with general self-efficacy (ρ=0.347), perceived physical activity triggers (ρ=0.539), and the ZTEx-IP (ρ=0.371). General self-efficacy was positively correlated with perceived physical activity triggers (ρ=0.349) and the ZTEx-IP (ρ=0.378). Perceived physical activity triggers showed the strongest correlation with the ZTEx-IP (ρ=0.420).

XSL•FO
RenderX

Table 2. Spearman rank correlations among exercise motivation, general self-efficacy, perceived physical activity triggers, and the ZTEx-IP[a] (n=790)[b].

Variable	Median (IQR)	1	2	3	4
Exercise motivation	51 (44-56)	—[c]			
General self-efficacy	29 (26-30)	.347[d]	—		
Perceived physical activity triggers	36 (28-43)	.539[d]	.349[d]	—	
ZTEx-IP	59 (50-65.75)	.371[d]	.378[d]	.420[d]	—

[a]ZTEx-IP: zero-time exercise–related intention proxy.

[b]Scale scores are presented as median (IQR). Values in the lower triangle are Spearman rank correlation coefficients (ρ). The ZTEx-IP was operationalized using the total score of the Adult Sedentary Behavior Reduction Intention Questionnaire.

[c]Self-correlations, which were not estimated.

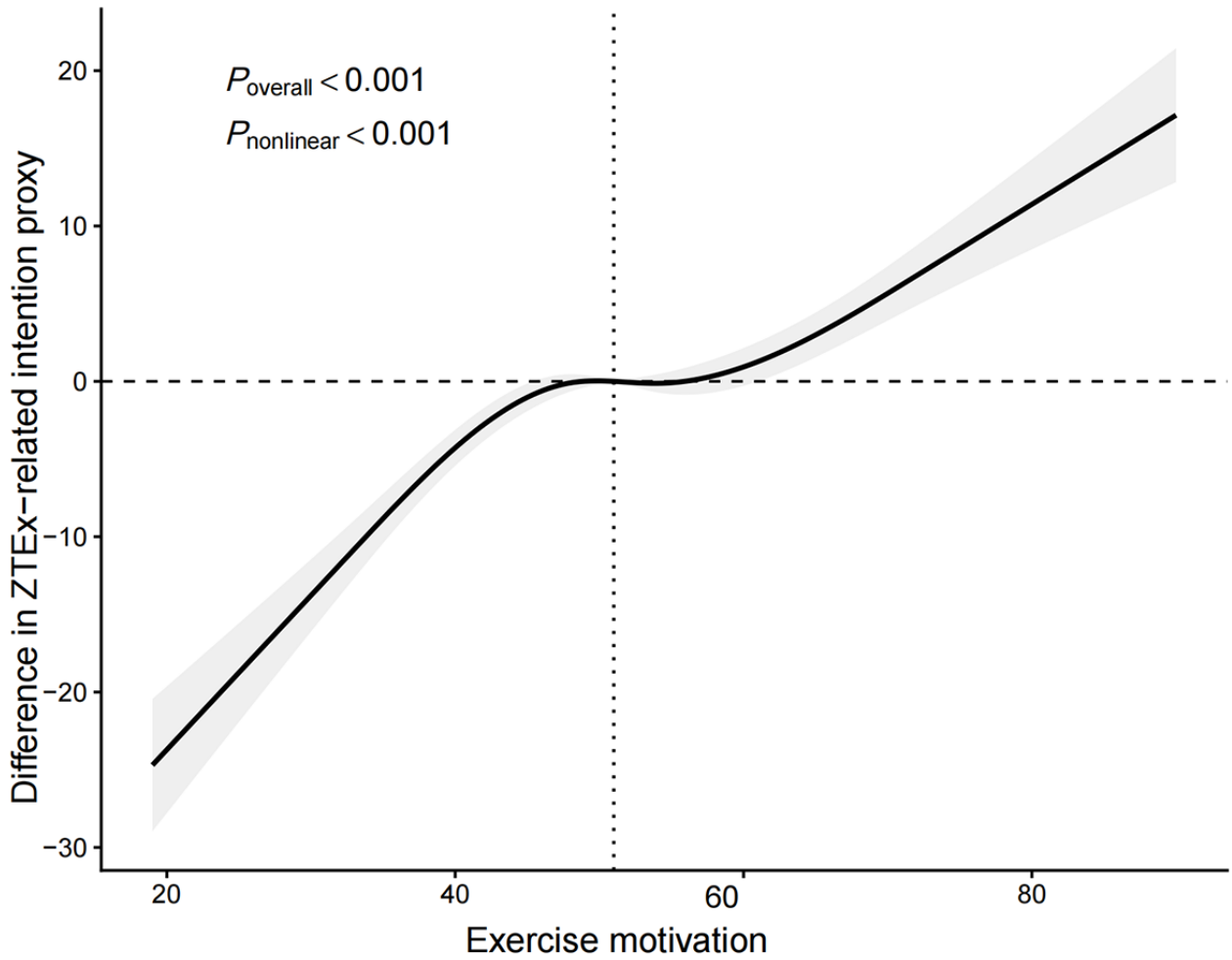
[d]$P<.001$.

Association Between Exercise Motivation and ZTEx-IP

Participants were categorized into tertiles according to their exercise motivation scores. In the unadjusted model, compared with participants in the lowest motivation tertile, those in the middle and highest tertiles had higher ZTEx-IP scores, with B coefficients of 5.58 (95% CI 3.60-7.56; $P<.001$) and 10.60 (95% CI 8.63-12.60; $P<.001$), respectively. After adjustment for age, income, incidental physical activity duration, structured exercise frequency, structured exercise duration, wearable-device use, and ZTEx familiarity, the corresponding B coefficients were 4.64 (95% CI 2.69-6.60; $P<.001$) and 8.49 (95% CI 6.43-10.60; $P<.001$), respectively (Multimedia Appendix 6). A significant positive trend was observed across increasing motivation tertiles (adjusted B for trend=4.25, 95% CI 3.22-5.28; $P<.001$). When exercise motivation was analyzed as a continuous variable, restricted cubic spline analysis showed a significant overall association with the ZTEx-IP (P overall<.001), with evidence of nonlinearity (P nonlinear<.001; Figure 1). Using an exercise motivation score of 51 as the reference value, the spline curve showed an overall increasing pattern, although the magnitude of the association varied across the motivation score range.

XSL•FO
RenderX

Figure 1. Restricted cubic spline analysis of the association between exercise motivation and the zero-time exercise–related intention proxy (ZTEx-IP) among Chinese adults in sedentary occupations. The solid line shows the multivariable-adjusted difference in the ZTEx-IP, and the shaded band shows the 95% CI. The vertical dotted line marks the reference motivation score of 51, at which the adjusted difference was set to 0; the horizontal dashed line indicates no difference from that reference. Negative values indicate lower adjusted ZTEx-IP scores relative to the reference and do not imply an inverse overall association. The model was adjusted for age, income, incidental physical activity duration, structured exercise frequency, structured exercise duration, wearable-device use, and zero-time exercise familiarity. Both the overall and nonlinear associations were statistically significant (*P*<.001 for each). ZTEx: zero-time exercise.



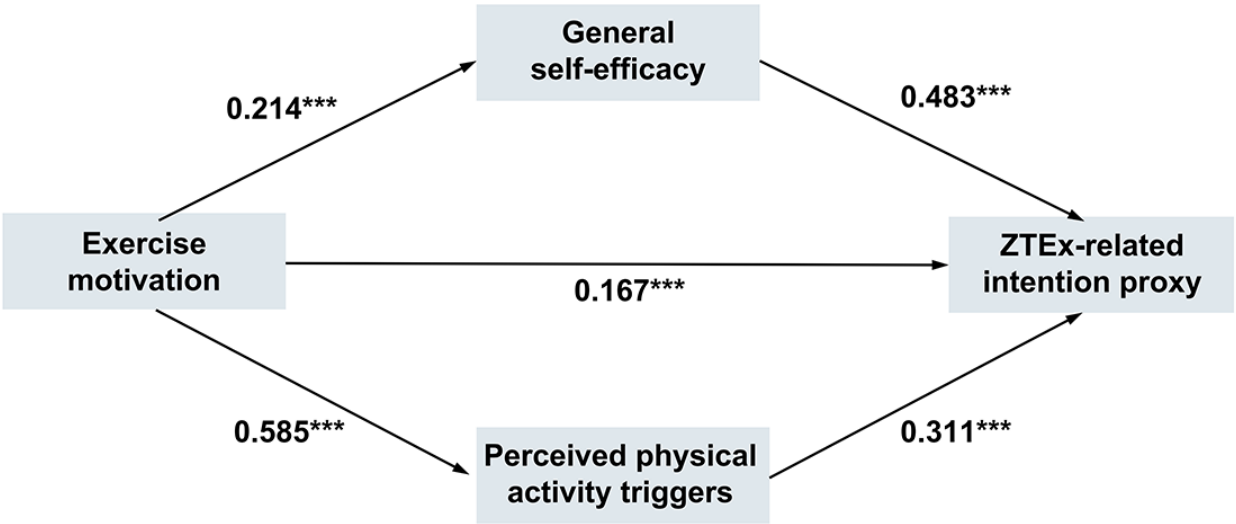
Parallel Indirect-Association Analysis (PROCESS Model 4)

Before the indirect-association analyses, common method variance was assessed using Harman's single-factor test. All 58 core items from the BREQ-2, GSES, PATQ, and Adult Sedentary Behavior Reduction Intention Questionnaire were included. The first unrotated factor accounted for 31.93% of the total variance. This result did not indicate a predominant single factor, although the test cannot rule out common method bias. Full results are reported in Multimedia Appendix 7.

A parallel indirect-association model was fitted to estimate model-implied indirect associations involving general self-efficacy and perceived physical activity triggers. The model was adjusted for age, income, incidental physical activity duration, structured exercise frequency, structured exercise duration, wearable-device use, and ZTEx familiarity. The adjusted parallel indirect-association model is illustrated in Figure 2. Exercise motivation was positively associated with both general self-efficacy (B=0.214, SE 0.015, 95% CI 0.184-0.245; *P*<.001) and perceived physical activity triggers (B=0.585, SE 0.031, 95% CI 0.524-0.645; *P*<.001). In the outcome model, both general self-efficacy (B=0.483, SE 0.075, 95% CI 0.336-0.631; *P*<.001) and perceived physical activity triggers (B=0.311, SE 0.038, 95% CI 0.237-0.386; *P*<.001) were positively associated with the ZTEx-IP after adjustment for exercise motivation and covariates. The direct association between exercise motivation and the ZTEx-IP remained significant (B=0.167, SE 0.04, 95% CI 0.088-0.246; *P*<.001). Bootstrap analysis with 5000 resamples yielded nonzero indirect-association estimates involving general self-efficacy (B=0.104, bootstrap SE (BootSE)=0.026, 95% CI 0.056-0.158) and perceived physical activity triggers (B=0.182, BootSE=0.026, 95% CI 0.129-0.233). The total indirect association was also significant (B=0.286, BootSE=0.034, 95% CI 0.220-0.353). The adjusted outcome model explained 37.6% of the variance in the ZTEx-IP. Detailed adjusted Model 4 results are provided in Multimedia Appendix 8.

XSL•FO
RenderX

Figure 2. Adjusted parallel indirect-association model involving exercise motivation, general self-efficacy, perceived physical activity triggers, and the zero-time exercise–related intention proxy. Values are unstandardized regression coefficients. The model was adjusted for age, income, wearable-device use, incidental physical activity duration, structured exercise frequency, structured exercise duration, and zero-time exercise familiarity. ZTEx: zero-time exercise. ***P<.001.



Ordered Indirect-Association Analysis (PROCESS Model 6)

PROCESS Model 6 was fitted to estimate model-implied indirect associations under the prespecified theory-informed ordering of exercise motivation, general self-efficacy, perceived physical activity triggers, and the ZTEx-IP. As shown in Table 3, exercise motivation was positively associated with both general self-efficacy (standardized regression coefficient, β=0.460, P<.001) and perceived physical activity triggers (β=0.487, P<.001), and general self-efficacy was positively associated with perceived physical activity triggers (β=0.170, P<.001). These coefficients were compatible with the prespecified statistical ordering, but they do not establish temporal precedence among the constructs. In the full model predicting the ZTEx-IP, exercise motivation (β=0.161, P<.001), general self-efficacy (β=0.216, P<.001), and perceived physical activity triggers (β=0.310, P<.001) all remained statistically significant. The adjusted outcome model explained 37.6% of the variance in the ZTEx-IP (R^2=0.376, $F_{10,779}$=46.995; P<.001). Compared with reduced models excluding each intermediate variable separately, general self-efficacy contributed an additional 3.3% of explained variance (ΔR^2=0.033, Cohen f^2=0.053), whereas perceived physical activity triggers contributed an additional 5.4% (ΔR^2=0.054, Cohen f^2=0.087). Complete model diagnostics and secondary technical robustness analyses are reported in Multimedia Appendix 4.

Table 3. Component equations for the prespecified ordered model (n=790)[a].

Regression		Model index		Coefficients	
Dependent variable	Independent variable	R^2	F test (df)	β	t test (df)
General self-efficacy	Exercise motivation	0.269	35.997 (8, 781)[b]	0.460	13.876 (781)[b]
Perceived physical activity triggers	Exercise motivation	0.436	66.928 (9, 780)[b]	0.487	14.960 (780)[b]
Perceived physical activity triggers	General self-efficacy	—[c]	—	0.170	5.404 (780)[b]
ZTEx-IP[d]	Exercise motivation	0.376	46.995 (10, 779)[b]	0.161	4.150 (779)[b]
ZTEx-IP	General self-efficacy	—	—	0.216	6.428 (779)[b]
ZTEx-IP	Perceived physical activity triggers	—	—	0.310	8.239 (779)[b]

[a]All 3 equations were adjusted for age, income, incidental physical activity duration, structured exercise frequency, structured exercise duration, wearable-device use, and ZTEx familiarity. β denotes a standardized coefficient. The focal continuous variables were analyzed using their original scale scores and standardized for presentation in this table. Ordered categorical covariates were entered as ordinal category scores, and wearable-device use was entered as a binary indicator.

[b]P<.001.

[c]Not applicable.

[d]ZTEx-IP: zero-time exercise–related intention proxy.

XSL•FO
RenderX

Bootstrap Estimation of the Ordered Indirect-Association Model

A percentile bootstrap analysis with 5000 samples was conducted to evaluate model-implied indirect associations under the prespecified ordering. As shown in Table 4, the total indirect association between exercise motivation and the ZTEx-IP was B=0.286 (BootSE=0.034, 95% CI 0.218-0.352). The three specific estimates involved (1) general self-efficacy alone (B=0.104, BootSE=0.025, 95% CI 0.057-0.155), (2) perceived physical activity triggers alone (B=0.157, BootSE=0.024, 95% CI 0.110-0.205), and (3) the prespecified ordered combination of general self-efficacy and perceived physical activity triggers (B=0.025, BootSE=0.007, 95% CI 0.013-0.039). All bootstrap CIs excluded 0. The direct association between exercise motivation and the ZTEx-IP also remained statistically significant (B=0.167, SE=0.040, 95% CI 0.088-0.246). Both the direct association and the model-implied indirect association estimates therefore differed from 0. The observed statistical pattern was consistent with the theory-informed ordering used in PROCESS Model 6. However, the cross-sectional estimates cannot determine whether self-efficacy precedes trigger perception, whether trigger perception enhances self-efficacy, or whether both are influenced by another unmeasured factor. Figure 3 presents standardized coefficients for the prespecified ordered model; its arrows indicate model specification and not temporal sequence.

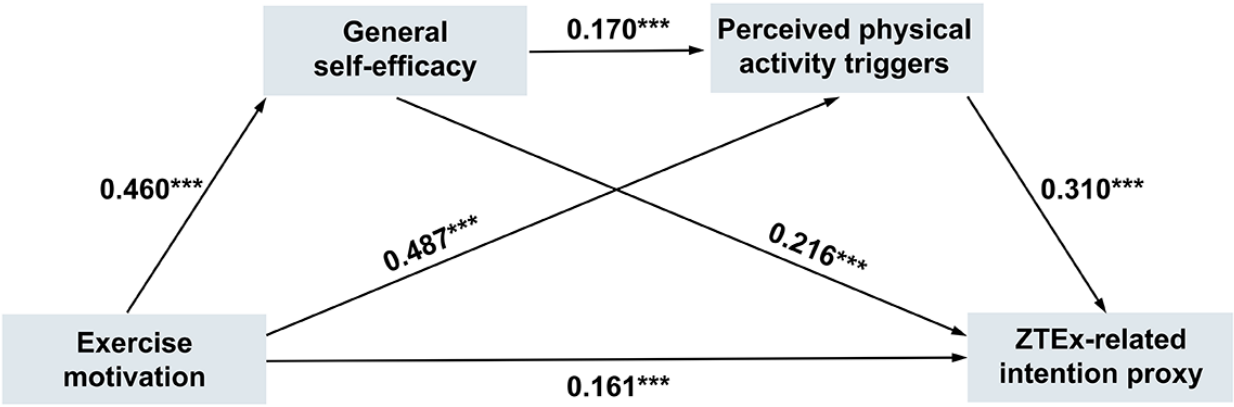
Table 4. Total, direct, and indirect associations between exercise motivation and the ZTEx-IP[a] involving general self-efficacy and perceived physical activity triggers (n=790)[b].

Association component	B (SE)	95% CI
Total association	0.453 (0.034)	0.385-0.520
Direct association	0.167 (0.040)	0.088-0.246
Total indirect association	0.286 (0.034)	0.218-0.352
X → M1 → Y	0.104 (0.025)	0.057-0.155
X → M2 → Y	0.157 (0.024)	0.110-0.205
X → M1 → M2 → Y	0.025 (0.007)	0.013-0.039

[a]ZTEx-IP: zero-time exercise–related intention proxy.

[b]X: exercise motivation; Y: ZTEx-IP; M1: general self-efficacy; M2: perceived physical activity triggers; B: unstandardized coefficient. Indirect associations were estimated using 5000 bootstrap resamples and percentile 95% CIs. All equations were adjusted for age, income, incidental physical activity duration, structured exercise frequency, structured exercise duration, wearable-device use, and zero-time exercise familiarity.

Figure 3. Standardized coefficient diagram for the prespecified ordered model involving exercise motivation, general self-efficacy, perceived physical activity triggers, and the zero-time exercise–related intention proxy. Coefficients were adjusted for age, income, incidental physical activity duration, structured exercise frequency, structured exercise duration, wearable-device use, and zero-time exercise familiarity. Arrows indicate the imposed model specification and do not establish temporal sequence. ZTEx: zero-time exercise. ***P<.001.



Theoretically Relevant Sensitivity Analyses

Analyses using the behavioral intention subscale or an outcome excluding perceived behavioral control items yielded positive direct and model-implied indirect associations, with the corresponding 95% CIs excluding 0. In the subgroup of participants who were at least somewhat familiar with ZTEx (n=187), the bootstrap CIs for both single-intermediate-variable associations and the total indirect association excluded 0, whereas the CIs for the direct and ordered indirect-association estimates included 0. An alternative variable ordering also yielded a nonzero ordered indirect-association estimate, reinforcing that the cross-sectional data could not identify a unique temporal order among the constructs. Complete numerical results and additional technical robustness analyses are reported in Multimedia Appendices 4 and 5.

XSL•FO
RenderX

Discussion

Principal Findings

In this study of 790 Chinese adults in sedentary occupations, higher exercise motivation was associated with higher ZTEx-IP scores. The association was graded and nonlinear across the observed motivation range. In both the parallel and ordered models, general self-efficacy and perceived physical activity triggers showed distinct model-implied indirect associations between exercise motivation and the ZTEx-IP. Perceived triggers had the larger specific indirect estimate and made the greater incremental contribution to explained variance, while the prespecified ordered term involving both intermediate variables yielded a smaller indirect estimate. Together, the adjusted variables explained 37.6% of the variance in the ZTEx-IP. The theoretically motivated sensitivity analyses retained the overall positive association pattern, although estimates in the ZTEx-familiar subgroup were less precise, and the alternative ordering confirmed that the cross-sectional data did not identify a unique temporal sequence. By considering these constructs in one model, the study extends prior work beyond isolated psychosocial correlates and highlights the joint relevance of motivational orientation, perceived agency, and cue-supportive contexts to readiness for brief, lifestyle-integrated movement.

Interpretation, Comparison With Prior Work, and Implications

The graded association between exercise motivation and the ZTEx-IP is most plausibly interpreted as a relationship between a more self-determined regulatory profile and readiness to reorganize sedentary routines. The BREQ-2 assesses exercise regulation along an internalization continuum, capturing the quality and self-endorsement of motivation rather than the intensity of a momentary impulse [27,28]. This interpretation is consistent with evidence that perceived competence and more autonomous forms of regulation are recurrent correlates of physical activity [12]. The Chinese working-age context makes the finding particularly relevant. Half of Chinese adults aged 18-44 years report sitting for at least 6 hours per day [5], and longer sitting duration has been associated with a higher risk of major cardiovascular events among both manual and nonmanual workers [33]. For workers who find it difficult to set aside dedicated exercise time, an internalized value placed on movement may make brief bouts incorporated into existing routines more acceptable. The nonlinear pattern adds nuance but does not identify a threshold or plateau. Instead, it suggests that the strength of the motivation–ZTEx-IP association varies across the observed motivation range, providing a reason to test support tailored to workers' motivational profiles without imposing arbitrary cutoffs. The 37.6% of variance explained by the adjusted model falls within the broad range reported for theory-based models of behavioral and physical activity intentions [34]. Direct comparison should nevertheless be cautious because the present outcome includes several Theory of Planned Behavior domains. The unexplained variance may reflect factors not captured in the model, including job control, workload, habit, time availability, organizational norms, and access to feasible opportunities for movement [35,36].

The model-implied indirect associations show that exercise motivation, generalized perceived agency, and recognizable movement opportunities covaried within the prespecified statistical model. General self-efficacy reflects confidence in handling challenges across situations rather than efficacy for a particular ZTEx task [30]. From an affordance perspective, workers reporting greater agency may also be more likely to regard a task transition, reminder, or scheduled break as a usable opportunity for movement [37]. This is one theoretically plausible interpretation, not evidence that self-efficacy temporally precedes trigger perception. The larger trigger-related estimate is also plausible for a low-threshold strategy such as ZTEx, which requires little time, space, equipment, or technical skill. A cue is useful only when movement is feasible within the task and acceptable in the workplace. This point is especially relevant in Chinese workplaces, where time pressure and limited social support have been reported as barriers to sustained activity, while locally tested programs increasingly combine digital delivery with managerial, team-based, or interactive support [38-41]. Perceived triggers may therefore reflect both cue exposure and the perceived legitimacy of responding to a cue. The co-occurrence of motivation, perceived agency, and salient opportunity is compatible with COM-B and the Fogg behavior model, but neither framework determines temporal order in these data. Trigger perception might enhance self-efficacy; both constructs might reflect unmeasured influences such as job control or workplace culture, or reciprocal relations might operate. The observed statistical pattern was consistent with the theory-informed ordering used in PROCESS Model 6; however, longitudinal or intensive repeated-measures research is required to determine whether and how these constructs develop over time.

The relative estimates identify perceived trigger exposure, exercise motivation, and general self-efficacy as constructs for hypothesis-driven intervention research rather than as intervention components ready for implementation. In the prespecified ordered model, the specific indirect estimate involving perceived triggers exceeded that involving general self-efficacy (B=0.157 vs B=0.104), and triggers made the larger incremental contribution to explained variance (ΔR^2=0.054 vs 0.033). Because these observational estimates relate perceived triggers to the ZTEx-IP, they do not demonstrate that delivered prompts or changes in workplace practices affect actual ZTEx behavior. A primary hypothesis for prospective evaluation is that context-sensitive prompts may help translate intention into objectively measured sitting interruptions and ZTEx performance. Future trials could also examine whether any such effect varies according to baseline motivation, self-efficacy, or occupational context. Evidence that workplace prompts have modest and heterogeneous effects [42-44] and that the effects of team communication vary according to message type and frequency [38] underscores the need to evaluate, rather than assume, the added value of particular delivery or organizational components. The present data cannot determine which components, delivery modes, or combinations are effective. Future workplace trials should prespecify occupation-specific

XSL•FO
RenderX

feasibility and safety criteria and assess objectively measured sitting interruptions, actual ZTEx frequency, adherence over time, and acceptability to workers and employers rather than intention alone [45].

Limitations

This study had several strengths, including a relatively large and geographically diverse sample, standardized Chinese-language instruments with satisfactory internal consistency in the present sample, and extensive sensitivity analyses. Several limitations should nevertheless be considered. First, all variables were measured at one time point, so temporal precedence and causal relationships among exercise motivation, general self-efficacy, perceived physical activity triggers, and the ZTEx-IP could not be established. The model-implied indirect associations should therefore be understood as theory-informed decompositions of contemporaneous associations, not evidence of causal mediation. Reverse or bidirectional associations and alternative variable orderings remain possible [46-48]. Although the candidate variables were substantively motivated, the final adjustment set partly relied on nominal associations with the ZTEx-IP and was not derived from a prespecified causal graph. Outcome-based screening cannot identify confounding and may omit relevant variables with weak unadjusted associations. Residual confounding by workload, job control, workplace culture, or other excluded or unmeasured factors therefore remains possible. Second, the Adult Sedentary Behavior Reduction Intention Questionnaire has not been validated as a ZTEx-specific intention measure. It captures general willingness and perceived control regarding sedentary behavior reduction rather than intention to perform a specified ZTEx movement, and it does not directly assess ZTEx-specific knowledge, task feasibility, workplace appropriateness, or action planning [26]. This limitation is particularly important because 76.3% (603/790) of participants were unfamiliar with ZTEx and may have responded with reference to standing, walking, or other ways of interrupting sitting. Familiarity was measured with a single self-report item, and neither the standardized definition nor the sensitivity analyses could confirm ZTEx-specific understanding or construct validity. The ZTEx-IP should therefore be interpreted as a broad indicator of intention relevant to ZTEx adoption rather than a direct measure of ZTEx-specific intention. Future studies should develop and validate ZTEx-specific intention and readiness instruments. Third, although the instruments were reliable and theoretically relevant, they were not one-to-one operationalizations of COM-B or Fogg constructs. The BREQ-2 assesses the regulation and internalization of exercise motivation, the GSES assesses generalized perceived agency rather than confidence in performing ZTEx under specific work conditions, and the PATQ assesses perceived exposure to broad activity-related cues rather than prompts immediately preceding behavior [31]. Their differing behavioral and temporal referents further limit their correspondence with the proposed frameworks. Future research should use behavior-specific measures with aligned behavioral and temporal referents. Fourth, all principal variables were self-reported by the same participants at the same time. Common method variance, recall error, social desirability, and acquiescent responding may therefore have contributed to the observed associations. Although the first factor accounted for only 31.93% of the variance in Harman's single-factor analysis, this diagnostic cannot exclude common method bias [49]. Temporally separated or objective assessments would provide stronger protection against this bias. Fifth, the study assessed an intention proxy rather than actual ZTEx uptake, the frequency of sitting interruptions, or behavioral maintenance. Higher ZTEx-IP scores therefore cannot be assumed to translate into ZTEx adoption or sustained practice [45]. Future studies should combine ZTEx-specific intention measures with objective sitting assessments, ZTEx performance records, and longitudinal follow-up. Finally, multichannel online convenience sampling did not provide a population-based sampling frame or a response-rate denominator, and the recruitment source was not recorded at the individual level. Individuals with greater interest in health, exercise, wearable devices, or digital health, as well as those more active in workplace WeChat groups or occupation-related online communities, may have been more likely to participate. This self-selection may have yielded more favorable psychosocial score profiles and may limit the transportability of the observed score distributions, reliability estimates, and associations. Individuals with limited digital access or less engagement with online workplace networks may also have been underrepresented. Generalizability is further limited by the inclusion of only 14 adults aged 60 years or older and by differences in occupational constraints on sitting interruption, particularly for drivers. Nonsignificant subgroup differences should not be interpreted as evidence of equivalence because some estimates were imprecise. Occupation-specific heterogeneity should be examined in adequately represented, independently recruited samples.

Conclusions

Among Chinese adults in sedentary occupations, higher exercise motivation, general self-efficacy, and more frequent perceived physical activity triggers were jointly associated with higher scores on the ZTEx-IP. Perceived triggers had the largest specific indirect estimate and incremental contribution, but the findings do not establish a unique causal or temporal sequence or demonstrate that modifying these factors would increase actual ZTEx behavior. More broadly, intention relevant to ZTEx adoption may reflect the alignment of motivational orientation, perceived agency, and recognizable opportunities within everyday work routines rather than motivation alone. By integrating these theoretically distinct factors within an occupational context, this study identifies an empirically supported pattern that can inform the formulation and prioritization of hypotheses for prospective intervention research. Specifically, the findings support testing whether multicomponent strategies combining autonomy-supportive motivational content, graded movement options, context-sensitive prompts, and appropriate organizational support can facilitate the translation of intention into behavior. Workplace trials should determine whether these components affect objectively measured sitting interruptions, actual ZTEx frequency, adherence and maintenance over time, and workplace acceptability across desk-based and safety-sensitive occupations. Thus, rather than providing evidence of intervention

XSL•FO
RenderX

effectiveness, the present study offers an empirically grounded and context-sensitive framework for advancing ZTEx research from intention-related associations to behavioral and intervention evaluation in Chinese occupational settings.

Acknowledgments

ChatGPT was used solely as an editorial aid during revision to review grammar, spelling, punctuation, and sentence clarity. It was not used to develop the study arguments, analyses, structure, interpretations, or substantive academic content. All ideas, source selection, scientific judgments, interpretations, and final decisions were made by the authors, who take full responsibility for the submitted work.

Data Availability

The data analyzed in this study may be made available from the corresponding author upon reasonable request.

Funding

This work was supported by the Hengyang Central Hospital-Hunan Normal University Health Science Center Joint Research Grant (grant number HYCHHNU202505) and the Hunan Province College Students Research Learning and Innovative Experiment Project (grant number S202510542438). The funders had no role in the study design, data collection, analysis, interpretation of the results, preparation of the manuscript, or decision to submit the manuscript for publication.

Authors' Contributions

ZX designed the study, performed data curation and formal analysis, and drafted the original manuscript. XL contributed to data collection, resource coordination, and project administration. L Chai and XH were involved in the conceptualization and investigation of the study, as well as data curation. QD participated in the investigation and contributed to data validation. YX was responsible for software implementation and data management. L Cong acquired funding, supervised the project, and critically revised the manuscript for important intellectual content. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
STROBE checklist.
[PDF File (Adobe PDF File), 152 KB - publichealth_v12i1e95707_app1.pdf]

Multimedia Appendix 2
Statistical methods and supplementary analyses.
[DOCX File , 91 KB - publichealth_v12i1e95707_app2.docx]

Multimedia Appendix 3
Benjamini-Hochberg–adjusted results for group comparisons of the zero-time exercise–related intention proxy.
[DOCX File , 18 KB - publichealth_v12i1e95707_app3.docx]

Multimedia Appendix 4
Regression diagnostics and technical robustness analyses for the adjusted regression models.
[DOCX File , 13317 KB - publichealth_v12i1e95707_app4.docx]

Multimedia Appendix 5
Theoretically motivated sensitivity analyses and secondary robustness checks for the indirect-association models.
[DOCX File , 4989 KB - publichealth_v12i1e95707_app5.docx]

Multimedia Appendix 6
Multivariable-adjusted associations between exercise motivation tertiles and the zero-time exercise–related intention proxy among Chinese adults in sedentary occupations.
[DOCX File , 74 KB - publichealth_v12i1e95707_app6.docx]

Multimedia Appendix 7

XSL•FO
RenderX

Factor loadings from Harman's single-factor test using all core questionnaire items.
[DOCX File , 21 KB - publichealth_v12i1e95707_app7.docx]

Multimedia Appendix 8
Adjusted parallel indirect-association model (PROCESS Model 4) results.
[DOCX File , 22 KB - publichealth_v12i1e95707_app8.docx]

References

1. Tremblay MS, Aubert S, Barnes JD, Saunders TJ, Carson V, Latimer-Cheung AE, SBRN Terminology Consensus Project Participants. Sedentary behavior research network (SBRN) - terminology consensus project process and outcome. Int J Behav Nutr Phys Act 2017;14(1):75 [FREE Full text] [doi: 10.1186/s12966-017-0525-8] [Medline: 28599680]
2. Gao W, Sanna M, Chen Y, Tsai M, Wen C. Occupational sitting time, leisure physical activity, and all-cause and cardiovascular disease mortality. JAMA Netw Open 2024;7(1):e2350680 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.50680] [Medline: 38241049]
3. Oye-Somefun A, Mirzadeh P, Gao-Kang J, Rotondi M, Kuk JL, Tamim H, et al. The relationships between physical activity, sedentary behaviour, sleep, and dementia: a systematic review and meta-analysis of cohort studies. PLoS One 2026;21(4):e0343621 [FREE Full text] [doi: 10.1371/journal.pone.0343621] [Medline: 41950182]
4. Patterson R, McNamara E, Tainio M, de Sá TH, Smith AD, Sharp SJ, et al. Sedentary behaviour and risk of all-cause, cardiovascular and cancer mortality, and incident type 2 diabetes: a systematic review and dose response meta-analysis. Eur J Epidemiol 2018;33(9):811-829 [FREE Full text] [doi: 10.1007/s10654-018-0380-1] [Medline: 29589226]
5. Ouyang Y, Wang H, He Y, Su C, Zhang J, Du W, et al. Physical activity and sedentary behavior among Chinese adults - 10 PLADs, China, 2022-2023. China CDC Wkly 2025;7(1):6-9. [doi: 10.46234/ccdcw2025.002] [Medline: 39801822]
6. Chan NPT, Lai AYK, Choy HK, Cheung DYT, Wan ANT, Cheng VYH, et al. Feasibility and potential effectiveness of a smartphone zero-time exercise intervention for promoting physical activity and fitness in patients with coronary heart disease: a pilot randomized controlled trial. Front Public Health 2022;10:865712 [FREE Full text] [doi: 10.3389/fpubh.2022.865712] [Medline: 35910893]
7. Yin M, Li Y, Aziz AR, Buffey A, Bishop DJ, Bao D, et al. Short bouts of accumulated exercise: review and consensus statement on definition, efficacy, feasibility, practical applications, and future directions. J Sport Health Sci 2025;15:101088 [FREE Full text] [doi: 10.1016/j.jshs.2025.101088] [Medline: 40972791]
8. Ahmadi MN, Hamer M, Gill JMR, Murphy M, Sanders JP, Doherty A, et al. Brief bouts of device-measured intermittent lifestyle physical activity and its association with major adverse cardiovascular events and mortality in people who do not exercise: a prospective cohort study. Lancet Public Health 2023;8(10):e800-e810 [FREE Full text] [doi: 10.1016/S2468-2667(23)00183-4] [Medline: 37777289]
9. Alexe DI, Saha S, Choudhary PK, Alexe CI, Choudhary S, Tohănean DI. Exercise snacks as a strategy to interrupt sedentary behavior: a systematic review of health outcomes and feasibility. Healthcare (Basel) 2025;13(24):3216 [FREE Full text] [doi: 10.3390/healthcare13243216] [Medline: 41464286]
10. Lai A, Stewart S, Wan A, Thomas C, Tse J, Ho D, et al. Development and feasibility of a brief Zero-time exercise intervention to reduce sedentary behaviour and enhance physical activity: a pilot trial. Health Soc Care Community 2019;27(4):e233-e245. [doi: 10.1111/hsc.12728] [Medline: 30821866]
11. Jekauc D, Gürdere C, Englert C, Strobach T, Bottesi G, Bray S, et al. The contribution and interplay of implicit and explicit processes on physical activity behavior: empirical testing of the physical activity adoption and maintenance (PAAM) model. BMC Public Health 2024;24(1):1239 [FREE Full text] [doi: 10.1186/s12889-024-18589-5] [Medline: 38711051]
12. Xu Z, Shamsulariffin S, Azhar Y, Xi M. Does self-determination theory associate with physical activity' a systematic review of systematic review. Int J Psychol 2025;60(3):e70044. [doi: 10.1002/ijop.70044] [Medline: 40256835]
13. Mackenzie K, Such E, Norman P, Goyder E. Sitting less at work: a qualitative study of barriers and enablers in organisations of different size and sector. BMC Public Health 2019;19(1):884 [FREE Full text] [doi: 10.1186/s12889-019-7148-8] [Medline: 31272484]
14. Ajzen I. The theory of planned behavior: frequently asked questions. Human Behav and Emerg Tech 2020;2(4):314-324. [doi: 10.1002/hbe2.195]
15. Porath J, Schmidt LI, Möckel J, Dold C, Hennerkes L, Haussmann A. What it takes to reduce sitting at work: a pilot study on the effectiveness and correlates of a multicomponent intervention. Int Arch Occup Environ Health 2024;97(1):9-21 [FREE Full text] [doi: 10.1007/s00420-023-02020-4] [Medline: 37950069]
16. Ojo SO, Bailey DP, Hewson DJ, Chater AM. Perceived barriers and facilitators to breaking up sitting time among desk-based office workers: a qualitative investigation using the TDF and COM-B. Int J Environ Res Public Health 2019;16(16):2903 [FREE Full text] [doi: 10.3390/ijerph16162903] [Medline: 31416112]
17. Morris A, Murphy R, Shepherd S, Graves L. Multi-stakeholder perspectives of factors that influence contact centre call agents' workplace physical activity and sedentary behaviour. Int J Environ Res Public Health 2018;15(7):1484 [FREE Full text] [doi: 10.3390/ijerph15071484] [Medline: 30011821]

XSL•FO
RenderX

18. Deci EL, Ryan RM. The "What" and "Why" of goal pursuits: human needs and the self-determination of behavior. Psychological Inquiry 2000;11(4):227-268. [doi: 10.1207/s15327965pli1104_01]
19. Bandura A. Health promotion by social cognitive means. Health Educ Behav 2004;31(2):143-164. [doi: 10.1177/1090198104263660] [Medline: 15090118]
20. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
21. Willmott TJ, Pang B, Rundle-Thiele S. Capability, opportunity, and motivation: an across contexts empirical examination of the COM-B model. BMC Public Health 2021;21(1):1014 [FREE Full text] [doi: 10.1186/s12889-021-11019-w] [Medline: 34051788]
22. Fogg B. A behavior model for persuasive design. 2009 Presented at: Persuasive 2009: Persuasive 2009; 4th International Conference on Persuasive Technology; 2009 April 26 - 29; Claremont, California, USA p. 1-7. [doi: 10.1145/1541948.1541999]
23. Fritz MS, Mackinnon DP. Required sample size to detect the mediated effect. Psychol Sci 2007;18(3):233-239 [FREE Full text] [doi: 10.1111/j.1467-9280.2007.01882.x] [Medline: 17444920]
24. Wolf EJ, Harrington KM, Clark SL, Miller MW. Sample size requirements for structural equation models: an evaluation of power, bias, and solution propriety. Educ Psychol Meas 2013;76(6):913-934 [FREE Full text] [doi: 10.1177/0013164413495237] [Medline: 25705052]
25. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep 1985;100(2):126-131 [FREE Full text] [Medline: 3920711]
26. Ye J, Li Y, Yang L. Development and validation of the Chinese version of the adult sedentary behavior reduction intention questionnaire based on the theory of planned behavior. BMC Public Health 2025;25(1):1093 [FREE Full text] [doi: 10.1186/s12889-025-22301-6] [Medline: 40119300]
27. Markland D, Tobin V. A modification to the behavioural regulation in exercise questionnaire to include an assessment of amotivation. J Sport Exerc Psychol 2004;26(2):191-196. [doi: 10.1123/jsep.26.2.191]
28. Liu JD, Chung P, Zhang C, Si G. Chinese-translated behavioral regulation in exercise questionnaire-2: evidence from university students in the mainland and Hong Kong of China. J Sport Health Sci 2015;4(3):228-234. [doi: 10.1016/j.jshs.2014.03.017]
29. Schwarzer R, Jerusalem M. Generalized self-efficacy scale. In: Measures in Health Psychology: A User's Portfolio. Causal and Control Beliefs. Windsor, UK: NFER-Nelson; 1995:35-37.
30. Wang CK, Hu ZF, Liu Y. Evidences for reliability and validity of the Chinese version of the General Self-Efficacy Scale. Chin J Appl Psychol 2001;7(1):37-40. [doi: 10.3969/j.issn.1006-6020.2001.01.007]
31. Wang Y, Kang HK. Development of a Physical Activity Triggers Questionnaire. Healthcare (Basel) 2022;11(1):25 [FREE Full text] [doi: 10.3390/healthcare11010025] [Medline: 36611487]
32. Chen N, Lee KJ, Kim J, Tsai CL. The effects of individual-level and area-level socioeconomic status on preferences and behaviour in leisure time physical activities: a cross-sectional analysis of Chinese adults. Leisure/Loisir 2022;46(4):623-643. [doi: 10.1080/14927713.2022.2054459]
33. Lan L, Lang X, McKee M, Tse LA, Rangarajan S, Qiang D, PURE-China Investigators. Association of sitting time with cardiovascular events among manual and non-manual workers: a prospective cohort study (PURE-China). BMC Public Health 2025;25(1):750 [FREE Full text] [doi: 10.1186/s12889-025-21948-5] [Medline: 39994596]
34. Hagger M, Chatzisarantis N, Biddle S. A meta-analytic review of the theories of reasoned action and planned behavior in physical activity: predictive validity and the contribution of additional variables. J Sport Exerc Psychol 2002;24(1):3-32. [doi: 10.1123/jsep.24.1.3]
35. Wood W, Rünger D. Psychology of habit. Annu Rev Psychol 2016;67:289-314. [doi: 10.1146/annurev-psych-122414-033417] [Medline: 26361052]
36. Wang L, Yang H, Song X, Wang C, Qiao M, Tao H, et al. Physical activity level and social-ecological influence factors in Chinese information technology professionals: a cross-sectional study. J Environ Public Health 2022;2022:4580589 [FREE Full text] [doi: 10.1155/2022/4580589] [Medline: 36060889]
37. Jenkins K, Buchan J, Rhodes RE, Hamilton K. Exploring environmental cues to instigate physical movement in the workplace. Health Psychol Behav Med 2024;12(1):2323433 [FREE Full text] [doi: 10.1080/21642850.2024.2323433] [Medline: 38476211]
38. Jia Y, Xiao Y, Chen H, Gebel K, Li C, Sun S, et al. Effects of group communication norms on daily steps in a team-based financial incentive mobile phone intervention in Shanghai, China. Int J Behav Nutr Phys Act 2025;22(1):9 [FREE Full text] [doi: 10.1186/s12966-025-01707-w] [Medline: 39827159]
39. Shang B, Duan Y, Brehm W, Liang W. Subjective theories of Chinese office workers with irregular physical activity: an interview-based study. Front Psychol 2022;13:854855 [FREE Full text] [doi: 10.3389/fpsyg.2022.854855] [Medline: 35529547]
40. Blake H, Lai B, Coman E, Houdmont J, Griffiths A. Move-it: a cluster-randomised digital worksite exercise intervention in China: outcome and process evaluation. Int J Environ Res Public Health 2019;16(18):3451 [FREE Full text] [doi: 10.3390/ijerph16183451] [Medline: 31533292]

XSL•FO
RenderX

41. Sun Y, Gao Y, Ou AYT, Wang X, Xie Y, Chen X, et al. Effectiveness of a blended intervention to promote physical activity among office employees: randomized controlled trial. J Med Internet Res 2026;28:e80249 [FREE Full text] [doi: 10.2196/80249] [Medline: 42172646]
42. Rouyard T, Yoda E, Akksilp K, Dieterich AV, Kc S, Dabak SV, et al. Effects of workplace interventions on sedentary behaviour and physical activity: an umbrella review with meta-analyses and narrative synthesis. Lancet Public Health 2025;10(4):e295-e308 [FREE Full text] [doi: 10.1016/S2468-2667(25)00038-6] [Medline: 40175011]
43. Leppe-Zamora J, Ramos-Fuster S, Muñoz-Monari B, Roa-Alcaino S, Sarmiento OL. The effect of computer prompt in breaks of sedentary behaviour among office workers: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2025;22(1):75 [FREE Full text] [doi: 10.1186/s12966-025-01781-0] [Medline: 40514667]
44. Iwakura M, Ozeki C, Jung S, Yamazaki T, Miki T, Nohara M, et al. An umbrella review of efficacy of digital health interventions for workers. NPJ Digit Med 2025;8(1):207 [FREE Full text] [doi: 10.1038/s41746-025-01578-2] [Medline: 40229460]
45. Feil K, Fritsch J, Rhodes RE. The intention-behaviour gap in physical activity: a systematic review and meta-analysis of the action control framework. Br J Sports Med 2023;57(19):1265-1271. [doi: 10.1136/bjsports-2022-106640] [Medline: 37460164]
46. Maxwell SE, Cole DA. Bias in cross-sectional analyses of longitudinal mediation. Psychol Methods 2007;12(1):23-44. [doi: 10.1037/1082-989X.12.1.23] [Medline: 17402810]
47. O'Laughlin KD, Martin MJ, Ferrer E. Cross-sectional analysis of longitudinal mediation processes. Multivariate Behav Res 2018;53(3):375-402. [doi: 10.1080/00273171.2018.1454822] [Medline: 29624079]
48. Lee H, Cashin AG, Lamb SE, Hopewell S, Vansteelandt S, VanderWeele TJ, AGReMA group, et al. A guideline for reporting mediation analyses of randomized trials and observational studies: the AGReMA statement. JAMA 2021;326(11):1045-1056 [FREE Full text] [doi: 10.1001/jama.2021.14075] [Medline: 34546296]
49. Podsakoff PM, Podsakoff NP, Williams LJ, Huang C, Yang J. Common method bias: it's bad, it's complex, it's widespread, and it's not easy to fix. Annu Rev Organ Psychol Organ Behav 2024;11(1):17-61. [doi: 10.1146/annurev-orgpsych-110721-040030]

Abbreviations

BootSE: bootstrap SE
BREQ-2: Behavioral Regulation in Exercise Questionnaire
COM-B: capability, opportunity, motivation-behavior
GSES: General Self-Efficacy Scale
MET: metabolic equivalent
PATQ: Physical Activity Triggers Questionnaire
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
ZTEx: zero-time exercise
ZTEx-IP: zero-time exercise–related intention proxy

Edited by A Mavragani, T Sanchez; submitted 19.Mar.2026; peer-reviewed by X Tao, T Yang; comments to author 07.Jul.2026; revised version received 13.Aug.2026; accepted 19.Aug.2026; published 04.Sep.2026.

Please cite as:
Xiong Z, Li X, Chai L, Duan X, Hu X, Xie Y, Cong L
Psychosocial Correlates of Sedentary Behavior Reduction Readiness Relevant to Zero-Time Exercise Among Chinese Adults in Sedentary Occupations: Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e95707
URL: https://publichealth.jmir.org/2026/1/e95707
doi:10.2196/95707
PMID:

©Ziyan Xiong, Xiangyuan Li, Lu Chai, Xiongqiang Duan, Xiang Hu, Yuan Xie, Li Cong. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 04.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Clinical Epidemiology of Cancer in People Living With HIV in Germany: Retrospective, Observational, Multicenter Federated Claims Data Analysis

Bastian Reiter[1,2]; Stefanie Andreas[1,3], MSc; Linnea Schumann[1], MSc; Bernd Bender[1,3], DPhil; Isabel Schnorr[1,2], Dr med; Andreas Heidenreich[4], Dr; Christoph Stephan[5], Prof Dr med; Andrea Prunotto[6], PhD; Andrea Laukhuf[6], MSc; Julius Wehrle[6,7], Dr; Matthias Müller[8,9], MSc, Dr med; Fabio Paul Aubele[10], MSc; Marlien Hagedorn[10], MSc; Fady Albashiti[10,11], Prof Dr sc hum; Ulrich Seybold[12], MSc, PD Dr med; Julia Roider[12], PD Dr med; Melanie Stecher[13,14], PhD; Daniel Maier[1,3*], Dr, MPH; Jörg Janne Vehreschild[1,2*], Prof Dr med

[1]Institute for Digital Medicine and Clinical Data Sciences, Faculty of Medicine, Goethe University Frankfurt, Marienburgstrasse 6 Haus 97E / 2. OG, Frankfurt am Main, Germany

[10]Medical Data Integration Center, LMU University Hospital, LMU Munich, Munich, Germany

[11]Neu-Ulm University of Applied Sciences, Neu-Ulm, Germany

[12]Division of Infectious Diseases, LMU University Hospital, LMU Munich, Munich, Germany

[13]Department I of Internal Medicine, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany

[14]German Center for Infection Research (DZIF), partner site Bonn-Cologne, Cologne, Germany

[2]Medical Department 2, Center for Internal Medicine, University Hospital Frankfurt, Frankfurt am Main, Germany

[3]German Cancer Consortium (DKTK), partner site Frankfurt/Mainz, a partnership between DKFZ and University Hospital Frankfurt, Frankfurt am Main, Germany

[4]Data Integration Center (DIC), University Hospital Frankfurt, Frankfurt am Main, Germany

[5]Medical HIV Treatment and Research Unit, HIVCenter, University Hospital Frankfurt, Frankfurt am Main, Germany

[6]Data Integration Center, Medical Center and Faculty of Medicine, University of Freiburg, Freiburg im Breisgau, Germany

[7]Department of Medicine I, Medical Center and Faculty of Medicine, University of Freiburg, Freiburg im Breisgau, Germany

[8]Department of Medicine II (Gastroenterology, Hepatology, Endocrinology and Infectious Diseases), Medical Center and Faculty of Medicine, University of Freiburg, Freiburg im Breisgau, Germany

[9]Department of Infection Medicine, Medical Service Centre Clotten, Freiburg im Breisgau, Germany

[*]these authors contributed equally

Corresponding Author:
Bastian Reiter
Institute for Digital Medicine and Clinical Data Sciences, Faculty of Medicine, Goethe University Frankfurt, Marienburgstrasse 6 Haus 97E / 2. OG, Frankfurt am Main, Germany

Abstract

Background: People living with HIV are at increased risk for developing cancer, a leading cause of death in this population. The management of cancer in people living with HIV is particularly challenging, necessitating specialized, interdisciplinary care. However, insights into cancer care provision for people living with HIV in Germany remain scarce.

Objective: This study analyzed inpatient cancer care for people living with HIV, comparing treatment patterns and complications with those of an HIV-negative control group. Using claims data from 3 German university hospitals related to admissions between 2005 and 2022, we aimed to identify care disparities and provide evidence to support improved cancer management.

Methods: A customized federated approach was used to analyze inpatient claims data of patients across the 3 data-holding institutions. The data included demographics, diagnoses, procedures, and treatment codes as well as discharge information. Using nearest-neighbor matching, we analyzed demographic features, cancer diagnoses, anticancer therapies, and outcomes in people living with HIV and cancer and for a control group of HIV-negative patients with cancer.

Results: Among 162,380 patients, 907 (0.6%) were people living with HIV and cancer. The count of cancer diagnoses declined over time, particularly for AIDS-defining cancers (total cancer diagnoses: P=.001; AIDS-defining cancers: P=.002), with a shift toward older age at diagnosis. Compared with matched controls, people living with HIV and cancer had longer hospital stays, experienced more postchemotherapy complications (cancer with HIV: 64/907, 15.6%; cancer without HIV: 20/907, 5.5%; P<.001), and showed higher rates of metastasis after initial diagnosis (cancer with HIV: 128/267, 47.9%; cancer without HIV: 97/287, 33.8%; P<.001). People living with HIV and cancer also showed increased in-hospital mortality, although mortality declined

XSL•FO
RenderX

over time (*P*=.02). Our data suggested differences in documented therapy modalities between the compared groups, with people living with HIV and cancer receiving more chemo- and immunotherapy and less surgery.

Conclusions: Using federated analysis techniques, we were able to show that cancer diagnoses and mortality among people living with HIV in Germany have decreased over time; however, disparities in treatment and outcomes persisted as compared with HIV-negative patients with cancer. Our findings underscore the need for tailored, multidisciplinary care strategies to improve cancer management for this population.

(JMIR Public Health Surveill 2026;12:e81092) doi:10.2196/81092

KEYWORDS

secondary data analysis; federated analysis; real-world data; HIV; cancer; malignancy; people living with HIV and cancer; inpatient care; claims data; Section 21 German Hospital Fees Act

Introduction

Despite advances in treatment and prevention, HIV remains a major global health threat, with 39.9 million cases recorded worldwide in 2023 [1]. In Germany, HIV incidence has stabilized, and treated individuals no longer exhibit increased mortality, which is seen as a contributing factor to a rise in HIV prevalence [2]. The decline in AIDS-related deaths is largely attributed to effective antiretroviral therapy (ART) [3-6]. However, people living with HIV still face a significantly higher cancer risk, particularly for infection-related cancers. These occur approximately 30 times more frequently in female people living with HIV and up to 100 times more frequently in male people living with HIV as compared with the general population [7-10].

HIV-associated cancers in people living with HIV can be categorized as AIDS-defining (AD) and non–AIDS-defining (NAD), with AD cancers including Kaposi sarcoma, non-Hodgkin lymphoma, and invasive cervical cancer [11]. NAD cancers are further divided into virus-associated (virus-NAD) and non–virus-associated NAD types. Virus-NAD cancers linked to coinfections include lung, anal, vulvar, penile, oral, and pharyngeal cancers; Hodgkin lymphoma; and hepatocellular carcinoma [8,9,12,13].

Since the introduction of ART, the incidence of AD cancers has decreased, whereas the incidence of virus-NAD and non–virus-associated NAD cancers has increased [7,14,15]. Due to increased life expectancy, both virus-associated and non–virus-associated NAD cancers have become leading causes of death among people living with HIV, surpassing deaths directly attributable to HIV infection [5,16,17]. Although combination ART has improved survival and enables full-dose chemotherapy [18], managing cancer in people living with HIV remains complex. Immunosuppression, drug interactions with ART, and increased treatment toxicity pose significant challenges, requiring a multidisciplinary care approach involving oncology, hematology, infectious diseases, pharmacy, and supportive care specialists [19]. Given these complexities, people living with HIV and cancer may need to travel longer distances to access appropriate, specialized care.

While cancer epidemiology in people living with HIV is well-documented [7-10,12-15,19], evidence on cancer care in people living with HIV—particularly regarding treatment and outcomes—remains limited. Existing data, primarily from US registries, suggest that people living with HIV may receive cancer treatment less frequently than HIV-negative patients [20,21].

Using claims data from 3 German university hospitals, we analyzed inpatient cancer care for people living with HIV in Germany and compared it to that of an HIV-negative control group.

Methods

Study Data

This multicenter retrospective study used claims data collected under Section 21 of the German Hospital Fees Act (German: Krankenhausentgeltgesetz, KHEntgG). This law requires hospitals to report individual patient billing records to the federal institute for the hospital payment system (Institut für das Entgeltsystem im Krankenhaus). The data were sourced from data integration centers [22,23] of university hospitals in 3 large German cities (Frankfurt, Freiburg, and Munich).

The raw data included patient demographics and basic diagnostic- and treatment-related information (*International Statistical Classification of Diseases and Related Health Problems, Tenth Revision* [*ICD-10*]), as well as operation and procedure codes (Operationen- und Prozedurenschlüssel [OPS]; for more details, refer to Table S1 in Multimedia Appendix 1).

Cohort Selection

Inclusion criteria required patients to be aged 18 years or older at the time of inpatient admission. The observation period was defined from January 1, 2005, to December 31, 2022. For inclusion, each patient also required at least 1 documented *ICD-10* code indicating either a cancer diagnosis or an HIV infection (Table S2 in Multimedia Appendix 1).

Federated Data Analysis Approach

To analyze sensitive patient data distributed across multiple university hospitals, we used a federated approach; this means individual patient data were not transferred to a central computer to conduct pooled analysis. Instead, analysis scripts were sent to and executed locally at each participating hospital. Only aggregated or anonymized data were subsequently transferred to be jointly analyzed.

XSL•FO
RenderX

Data Preprocessing and Patient Matching

A harmonized input data model (Table S3 in Multimedia Appendix 1) was developed to enable consistent and standardized analysis across study sites. Uniform value formats and restrictions were defined to ensure that only valid and consistent cases and data entries were included. A case was defined as a single hospitalization for a given patient. Using pseudonymized patient identifiers, individual hospital stays were linked to the same patient, allowing for multiple documented cases per patient during the observation period. Special focus was placed on plausibility checks to correctly classify main cancer and HIV diagnoses as well as therapeutic interventions (Tables S2, S4, and S5 in Multimedia Appendix 1). The data preprocessing resulted in a ready-for-analysis data model customized for the purpose of our study (Table S6 in Multimedia Appendix 1).

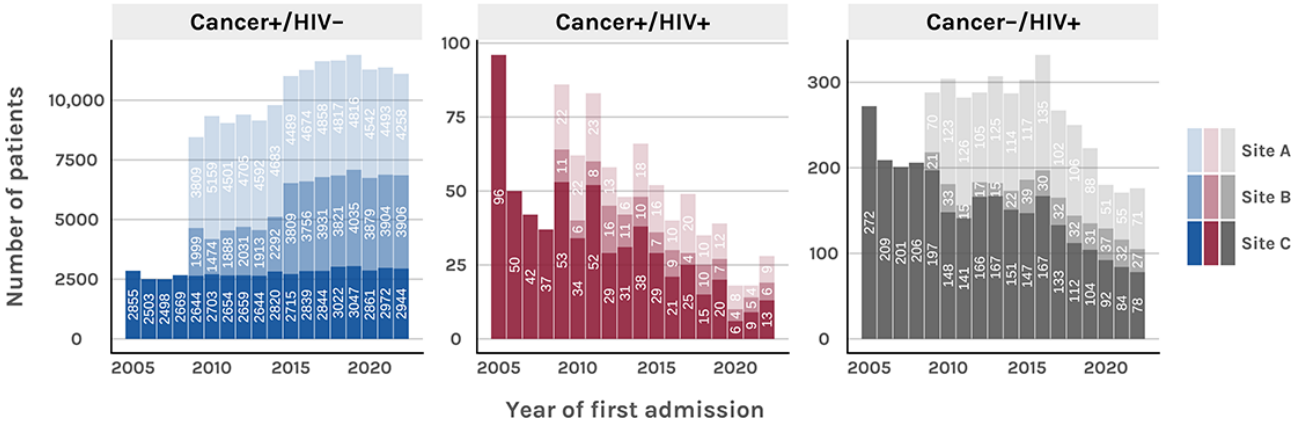
Included patients were categorized into 1 of 3 groups: cancer with HIV (cancer+/HIV+), cancer without HIV (cancer+/HIV–), and HIV without cancer (cancer–/HIV+) (Tables S2 and S5 in Multimedia Appendix 1). For patients with cancer, the first recorded cancer diagnosis was designated as the primary cancer diagnosis. For each patient group, frequencies of demographic and cancer-related, case-related, and therapy-related variables were summarized annually.

Patients in the cancer+/HIV+ group were matched with patients in the cancer+/HIV– group using 1:1 nearest-neighbor matching [24], based on age at cancer diagnosis, sex, year of cancer diagnosis, cancer topography group, presence of carcinoma in situ, and comorbidity score at the time of diagnosis (Table S7 in Multimedia Appendix 1). The comorbidity score was calculated using van Walraven weights for the Elixhauser comorbidity groups [25]. HIV infection and AIDS were assigned to a weight of zero, and cancer diagnoses beyond the primary diagnosis were considered comorbidities. Patient matching was performed to compare therapy-related data, including anticancer therapy frequencies, complications following chemotherapy (eg, transfer to the intensive care unit, dialysis, and mechanical ventilation), and outcomes (derived from the last recorded discharge reason) between the cancer+/HIV+ and cancer+/HIV– groups.

Statistical Analysis

Relative frequencies of individual characteristics were calculated for each group across the entire study period (2005 - 2022). For matched group comparisons, *P* values from the Fisher exact test were calculated. Trend analyses were based on data from 2009 to 2022, the period for which data were available from all 3 sites (Figure 1). Poisson regression models were used to assess temporal trends in age at cancer diagnosis and cancer-related diagnoses within the cancer+/HIV+ group. When overdispersion was detected, negative binomial regression models were used as an alternative [26].

Figure 1. Number of patients in each defined group per year and per site (2005 - 2022).



For each regression model, the number of patients or diagnoses was used as the dependent variable, with the year of cancer diagnosis as the predictor. In contrast, for therapy-related trends, linear regression models were fitted using proportions as the dependent variable.

To assess travel distances for each patient group, the geographical distance between the centroid of each patient's postal code area at the time of first admission and the hospital location was estimated using the Haversine formula [27], providing an approximation of how far patients in the 3 subgroups traveled to receive care. Additionally, the catchment areas of the individual sites were visualized by mapping the distribution of patients to 3-digit postal code areas for each group (Figure S1 in Multimedia Appendix 1).

All statistical analyses were conducted using R (version 4.3.2; R Foundation for Statistical Computing) [28]. Nearest-neighbor matching was performed with the MatchIt package [24]. The Fisher exact test as well as linear and Poisson regression models were computed with the stats package [28], while negative binomial regressions were modeled using the MASS package [29]. Travel distances for the patient subgroups were calculated with the geosphere [27] and sf [30] packages.

Ethical Considerations

This retrospective study was conducted after consultation with the responsible ethics committees (research ethics committee, Faculty of Medicine, Goethe University Frankfurt; 274/18; research ethics committee, University of Freiburg: 22-1279-S1-retro). In accordance with national law, no informed consent was obtained since only pseudonymized patient data was used for analysis. Furthermore, the federated data analysis

XSL•FO
RenderX

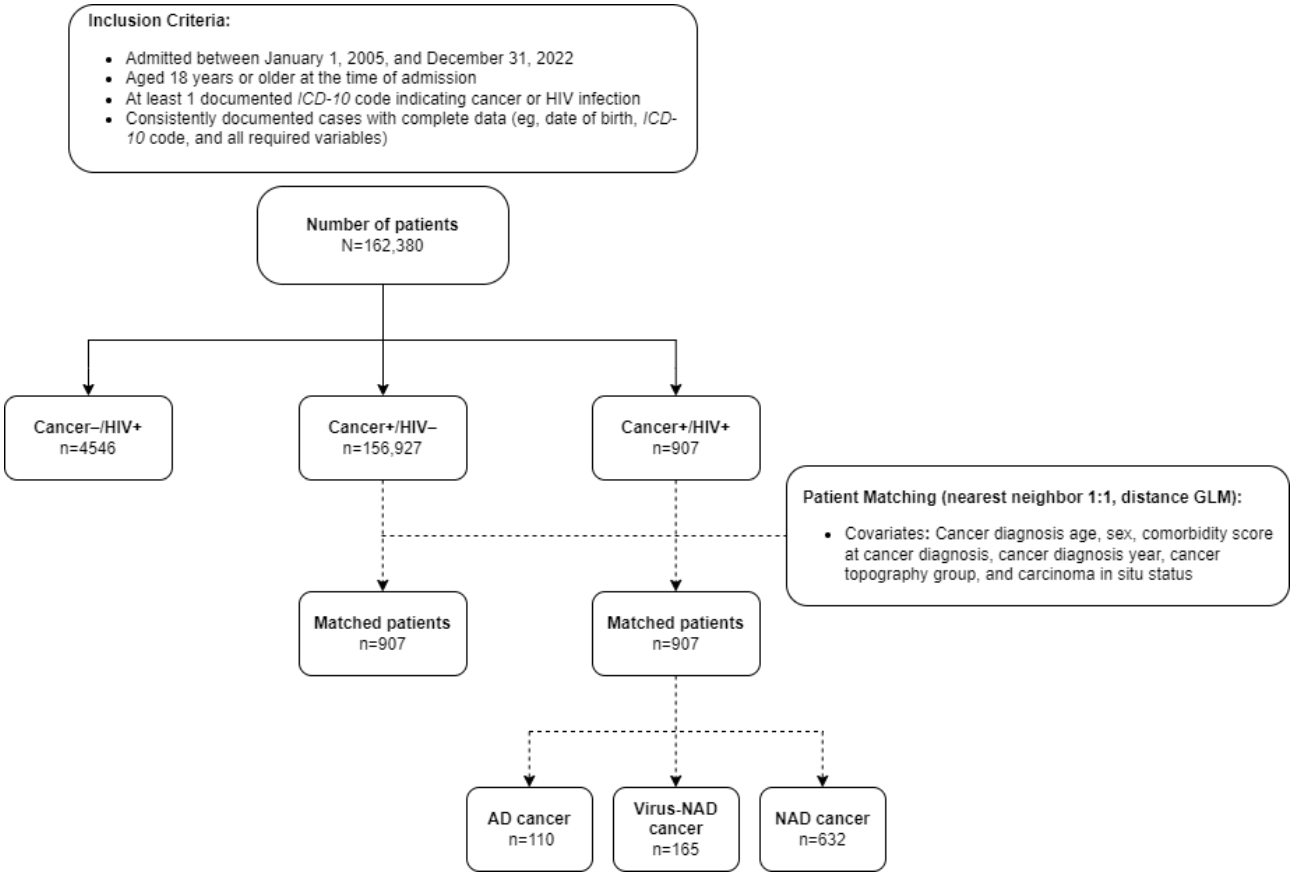
approach ensured data privacy through aggregation of individual-level data.

Results

Study Cohort

Between 2005 and 2022, a total of 162,380 eligible inpatients were recorded across all 3 sites and patient groups (Figure 2). Among these, 907 (0.6%) patients were classified as people living with HIV and cancer (cancer+/HIV+), exhibiting at least 1 *ICD-10* code from both diagnosis groups. A total of 156,927 (96.8%) patients had at least 1 cancer-related *ICD-10* code without an HIV diagnosis (cancer+/HIV–), while 4546 (2.8%) patients were identified as people living with HIV without cancer (cancer–/HIV+). Figure 1 shows the annual distribution of patients in each group by site.

Figure 2. Inclusion, grouping, and matching of patients with the respective number of patients. AD: AIDS-defining; *ICD-10: International Statistical Classification of Diseases and Related Health Problems, Tenth Revision*; GLM: generalized linear model; NAD: non–AIDS-defining; Virus-NAD: cancer associated with HIV infection but not considered AIDS-defining.



Patient demographics and admission-related characteristics for the 3 patient groups are displayed in Table 1. At the time of first admission, people living with HIV and cancer (cancer+/HIV+) were younger than HIV-negative patients with cancer (cancer+/HIV–), but older than people living with HIV without a cancer diagnosis (cancer–/HIV+). Fewer than one-quarter (214/907, 23.6%) of people living with HIV who have cancer were female.

XSL•FO
RenderX

Table . Demographic and admission-related characteristics of the 3 patient groups (2005-2022).

Characteristics	Cancer+/HIV– (n=156,927), n (%)	Cancer+/HIV+ (n=907), n (%)	Cancer–/HIV+ (n=4546), n (%)
Female patients	72,231 (46.0)	214 (23.6)	1441 (31.7)
Age (years) at first relevant (cancer or HIV associated) admission			
18 - 39	13,373 (8.5)	176 (19.4)	1781 (39.2)
40 - 59	44,984 (28.7)	536 (59.1)	2204 (48.5)
60 - 79	80,335 (51.2)	192 (21.2)	547 (12.0)
>80	18,235 (11.6)	3 (0.3)	14 (0.3)
Admission count per patient			
1	58,444 (37.2)	144 (15.9)	2438 (53.6)
2 - 4	62,840 (40.0)	351 (38.7)	1674 (36.8)
5 - 10	27,877 (17.8)	301 (33.2)	356 (7.8)
More than 10	7766 (4.9)	111 (12.2)	78 (1.7)
Average length of stay per admission (days)			
Up to 7	77,096 (49.1)	307 (33.8)	2204 (48.5)
7 - 14	49,898 (31.8)	298 (32.9)	1191 (26.2)
14 - 30	23,194 (14.8)	233 (25.7)	835 (18.4)
More than 30	6739 (4.3)	69 (7.6)	316 (7.0)

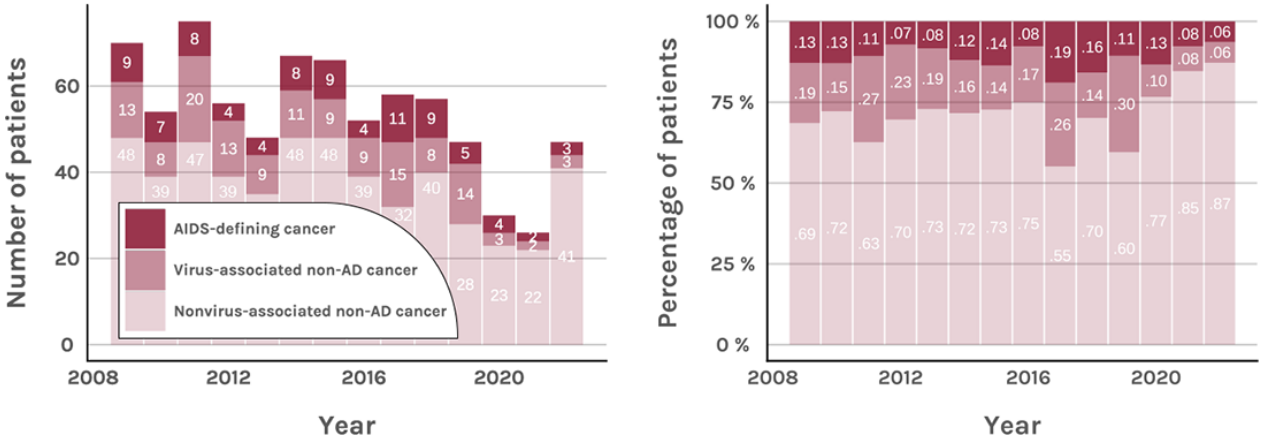
On average, patients in the cancer+/HIV+ group exhibited the highest admission frequencies among the 3 groups: 2 to 4 admissions in 38.7% (351/907), 5 to 10 admissions in 33.2% (301/907), and >10 admissions in 12.2% (111/907) of all observed patient histories.

Cancer Incidence and Cancer Types in People Living With HIV

Among 907 patients with both diagnoses (cancer+/HIV+), the data suggest that 504 (55.6%) patients received HIV and cancer diagnoses simultaneously or within a short time interval. In 332 (36.6%) patients, the HIV diagnosis preceded the cancer diagnosis, whereas in 71 (7.8%) patients, the first recorded cancer diagnosis was documented before the HIV diagnosis (Table S8 in Multimedia Appendix 1).

While fluctuating over the observation period, the total number of cancer diagnoses in people living with HIV decreased significantly (*P*=.002). Among the cancer groups (AD, virus-NAD, and non–virus-associated NAD; Figure 3), the most notable decline was observed in the frequency of virus-associated NAD cancer diagnoses. In total, 110 (12.1%) diagnoses were classified as AD, 165 (18.2%) as virus-associated NAD, and 632 (69.7%) as non–virus-associated NAD. Among specific cancer subtypes, a significant reduction was observed in diagnoses involving lymphoid and hematopoietic tissues (*P*=.005), especially in non-Hodgkin lymphomas (*P*=.002) and nonfollicular lymphomas (*P*=.008).

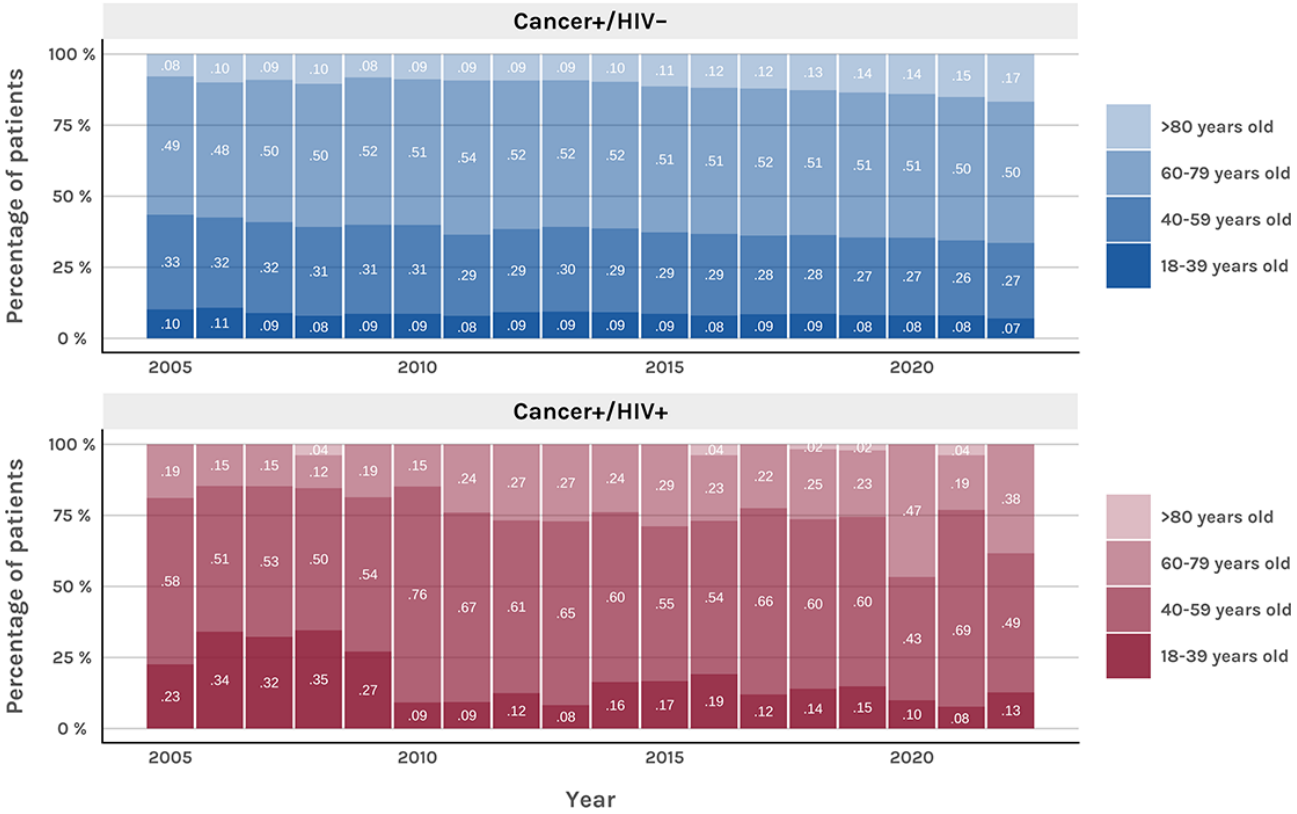
Figure 3. Absolute frequencies and proportions of cancer categories in people living with HIV and cancer (2009-2022). AD: AIDS defining.



We observed an increase in the age at cancer diagnosis in both people living with HIV and HIV-negative patients with cancer (Figure 4). Consistent with this finding, regression analysis indicated a significant decrease in the proportion of patients younger than 60 years between 2009 and 2022 (*P*=.002).

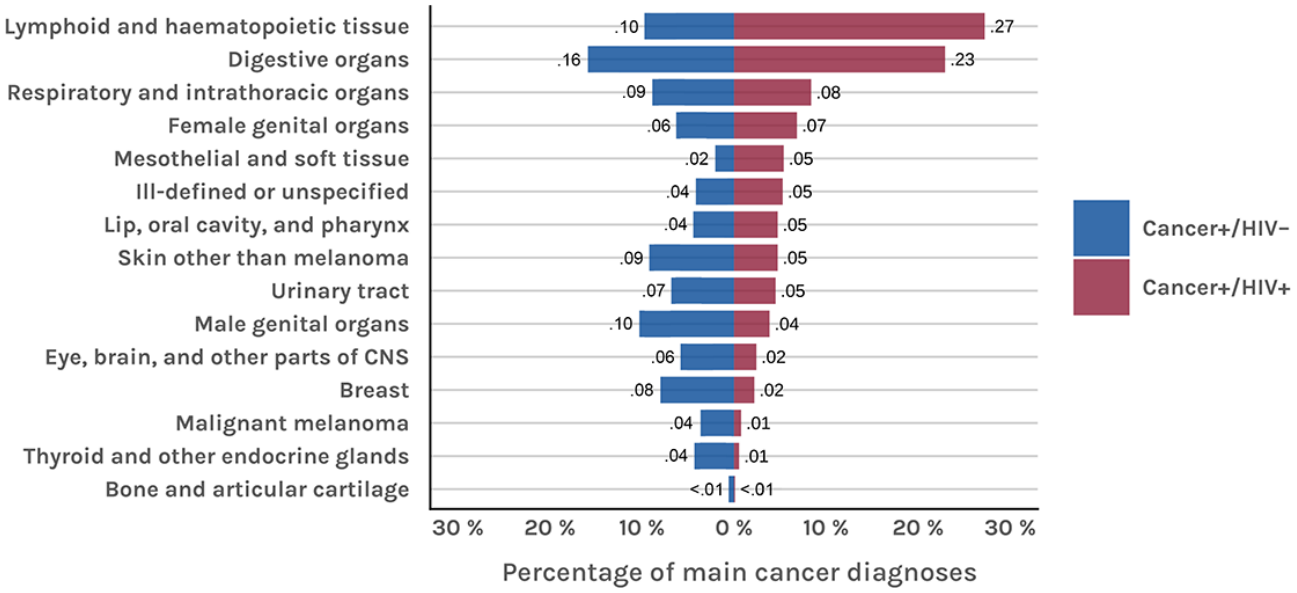
XSL•FO
RenderX

Figure 4. Proportions of grouped age at cancer diagnosis in people living with HIV and cancer and (unmatched) HIV-negative patients with cancer (2005-2022).



A comparison of cancer topography frequencies between people living with HIV and cancer (cancer+/HIV+) and HIV-negative patients with cancer (cancer+/HIV–) is presented in Figure 5.

Figure 5. Most common cancer types in people living with HIV and cancer, sorted by frequency. CNS: central nervous system.



In the matched comparison between cancer+/HIV+ and cancer+/HIV–, results showed an overall comparable prevalence of metastasis; however, metastasis was present at time of cancer diagnosis more often in the cancer+/HIV– group, whereas patients in the cancer+/HIV+ group seemed to develop metastasis after initial diagnosis more often than their HIV-negative matches (cancer+/HIV+: 128/267, 47.9%; cancer+/HIV–: 97/287, 33.8%; P<.001).

Anticancer Treatment of People Living With HIV

People living with HIV and cancer (cancer+/HIV+) received anticancer therapies at rates comparable to those of HIV-negative patients with cancer (cancer+/HIV–) but were more frequently treated with chemotherapy and immunotherapy, whereas patients in the cancer+/HIV– group seemed to have received more stem cell therapy (Table 2). Our results show a significant decrease in the use of chemotherapy in the

XSL•FO
RenderX

cancer+/HIV+ group between 2009 and 2022 (*P*=.02). People living with HIV and cancer presumably showed a higher rate of adverse events following chemotherapy than patients in the control group (cancer+/HIV+: 64/409, 15.6%; cancer+/HIV–: 20/361, 5.5%; *P*<.001). Additionally, the cancer+/HIV+ group was less frequently discharged home (516/713, 72.4%; *P*<.001) compared with the cancer+/HIV– group (583/714, 81.7%) and showed higher in-hospital mortality (cancer+/HIV+: 135/713, 18.9%; cancer+/HIV–: 90/714, 12.6%; *P*=.001).

Table . Comparison between people living with HIV and cancer and the matched patients with cancer without HIV.

Characteristics	Cancer+/HIV–, n (%)	Cancer+/HIV+, n (%)	*P* value[a]
Metastasis occurrence (n=907)			
Metastasis documented	287 (31.6)	267 (29.4)	.33
Metastasis at time of cancer diagnosis	190 (66.2)	139 (52.1)	<.001
Metastasis after cancer diagnosis	97 (33.8)	128 (47.9)	<.001
Therapy modalities (n=907)			
Any major therapy documented	631 (69.6)	639 (70.5)	.72
Surgery	285 (31.4)	251 (27.7)	.09
Chemotherapy	361 (39.8)	409 (45.1)	.03
Immunotherapy	90 (9.9)	138 (15.2)	<.001
Radiotherapy	158 (17.4)	169 (18.6)	.54
Stem cell therapy	84 (9.3)	33 (3.6)	<.001
Bone marrow transplant	12 (1.3)	0 (0.0)	.001
Complications after chemotherapy (n=907)			
Chemotherapy documented	361 (39.8)	409 (45.1)	.03
Complication after chemotherapy	20 (5.5)	64 (15.6)	<.001
Average length of stay per admission (days; n=714)[b]			
Up to 7	319 (44.7)	228 (31.9)	<.001
7-14	207 (29.0)	232 (32.5)	.17
14-30	133 (18.6)	198 (27.7)	<.001
More than 30	55 (7.7)	56 (7.8)	>.99
Last documented discharge category[c]			
Home	583 (81.7)	516 (72.4)	<.001
Deceased	90 (12.6)	135 (18.9)	.001
Other hospital	24 (3.4)	28 (3.9)	.58
Rehabilitation or residential care	11 (1.5)	20 (2.8)	.11
Hospice care	6 (0.8)	14 (2.0)	.08

[a]*P* values were calculated using the Fisher exact test.

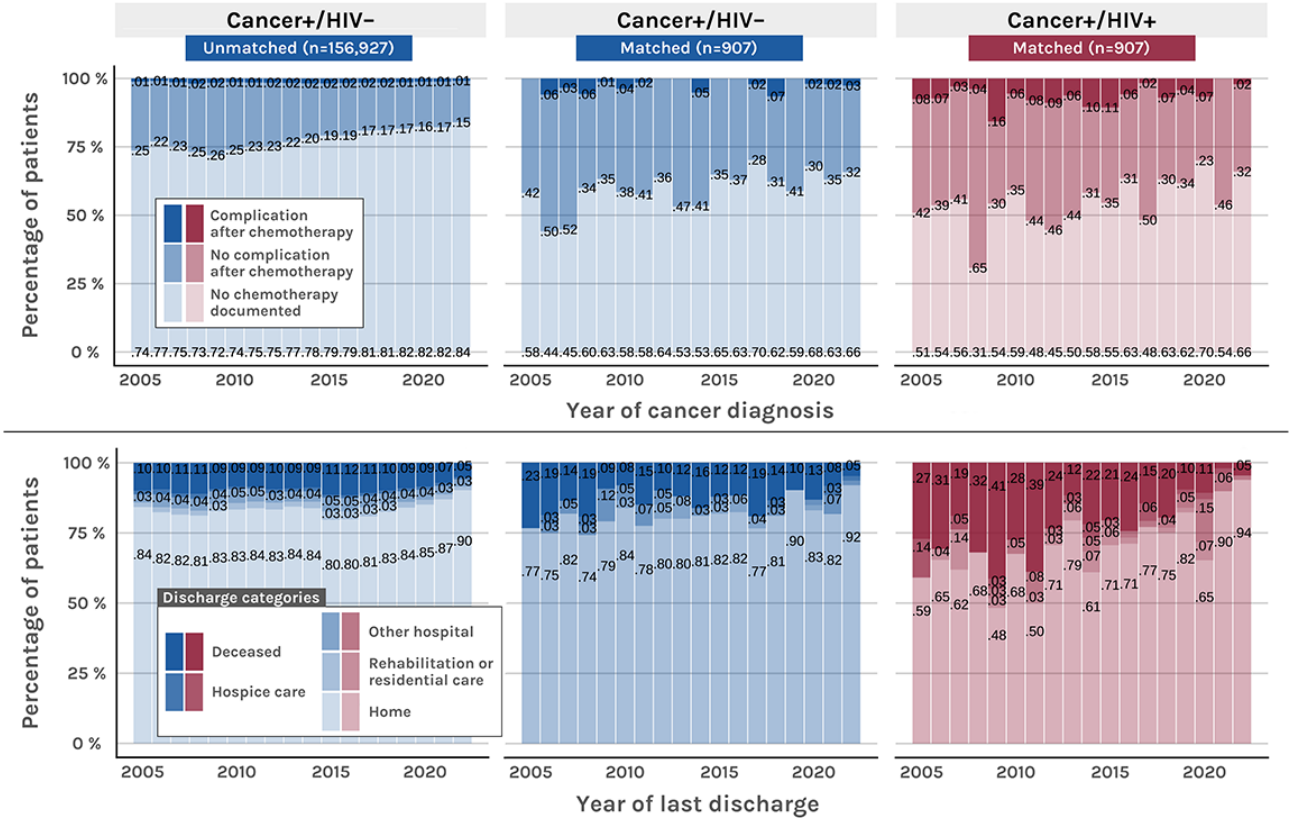
[b]Data were available from only 2 sites (n=714 for both groups).

[c]Data were available from only 2 sites (n=714 for cancer+/HIV– and n=713 for cancer+/HIV+).

Figure 6 depicts a comparison of temporal trends in complication rates after chemotherapy and discharge reasons across the total observed HIV-negative cancer population and the matched cancer groups.

XSL•FO
RenderX

Figure 6. Comparison of complication rates after chemotherapy and last documented discharge categories between unmatched cancer+/HIV– and matched cancer+/HIV– and cancer+/HIV+.



For more detailed results, refer to Tables S7, S8, and S9 in Multimedia Appendix 1. A visual abstract of the study is presented in Multimedia Appendix 2.

Discussion

Principal Findings

In this study, we examined epidemiological as well as treatment- and outcome-related characteristics of inpatients affected by HIV and cancer in 3 selected German university hospitals.

The overall number of cancer diagnoses in the cancer+/HIV+ group decreased between 2009 and 2022, primarily driven by a decline in AD cancers. This reduction was largely attributable to fewer cases of non-Hodgkin lymphomas, particularly within the nonfollicular lymphoma subgroup. Previous studies examining earlier observation periods found that following the introduction of ART, diagnoses of AD cancers, which had predominated in the early HIV era, began to decline, whereas NAD cancers became increasingly common over time [14,15].

The most frequently observed topographical categories in our study are consistent with those reported in previous research [8,10]. Lymphoma was the most frequent cancer diagnosis among people living with HIV, followed by cancers of the lung and bronchi, and the anus and anal canal. Previous studies have shown that, among NAD cancers, lung and anal cancers represent the greatest disease burden in high-income countries [31], which is in line with our findings. No increase over time was observed for typical age-related cancers, such as prostate cancer or malignant melanoma of the skin. However, it remains unclear whether this reflects stable incidence rates or a shift toward outpatient care.

The results showed an increase in the age at cancer diagnosis among people living with HIV, with fewer cases occurring in younger adults (<60 years). This trend could be explained by the decline in AD cancer diagnoses and is consistent with findings from previous studies [32]. Despite age matching, people living with HIV were diagnosed with cancer at younger ages compared with HIV-negative controls. This observation reflects the need for targeted screening strategies and emphasizes the importance of preventive measures for cancers linked to modifiable risk factors, such as smoking in the case of lung cancer [19,31] and human papillomavirus infection in cervical cancer [33].

Despite similar overall cancer treatment rates, people living with HIV received chemotherapy and immunotherapy more frequently but underwent slightly fewer surgical interventions. This disparity may be explained by differences in cancer type, stage at diagnosis, and tumor operability, as well as HIV-related factors such as immunosuppression, elevated postoperative infection risk, and potential interactions between ART and anesthesia [32,34]. Deviating from previous findings [19], cancers in our observed cohort of people living with HIV were more often diagnosed at premetastatic stages, which contradicts the lower prevalence of surgical therapy in this group.

Our findings highlight the need for improved posttreatment care. The longer and more frequent hospitalizations observed among people living with HIV and cancer require further investigation to determine whether they stem from greater

XSL•FO
RenderX

comorbidity burden or distinct clinical challenges associated with HIV infection.

Chemotherapy use during hospitalization declined among people living with HIV, suggesting a shift toward outpatient care [35,36]. To confirm this hypothesis, further analyses of outpatient care data are necessary. The study also showed a decrease in mortality among people living with HIV over time. Together, these trends suggest that cancer care and outcomes for people living with HIV have improved over the past 2 decades; however, additional research focusing on outpatient care is needed.

Contrary to our results, we initially hypothesized that people living with HIV and cancer would need to travel longer distances to access complex, multidisciplinary care. According to a recent study [37], HIV prevalence in Germany tends to be higher in urban compared with rural areas, which may explain shorter travel distances to urban university hospitals for people living with HIV in our study. Further research is needed to examine rural health care settings and to determine whether local care delivers comparable treatment outcomes or if traveling longer distances offers substantial benefits for specific patient groups like people living with HIV and cancer.

Federated Data Analysis

The use of federated data analysis offered significant advantages, including improved data privacy by keeping sensitive patient information at its original site while enabling collaboration across multiple institutions. Nonetheless, challenges such as data heterogeneity, data inconsistencies, and technical limitations required repeated script adjustments for each center. This led to increased time and staffing costs, as well as minor discrepancies between site-specific results. Specialized federated analysis software such as DataSHIELD (DataSHIELD Research Project) [38] could mitigate these drawbacks but was not available in this study. All script modifications were recorded in a GitLab repository to maintain transparency and traceability.

Limitations

A key limitation of this study is that the results cannot be generalized to the overall care of people living with HIV and cancer in Germany, as the analysis was based on data from 3 university hospitals representing the tertiary care sector. Further studies are needed to include nonuniversity hospitals and the outpatient sector.

Our study used large retrospective claims data, which the hospitals primarily documented for the purpose of billing services. Although data processing involved thorough plausibility checks, the risk that procedures or diagnoses were missing is high; furthermore, relevant information such as cancer staging and systemic therapy substances or dosage information was not part of the available data.

Sample size variations arose due to technical issues, leading to stratified analyses for the periods 2005 to 2008, 2009 to 2014, and 2015 to 2022 (Tables S6 and S7 in Multimedia Appendix 1). Other limitations include potential misclassification of disease onset and the inability to associate specific therapies with cancer types in patients with multiple diagnoses, both of which may have influenced the findings.

In our study, we chose to categorize malignancies as AD vs NAD. Different categorizations such as hematological vs solid tumors could have led to more nuanced insights. Further research should critically consider such decisions and potentially include multiple categorizations.

Conclusions

Our study offers valuable insights into the inpatient care of people living with HIV and cancer in Germany, revealing shifts in cancer epidemiology and an aging patient population. Although the incidence of AD cancers has declined, people living with HIV continue to experience longer hospital stays and higher rates of posttreatment complications. Reduced mortality and inpatient chemotherapy use suggest that cancer care and outcomes for people living with HIV have improved over the past 2 decades; however, additional research focusing on outpatient care is needed.

The study also highlights the potential of federated data analysis for multicenter research, while emphasizing the need for standardized data collection to overcome technical challenges and data inconsistencies across sites.

Acknowledgments

The authors would like to thank the participating university hospitals and their data integration centers for providing data and for their excellent cooperation. They also thank the German Cancer Consortium (DKTK) for partially funding this study.

During the preparation of this work, the authors used ChatGPT from OpenAI for proofreading. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Funding

This study was funded in part by the German Cancer Consortium and in part by grant 01KX2121 (Network University Medicine). These funding sources were not involved in study design or the analysis and interpretation of the data.

Data Availability

The datasets generated or analyzed during this study are not publicly available. We conducted a federated approach to data analysis, which means all individual-level patient data remain with the participating sites and are not publicly available due to health data privacy regulations. The aggregated data returned from site-specific analyses can be considered sufficiently anonymized.

XSL•FO
RenderX

Due to unclear use and access policies at the participating hospitals, these datasets are not made publicly available at this time; however, they can be provided by the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: BR, DM, JJV, MS, SA
Data curation: AH, AL, AP, BR, FPA, JW, MH, SA
Formal analysis: BB, BR, DM, LS, SA
Funding acquisition: JJV
Investigation: BB, BR, DM, LS, SA
Methodology: BB, BR, DM, LS, SA
Project administration: BR, DM, MM, SA
Resources: FA, JJV, JW
Software: BR
Supervision: CS, DM, JJV, JR, JW, MM, US
Visualization: BB, BR, LS, SA
Writing – original draft: BB, BR, DM, IS, LS, SA
Writing – review and editing: AH, AL, AP, BB, BR, CS, DM, FA, FPA, IS, JJV, JR, JW, LS, MH, MM, MS, SA, US

Conflicts of Interest

SA owns stock in HAVN Life Sciences, Heidelberg Pharma AG, and Pfizer. LS has received travel compensation from IQVIA. CS has received honorarium fees from the following pharmaceutical companies for scientific lectures, presentations, manuscript writing, educational events, or participation in scientific advisory boards: Gilead Sciences, GlaxoSmithKline, Janssen-Cilag, Merck Sharp & Dohme (MSD), Shionogi, Thera-Technologies, and ViiV Healthcare. Furthermore, the pharmaceutical companies AbbVie, Gilead Sciences, and Janssen have covered his travel, accommodation, and registration fees for scientific conference participation. DM has received speaker honoraria from the Free University Berlin and travel compensation from IQVIA. JJV has received personal fees from Merck/MSD, Gilead, Pfizer, Astellas Pharma, Basilea, German Centre for Infection Research (DZIF), University Hospital Freiburg/Congress and Communication, Academy for Infectious Medicine, University Manchester, German Society for Infectious Diseases (DGI), Ärztekammer Nordrhein, University Hospital Aachen, Back Bay Strategies, German Society for Internal Medicine (DGIM), Shionogi, Molecular Health, Netzwerk Universitätsmedizin, Janssen, NordForsk, Biontech, and APOGEPHA, and has received grants from Merck/MSD, Gilead, Pfizer, Astellas Pharma, Basilea, German Centre for Infection Research (DZIF), German Federal Ministry of Education and Research (BMBF), Deutsches Zentrum für Luft- und Raumfahrt (DLR), University of Bristol, Rigshospitalet Copenhagen, and Network University Medicine. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
Supplementary tables containing structural and semantic metadata as well as detailed study results.
[DOC File, 862 KB - publichealth_v12i1e81092_app1.doc]

Multimedia Appendix 2
Visual abstract.
[PNG File, 338 KB - publichealth_v12i1e81092_app2.png]

References

1. UNAIDS data 2024. : Joint United Nations Programme on HIV/AIDS; 2024 URL: https://www.unaids.org/en/resources/documents/2024/2024_unaids_data [accessed 2026-01-24]
2. Bätzing-Feigenbaum J, Kollan C, Kühne A, et al. Cohort profile: the German ClinSurv HIV project--a multicentre open clinical cohort study supplementing national HIV surveillance. HIV Med 2011 May;12(5):269-278. [doi: 10.1111/j.1468-1293.2010.00879.x] [Medline: 20955355]
3. Menéndez-Arias L, Delgado R. Update and latest advances in antiretroviral therapy. Trends Pharmacol Sci 2022 Jan;43(1):16-29. [doi: 10.1016/j.tips.2021.10.004] [Medline: 34742581]
4. Schmidt D, Kollan C, Stoll M. Long-acting prescriptions and therapy for HIV-1 from market launch to the present in Germany (May 2021 to December 2023). Front Public Health 2024 May 30;12:1404255. [doi: 10.3389/fpubh.2024.1404255] [Medline: 38873299]
5. Palella FJ, Baker RK, Moorman AC, et al. Mortality in the highly active antiretroviral therapy era: changing causes of death and disease in the HIV outpatient study. J Acquir Immune Defic Syndr 2006 Sep;43(1):27-34. [doi: 10.1097/01.qai.0000233310.90484.16] [Medline: 16878047]

XSL•FO
RenderX

6. Lau B, Gange SJ, Moore RD. Risk of non-AIDS-related mortality may exceed risk of AIDS-related mortality among individuals enrolling into care with CD4+ counts greater than 200 cells/mm3. J Acquir Immune Defic Syndr 2007 Feb 1;44(2):179-187. [doi: 10.1097/01.qai.0000247229.68246.c5] [Medline: 17075385]
7. Engels EA, Biggar RJ, Hall HI, et al. Cancer risk in people infected with human immunodeficiency virus in the United States. Int J Cancer 2008 Jul 1;123(1):187-194. [doi: 10.1002/ijc.23487] [Medline: 18435450]
8. Silverberg MJ, Abrams DI. AIDS-defining and non-AIDS-defining malignancies: cancer occurrence in the antiretroviral therapy era. Curr Opin Oncol 2007 Sep;19(5):446-451. [doi: 10.1097/CCO.0b013e3282c8c90d] [Medline: 17762569]
9. Clifford GM, Polesel J, Rickenbach M, et al. Cancer risk in the Swiss HIV Cohort Study: associations with immunodeficiency, smoking, and highly active antiretroviral therapy. J Natl Cancer Inst 2005 Mar 16;97(6):425-432. [doi: 10.1093/jnci/dji072] [Medline: 15770006]
10. Vogel M, Friedrich O, Lüchters G, et al. Cancer risk in HIV-infected individuals on HAART is largely attributed to oncogenic infections and state of immunocompetence. Eur J Med Res 2011 Mar 28;16(3):101-107. [doi: 10.1186/2047-783x-16-3-101] [Medline: 21486722]
11. 1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. MMWR Recomm Rep 1992 Dec 18;41(RR-17):1-19. [Medline: 1361652]
12. Grulich AE, van Leeuwen MT, Falster MO, Vajdic CM. Incidence of cancers in people with HIV/AIDS compared with immunosuppressed transplant recipients: a meta-analysis. Lancet 2007 Jul 7;370(9581):59-67. [doi: 10.1016/S0140-6736(07)61050-2] [Medline: 17617273]
13. Shiels MS, Cole SR, Kirk GD, Poole C. A meta-analysis of the incidence of non-AIDS cancers in HIV-infected individuals. J Acquir Immune Defic Syndr 2009;52(5):611-622. [doi: 10.1097/QAI.0b013e3181b327ca] [Medline: 19770804]
14. Engels EA, Pfeiffer RM, Goedert JJ, et al. Trends in cancer risk among people with AIDS in the United States 1980-2002. AIDS 2006 Aug 1;20(12):1645-1654. [doi: 10.1097/01.aids.0000238411.75324.59] [Medline: 16868446]
15. Crum-Cianflone N, Hullsiek KH, Marconi V, et al. Trends in the incidence of cancers among HIV-infected persons and the impact of antiretroviral therapy: a 20-year cohort study. AIDS 2009 Jan 2;23(1):41-50. [doi: 10.1097/QAD.0b013e328317cc2d] [Medline: 19050385]
16. Teeraananchai S, Kerr SJ, Amin J, Ruxrungtham K, Law MG. Life expectancy of HIV-positive people after starting combination antiretroviral therapy: a meta-analysis. HIV Med 2017 Apr;18(4):256-266. [doi: 10.1111/hiv.12421] [Medline: 27578404]
17. Masiá M, Gutiérrez-Ortiz de la Tabla A, Gutiérrez F. Cancer screening in people living with HIV. Cancer Med 2023 Nov;12(21):20590-20603. [doi: 10.1002/cam4.6585] [Medline: 37877338]
18. Re A, Cattaneo C, Rossi G. HIV and lymphoma: from epidemiology to clinical management. Mediterr J Hematol Infect Dis 2019;11(1):e2019004. [doi: 10.4084/MJHID.2019.004] [Medline: 30671210]
19. Rubinstein PG, Aboulafia DM, Zloza A. Malignancies in HIV/AIDS: from epidemiology to therapeutic challenges. AIDS 2014 Feb 20;28(4):453-465. [doi: 10.1097/QAD.0000000000000071] [Medline: 24401642]
20. Suneja G, Shiels MS, Angulo R, et al. Cancer treatment disparities in HIV-infected individuals in the United States. J Clin Oncol 2014 Aug 1;32(22):2344-2350. [doi: 10.1200/JCO.2013.54.8644] [Medline: 24982448]
21. Coghill AE, Han X, Suneja G, Lin CC, Jemal A, Shiels MS. Advanced stage at diagnosis and elevated mortality among US patients with cancer infected with HIV in the National Cancer Data Base. Cancer 2019 Aug 15;125(16):2868-2876. [doi: 10.1002/cncr.32158] [Medline: 31050361]
22. Prokosch HU, Acker T, Bernarding J, et al. MIRACUM: Medical Informatics in Research and Care in University Medicine. Methods Inf Med 2018 Jul;57(S 01):e82-e91. [doi: 10.3414/ME17-02-0025] [Medline: 30016814]
23. Winter A, Stäubert S, Ammon D, et al. Smart Medical Information Technology for Healthcare (SMITH). Methods Inf Med 2018 Jul;57(S 01):e92-e105. [doi: 10.3414/ME18-02-0004] [Medline: 30016815]
24. Ho D, Imai K, King G, Stuart EA. MatchIt: nonparametric preprocessing for parametric causal inference. J Stat Softw 2011;42(8):1-28. [doi: 10.18637/jss.v042.i08]
25. van Walraven C, Austin PC, Jennings A, Quan H, Forster AJ. A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. Med Care 2009 Jun;47(6):626-633. [doi: 10.1097/MLR.0b013e31819432e5] [Medline: 19433995]
26. Schober P, Vetter TR. Count data in medical research: Poisson regression and negative binomial regression. Anesth Analg 2021 May 1;132(5):1378-1379. [doi: 10.1213/ANE.0000000000005398] [Medline: 33857979]
27. Hijmans RJ. Rspatial/geosphere. GitHub. 2024. URL: https://github.com/rspatial/geosphere [accessed 2026-01-25]
28. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2024. URL: https://www.R-project.org [accessed 2026-01-25]
29. Venables WN, Ripley BD. Modern Applied Statistics with S-PLUS, 4th edition: Springer; 2002. [doi: 10.1007/978-0-387-21706-2]
30. Pebesma EJ, Bivand R. Spatial Data Science: With Applications in R: Chapman and Hall/CRC; 2023.
31. Brickman C, Palefsky JM. Cancer in the HIV-infected host: epidemiology and pathogenesis in the antiretroviral era. Curr HIV/AIDS Rep 2015 Dec;12(4):388-396. [doi: 10.1007/s11904-015-0283-7] [Medline: 26475669]
32. Wing EJ. HIV and aging. Int J Infect Dis 2016 Dec;53:61-68. [doi: 10.1016/j.ijid.2016.10.004] [Medline: 27756678]

XSL•FO
RenderX

33. Staadegaard L, Rönn MM, Soni N, et al. Immunogenicity, safety, and efficacy of the HPV vaccines among people living with HIV: a systematic review and meta-analysis. EClinicalMedicine 2022 Aug 3;52:101585. [doi: 10.1016/j.eclinm.2022.101585] [Medline: 35936024]
34. Gu A, Kaar S, Corey R, et al. Rotator cuff repair in HIV-infected patients: an analysis of postoperative complications. Orthop J Sports Med 2020 Jul 31;8(7_suppl6). [doi: 10.1177/2325967120S00412]
35. King MT, Hall J, Caleo S, Gurney HP, Harnett PR. Home or hospital? An evaluation of the costs, preferences, and outcomes of domiciliary chemotherapy. Int J Health Serv 2000;30(3):557-579. [doi: 10.2190/CY03-EV15-K38Y-X4AA] [Medline: 11109181]
36. Lau PK, Watson MJ, Hasani A. Patients prefer chemotherapy on the same day as their medical oncology outpatient appointment. J Oncol Pract 2014 Nov;10(6):e380-e384. [doi: 10.1200/JOP.2014.001545] [Medline: 25248724]
37. Akmatov MK, Hu E, Rüsenberg R, et al. Areas with high HIV prevalence: a spatial analysis of nationwide claims data in Germany. HIV Med 2024 Apr;25(4):498-503. [doi: 10.1111/hiv.13601] [Medline: 38072823]
38. Gaye A, Marcon Y, Isaeva J, et al. DataSHIELD: taking the analysis to the data, not the data to the analysis. Int J Epidemiol 2014 Dec;43(6):1929-1944. [doi: 10.1093/ije/dyu188] [Medline: 25261970]

Abbreviations

AD: AIDS-defining
ART: antiretroviral therapy
***ICD-10*:** *International Statistical Classification of Diseases and Related Health Problems, Tenth Revision*
KHEntgG: German Hospital Fees Act (German: Krankenhausentgeltgesetz)
NAD: non–AIDS-defining
OPS: operation and procedure classification system (German: Operationen- und Prozedurenschlüssel)
virus-NAD: cancer associated with HIV infection but not considered AIDS-defining

Edited by GJB Sousa; submitted 22.Jul.2025; peer-reviewed by M Christiansen, P Wang, S Sahai; revised version received 27.Nov.2025; accepted 16.Dec.2025; published 13.Feb.2026.

Please cite as:
Reiter B, Andreas S, Schumann L, Bender B, Schnorr I, Heidenreich A, Stephan C, Prunotto A, Laukhuf A, Wehrle J, Müller M, Aubele FP, Hagedorn M, Albashiti F, Seybold U, Roider J, Stecher M, Maier D, Vehreschild JJ
Clinical Epidemiology of Cancer in People Living With HIV in Germany: Retrospective, Observational, Multicenter Federated Claims Data Analysis
JMIR Public Health Surveill 2026;12:e81092
URL: https://publichealth.jmir.org/2026/1/e81092
doi:10.2196/81092

© Bastian Reiter, Stefanie Andreas, Linnea Schumann, Bernd Bender, Isabel Schnorr, Andreas Heidenreich, Christoph Stephan, Andrea Prunotto, Andrea Laukhuf, Julius Wehrle, Matthias Müller, Fabio Paul Aubele, Marlien Hagedorn, Fady Albashiti, Ulrich Seybold, Julia Roider, Melanie Stecher, Daniel Maier, Jörg Janne Vehreschild. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 13.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Acceptability and Feasibility of Longitudinal Sampling for Sexually Transmitted Enteric Infections in Gay, Bisexual, and Other Men Who Have Sex With Men (GBMSM): Prospective Cohort Pilot Study Conducted in 2022 in South East England

Holly Fountain[1], MSc; Katie Thorley[1], MSc; David Reid[2,3], MSc; Dana Ogaz[1,2], MSc; Daniel Richardson[4,5], FRCP, FAChSHM; Hannah Charles[1], MSc; Kate S Baker[6,7], PhD; Claire Jenkins[7,8], PhD; Caisey V Pulford[1,2], PhD; Gwenda Hughes[2,9], PhD; Nigel Field[3], MB, PhD; Catherine H Mercer[2,3], PhD; Xavier Didelot[7,10], PhD; Noel McCarthy[7,11], DPhil; Hamish Mohammed[1,2], PhD; Holly D Mitchell[1,2], PhD

[1]Blood Safety, Hepatitis, STI and HIV Division, UK Health Security Agency, London, United Kingdom
[10]School of Life Sciences and Department of Statistics, University of Warwick, Coventry, United Kingdom
[11]School of Medicine, Trinity College Dublin, Dublin, Ireland
[2]University College London, NIHR Health Protection Research Unit in Blood Borne and Sexually Transmitted Infections, London, United Kingdom
[3]Institute for Global Health, University College London, Mortimer Market Centre, London, United Kingdom
[4]University Hospitals Sussex NHS Foundation Trust, Brighton, United Kingdom
[5]Brighton and Sussex Medical School, Brighton, United Kingdom
[6]Department of Genetics, University of Cambridge, Cambridge, United Kingdom
[7]University of Liverpool, NIHR Health Protection Research Unit in Gastrointestinal Infections, Liverpool, United Kingdom
[8]Gastrointestinal Bacteria Reference Unit, UK Health Security Agency, London, United Kingdom
[9]Department of Infectious Disease Epidemiology and Dynamics, London School of Hygiene and Tropical Medicine, London, United Kingdom

Corresponding Author:
David Reid, MSc
University College London, NIHR Health Protection Research Unit in Blood Borne and Sexually Transmitted Infections, London, United Kingdom

Abstract

Background: In the last 2 decades, there has been an increasing number of sexually transmissible enteric infection (STEI) outbreaks among gay, bisexual, or other men who have sex with men (GBMSM). There remain important gaps in our understanding of how STEI transmission is sustained that repeated collection of samples could help to address.

Objective: This study aimed to assess the feasibility and acceptability of longitudinal samples and epidemiological data collection among GBMSM accessing sexual health services (SHS) through a prospective cohort study.

Methods: GBMSM (≥16 years) accessing 2 SHS in Brighton and Sussex were recruited between May and October 2022. Participants provided an initial rectal swab and optional fecal sample and completed an online baseline questionnaire. Weekly follow-up questionnaires and rectal swabs were collected for a further 11 weeks. Sexually transmitted infection (STI) surveillance data were pseudonymously linked to provide additional clinical and demographic information. Selected participants were invited to take part in an optional one-to-one interview. We assessed the number completing study procedures, characteristics of those completing procedures and not, and representativeness of the study sample alongside facilitators and barriers from interviews. A chi-square test was used to compare groups.

Results: Overall, 193 participants were recruited. Half (100/193, 51.8%) provided a baseline rectal swab, with a third (34/100, 34.0%) of them providing all 12 swabs. Alongside the baseline swab, 76.0% (76/100) provided the optional fecal sample. Just over a third (71/193, 36.8%) of participants completed a baseline questionnaire, with a fifth (15/71, 21.1%) providing all follow-up questionnaires. Interviews (n=21) found that participation was motivated by the feeling of giving something back for services received and a perceived indirect benefit to self. The study was generally accepted, with over half reflecting on the perceived ease of participation and relatively simple tasks that could be easily integrated into normal routine with an element of flexibility. Most participants were satisfied with the 12-week study length, and having a definitive end point aided the ability to assess if they would be able to participate. Barriers to completing the study procedures included not being aware of what was required, particularly for the follow-up questionnaires. Suggested improvements included concise and easier-to-read instructions, with a

XSL•FO
RenderX

section to clearly list out the key procedures of the study. SMS text messaging reminders were sent, but these were seen with variable utility and interpreted in different ways (eg, personal or generic reminders or thank you messages).

Conclusions: This study has provided evidence that longitudinal rectal swab sampling and data collection for research purposes are feasible and acceptable among GBMSM attending SHS. This provides an innovative way to address important knowledge gaps about STEI transmission, which will help inform public health measures for infection control. Tangible insights from this pilot can also inform the design of similar studies in other settings.

(JMIR Public Health Surveill 2026;12:e73762) doi:10.2196/73762

KEYWORDS

sexually transmitted diseases; sexually transmitted enteric infection; gay, bisexual, or other men who have sex with men; longitudinal sampling; acceptability; feasibility

Introduction

Enteric pathogens are classically transmitted by ingesting food or water contaminated by fecal matter, causing gastroenteritis, hepatitis, and malabsorption [1]. Over the past 50 years, it has been recognized that these pathogens can also be transmitted directly from person to person during sexual activity. Gay, bisexual, or other men who have sex with men (GBMSM) are particularly at risk of sexually transmissible enteric infections (STEIs) through sexual behaviors such as oral-anal sex (rimming), oral sex, scat play, or indirectly through oral sex after anal sex [1,2]. Transmission is likely also facilitated amongst groups of GBMSM in dense sexual networks through chemsex, group sex, and the use of sexual networking applications [3,4].

The number of reported STEI outbreaks among GBMSM globally has increased [1,5-7], with circulating strains often associated with antimicrobial resistance (AMR) [8-10], causing a growing public health concern. Antimicrobial-resistant STEIs, including *Campylobacter* spp. and Shiga toxin–producing *Escherichia coli*, and extensively drug-resistant *Shigella* spp., have all been reported, compromising management with antimicrobials [11]. Previous studies have suggested that the acquisition of AMR determinants could be linked to selection pressure from high levels of antimicrobial use in GBMSM for treating bacterial sexually transmitted infections (STIs) [8,12]. However, these mechanisms are not fully understood, and other factors, such as the gut microbiota, may also play a role in the development and severity of gastrointestinal disease and the transfer of AMR genes [13].

There also still remain important gaps in our understanding about the duration of STEI carriage and infectiousness and how this might contribute to sustained transmission of STEIs in GBMSM sexual networks. An analysis of *Shigella* spp. isolates from symptomatic adult men from the general population reported a median time interval of 23.5 days (range 6 - 176 days; IQR 8 - 70) between isolates recovered from the same individual [14], suggesting possible persistent, recurrent, or recrudescent infection or carriage (as no intervening clinical data were available). This highlights the need for longitudinal studies with matched questionnaire data, such as symptom status, the type or duration of reported symptoms, any antimicrobials taken, and any sexual partners or specific sexual behaviors.

Other studies have shown that asymptomatic carriage may play a key role in sustaining transmission of enteric pathogens among specific sexual networks of GBMSM [12,15,16] and may be a significant barrier to effective control. A meta-analysis provided estimates of prevalence of STEIs among asymptomatic GBMSM, ranging from 0.3% to 3.8% [17]. The analysis included 6 studies, of which only 2 explored potential risk factors, one from Melbourne, Australia, where STEI detection was associated with insertive oral-anal sex and group sex [16], and one from London, United Kingdom, where STEI detection was associated with having 5-9 new sexual partners in the last 3 months (as compared to 0-1 new partners), a bacterial STI in the last year, and the use of HIV pre-exposure prophylaxis (PrEP) or living with HIV [12]. While these studies contributed to our wider understanding of STEI transmission and the importance of certain sexual practices, they were cross-sectional and did not provide information on duration of carriage and did not include detailed information on gastrointestinal symptoms, antimicrobial treatment, or travel history.

Thus, there is a need to better understand the factors that might be sustaining transmission of, and the development of AMR in, enteric pathogens among sexual networks of GBMSM to inform the development, targeting, and delivery of clinical and public health measures that control transmission. To address these knowledge gaps, a possible study design is the repeated collection of fecal samples and/or rectal swabs, in addition to clinical and behavioral data, over a period of several months. The longitudinal samples could be sequenced and compared to study within-host evolution [18] and transmission pathways [19]. As a first step toward addressing the knowledge gaps, we conducted a mixed methods observational pilot study of sexually transmitted enteric infections among GBMSM (STEIM). We aimed to assess the feasibility and acceptability of longitudinal rectal swab sample and epidemiological data collection among GBMSM attending sexual health services (SHSs). Here we describe the characteristics of participants, completion of study procedures, and facilitators and barriers to completing the study.

Methods

Study Design

We conducted a longitudinal, observational prospective cohort study among GBMSM that accessed 2 SHSs in Brighton and Sussex between May 23 and October 11, 2022, with samples collected from May 26, 2022, to January 24, 2023. This region

XSL•FO
RenderX

of England has high numbers of diagnoses of STEIs in GBMSM [20], making it a suitable setting for the pilot. Participants were asked to consent to provide an initial rectal swab and optional fecal sample and complete a baseline questionnaire upon enrollment. Follow-up questionnaires and rectal swabs were collected at weekly intervals for a further 11 weeks. Samples were sent to the Gastrointestinal Bacteria Reference Unit at the UK Health Security Agency (UKHSA) and tested for a range of enteric pathogens, and results were linked to questionnaire responses. Where participants provided specific informed consent, study data were linked pseudonymously (without using direct identifiers such as name or date of birth) to the national GUMCAD STI Surveillance System using unique clinic and patient numbers [21]. Selected participants were invited to take part in an optional one-to-one interview to discuss their experience of the study. These were conducted between October 2022 and March 2023.

Study Setting

The study took place within Brighton and Sussex at SHSs provided by University Hospitals Sussex NHS Foundation Trust. In England, SHSs are free and confidential, and they are available to anyone regardless of whether the person is resident in the local area.

Study Population

All men (cis or transgender), transgender women, or gender-diverse people, aged 16 years or older, who attended the participating SHSs during the study period and reported sex with a man (cis or transgender) or nonbinary person assigned male at birth in the previous 3 months were eligible. The study excluded those not able to provide informed consent and/or complete the study procedures due to English language or literacy. Those unable to use or access a suitable mobile device or internet were also excluded. The study aimed to recruit a representative sample of GBMSM attending the SHSs for routine STI screening or care, regardless of symptoms.

Recruitment

All eligible participants who attended for routine sexual health care during the study period were to be approached for recruitment by SHS staff. The participation information sheet (PIS) provided for recruitment was available online during recruitment, along with the study protocol, to provide more information to potential participants.

Where eligible SHS attendees declined to participate, SHS staff were asked to complete a separate screening log to document the reasons for declining without recording personal data. Informal insights from SHS staff provided further information on the feasibility and acceptability of recruitment.

SHS staff contacted participants who agreed to be contacted about participating in a one-to-one interview. Purposive sampling was used initially to ensure representation across ages and different levels of participation in study procedures. The invite was extended to all consenting participants in the latter stages of the study.

Sample Size

A recruitment target of 200 participants was set, based on our estimates of how many GBMSM attend the SHSs and how many will agree to participate in the time frame set for the study. No formal sample size calculation was completed, as the pilot was designed to inform sample size calculations for future larger studies. We aimed to conduct up to 20 interviews, with this estimated to be a sufficient number to reach saturation of issues related to acceptability and feasibility [22,23].

Study Procedures

At enrollment, participants were provided with 12 rectal swabs and a fecal sample collection kit, including prepaid envelopes. For completion of the first rectal swab (the baseline rectal swab), there was the option to provide a concurrent fecal sample on the same day. After baseline, a rectal swab was to be returned each week for a further 11 weeks. Participants recorded the date of collection on the sample tube before posting.

Separate web-based questionnaires were completed at enrollment (baseline) and at each weekly follow-up. The surveys were hosted on a secure remote server (Snap Surveys Ltd) and were completed using personal digital devices. The access links to the questionnaires were available through the study web page; participants were required to enter their unique study ID only. For those who agreed to it, one-to-one interviews were also conducted to understand how best to minimize barriers and better facilitate participation and retention in the study.

For participants who gave consent, a weekly email or SMS text message reminder about the study procedures was sent on the same day every week (the day that they were recruited). Participants could choose the day of the week they completed the procedures and could continue to take part in the study even if they did not provide all the requested samples. Providing an incentive to participants was explored during the study design phase, and the decision was made not to pursue this given the study was a pilot and the aim was to assess feasibility and acceptability, including barriers and motivators for taking part. The concept of incentives was explored during the interviews.

Data Sources and Outcomes

The data sources included behavioral and clinical data from the baseline and follow-up questionnaires; the baseline questionnaire was longer (median 8.3 minutes to complete) and had a longer look-back period than the weekly follow-up questionnaire (median 2.8 minutes to complete; Table 1 and Multimedia Appendices 1 and 2). The questionnaires did not contain any standardized tools or instruments; however, they were developed by drawing upon questionnaires used in other studies [24,25], and comprehension and clarity of the questions were checked during the study design phase.

XSL•FO
RenderX

Table . Overview of data collected and look-back periods for the baseline and follow-up questionnaires for the sexually transmitted enteric infections among gay, bisexual, or other men who have sex with men (GBMSM; sexually transmitted enteric infections among GBMSM [STEIM]) study.

Data collected	Baseline questionnaire	Look-back period	Follow-up questionnaire	Look-back period
Reason for attending SHS[a] when recruited	Yes	N/A[b]	N/A	N/A
HIV-related factors				
HIV status and date of most recent test	Yes	No time limit	No	N/A
Antiretroviral treatment (living with HIV)	Yes	No time limit	No	N/A
Viral load count (living with HIV)	Yes	No time limit	No	N/A
HIV Pre-exposure prophylaxis (PrEP; HIV negative)	Yes	No time limit	No	N/A
Sexual behavior				
Number of sexual contacts (and additional information on partners, for example new or casual)	Yes	3 months	Yes	1 week
Type of sex	Yes	3 months	Yes	1 week
Frequency of condom use	Yes	3 months	Yes	1 week
Drug use before or after sex				
Use and frequency of listed drugs before or during sex	Yes	3 months	Yes	1 week
Injection drug use	Yes	3 months	Yes	1 week
Antibiotic use				
Consumption of antibiotics for any reason	Yes	3 months	Yes	1 week
Use of antibiotics as STI[c] prophylaxis	Yes	3 months	Yes	1 week
Gastrointestinal symptoms				
Gastrointestinal symptoms and duration	Yes	3 months	Yes	1 week
Sexual contacts with gastrointestinal symptoms	Yes	3 months	Yes	1 week
Other contacts (eg, household) with gastrointestinal symptoms	Yes	3 months	Yes	1 week
Travel				
Travel outside the United Kingdom	Yes	3 months	Yes	1 week
New sexual partners while abroad	Yes	3 months	Yes	1 week

[a]SHS: sexual health service.

[b]N/A: not applicable.

[c]STI: sexually transmitted infection.

One-to-one interviews were conducted using a semistructured topic guide. This was developed collaboratively and designed to allow for data on participation motivation, research experience, understanding of compliance with and the acceptability of study requirements and procedures, barriers and facilitators to study participation and retention, and suggestions for improvement for subsequent research to be collected. Interviews were conducted by one member of the STEIM study team (D Reid) and were conducted as remote audio or video interviews using Microsoft Teams. All interviews

XSL•FO
RenderX

were also transcribed through Microsoft Teams, with transcriptions reviewed by the interviewer. As this was a pilot study, interviews were not piloted beforehand.

Additional clinical and demographics information was sourced from GUMCAD. GUMCAD is the mandatory surveillance system for STIs in England and collects data on STI tests and diagnoses from all publicly commissioned SHSs [21]. For consenting participants, pseudonymous linkage to GUMCAD data was conducted to provide additional clinical and demographic information, including recent STI diagnoses (*Chlamydia trachomatis*, *Neisseria gonorrhoeae*, or *Treponema pallidum* [primary, secondary, or early latent] at enrollment and in the past 12 months); HIV status; HIV PrEP use; sexual orientation; ethnicity and country of birth; and patient residence to allow for determination of index of multiple deprivation (IMD).

Outcomes of the study included the number and proportion of participants providing samples and questionnaires at various time points throughout the study, for example, how many completed baseline procedures and how many completed all weeks; description of the characteristics of the study population; comparison of the characteristics of participants who did and did not return samples and questionnaires; whether the study population was representative of all GBMSM attending the SHSs during the study period; and facilitators and barriers to the study as drawn out from the one-to-one interviews. Qualitative interview results were used to contextualize the quantitative results and to provide a full picture of the acceptability and feasibility of the pilot study.

Data Management and Analysis

All quantitative data were managed and analyzed using Microsoft Excel and STATA (v17; StataCorp LLC). Qualitative data were analyzed using NVivo (20; Lumivero).

GUMCAD records were linked to the enrollment log using a deterministic linkage approach based on participants' patient number, SHS attended, and SHS attendance date. Data matches were accepted where the GUMCAD SHS attendance date was within 30 days of the participant enrollment date, but information on concurrent STI diagnosis at baseline was only included where the dates matched exactly. After records were linked, the GUMCAD patient number was dropped. The GUMCAD data were subsequently linked to the laboratory testing results and questionnaire responses using the study ID. Where responses from the questionnaire differed from GUMCAD data on the same topic, the questionnaire response was used for analysis (being the most up-to-date and less subjective to clinician or coder bias) and replaced the input from GUMCAD. When the date of collection of rectal swabs was not provided (254/949, 26.8%), the date of receipt at UKHSA was used.

Summary metrics, such as the median number of samples returned, were used to assess the feasibility of longitudinal rectal swab collection and completion of self-administered questionnaires. Participant characteristics were analyzed and summarized using data from GUMCAD and questionnaires. To minimize the risk of deductive disclosure, in some cases, small numbers (1-4 inclusive) were masked and referred to as "<5" or groups containing small numbers have been aggregated. The characteristics of participants enrolled in the study who did and did not return at least one sample or completed questionnaire were compared using the chi-square test.

The representativeness of the study sample was ascertained by comparing the demographic characteristics, obtained from the GUMCAD data, to participating SHS attendees during the study period, excluding the study participants, who reported as GBMSM; the chi-square test was used for statistical comparisons.

An analytic "framework" approach [26] was used to analyze the interview transcripts. Initial deductive and inductive thematic codes were derived using the interview topic guides and 3 transcripts by 2 members of the STEIM study team (D Reid and DO). Wider themes were discussed to agree on all themes and coding structure before the agreed codes were then applied to all transcripts (by D Reid) and data were charted, summarized, and interpreted. As this was a pilot study, we did not formally assess saturation, and the number of interviews was set in advance.

Ethical Considerations

This study was reviewed and approved by the London and South East NHS Research Ethics Committee (Reference 21/LO/0891) and the University Hospitals Sussex NHS Foundation Trust Research and Development department. Written informed consent was obtained from all participants prior to participation (Multimedia Appendix 3), and participants could withdraw from the study at any time. Those willing to take part in an optional one-to-one interview completed a separate interview consent form (Figure S1 in Multimedia Appendix 4). In addition, consent for (1) linkage to GUMCAD data, (2) residual study samples to be stored for future ethically approved research studies, (3) text or email reminders to be sent about completing the follow-up procedures each week, and (4) participating in the one-to-one interview was recorded. All four of these tasks were optional to agree to. All data were pseudonymized using a unique study identifier before sending to researchers, so they did not have access to personal identifying information. No images are provided in the manuscript or supplementary material that would allow identification of individual participants. No type of incentive was given to participants.

During the study design phase, cognitive interviews with GBMSM were used to explore general interest and acceptability of the study proposal and methodology as part of the patient and public engagement and involvement strategy. The interviews also explored acceptability, comprehension, and clarity of the participant information documents and questionnaires. Suggestions for improvement to methods and study materials were incorporated.

XSL•FO
RenderX

Results

Quantitative Results

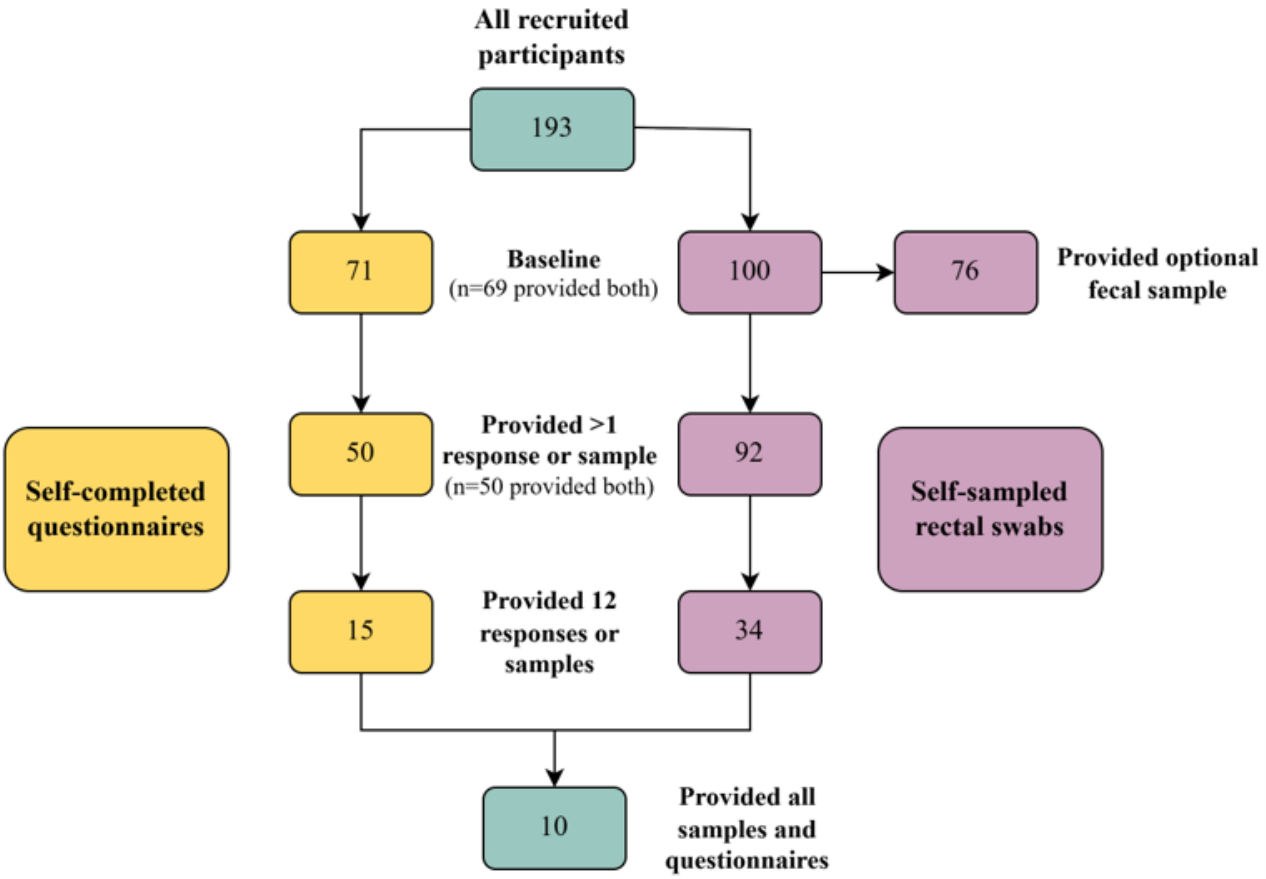
Recruitment and Study Completion Rates

Overall, 193 people were recruited into the STEIM study between May 23, 2022, and October 11, 2022 (96.5% of the planned 200 to be recruited). Screening logs were completed for 23 eligible SHS attendees who did not agree to take part; not having the time or capacity to participate was cited as a reason by half (12/23), often due to travel abroad or relocating (7/12). Others perceived they had too few recent sexual partners or had not engaged in specific sexual behaviors to be suitable for the study (7/23), lacked access or proficiency using the internet (n<5), or did not want to provide the samples required (n<5). Half (12/23, 50%) of the screening log entries were from the Sussex SHS. Informal discussions with SHS staff indicated that the screening log was not routinely completed, meaning a lack of accurate records of those declining to participate. They also highlighted that the most frequently heard refusal was a lack of interest in participating. Other reasons included being put off by rectal swabbing and the lack of returned diagnostic results.

Most participants were recruited at the SHS in Brighton (168/193, 87.0%), the largest SHS in the study area. Baseline questionnaires were completed by 71 participants (71/193, 36.8%), among which most were recruited while attending for a vaccination (28/71, 39.4%), a general sexual health check (22/71, 31.0%), and/or to receive HIV PrEP (20/71, 28.2%). Informal discussions with SHS staff suggested that recruitment was often conducted during mpox vaccination clinics (the opening of recruitment coincided with the start of the mpox outbreak in the United Kingdom in the summer of 2022).

Of those recruited, 51.8% (100/193) provided a baseline rectal swab (Figure 1), among which 34.0% (34/100) provided all 12 swabs (Figure S2 in Multimedia Appendix 4 shows more details), with a median of 11 rectal swabs submitted (IQR 5-12). Those providing all 12 swabs did so over 11-23 weeks, with 9 participants providing swabs for 14 or more weeks. Among these participants, 76.0% (76/100) also provided an optional fecal sample. Of those who completed a baseline questionnaire, 21.1% (15/71) completed all follow-up questionnaires (Figure 1) over 11-15 weeks. There were 11 participants who submitted follow-up questionnaires but no baseline; these participants were not included in completed questionnaire counts, and most submitted 3 or fewer questionnaires (7/11, 63.6%), with only one submitting all follow-up questionnaires.

Figure 1. Flowchart of sexually transmitted enteric infections among gay, bisexual, or other men who have sex with men (GBMSM; sexually transmitted enteric infections among GBMSM [STEIM]) study participant recruitment and participation by study component (self-completed questionnaires and self-sampled rectal swabs and fecal sample).



XSL•FO
RenderX

Feasibility of Data Linkage

Most participants (188/193, 97.4%) agreed to GUMCAD data linkage, and linkage was successful for 97.9% (184/188). Of the 102 participants with a baseline sample and/or questionnaire, 98.0% (100/102) agreed to linkage, and 98% (98/100) were successfully linked to a GUMCAD record. Demographic data were obtained for all of those successfully linked; concurrent STI diagnosis at baseline was available for 97.3% (179/184) of participants.

Characteristics of Study Participants

Most study participants were UK-born (141/184, 76.6%) and of white ethnicity (171/184, 92.9%; Table 2), with a median age of 43 years (IQR 33-52). At baseline, most participants were not known to be living with HIV (134/184, 72.8%), of which 47.8% (64/134) were taking HIV PrEP. While only 5.0% (9/179) of participants were diagnosed with a bacterial STI on the day of enrollment, 38.0% (70/184) had a bacterial STI diagnosed in the previous year.

XSL•FO
RenderX

Table . Characteristics of sexually transmitted enteric infections among gay, bisexual, or other men who have sex with men (GBMSM; sexually transmitted enteric infections among GBMSM [STEIM]) study participants compared to all GBMSM sexual health service (SHS) attendees (GUMCAD national sexually transmitted infection [STI] data) between May 2022 and October 2022.

Characteristics	Study participants[a,b] (n=184)	SHS[c] attendees, excluding study participants (n=4238)	*P* value
Age (years), n (%)			.06
16-24	6 (3.3)	339 (8.0)	
25-34	47 (25.5)	913 (21.5)	
35-44	44 (23.9)	987 (23.3)	
45-64	79 (42.9)	1688 (39.8)	
≥65	8 (4.3)	311 (7.3)	
Unknown/missing	0 (0)	0 (0)	
Sexual orientation, n (%)			N/A[d]
Heterosexual	1 (0.5)	N/A	
Gay	164 (89.1)	N/A	
Bisexual	9 (4.9)	N/A	
Other, not listed	10 (5.4)	N/A	
Unknown	0 (0.0)	N/A	
Ethnicity[e], n (%)			.12
White	171 (92.9)	3736 (88.2)	
All other ethnic groups	9 (4.9)	387 (9.1)	
Unknown/missing	4 (2.2)	115 (2.7)	
Region of birth[f], n (%)			.67
United Kingdom	141 (76.6)	3222 (76.0)	
All other regions of birth	43 (23.4)	998 (23.5)	
Unknown	0 (0.0)	18 (0.4)	
Index of multiple deprivation (IMD) quintile, n (%)			.94
Most deprived	34 (18.5)	807 (19.0)	
2	43 (23.4)	1041 (24.6)	
3	52 (28.3)	1126 (26.6)	
4	34 (18.5)	725 (17.1)	
Least deprived	16 (8.7)	444 (10.5)	
Unknown/missing	5 (2.7)	95 (2.2)	
HIV status with PrEP[g] status, n (%)			<.001
Not known to be living with HIV, taking PrEP	64 (34.8)	608 (14.3)	
Not known to be living with HIV, not taking PrEP	70 (38.0)	1854 (43.7)	
Living with HIV	50 (27.2)	1776 (41.9)	

XSL•FO
RenderX

Characteristics	Study participants[a,b] (n=184)	SHS[c] attendees, excluding study participants (n=4238)	*P* value
STI in the previous year[h], n (%)			<.001
Yes	70 (38.0)	862 (20.3)	
No/unknown	114 (62.0)	3376 (79.7)	
Concurrent STI[i] (ie, STI diagnosed on day of enrollment)[h], n (%)	n=179		N/A
Yes	9 (5.0)	N/A	
No/unknown	170 (95.0)	N/A	

[a]Data obtained from routine returns to the GUMCAD STI surveillance system and baseline questionnaire entries, where eligible for the study participants only.

[b]Study participants were eligible if reporting sex with a man (cis/transgender) or nonbinary person assigned male at birth in the past 3 months; as this same level of detail on sexual partners is not available in GUMCAD, the comparison is made to all male individuals with sexual orientation reported as gay or bisexual in GUMCAD data rather than explicitly those reporting sex with a man in the past 3 months.

[c]SHS: sexual health service.

[d]N/A: not applicable.

[e]All other ethnic groups include groupings for those of Asian ethnicity, Black ethnicity, mixed ethnicity, and any other ethnicity. These data were aggregated into one group due to small data points.

[f]All other regions of birth include groupings for those born in the Caribbean, Central and South America, North America, Europe, South Asia, sub-Saharan Africa, and any other region. These data were aggregated into one group due to small data points.

[g]PrEP: pre-exposure prophylaxis.

[h]Concurrent or previous STI was considered as *Chlamydia trachomatis*, *Neisseria gonorrhoeae*, or *Treponema pallidum* (primary, secondary, or early latent) diagnosis.

[i]STI: sexually transmitted infection.

From the baseline questionnaires, the most common number of sexual partners in the previous 3 months was between 2 and 4 (29/71, 40.8%; Table 3). During that time, very few participants had used STI prophylaxis (taken antimicrobials immediately before or after sex to prevent STIs other than HIV; 2/71, 2.8%; missing data n=1, 1.4%), and a majority of participants (69.0%, 49/71) reported gastrointestinal symptoms, with diarrhea being most common (42/49, 85.7%) followed by mucus and/or blood in feces (23/49, 46.9%). Just under half (31/71, 43.7%) had traveled abroad, and about a third (10/31, 32.3%) of those reported sex abroad.

XSL•FO
RenderX

Table . Sexually transmitted enteric infections among gay, bisexual, or other men who have sex with men (GBMSM; sexually transmitted enteric infections among GBMSM [STEIM]) study participant behavioral characteristics, from baseline questionnaires between May 2022 and January 2023.

Characteristics	Number of participants (n=71), n (%)
Number of sexual partners (past 3 months)	
1	7 (9.9)
2 - 4	29 (40.8)
5 - 10	18 (25.4)
11 - 19	4 (5.6)
20 - 99	11 (15.5)
Don't know	2 (2.8)
STI[a] prophylaxis	
No	68 (95.8)
Yes	2 (2.8)
Don't know	1 (1.4)
Travel abroad	
No	40 (56.3)
Yes	31 (43.7)
Don't know	0 (0.0)
Sex abroad	
No	19 (61.3)
Yes	10 (32.3)
Not answered	2 (6.5)
Not applicable	40
GI[b] symptoms (past 3 months)	
No	22 (31.0)
Yes	49 (69.0)
Don't know	0 (0.0)

[a]STI: sexually transmitted infection.

[b]GI: gastrointestinal

Nearly half (89/193, 46.1%) of participants did not complete questionnaires or provide samples after initial recruitment. These participants had similar characteristics to those who did provide samples and/or questionnaires (Table S1 in Multimedia Appendix 4). However, participants who provided completed samples or questionnaires were more likely to be from the least deprived IMD quintiles (*P*=.004). Participants returning samples or questionnaires, and who were not known to be living with HIV, were more likely to be taking HIV PrEP (46.5% vs 21.2%; *P*=.001) and less likely to be diagnosed with a concurrent bacterial STI at enrollment (2.1% vs 8.3%; *P*=.06).

Representativeness of Study Population

Recruited participants were representative of GBMSM attending the participating SHSs within the study period for most characteristics. Most GBMSM attending the participating SHSs were also UK-born (76.0%) and of White ethnicity (88.2%; Table 2), with a median age of 43 (IQR 33-54) years. However, more study participants were taking HIV PrEP (34.8% vs 14.3%; *P*<.001) and more had experienced a bacterial STI in the previous year (38.0% vs 20.3%; *P*<.001) compared to all GBMSM attending the SHSs.

Qualitative Results

Characteristics of Interview Participants

Most participants (153/193, 79.3%) agreed to be contacted for an interview, with 21 participants completing interviews. Interviewees were slightly older on average than all participants, ranging from 28 to 76 years old with a median of 49 (IQR 43-55) years, compared to a median of 43 (IQR 33-52) years for all participants. Most interviewees reported some pre-STEIM research participation (17/21), and they had better completion of study procedures than the overall study participants, with almost all interviewees submitting a fecal sample (20/21) and all submitting one or more rectal swabs.

Motivations for Study Participation

Qualitative findings from interviews showed that the context in which participants were recruited may have motivated many

XSL•FO
RenderX

to consider joining the study, with most interviewees mentioning concurrent receipt of mpox vaccination or other valued services:

> *I'm really grateful for having the monkeypox vaccine so quickly and because of that, I'm going to take part in this slightly uncomfortable study as my way of giving something back.*

Taking part in STEIM was perceived as a welcome opportunity for many to give back to SHSs.

Almost all participants referred to indirect benefits to self through helping others or achieving a collective benefit for peers. More than half (12/21) of the interviewees suggested they were interested in the STEIM study or studies related to gut health or the sexual health of gay men more generally. The relative novelty of a sexual health study focused on gut infections was appealing and interesting to many. Two participants did mistakenly expect that tests on their samples that identified an infection would be communicated either to medical staff or themselves and therefore incorrectly inferred a potential direct benefit in early diagnosis and potential treatment.

All interviewees had some initial discussion with one or more SHS staff members before agreeing to participate. Staff members who were known, trusted, and enthusiastic about the study encouraged participation. Reassurances that tasks were not onerous and perceptions of staff as knowledgeable and able to provide ongoing support or advice also promoted a decision to participate.

Many participants noted that an incentive to take part would not personally be a motivator to take part in the study, but the acceptability of an incentive offer was reported by a majority. Financial incentives were mentioned, and it was suggested that these should be easy to use, useful, small, and targeted at those in need. Some participants suggested they may also be incentivized by receiving study results instead of a monetary reward; from one participant,

> *I'm never going to say no to money, but it's ok if someone wants to give you money for the thing. But I'd be happy if it was just a general, either like a feedback, personal health monitoring thing or seeing the sort of results of the study. So, seeing what it's accomplishing, hearing sort of updates on what the research is finding.*

Facilitators and Barriers to Participation and Completion of Procedures

Over half of interview participants expressed no reservations about participating in the study, reflecting on the perceived ease of participation and relatively simple and not onerous tasks. From one participant,

> *Once you get into it, it was very straightforward and easy. You know, I think for me; do the sample, fill in the survey, 5 minutes, not an issue, you know and, tick the box and job done. Really.*

Study time requirements and potential disruption to normal life were important considerations for more than half, though generally assessed as acceptable. A few participants related that a definite study end point aided them in assessing the time required when considering other commitments such as work or holidays. Most were highly satisfied with the 12-week pilot length, a majority (16/21) of whom felt it was a preferable maximum. Three months is a period most felt they could commit to participate and in which interest and remembering tasks could be maintained. From one participant, "*Probably about this [length] would have been the longest. I wouldn't have done it if it was much more*." However, half the participants (11/21) would consider a longer study duration, and a third reported that they would consider 6 months to 1 year, and a small minority may consider an indefinite period.

The ability to integrate and routinize the study into participants' lives facilitated participation. The flexibility to choose a preferred day of the week to complete the study procedures, with allowances and leeway of a day or 2, was appreciated and reported as useful. Many participants recounted the ease of taking rectal swabs as part of their morning routine or in posting samples soon after on the way to work, during a walk, or when passing a postbox. The weekly repetition of tasks helped to aid memory, and over half the participants discussed the utility of setting their own reminders, alarms, or in-diary tasks positively.

Many discussed the variable or limited utility of text reminders sent by the clinic. These texts sometimes came on a different day from that set by the participant and were perceived in different ways as reminders, either targeted personally or automatic and generic or as thanks or acknowledgment of completion and receipt of samples. These interpretations were influenced by whether they arrived before or after the day that samples were taken, posted, or received and by previous interpretations or expectations. The utility of these texts was enhanced if they were seen to improve a sense of study connection and engagement, which participants suggested might be increased by a personal rather than generic focus, being tailored to tasks and including messages of appreciation for participation as well as feedback on study findings.

Reception to the PIS and other written instructions was mixed, with participants often suggesting improvements to these. In reading the PIS, some participants reported that they missed aspects of the study methodology or were left uncertain and made suggestions for step-by-step instructions or bullet points of the key tasks.

> *I think the way that it's all presented, a patient information sheet, it's quite sort of quite a lot of dense information and lots of separate paragraphs. […] Just too much text and I think I could have done with a little reminder right at the top, maybe at the top of that just bullet points like every week, one bullet saying take your swab and send it off and then another bullet point saying complete the questionnaire.*

Facilitators and Barriers to Rectal Swab and Stool Sample Collection

The ability to provide the specimens at home, allowing for privacy and saving time and potential SHS visits, was appreciated and desired. One interviewee stated,

XSL•FO
RenderX

> *… I think the fact that I could do it myself and post them in was one of the things that convinced me it would be fine to do it.*

Interviews suggested that there were differences in the interpretation of the end point of the study; most interviewees reported believing that the end point was when all swab samples were submitted, but others identified 12 weeks as an end point even if some samples were missed, which was the intended outcome.

Instructions for rectal swab collection were described as useful for understanding procedures, and those for stool sample collection detailed best practices, though many men described using methods of collection that were slightly adapted to suit their own situation and resources, for example, using a toilet paper nest or a Tupperware container. Most interview participants were happy to provide rectal swabs, relating a perception of rectal swab collection that was relatively simple.

There was less general acceptance for the fecal samples. Interview participants perceived stool sample collection as more difficult than rectal swabbing. They discussed difficulties, such as disgust in handling feces, less familiarity with collection, requirements to follow body processes and timings, and more complexity in planning and execution. This included needing a private toilet and more complex requirements for preparing and taking the sample, getting the sample into the correct container in the required amount without contamination, and disposing of or cleaning equipment or containers. Overall, a quarter of participants (5/21) felt it was a difficult thing to do; the rest were relatively comfortable while also acknowledging less ease and greater effort were required than in taking rectal swabs [1].

Facilitators and Barriers to Questionnaire Completion

Interviews confirmed that the aspect of the study that proved the most challenging or difficult to complete as instructed was the online weekly questionnaires. No participant reported returning a questionnaire on the same day every week for a consecutive 12-week period, and almost a third (6/21) reported being unaware of a requirement to complete a weekly questionnaire at all. From one participant,

> *Oh, I didn't get that because I think when I read and I got information, I thought there only was one questionnaire at the start of the study and then you only send samples.*

Some of the men interviewed were not aware of online questionnaire requirements, though they reported reading participant information detailing them. Suggestions were made for improved instructions, including greater clarity on requirements for weekly questionnaire completion, less densely presented but more concise information, bullet points with study requirements, and use of less formal and more readable language.

Among participants who completed the questionnaires, this was facilitated by online links or QR codes in study information, text reminders, or self-bookmarking the site of the weekly questionnaire. From one interviewee,

> *[…] But then I got a text that day thanking me for taking part in the study which had the link embedded which is really useful because I couldn't remember what the hell the name of it was.*

The content and length of the questionnaires were largely acceptable to those who were aware of and completed one or more questionnaires.

> *The length as in. Uh, yeah [...] I found it pretty straightforward and easy, in fact, I wish I'd known about it earlier. I would have probably done it more regularly if I'd known how easy it was.*

Overall, participants expressed comfort with the content of questions, and their assessments suggested a reasonable understanding or acceptance that questions were included because they would be useful and appropriate to the study.

A couple of interviewees related some irritating repetition in individual surveys or in having to answer questions when little had changed in their behavior from preceding weeks and may prefer a more concise way of indicating as such, for example, a "No change" option that would result in skipping most of the questions. Others suggested they would appreciate improvements to layout and legibility, particularly on mobile screens.

> *[…] I tried to do [the questionnaire] on my phone again, but it was just impossible. It was just far too small.*

Discussion

The STEIM study design was feasible to implement and largely acceptable in an SHS-attending GBMSM population in Brighton and Sussex. The study team was successful in recruiting a study population of the target size that was broadly representative of GBMSM attending the participating SHSs. Just under half of participants provided more than one rectal swab, and many provided weekly or nearly weekly swabs. Questionnaire completion was less complete than swab return. Insights from qualitative interviews identified barriers and facilitators that could be addressed in a future implementation of the study design, and many made suggestions for improving participant-facing study materials. For example, participants reported not being aware of the need for regular questionnaire completion and being surprised about how easy they were to complete, which could be addressed by improved study information and communication.

Our findings align with other research involving the recruitment of GBMSM to provide self-completed questionnaires and a biological sample; one study recruiting GBMSM at a central London SHS to complete a questionnaire and collect biological samples for human papillomavirus testing had a high acceptance rate of 81.4% [27], while the Gay Men's Sexual Health Survey in Scotland recorded a response rate of 65.2% for the survey aspect, with 80.4% of those surveyed providing an optional oral fluid specimen [28]. Our findings are particularly encouraging given the longitudinal sampling component of STEIM.

XSL•FO
RenderX

Qualitative interviews highlighted that the innovation of the study, especially given the focus on gut health, was appealing to many and may have helped with recruitment. Altruism is a common facilitator for participation in a study [29], and the feeling of giving something back was also a driver for participation in this study, with some recruitment happening in vaccination clinics responding to the 2022 mpox outbreak in England. This was a special circumstance, but linking recruitment to valued services may be beneficial for future studies.

Perceived personal benefits are also a common facilitator [29]. However, the study design did not provide individual sample test results, which could present a possible barrier to participation. Two participants reported being motivated to participate because they expected to receive diagnostic test results. It will be important to better emphasize that this is not the case for future recruitment or to develop mechanisms to explore feedback in future studies, especially as it was suggested that receiving results, both individual sample and overall study results, may incentivize participants to complete the study.

A large proportion of participants did not complete any study procedures after enrollment. Other studies have documented similar problems. For example, in Positive Voices, a national survey of people living with HIV, 90% of individuals accepted the initial survey invitation, but only 23% completed it [24]. Unfortunately, limited data were available to understand the drop-off in this study, with no qualitative interviews conducted with these participants. Demographic analysis showed that those who completed study procedures were more likely to have taken HIV PrEP or to live in the least deprived IMD quintile, highlighting more routine use of SHSs; a desire for reciprocity of valued services or high socioeconomic status might facilitate taking part. Future studies must address this challenge, potentially through better public engagement work or through improving the study procedures based on feedback from this pilot. While the lack of financial offers was not highlighted as a barrier to participation in the interviews, with no interviews conducted with those who did not complete procedures, it is possible that financial incentives could help give higher overall participation and retention for future studies, especially for those living in more deprived areas, who were less likely to complete study procedures. The potential drop-off should also be accounted for in future sample size calculations.

Participant engagement with questionnaires was challenging, with 36.8% of participants completing a baseline and only 15 participants completing all follow-up questionnaires, less than half the number who provided rectal swabs for the whole study. Qualitative interviews revealed low awareness of questionnaires, especially the requirement for weekly completion, and this will be a focus for future iterations. Concise, bulleted instructions of the key steps might help improve participant understanding and completion rates. Optimizing mobile access and completion of the questionnaire could also improve engagement, for example, by using the mobile surveys feature of Snap Surveys Ltd.

SMS text messaging sent to participants as reminders to complete study procedures had variable utility, with mixed understanding as to whether they were reminders or were a thank you instead. A meta-analysis showed varied usefulness of reminders in promoting participant engagement with online surveys [30]. The utility of SMS text messaging increased if they were seen to improve a sense of study connection and engagement, which participants suggested might be done by a more personal focus and tailoring to specific tasks. In other studies, personalized reminder messages have been discussed as a useful tool to generate extra responses and encourage reluctant participants [31].

There was better engagement with the return of rectal swabs, with over half of those returning one swab returning 10 or more swabs. This might be because rectal swab sampling was easier to incorporate into daily routines, as highlighted by qualitative interviews. Qualitative interviews also revealed that most participants were satisfied with the 12-week study length and that the definitive end point supported decision-making on the ability to participate, further suggesting that this type of longitudinal sampling is feasible. Patient familiarity with the rectal swab procedure, a routine form of testing in SHS [32], may also have contributed to acceptability in this cohort.

Even so, few of the interviewees at the time of the interview had fully completed the study procedures, and none completed them on the same day every week. This indicates that, although feasible and acceptable, competing life tasks and the unpredictability of weekly routines still limit participation in a longitudinal study. It will be important to ensure that participants are aware there is flexibility and that they can complete weekly procedures on different days or miss one week (or a couple) and still provide important participation.

There are some limitations to the STEIM study. Regarding recruitment, estimating recruitment success relied on the completion of study screening logs, which were not routinely done, leading to possibly overestimated recruitment success. Participants were required to have sufficient English language and literacy skills and access to a mobile device and the internet to be able to complete the study procedures, and most participants were recruited from Brighton SHS, making this essentially a single-center study. Together, this limits the generalizability of the study, the demographics of the recruited individuals, and the participant experience. Based on experience from other studies [25], we were confident that online data collection methods were acceptable to most SHS attendees in this population, and within the time and financial constraints of our pilot study, it was not possible to translate materials into different languages or supply alternatives to digital devices. However, these options should be explored for future studies to increase the generalizability of the study and the representativeness of those who complete study procedures. Additionally, the STEIM study had specific eligibility criteria, such as having had a recent sexual partner, which, at that time, could not be compared to national STI surveillance data. Therefore, representativeness was compared to the overall population and may explain the higher frequency of recent STIs and taking HIV PrEP in the study cohort. Interview participants had all returned at least one questionnaire and rectal swab, which limited our qualitative insights on barriers to completing study procedures.

XSL•FO
RenderX

Overall, based on the findings from this pilot study, we recommend that future studies (1) provide a user-friendly outline of the study requirements at the beginning of the participant instruction sheet that clearly lays out the expectation of the participant over the study period through concise bulleted instructions, with an emphasis on the weekly questionnaire alongside the swab; (2) modify SMS text messaging reminders so that they improve a sense of connection to the study, for example, are tailored to specific tasks, and are clearly phrased to serve as reminders of study procedures and what those procedures are; (3) optimize the questionnaires for use on mobile phones; (4) consider participation incentives, which could include financial incentives and/or reports on study results; and (5) keep and emphasize the element of flexibility to the study procedures so that they can be integrated into normal routine.

In conclusion, this phase of the STEIM study has provided evidence of the feasibility and acceptability of the study methods, as well as tangible insights to inform the design of future similar studies involving longitudinal rectal swabs and self-reported data collection for STEI-related studies as well as for studies in other sexual health patient populations and outside of this setting. This pilot has provided innovative results in the field of STEIs; previous research studies for exploring risk factors for and transmission of STEIs (in the United Kingdom) have been cross-sectional in nature [12,15] and, to our knowledge, this is the first study that is longitudinal. This type of study is needed to address important gaps in our understanding of STEI transmission, which can then be used to inform appropriately targeted clinical and public health interventions for infection control.

Acknowledgments

The authors thank all staff at the participating clinics that enabled recruitment, and the study participants. We would also like to acknowledge the support of the NIHR Clinical Research Network. We acknowledge members of the National Institute for Health and Care Research Health Protection Research Unit (NIHR HPRU) in Blood Borne and Sexually Transmitted Infections (BBSTI) Steering Committee: Professor Caroline Sabin (HPRU Director), Dr John Saunders (UKHSA Lead), Professor Catherine Mercer, Dr Hamish Mohammed, Professor Greta Rait, Dr Ruth Simmons, Professor William Rosenberg, Dr Tamyo Mbisa, Professor Rosalind Raine, Dr Sema Mandal, Dr Rosamund Yu, Dr Samreen Ijaz, Dr Fabiana Lorencatto, Dr Rachel Hunter, Dr Kirsty Foster and Dr Mamooma Tahir.

The views expressed are those of the authors and not necessarily those of the NIHR, the Department of Health and Social Care, or the UK Health Security Agency.

The copyright of this manuscript is held under the provisions of the UK Governement's contract rights (Open Government Licence).

Funding

This study was funded by the NIHR HPRU in BBSTI at University College London in partnership with the UK Health Security Agency, and the NIHR HPRU in Gastrointestinal Infections, a partnership between the UK Health Security Agency, the University of Liverpool, and the University of Warwick.

Data Availability

The data that support the findings of this study have been assessed by the UK Health Security Agency's Office for Data Acquisition and Release as having sensitive personal information and are therefore not publicly available to protect participant privacy. However, some aggregate data may be available upon reasonable request from the UKHSA. Requests can be directed to DataAccess@ukhsa.gov.uk.

Authors' Contributions

HDM, D Richardson, CJ, KSB, NF, and GH conceptualized the study. HDM, D Richardson, CJ, KSB, NF, GH, CVP, HM, KT, HC, XD, NM, and CHM developed the study design and methodology. D Reid conducted interviews. HF, KT, D Reid and DO conducted formal analysis. HF and KT wrote the original draft manuscript. All authors reviewed, edited, and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study baseline questionnaire.
[PDF File, 248 KB - publichealth_v12i1e73762_app1.pdf]

Multimedia Appendix 2

XSL•FO
RenderX

Study follow-up questionnaire.
[PDF File, 241 KB - publichealth_v12i1e73762_app2.pdf]

Multimedia Appendix 3
Study participant consent form.
[PDF File, 136 KB - publichealth_v12i1e73762_app3.pdf]

Multimedia Appendix 4
Additional figures and table.
[PDF File, 134 KB - publichealth_v12i1e73762_app4.pdf]

References

1. Mitchell H, Hughes G. Recent epidemiology of sexually transmissible enteric infections in men who have sex with men. Curr Opin Infect Dis 2018 Feb;31(1):50-56. [doi: 10.1097/QCO.0000000000000423] [Medline: 29251673]
2. McNeil CJ, Kirkcaldy RD, Workowski K. Enteric infections in men who have sex with men. Clin Infect Dis 2022 Apr 13;74(Suppl_2):S169-S178. [doi: 10.1093/cid/ciac061] [Medline: 35416972]
3. Gilbart VL, Simms I, Jenkins C, et al. Sex, drugs and smart phone applications: findings from semistructured interviews with men who have sex with men diagnosed with Shigella flexneri 3a in England and Wales. Sex Transm Infect 2015 Dec;91(8):598-602. [doi: 10.1136/sextrans-2015-052014] [Medline: 25921020]
4. Thorley K, Charles H, Greig DR, et al. Emergence of extensively drug-resistant and multidrug-resistant Shigella flexneri serotype 2a associated with sexual transmission among gay, bisexual, and other men who have sex with men, in England: a descriptive epidemiological study. Lancet Infect Dis 2023 Jun;23(6):732-739. [doi: 10.1016/S1473-3099(22)00807-6] [Medline: 36731481]
5. Borg ML, Modi A, Tostmann A, et al. Ongoing outbreak of Shigella flexneri serotype 3a in men who have sex with men in England and Wales, data from 2009-2011. Euro Surveill 2012 Mar 29;17(13):20137. [doi: 10.2807/ese.17.13.20137-en] [Medline: 22490381]
6. European Centre for Disease Prevention and Control. European centre for disease prevention and control. hepatitis a outbreaks in the EU/EEA mostly affecting men who have sex with men – third update, 28 june 2017. : Stockholm: ECDC; 2017 URL: https://www.ecdc.europa.eu/sites/default/files/documents/RRA%20hep%20A%20outbreak%20EU%20EEA%20in%20MSM%20third%20update%2028%20June%202017_0.pdf [accessed 2026-03-14]
7. Simms I, Field N, Jenkins C, et al. Intensified shigellosis epidemic associated with sexual transmission in men who have sex with men--Shigella flexneri and S. sonnei in England, 2004 to end of February 2015. Euro Surveill 2015 Apr 16;20(15):21097. [doi: 10.2807/1560-7917.es2015.20.15.21097] [Medline: 25953129]
8. Baker KS, Dallman TJ, Ashton PM, et al. Intercontinental dissemination of azithromycin-resistant shigellosis through sexual transmission: a cross-sectional study. Lancet Infect Dis 2015 Aug;15(8):913-921. [doi: 10.1016/S1473-3099(15)00002-X] [Medline: 25936611]
9. Mook P, McCormick J, Bains M, et al. ESBL-producing and macrolide-resistant Shigella sonnei infections among men who have sex with men, England, 2015. Emerg Infect Dis 2016 Nov;22(11):1948-1952. [doi: 10.3201/eid2211.160653] [Medline: 27767929]
10. Mason LCE, Richardson D, Charles H, et al. Using demographics of patients to inform treatment of shigellosis in England. Lancet Microbe 2025 Apr;6(4):101026. [doi: 10.1016/j.lanmic.2024.101026] [Medline: 39547241]
11. Richardson D, Pakianathan M, Ewens M, et al. British Association of Sexual Health and HIV (BASHH) United Kingdom national guideline for the management of sexually transmitted enteric infections 2023. Int J STD AIDS 2023 Aug;34(9):588-602. [doi: 10.1177/09564624231168217] [Medline: 37247427]
12. Mitchell HD, Whitlock G, Zdravkov J, et al. Prevalence and risk factors of bacterial enteric pathogens in men who have sex with men: a cross-sectional study at the UK's largest sexual health service. J Infect 2023 Jan;86(1):33-40. [doi: 10.1016/j.jinf.2022.10.033] [Medline: 36343788]
13. Baker KS, Dallman TJ, Field N, et al. Genomic epidemiology of Shigella in the United Kingdom shows transmission of pathogen sublineages and determinants of antimicrobial resistance. Sci Rep 2018 May 9;8(1):7389. [doi: 10.1038/s41598-018-25764-3] [Medline: 29743642]
14. Allen H, Mitchell HD, Simms I, et al. Evidence for re-infection and persistent carriage of Shigella species in adult males reporting domestically acquired infection in England. Clin Microbiol Infect 2021 Jan;27(1):126. [doi: 10.1016/j.cmi.2020.03.036] [Medline: 32247893]
15. Hughes G, Silalang P, Were J, et al. Prevalence and characteristics of gastrointestinal infections in men who have sex with men diagnosed with rectal chlamydia infection in the UK: an "unlinked anonymous" cross-sectional study. Sex Transm Infect 2018 Nov;94(7):518-521. [doi: 10.1136/sextrans-2016-053057] [Medline: 28360379]

XSL•FO
RenderX

16. Williamson DA, Chow EPF, Lee D, et al. Risk factors for asymptomatic enteric pathogen detection among men who have sex with men. Open Forum Infect Dis 2019 Sep;6(9):ofz326. [doi: 10.1093/ofid/ofz326] [Medline: 31660405]
17. Richardson D, Savary-Trathen A, Fitzpatrick C, Williams D. Estimated prevalence and associations of sexually transmissible bacterial enteric pathogens in asymptomatic men who have sex with men: a systematic review and meta-analysis. Sex Transm Infect 2024 Nov 18;100(8):532-537. [doi: 10.1136/sextrans-2024-056183] [Medline: 38902026]
18. Didelot X, Walker AS, Peto TE, Crook DW, Wilson DJ. Within-host evolution of bacterial pathogens. Nat Rev Microbiol 2016 Mar;14(3):150-162. [doi: 10.1038/nrmicro.2015.13] [Medline: 26806595]
19. Carson J, Keeling M, Wyllie D, Ribeca P, Didelot X. Inference of infectious disease transmission through a relaxed bottleneck using multiple genomes per host. Mol Biol Evol 2024 Jan 3;41(1):msad288. [doi: 10.1093/molbev/msad288] [Medline: 38168711]
20. Thorley K, Charles H, Mitchell H, Jenkins C, Godbole G, Sinka K. Sexually transmitted shigella spp in england: data up to quarter 2, 2022. : UK Health Security Agency; 2022 URL: https://www.gov.uk/government/publications/non-travel-associated-shigella-infections/sexually-transmitted-shigella-spp-in-england-data-up-to-quarter-2-2022 [accessed 2026-03-14]
21. Savage EJ, Mohammed H, Leong G, Duffell S, Hughes G. Improving surveillance of sexually transmitted infections using mandatory electronic clinical reporting: the genitourinary medicine clinic activity dataset, England, 2009 to 2013. Euro Surveill 2014 Dec 4;19(48):20981. [doi: 10.2807/1560-7917.es2014.19.48.20981] [Medline: 25496573]
22. Hennink MM, Kaiser BN, Marconi VC. Code saturation versus meaning saturation: how many interviews are enough? Qual Health Res 2017 Mar;27(4):591-608. [doi: 10.1177/1049732316665344] [Medline: 27670770]
23. Hennink M, Kaiser BN. Sample sizes for saturation in qualitative research: a systematic review of empirical tests. Soc Sci Med 2022 Jan;292:114523. [doi: 10.1016/j.socscimed.2021.114523]
24. Kall M, Shahmanesh M, Nardone A, Gilson R, Delpech V, editors. “Positive voices” a RCT pilot survey of the behaviour and healthcare needs of people with HIV: study methods and respondent characteristics. 2015 Presented at: BHIVA conference; Apr 21-24, 2015 URL: https://bhiva.org/wp-content/uploads/2024/10/AbstractBook2015.pdf [accessed 2026-03-14]
25. Wayal S, Reid D, Blomquist PB, Weatherburn P, Mercer CH, Hughes G. The acceptability and feasibility of implementing a bio-behavioral enhanced surveillance tool for sexually transmitted infections in England: mixed-methods study. JMIR Public Health Surveill 2018 May 4;4(2):e52. [doi: 10.2196/publichealth.9010] [Medline: 29728348]
26. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013 Sep 18;13:117. [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
27. King EM, Gilson R, Beddows S, et al. Human papillomavirus DNA in men who have sex with men: type-specific prevalence, risk factors and implications for vaccination strategies. Br J Cancer 2015 Apr 28;112(9):1585-1593. [doi: 10.1038/bjc.2015.90] [Medline: 25791874]
28. Wallace LA, Li J, McDaid LM. HIV prevalence and undiagnosed infection among a community sample of gay and bisexual men in Scotland, 2005-2011: implications for HIV testing policy and prevention. PLoS ONE 2014;9(3):e90805. [doi: 10.1371/journal.pone.0090805] [Medline: 24621479]
29. Sheridan R, Martin-Kerry J, Hudson J, Parker A, Bower P, Knapp P. Why do patients take part in research? An overview of systematic reviews of psychosocial barriers and facilitators. Trials 2020 Mar 12;21(1):259. [doi: 10.1186/s13063-020-4197-3] [Medline: 32164790]
30. Wu MJ, Zhao K, Fils-Aime F. Response rates of online surveys in published research: a meta-analysis. Computers in Human Behavior Reports 2022 Aug;7:100206. [doi: 10.1016/j.chbr.2022.100206]
31. Shiyab W, Ferguson C, Rolls K, Halcomb E. Solutions to address low response rates in online surveys. Eur J Cardiovasc Nurs 2023 May 25;22(4):441-444. [doi: 10.1093/eurjcn/zvad030] [Medline: 36827086]
32. BASHH summary guidance on testing for sexually transmitted infections. British association for sexual health and HIV (BASHH). 2023. URL: https://www.bashh.org/_userfiles/pages/files/resources/bashh_summary_guidance_on_stis_testing_2023.pdf [accessed 2026-03-14]

Abbreviations

AMR: antimicrobial resistance
GBMSM: gay, bisexual, and other men who have sex with men
IMD: index of multiple deprivation
PIS: participant information sheet
PrEP: pre-exposure prophylaxis
SHS: sexual health service
STEI: sexually transmitted enteric infection
STEIM: sexually transmitted enteric infections among GBMSM
STI: sexually transmitted infection

XSL•FO
RenderX

UKHSA: UK Health Security Agency

Edited by S Brini; submitted 21.Mar.2025; peer-reviewed by M Ewens, Y Mowat; revised version received 10.Feb.2026; accepted 12.Feb.2026; published 30.Mar.2026.

Please cite as:
Fountain H, Thorley K, Reid D, Ogaz D, Richardson D, Charles H, Baker KS, Jenkins C, Pulford CV, Hughes G, Field N, Mercer CH, Didelot X, McCarthy N, Mohammed H, Mitchell HD
Acceptability and Feasibility of Longitudinal Sampling for Sexually Transmitted Enteric Infections in Gay, Bisexual, and Other Men Who Have Sex With Men (GBMSM): Prospective Cohort Pilot Study Conducted in 2022 in South East England
JMIR Public Health Surveill 2026;12:e73762
URL: https://publichealth.jmir.org/2026/1/e73762
doi:10.2196/73762

© Holly Fountain, Katie Thorley, David Reid, Dana Ogaz, Daniel Richardson, Hannah Charles, Kate S Baker, Claire Jenkins, Caisey V Pulford, Gwenda Hughes, Nigel Field, Catherine H Mercer, Xavier Didelot, Noel McCarthy, Hamish Mohammed, Holly D Mitchell. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 30.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Moderating Role of Condom-Use Inertia on the Association Between Status Quo Bias and Pre-Exposure Prophylaxis Resistance Intention Among Chinese Men Who Have Sex With Men: Cross-Sectional Study

Min Zhao[1,2], PhD; Shanshan Li[2], MPH; Xiang Zhao[3], MD; Jincheng Hu[2], MSc; Zhuoheng Yin[4], MA; Fanpu Ji[1], Prof Dr; Weiming Tang[5], Prof Dr; Lei Zhang[1,2,6,7], Prof Dr

[1]Department of Hepatology, Second Affiliated Hospital of Xi'an Jiaotong University, No. 157 Xiwu Rd, Xi'an, Shaanxi, China

[2]China-Australia Joint Research Centre for Infectious Diseases, School of Public Health, Xi'an Jiaotong University, Xi'an, Shaanxi, China

[3]Second Department of Coronary Heart Disease, First Affiliated Hospital of Xinjiang Medical University, Urumqi, Xinjiang, China

[4]University of North Carolina at Chapel Hill Project-China, Guangzhou, Guang dong, China

[5]Institute of Global Health and Infectious Diseases, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[6]Melbourne Sexual Health Centre, Alfred Health, Melbourne, Victoria, Australia

[7]School of Translational Medicine, Faculty of Medicine, Monash University, Melbourne, Victoria, Australia

Corresponding Author:
Lei Zhang, Prof Dr
Department of Hepatology, Second Affiliated Hospital of Xi'an Jiaotong University, No. 157 Xiwu Rd, Xi'an, Shaanxi, China

Abstract

Background: While prior studies have examined structural and individual-level barriers to pre-exposure prophylaxis (PrEP) uptake, little is known about the psychological mechanisms underlying resistance to PrEP, particularly among high-risk groups. The status quo bias (SQB) theory provides a theoretical framework for understanding why individuals may resist beneficial health innovations.

Objective: The aim of this study was to examine the associations between SQB and PrEP resistance intention among Chinese men who have sex with men (MSM) and to test whether condom-use inertia moderated these relationships.

Methods: We conducted a cross-sectional online survey among 1022 MSM in China from November 2024 to February 2025. Theory-guided multi-item measures were constructed to capture key dimensions of SQB. Their internal consistency and construct validity were examined using confirmatory factor analysis prior to regression modeling. Multiple linear regression models assessed main effects and moderation effects, adjusting for sociodemographic and behavioral covariates.

Results: A total of 1022 MSM were included in the final analysis (mean age 29.6 y); the majority identified as homosexual (767/1022, 75.1%) and were unmarried (896/1022, 87.7%). Regression analyses revealed that transition costs were positively associated with PrEP resistance intention (β=0.26, 95% CI 0.17-0.35; *P*<.001), while social norms were negatively associated (β=–0.19, 95% CI –0.27 to –0.12; *P*<.001). Condom-use inertia significantly moderated both associations, amplifying the positive relationship between transition costs and resistance (β=0.04, 95% CI 0.01-0.09; *P*=.03) and enhancing the negative association of social norms (β=–0.05, 95% CI –0.09 to 0.00; *P*=.04). Subgroup analyses showed that the amplifying effect of condom-use inertia on transition costs was particularly evident among participants with postgraduate education (β=0.13, 95% CI 0.01-0.26; *P*=.04). Conversely, its strengthening effect on social norms was more pronounced among MSM 30 years or younger (β=–0.11, 95% CI –0.17 to –0.04; *P*=.001) and those identifying as homosexual (β=–0.07, 95% CI –0.12 to –0.02; *P*=.01).

Conclusions: This study suggests that SQB is an important psychological barrier to PrEP adoption among Chinese MSM. The moderating role of condom-use inertia highlights the relevance of habitual condom-use routines in shaping PrEP resistance. Behaviorally informed strategies that reduce transition costs and leverage social norms may help mitigate resistance and improve PrEP uptake.

(JMIR Public Health Surveill 2026;12:e88806) doi:10.2196/88806

KEYWORDS

HIV; pre-exposure prophylaxis; men who have sex with men; social norms; condoms; cross-sectional studies; status quo bias; China

XSL•FO
RenderX

Introduction

According to the latest statistics from the World Health Organization, an estimated 40.8 million people were living with HIV at the end of 2024, and approximately 1.3 million people acquired HIV in the same year [1]. Men who have sex with men (MSM) continue to be among the populations at highest risk for HIV infection, with a relative risk estimated to be 23 times higher than that of the general population [2,3]. In the ongoing global public health effort to end the HIV epidemic, pre-exposure prophylaxis (PrEP) has emerged as a milestone biomedical intervention. Evidence shows that when taken as prescribed, PrEP can reduce the risk of HIV acquisition among MSM by approximately 99% [4]. Nevertheless, despite its enormous potential to reshape the trajectory of the HIV epidemic, the implementation and scale-up of PrEP continue to encounter considerable challenges.

Implementing and promoting PrEP uptake among MSM remains a considerable challenge. Although studies estimate that over 66.7% of MSM in China express willingness to use PrEP, the actual uptake rate remains below 10% [5,6]. Existing research has focused on 2 broad categories of barriers to PrEP willingness or uptake. On the structural level, a large body of literature highlights obstacles such as limited health care access, high out-of-pocket costs, and service delivery gaps [7,8]. On the psychosocial level, researchers have emphasized the role of factors such as stigma, perceived HIV risk, and self-efficacy [9-11]. While crucial metrics like PrEP willingness and uptake measure the presence of a positive inclination or a final behavior, they may not fully capture the active psychological processes underlying the rejection of PrEP. This specific failure to adopt PrEP among those who need it most underscores the limitations of “willingness” as a construct and necessitates a new focus on the cognitive and behavioral factors driving active resistance intention.

From the perspective of status quo bias (SQB) theory, individuals often prefer preserving their current circumstances, even when alternative options offer greater benefits [12]. This study focuses on 2 principal mechanisms, drawn from SQB theory, that may explain this phenomenon: rational decision-making and psychological commitments [13]. Rational decision-making implies that individuals conduct a cost-benefit analysis before transitioning to a new HIV prevention strategy [13]. Transition costs refer to the burdens associated with adopting a new practice; when such costs outweigh the perceived benefits, individuals are likely to maintain the status quo [14]. In the context of PrEP, these transition costs may present a significant barrier to adoption, including the time and financial investments required for behavioral adjustments, such as medical consultations, regular HIV testing, and the establishment of consistent medication routines [15-17]. For example, Katz et al [18] reported that high-risk women in Kenya experienced difficulties initiating PrEP due to professional scheduling conflicts, educational disruptions, and prohibitive transportation expenses.

Psychological commitment reflects prevailing social norms [19]. In HIV prevention, social norms refer to the collective expectations and shared attitudes within MSM communities regarding PrEP usage. Previous studies suggest that the willingness of MSM to use PrEP is significantly influenced by perceived social approval or disapproval from significant others [20,21]. Specifically, the attitudes and opinions of family members, peers, and healthcare providers can influence the perceptions and decisions of MSM regarding PrEP [22,23]. When individuals are embedded in social networks that actively endorse or discuss PrEP use, these positive normative cues can help counteract status quo tendencies and encourage PrEP uptake [24,25].

Rather than arising from reflective cost-benefit analyses, inertia reflects an automatic reliance on established behavioral patterns and a natural resistance to change [26,27]. While condoms remain a cornerstone of HIV prevention, their effectiveness is often compromised by inconsistent use—a well-documented risk factor, particularly among high-risk MSM [28]. PrEP is a critical supplement to HIV prevention for this high-risk group. However, adopting this biomedical paradigm triggers cognitive resistance. For example, Han et al [29] found that individuals who preferred and consistently used condoms as a method of HIV prevention were significantly less willing to consider PrEP adoption. Similarly, previous research has shown that MSM with high self-efficacy regarding condom use were less likely to contemplate using PrEP [11]. Building on these insights, while recognizing the underexplored relationship between SQB and behavioral inertia, we hypothesize that condom-use inertia moderates this interplay, such that the influence of factors like transition costs and social norms becomes more pronounced among individuals with higher levels of inertia.

Although a growing body of research has explored barriers to PrEP adoption and willingness, few studies have explicitly examined PrEP resistance intention through a behavioral economics lens. Moreover, the potential influence of condom-use inertia has received little empirical attention. Guided by SQB theory, this study examines how transition costs and perceived social norms are associated with PrEP resistance intention among Chinese MSM and tests whether condom-use inertia conditions these associations.

Methods

Study Design and Participant Recruitment

We conducted this cross-sectional study using an anonymous, web-based questionnaire deployed via Sojump, a widely used online survey platform in China, and reported the online survey in accordance with the CHERRIES (Checklist for Reporting Results of Internet E-Surveys) (Checklist 1). We collected data from November 29, 2024, to February 11, 2025. Participants were recruited through convenience and snowball sampling. The research team developed recruitment materials, including digital posters and standardized outreach messages, which were disseminated through the WeChat official accounts of collaborating community-based organizations (CBOs) and local Centers for Disease Control and Prevention. Guangdong served as the initial dissemination hub because the core CBO partner responsible for launching recruitment was based there. To implement snowball sampling, participants were encouraged

XSL•FO
RenderX

to share the survey link with their peers. As a result, recruitment extended beyond the initiating hub and ultimately reached multiple regions across mainland China, including Guangdong, Hubei, Shandong, Jiangsu, Sichuan, and Shanghai.

Eligible participants were those who met the following inclusion criteria: aged 18 years or older, identified as cisgender male, had a history of sexual activity with other men, were self-reported HIV-negative or unaware of their HIV status, and provided electronic informed consent before beginning the survey. We excluded participants who declined to consent or were deemed unable to complete the survey due to severe psychiatric illness or significant cognitive impairment, as determined by self-report or investigator assessment. We restricted participation to cisgender MSM to preserve conceptual and measurement consistency because the study constructs and items were developed to capture PrEP-related resistance and condom-use routines in this population and have not been validated for transgender and gender-diverse participants.

Before accessing the questionnaire, potential participants viewed an electronic informed consent page that described the study investigator, the study's purpose, and the estimated survey duration (about 15 min). The consent page also explained what information would be collected, that responses would be stored on a secure, password-protected server managed by the study team, that only authorized investigators would have access, and that the data would be used solely for scientific research. Participants could proceed to the survey only after actively indicating consent (eg, selecting an "I agree" option). Participants whose responses passed the logic-consistency and attention-check procedures received 30 yuan (about US $5) as a token of appreciation for their time.

To ensure data quality, the questionnaire incorporated embedded attention-check items and logic-consistency checks. Of 1311 submitted responses, 257 failed quality control and were excluded, leaving 1054 responses. We then removed 32 duplicate entries identified using the phone numbers submitted with the questionnaire, resulting in 1022 unique participants included in the final analysis (Multimedia Appendix 1). Phone numbers were collected solely for duplicate checking and were removed from the analytic dataset after screening. The survey platform also required completion of mandatory items before submission, minimizing missing data.

The present study focuses on PrEP resistance intention and incorporates multiple covariates and interaction testing. An a priori power analysis using G*Power 3.1.9.7 indicated that, for a multiple linear regression model with α=.05 and up to 20 predictors, a sample of 157 participants would be required to detect a medium effect (f^2=0.15) with 80% power, whereas 1064 participants would be required to detect a small effect (f^2=0.02). The final analytic sample of 1022 participants fell between these 2 benchmarks, suggesting adequate power to detect effects in the small-to-medium range in the primary regression models. However, smaller interaction effects and exploratory subgroup analyses should be interpreted cautiously.

Ethical Considerations

We obtained ethical approval for this study from the Institutional Review Board of Shenzhen University (protocol PN-202400070). All participants provided informed consent electronically before accessing the survey. The survey was administered anonymously online. Phone numbers were collected only for duplicate screening and compensation purposes and were removed from the analytic dataset after screening. All data were stored on a password-protected server accessible only to the research team and were used solely for research purposes. Participants whose responses passed the quality-control procedures received 30 Yuan (about US $5) as compensation.

Measures

Study Constructs and Measurement

We operationalized the key study constructs using theory-guided, multi-item measures informed by the behavioral economics framework of SQB [13] and by prior empirical literature on PrEP adoption barriers among MSM [15-17,21-25,30,31]. We adapted all items to suit the cultural and behavioral context of the target population and refined them through expert consultation and pilot testing [32]. Participants rated all items on a 7-point Likert scale, ranging from 1 (strongly disagree) to 7 (strongly agree). We then calculated the final score for each dimension as the arithmetic mean of its corresponding items. Table 1 summarizes the study constructs and their items. These constructs were analyzed as distinct measures (outcome, predictors, and moderator) rather than as subscales of a single instrument.

XSL•FO
RenderX

Table . Study constructs and items.

Study constructs	Items
Transaction costs	
TC1	It would take a lot of time and effort for PrEP[a]-related physical exams.
TC2	I would spend a lot of money if I were to switch to using PrEP.
TC3	There would be health-related side effects if I were to switch to using PrEP.
TC4	Switching to using PrEP could result in unexpected hassles.
Social norms	
SN1	My doctors/community outreach workers think I should use PrEP. [reverse-coded item]
SN2	My family/friends think I should use PrEP. [reverse-coded item]
SN3	My sex partners think I should use PrEP. [reverse-coded item]
Condom-use inertia	
I will continue using condoms for preventing HIV…	
CI1	…simply because it is what I have always done.
CI2	…I enjoy doing so.
CI3	…even though I know it is not the best way of doing things.
Resistance intention	
RI1	I oppose the idea of using PrEP as a strategy to prevent HIV.
RI2	I will not start using PrEP for preventing HIV.
RI3	I will not comply with the medical instructions to use PrEP for preventing HIV.

[a]PrEP: pre-exposure prophylaxis.

Primary Outcome

The primary outcome variable, PrEP resistance intention, was assessed using 3 items designed to capture individuals' intentions to resist the adoption and use of PrEP. These items reflect a spectrum of resistance, ranging from general attitudinal opposition to the change to specific behavioral intentions to reject both the adoption process and the medication regimen. We calculated the outcome score as the mean of the 3 items, with higher scores indicating stronger resistance intentions.

Explanatory Variables

We focused on 2 theoretically salient dimensions of SQB as the main explanatory variables: transition costs and perceived social norms. These dimensions reflect distinct psychological mechanisms that may contribute to resistance to behavioral change. Transition costs included 4 items covering concerns about time, financial burden, side effects, and potential hassles associated with PrEP uptake [15-17]. We assessed social norms using 3 reverse-coded items that evaluated perceived approval of PrEP use from key referents, such as health care workers, friends, and sexual partners [21-25]. We reverse-coded these items so that higher values reflected more supportive norms, consistent with the hypothesized protective association.

Moderator Variable

Condom-use inertia, reflecting an inertial reliance on behavioral persistence, was measured using 3 items. These items explored the behavioral, affective, and cognitive tendencies of individuals to continue using condoms for HIV prevention, even in cases where condom use might be perceived as suboptimal [9,29,33]. Higher scores indicated more substantial condom-use inertia.

Covariates

We selected covariates based on prior empirical evidence identifying their relevance to PrEP-related decision-making and their potential roles as confounding factors [34,35]. Demographic characteristics included age, sexual orientation, marital status, education level, and monthly income. Sexual behavior–related variables included the number of regular male sexual partners and the number of casual male sexual partners reported in the past 6 months. We also included substance use as a behavioral covariate. We asked participants whether they had used any of a list of commonly reported substances within the previous 12 months, specifically in sexual contexts. The checklist included marijuana, cocaine, opium ("white powder"), methamphetamine, poppers, ecstasy, lysergic acid diethylamide, psilocybin, and gamma-hydroxybutyrate or gamma-butyrolactone. We dichotomized responses, coding any substance use during sex as 1 ("yes") and no reported use as 0 ("no").

Statistical Analysis

We used descriptive statistics to summarize participants' demographic and behavioral characteristics. We reported means and SDs for continuous variables and presented categorical variables as frequencies and proportions. Psychometric evaluation and regression analyses were conducted in the same analytic sample (N=1022). To support the subsequent regression analyses, we evaluated the measurement properties of the

XSL•FO
RenderX

theory-guided multi-item constructs. Because the factor structure was specified a priori based on theory, we conducted confirmatory factor analysis (CFA) to evaluate construct validity. We assessed internal consistency using Cronbach α and composite reliability. We examined convergent validity using standardized factor loadings and average variance extracted (AVE) from CFA, and we tested discriminant validity using the Fornell-Larcker criterion by comparing the square root of each construct's AVE with its correlations with other constructs. We then computed mean scores for all constructs and conducted Pearson correlation analyses to examine the bivariate associations between transition costs, social norms, condom-use inertia, and PrEP resistance intention.

To identify factors associated with PrEP resistance intention, we conducted hierarchical linear regression analyses in 3 steps. Model 1 estimated the main associations between transition costs, social norms, and condom-use inertia and PrEP resistance intention, while adjusting for sociodemographic and behavioral covariates. Model 2 added the transition costs × condom-use inertia interaction, and Model 3 added the social norms × condom-use inertia interaction. We mean-centered continuous predictors before creating interaction terms to reduce multicollinearity and facilitate interpretation. We reported standardized coefficients (β), 95% CIs, and *P* values. We assessed model fit using multiple criteria, including R^2, adjusted R^2, and nested model *F* tests, to determine whether interaction terms significantly improved the model fit. We performed all analyses using Stata 17 (StataCorp LLC). Statistical significance was set at $P<.05$ (2-tailed). For exploratory 3-way interaction analyses, *P* values between .05 and .10 were treated as marginal evidence and interpreted cautiously [36,37].

Results

Participants' Sociodemographic Characteristics

The final analytical sample consisted of 1022 Chinese MSM with a mean age of 29.6 (SD 7.6) years. The sample was predominantly homosexual-identified (767/1022, 75.1%), never married (896/1022, 87.7%), and well educated, with 87.8% (897/1022) holding a bachelor's degree or higher. Regarding monthly income, over one-third (369/1022, 36.1%) earned more than 8000 Yuan (about US $1157) monthly, while 13.3% (136/1022) earned less than 3000 Yuan (about US $434). Participants reported diverse sexual partnership patterns: 57.4% (587/1022) had 1 regular male partner, while 69.8% (714/1022) reported having casual partners in the past 6 months. Nearly one-quarter (253/1022, 24.8%) reported substance use during sexual activities (Table 2).

XSL•FO
RenderX

Table . Participants' sociodemographics.

Characteristic	Participants
Age (y)	
Mean (SD)	29.6 (7.6)
≤30, n (%)	658 (64.4)
>30, n (%)	364 (35.6)
Sexual orientation, n (%)	
Homosexual	767 (75.1)
Bisexual	223 (21.8)
Heterosexual	28 (2.7)
Unsure or other	4 (0.4)
Marital status, n (%)	
Never married	896 (87.7)
Engaged or married	78 (7.6)
Separated, divorced, or widowed	48 (4.7)
Highest education level, n (%)	
High school or below	125 (12.2)
Bachelor's degree	739 (72.3)
Graduate degree or above	158 (15.5)
Monthly income (Yuan)[a], n (%)	
≤3000	136 (13.3)
3001 - 5000	192 (18.8)
5001 - 8000	325 (31.8)
>8000	369 (36.1)
Number of regular male partners, n (%)	
0	163 (16.0)
1	587 (57.4)
≥2	272 (26.6)
Number of casual male partners, n (%)	
0	308 (30.2)
1	398 (38.9)
≥2	316 (30.9)
Substance use during sexual activities, n (%)	
No	769 (75.2)
Yes	253 (24.8)

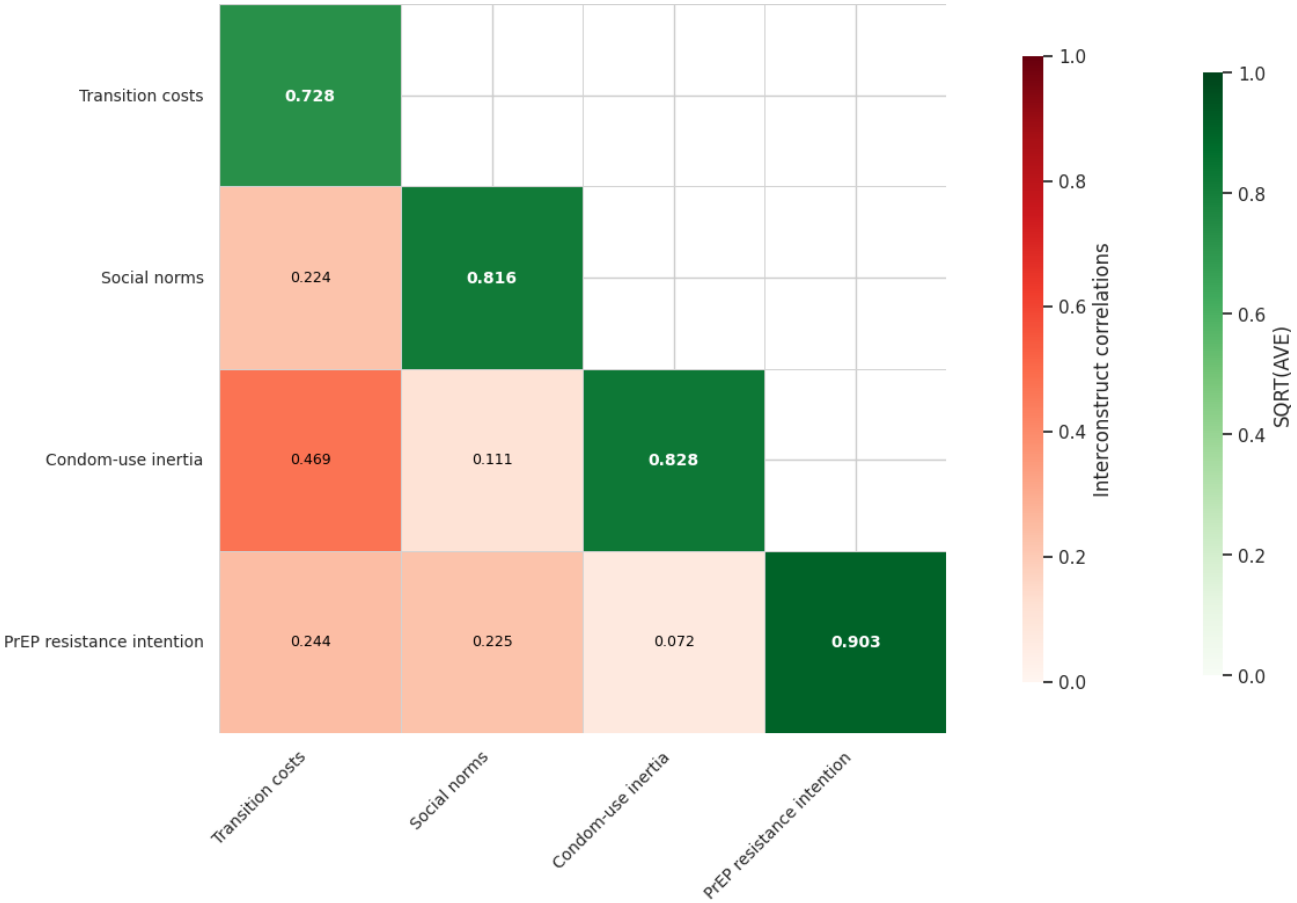
[a]At the time of analysis, CNY 1≈US $0.145.

Measurement Properties of Study Constructs

To support the regression analyses, we assessed the measurement properties of the theory-guided constructs. CFA demonstrated acceptable model fit for the measurement model (χ^2_{79}=523.8, P<.001; root mean square error of approximation=0.074, 90% CI 0.068 to 0.08; standardized root mean square residual=0.058; comparative fit index=0.947; Tucker-Lewis index=0.930) (Multimedia Appendix 2). All constructs demonstrated acceptable internal consistency and convergent validity. Internal consistency was excellent for PrEP resistance intention (Cronbach α=0.93) and good for condom-use inertia (α=.86), social norms (α=.85), and transition costs (α=.83). Composite reliability values exceeded 0.80 for all constructs. AVE values surpassed the 0.50 threshold, supporting convergent validity (Figure 1).

XSL•FO
RenderX

Figure 1. Discriminant validity assessment based on the Fornell-Larcker criterion. AVE: average variance extracted; PrEP: pre-exposure prophylaxis; SQRT: square root.



SQB Variables by PrEP Resistance Intention

Bivariate correlations revealed consistent associations (Multimedia Appendix 3). Transition costs (mean 4.7, SD 1.1) showed a small-to-medium positive correlation with PrEP resistance intention (mean 3.2, SD 1.6; r=0.21; P<.001), indicating that higher perceived switching costs to PrEP were associated with greater resistance. Social norms (mean 3.8, SD 1.2) showed a negative correlation of similar magnitude (r=–0.21; P<.001), suggesting that supportive social environments were associated with lower resistance. Condom-use inertia (mean 4.9, SD 1.3) showed a statistically significant but practically small association with resistance (r=0.08; P=.008), suggesting that its primary role might be as a moderator rather than a direct predictor.

Main and Moderation Analysis of PrEP Resistance Intention

Table 3 presents the results of the multiple linear regression models. In model 1, both SQB dimensions were significantly associated with PrEP resistance intention after adjusting for covariates. Transition costs were positively associated with PrEP resistance intention (β=0.26, 95% CI 0.17-0.35; P<.001). Social norms showed a comparable negative association (β=–0.19, 95% CI –0.27 to –0.12; P<.001), with more supportive normative environments linked to lower PrEP resistance intention. Condom-use inertia showed no significant direct association (β=–0.01, 95% CI –0.08 to 0.07; P=.80), supporting its hypothesized role as a moderator rather than a direct predictor of resistance intention.

XSL•FO
RenderX

Table . Main associations of status quo bias with pre-exposure prophylaxis resistance intentions and the moderating role of condom-use inertia[a].

Variables	Model 1: β (95% CI)[b]	Model 2: β (95% CI)[c]	Model 3: β (95% CI)[d]
Age (y)	–0.02 (–0.03 to –0)[e]	–0.02 (–0.03 to –0)[e]	–0.02 (–0.03 to –0)[e]
Sexual orientation (ref: Heterosexual)			
Homosexual	–1.57 (–2.14 to –1)[f]	–1.54 (–2.11 to –0.97)[f]	–1.58 (–2.15 to –1.01)[f]
Bisexual	–1.26 (–1.86 to –0.67)[f]	–1.24 (–1.83 to –0.65)[f]	–1.29 (–1.88 to –0.70)[f]
Unsure or other	–1.54 (–3.07 to –0.02)[e]	–1.51 (–3.04 to 0.01)[g]	–1.59 (–3.12 to –0.07)[e]
Marital status (ref: never married)			
Engaged or married	0.14 (–0.23 to 0.51)	0.12 (–0.25 to 0.49)	0.14 (–0.23 to 0.50)
Separated, divorced, or widowed	0.69 (0.22 to 1.15)[h]	0.66 (0.19 to 1.12)[h]	0.68 (0.22 to 1.15)[h]
Highest education level (ref: high school or below)			
Bachelor's degree	–0.76 (–1.05 to –0.47)[f]	–0.76 (–1.05 to –0.46)[f]	–0.75 (–1.04 to –0.46)[f]
Graduate degree or above	–0.79 (–1.16 to –0.43)[f]	–0.79 (–1.15 to –0.43)[f]	–0.78 (–1.15 to –0.42)[f]
Monthly income (Yuan)[i] (ref: ≤3000)			
3001 - 5000	–0.07 (–0.39 to 0.25)	–0.08 (–0.39 to 0.24)	–0.08 (–0.40 to 0.24)
5001 - 8000	0.08 (–0.22 to 0.37)	0.07 (–0.22 to 0.37)	0.07 (–0.22 to 0.37)
>8000	–0.01 (–0.31 to 0.29)	–0.01 (–0.31 to 0.29)	–0.02 (–0.32 to 0.28)
Number of regular male partners (ref: 0)			
1	–0.15 (–0.41 to 0.11)	–0.15 (–0.41 to 0.11)	–0.16 (–0.42 to 0.10)
≥2	0.06 (–0.23 to 0.36)	0.06 (–0.23 to 0.35)	0.06 (–0.23 to 0.35)
Number of casual male partners (ref: 0)			
1	0.31 (0.09 to 0.53)[h]	0.30 (0.08 to 0.52)[h]	0.31 (0.09 to 0.53)[h]
≥2	–0.47 (–0.72 to –0.21)[f]	–0.48 (–0.73 to –0.22)[f]	–0.46 (–0.72 to –0.21)[f]
Substance use (ref: no)			
Yes	–0.17 (–0.38 to 0.04)	–0.18 (–0.39 to 0.04)	–0.16 (–0.37 to 0.05)
Transition costs	0.26 (0.17 to 0.35)[f]	0.26 (0.17 to 0.35)[f]	0.27 (0.18 to 0.35)[f]
Social norms	–0.19 (–0.27 to –0.12)[f]	–0.19 (–0.27 to –0.12)[f]	–0.17 (–0.25 to –0.09)[f]
Inertia	–0.01 (–0.08 to 0.07)	0.00 (–0.07 to 0.08)	–0.00 (–0.08 to 0.08)
Transition costs × inertia	N/A[j]	0.04 (0.01 to 0.09)[e]	N/A
Social norms × inertia	N/A	N/A	–0.05 (–0.09 to –0)[e]

[a]Coefficients are shown with 95% CIs in parentheses. The F test for ΔR^2 indicates whether the variables added in each step significantly improved the model's explanatory power compared with the previous model.

[b]Model 1: Adjusted R^2=0.234.

[c]Model 2: Adjusted R^2=0.247; ΔAdjusted R^2=0.013; F test for ΔR^2=16.3[g].

[d]Model 3: Adjusted R^2=0.243; ΔAdjusted R^2=0.009; F test for ΔR^2=10.8[f].

[e]P<.05.

[f]P<.001.

[g]P<.10.

[h]P<.01.

[i]At the time of analysis, CNY 1≈US $0.145.

[j]N/A: not applicable.

Both hypothesized interactions were statistically significant, supporting the moderating role of condom-use inertia in SQB mechanisms. The transition costs × condom-use inertia interaction was significant (β=0.04, 95% CI 0.01-0.09; P=.03),

XSL•FO
RenderX

indicating that the positive association between transition costs and resistance intention was stronger among individuals with higher condom-use inertia. The social norms × condom-use inertia interaction was also significant (β=–0.05, 95% CI –0.09 to 0; *P*=.04), indicating that the negative association between supportive social norms and PrEP resistance intention was more pronounced among MSM with high condom-use inertia.

Moderating Role of Condom-Use Inertia Across Subgroups

Table 4 summarizes the patterns of these moderation effects across various subgroups. The amplifying influence of condom-use inertia on transition costs was particularly pronounced among several subgroups, including participants who were divorced, separated, or widowed (β=0.28, 95% CI 0.13-0.43; *P*<.001), those with a postgraduate degree (β=0.13, 95% CI 0.01-0.26; *P*=.04), individuals with 1 regular male partner (β=0.08, 95% CI 0.01-0.15; *P*=.02), and participants reporting no substance use (β=0.07, 95% CI 0.01-0.12; *P*=.01).

XSL•FO
RenderX

Table . Exploratory analysis of 3-way interactions among status quo bias, condom-use inertia, and covariates[a].

Characteristic	Transition costs × condom inertia	*P* value	Social norms × condom inertia	*P* value
Age (y)				
≤30	0.07 (–0.00 to 0.13)[b]	.05	–0.11 (–0.17 to –0.04)[c]	.001
>30	0.03 (–0.03 to 0.09)	.37	0.02 (–0.05 to 0.08)	.61
Sexual orientation				
Homosexual	0.04 (–0.02 to 0.09)	.21	–0.07 (–0.12 to –0.02)[d]	.01
Bisexual	0.05 (–0.03 to 0.14)	.23	0.01 (–0.08 to 0.10)	.85
Heterosexual	0.18 (–0.17 to 0.54)	.32	0.05 (–0.23 to 0.33)	.72
Unsure or other	–1.69 (–4.85 to 1.47)	.29	0.12 (–1.08 to 1.32)	.85
Marital status				
Never married	0.04 (–0.01 to 0.10)	.13	–0.06 (–0.11 to –0.01)[d]	.01
Engaged or married	–0.05 (–0.16 to 0.06)	.35	0.06 (–0.06 to 0.18)	.33
Separated, divorced, or widowed	0.28 (0.13 to 0.43)[c]	<.001	–0.08 (–0.26 to 0.11)	.42
Highest education level				
High school or below	0.01 (–0.10 to 0.12)	.84	0.02 (–0.10 to 0.14)	.73
Bachelor's degree	0.04 (–0.02 to 0.09)	.21	–0.06 (–0.11 to –0)[d]	.04
Graduate degree or above	0.13 (0.01 to 0.26)[d]	.04	–0.07 (–0.18 to 0.03)	.18
Monthly income (Yuan)				
≤5000	0.07 (–0.01 to 0.15)	.10	–0.17 (–0.38 to 0.04)	.11
>5000	0.03 (–0.02 to 0.09)	.24	–0.05 (–0.13 to 0.03)	.21
Number of regular male partners				
0	0.03 (–0.07 to 0.13)	.54	–0.01 (–0.10 to 0.08)	.76
1	0.08 (0.01 to 0.15)[d]	.02	–0.03 (–0.10 to 0.04)	.35
≥2	0.01 (–0.07 to 0.09)	.86	–0.10 (–0.18 to –0.02)[d]	.01
Number of casual male partners				
0	0.09 (–0.02 to 0.19)	.10	–0.12 (–0.23 to –0.02)[d]	.03
1	0.02 (–0.04 to 0.08)	.50	0 (–0.06 to 0.06)	.96
≥2	0.06 (–0.02 to 0.15)	.15	–0.09 (–0.17 to –0.01)[d]	.03
Substance use				
No	0.07 (0.01 to 0.12)[d]	.01	–0.04 (–0.09 to 0.02)	.18
Yes	–0.02 (–0.11 to 0.06)	.60	–0.07 (–0.15 to 0.01)[b]	.09

[a]The results are shown after controlling for all covariates and status quo bias variables.

[b]*P*<.1.

[c]*P*<.001.

[d]*P*<.05.

Conversely, the enhancing influence of condom-use inertia on the negative association with social norms was most evident among younger MSM (aged ≤30 y: β=–0.11, 95% CI –0.17 to –0.04; *P*=.001), homosexual-identified men (β=–0.07, 95% CI –0.12 to –0.02; *P*=.01), unmarried individuals (β=–0.06, 95% CI –0.11 to –0.01; *P*=.01), those with a bachelor's degree (β=–0.06, 95% CI –0.11 to 0; *P*=.04), individuals with 2 or more regular partners (β=–0.10, 95% CI –0.18 to –0.02; *P*=.01), and those with no casual partners (β=–0.12, 95% CI –0.23 to –0.02; *P*=.03).

XSL•FO
RenderX

Discussion

Principal Findings

This study examines the psychological mechanisms underlying resistance to PrEP adoption among Chinese MSM, focusing on key dimensions of SQB—transition costs and social norms—and the moderating role of condom-use inertia. Our findings show that both dimensions of SQB are significantly associated with PrEP resistance intention, with condom-use inertia playing a notable moderating role across psychological pathways. This study also highlights the heterogeneity of decision-making processes within MSM populations and may help explain why resistance to PrEP adoption persists despite increasing accessibility. These insights move beyond identifying descriptive barriers to illuminating the specific cognitive and behavioral processes at play, informing the design of behaviorally informed public health strategies.

This study provides empirical evidence on associations between SQB-related factors and PrEP resistance intention among MSM. Our findings suggest that PrEP resistance intention was higher among individuals reporting high transition costs, specifically in terms of time investment, financial costs, and potential side effects. National guidelines recommend HIV testing 1 month after PrEP initiation and quarterly follow-up thereafter, including renal and liver function monitoring [38]. Previous studies indicate that obtaining PrEP through licensed clinics in China is often perceived as time-consuming, requiring multiple visits for testing, prescription, and follow-up [8,39]. Limited provider availability has led some MSM to seek PrEP through informal channels (eg, Dài-Gòu), raising concerns about drug safety [39]. In addition, PrEP in China is largely paid for out of pocket, including medication and required clinical monitoring, and cost has consistently been reported as a major barrier among MSM [40]. Conversely, a supportive social environment was associated with lower resistance, highlighting the inverse relationship with perceived social norms. This pattern aligns with previous literature linking transition costs and normative factors to PrEP-related attitudes and behaviors [24,25,41], and our study extends this work by explicitly framing the outcome as PrEP resistance intention, providing more comprehensive evidence on the active rejection process.

We also identified condom-use inertia as a significant psychological moderator in PrEP decision-making. Although habitual condom use is protective and remains essential for preventing a wide range of sexually transmitted infections, a strong reliance on this established routine may be associated with psychological friction when individuals consider adding or transitioning to PrEP. This friction may relate to the cognitive stability and predictability associated with long-standing habits [42,43]. In this context, condom-use inertia may act as a powerful cognitive filter that reinforces the status quo and amplifies the perceived costs of behavioral change [44,45]. Our findings indicated that men with entrenched condom-use habits perceived the costs and hassles of initiating PrEP as significantly more burdensome, which was associated with greater resistance. Therefore, addressing inertia may help reduce the extent to which established condom-use routines are associated with PrEP resistance.

Paradoxically, this same inertia also amplified the positive association with social norms, suggesting a dual-amplifier mechanism. Individuals with strong habits appear particularly responsive to normative cues, potentially because such cues can penetrate habitual cognitive filters more effectively than informational messages alone [46]. This finding suggests the moderating role of inertia: it strengthens resistance to change when signals support maintaining existing condom routines, yet enhances receptiveness when signals favor PrEP adoption. This dual-amplifier mechanism is particularly relevant for MSM who meet clinical criteria for PrEP, including those with higher sexual activity, multiple recent partners, or increased HIV exposure risk. For these groups, PrEP does not replace condoms but provides supplementary HIV-specific protection where condom effectiveness may be inconsistent or challenged by real-world sexual practices [47].

Exploratory subgroup analyses suggested that the moderating role of condom-use inertia varies across demographic and behavioral profiles. The interaction between condom-use inertia and transition costs was more pronounced among participants with a postgraduate degree, MSM with 1 regular partner, and those reporting no substance use. For example, among participants with higher education, prior research suggests a higher likelihood of reporting consistent condom use [48]. When condom-use inertia is high, PrEP may be perceived as unnecessary and more burdensome to initiate and maintain, magnifying perceived transition burdens and being linked to higher PrEP resistance intention. Along similar lines, MSM reporting one regular partner or no substance use may perceive condoms as sufficiently protective for their current risk profile; high condom-use inertia may therefore make PrEP-related transition costs seem less justified, reinforcing PrEP resistance intention. This suggests that rigid behavioral routines may heighten the psychological burden of initiating PrEP among these groups. Strategies consistent with the MINDSPACE (Messenger, Incentives, Norms, Defaults, Salience, Priming, Affect, Commitments, Ego) framework, such as simplified defaults for PrEP initiation, streamlined refill procedures, and increased salience of concrete PrEP benefits, may help mitigate perceived transition barriers [49,50]. For instance, PrEP eligibility screening or counseling could be embedded within routine HIV and sexual health services using an opt-out default, which may reduce initiation friction. Follow-up and refills could be streamlined (eg, 1-click scheduling, prefilled forms, and automated reminders), which may lower perceived time and hassle costs [51,52].

In contrast, the interaction between condom-use inertia and social norms was most evident among younger MSM, homosexual-identified men, unmarried individuals, those with a bachelor's degree, and MSM with multiple regular partners, with inertia appearing to strengthen the negative association with social norms. Younger MSM and homosexual-identified men are often more embedded in MSM peer networks and community interactions, where HIV prevention information, attitudes, and behavioral modeling circulate more frequently [53]. Condom-use inertia may reflect the cumulative effects of

XSL•FO
RenderX

prior prevention socialization, including repeated reinforcement of condom-related norms, skills, and confidence within peer and sexual networks [54,55]. Within this normative context, individuals anchored in condom-use routines may be more responsive to supportive social endorsement, which may increase the perceived acceptability and feasibility of PrEP and lower resistance intention. Accordingly, norm-based nudges could increase the visibility of supportive norms in MSM social spaces by communicating descriptive and injunctive norms that normalize PrEP consideration. Credible messengers, including peer leaders, PrEP-experienced MSM, and trusted clinicians, may further strengthen normative influence and enhance the perceived feasibility of initiation and persistence [56,57].

Limitations

Several limitations of this study should be considered when interpreting the results, and these also point to directions for future research. First, this study employed a cross-sectional design, with all variables measured at a single time point. Although this design can identify associations between variables, it does not allow for temporal ordering or causal inference. Second, the sample for this study was drawn from the Chinese cultural and demographic context. While the findings provide valuable localized insights into the PrEP decision-making process among this group, their generalizability may be limited. The factors influencing SQB, such as the nature of social norms and the specific components of transition costs, may differ significantly across other cultural, social, and healthcare system contexts. Third, our reliance on online convenience and snowball sampling limits generalizability to all MSM in China. Recruitment via partner CBO or Centers for Disease Control and Prevention channels and peer sharing likely over-sampled younger, urban, and highly educated individuals, potentially under-representing MSM with lower educational attainment who may differ in HIV prevention knowledge, digital health literacy, and access to PrEP services. Finally, we assessed the prespecified measurement structure using CFA and standard discriminant validity checks; however, we did not conduct exploratory factor analysis or validate the measurement model in an independent sample. Future studies should replicate the factor structure in separate samples and consider incorporating exploratory factor analysis to further evaluate construct dimensionality and strengthen the psychometric evidence base.

Conclusion

This study applied the theory of SQB to explain PrEP resistance intention, offering a novel theoretical perspective and empirical tools for understanding the intention-behavior gap in public health. Crucially, we identified the moderating role of condom-use inertia, highlighting the cognitive friction involved in shifting from established condom-based routines to new prevention methods. In conclusion, our findings confirm that SQB is an important yet often overlooked factor associated with PrEP adoption among MSM. Integrating insights from behavioral economics into HIV prevention strategies has significant theoretical and practical importance for addressing challenges in PrEP promotion and for informing more effective public health interventions.

Acknowledgments

We express our sincere gratitude to the Social Entrepreneurship for Sexual Health team for their invaluable efforts in data collection for this study. We are also deeply grateful to all the participants who generously gave their time and shared their perspectives; this research would not have been possible without their participation. Generative artificial intelligence was used only for limited English grammar and language editing. All scientific content, interpretation, and references were reviewed and verified by the authors, who take full responsibility for the manuscript.

Funding

This work was supported by the Ministry of Education of the People’s Republic of China Project of Humanities and Social Sciences (grant 24YJC840051), the Natural Science Foundation of Shaanxi Province (grant 2025JC-YBQN-1106), the Fundamental Research Funds for the Central Universities (grant SK2026027), the Ministry of Science and Technology of the People’s Republic of China (grants 2022YFC2304900 and 2022YFC2304905), and the National Key Research and Development Program of China (grants 2022YFC2505100 and 2022YFC2505103). The funders had no role in the study design, data collection, analysis, data interpretation, manuscript writing, or the decision to submit the paper for publication.

Data Availability

The data presented in this study are not publicly available due to privacy and ethical restrictions related to the sensitive nature of the information collected from marginalized populations. Deidentified data may be available upon reasonable request from the corresponding author.

Authors' Contributions

Conceptualization: MZ
Data curation: WT
Formal analysis: MZ, XZ
Funding acquisition: LZ
Investigation: WT, ZY

XSL•FO
RenderX

Methodology: ZY
Project administration: WT
Resources: WT
Supervision: LZ, FJ
Validation: MZ, SL
Visualization: XZ
Writing – original draft: MZ
Writing – review & editing: MZ, WT, XZ, LZ, FJ, ZY, SL, JH

Conflicts of Interest

None declared.

Multimedia Appendix 1
Participant flow diagram.
[PNG File, 27 KB - publichealth_v12i1e88806_app1.png]

Multimedia Appendix 2
Reliability and convergent validity.
[DOCX File, 19 KB - publichealth_v12i1e88806_app2.docx]

Multimedia Appendix 3
Bivariate associations of study constructs with pre-exposure prophylaxis resistance intention.
[PNG File, 406 KB - publichealth_v12i1e88806_app3.png]

Checklist 1
CHERRIES checklist.
[DOCX File, 21 KB - publichealth_v12i1e88806_app4.docx]

References

1. HIV data and statistics. World Health Organization. URL: https://www.who.int/teams/global-hiv-hepatitis-and-stis-programmes/hiv/strategic-information/hiv-data-and-statistics [accessed 2026-03-30]
2. Leal M, Kerr L, Mota RMS, et al. Increasing HIV prevalence rate among men who have sex with men: results of a comparison of two national surveys. AIDS 2024 Oct 1;38(12):1799-1801. [doi: 10.1097/QAD.0000000000003934] [Medline: 39206885]
3. HIV and gay men and other men who have sex with men—thematic briefing note—2024 global AIDS update the urgency of now: AIDS at a crossroads. : UNAIDS (Joint United Nations Programme on HIV/AIDS); 2024 URL: https://www.unaids.org/sites/default/files/media_asset/2024-unaids-global-aids-update-gay-men_en.pdf [accessed 2026-03-25]
4. Anderson PL, Glidden DV, Liu A, et al. Emtricitabine-Tenofovir concentrations and pre-exposure prophylaxis efficacy in men who have sex with men. Sci Transl Med 2012 Sep 12;4(151):151ra125. [doi: 10.1126/scitranslmed.3004006] [Medline: 22972843]
5. Liu Y, Xian Y, Liu X, et al. Significant insights from a national survey in China: PrEP awareness, willingness, uptake, and adherence among YMSM students. BMC Public Health 2024;24(1):1009. [doi: 10.1186/s12889-024-18512-y]
6. Zhou J, Xu Y, Li Q, et al. Understanding awareness, utilization, and the awareness-utilization gap of HIV PrEP and nPEP among young MSM in China. AIDS Behav 2025 Apr;29(4):1327-1339. [doi: 10.1007/s10461-024-04606-6] [Medline: 39779626]
7. Zimmermann HM, Eekman SW, Achterbergh RC, et al. Motives for choosing, switching and stopping daily or event-driven pre-exposure prophylaxis - a qualitative analysis. J Int AIDS Soc 2019 Oct;22(10):e25389. [doi: 10.1002/jia2.25389] [Medline: 31612621]
8. Lin C, Li L, Liu J, et al. HIV PrEP services for MSM in China: a mixed-methods study. AIDS Care 2022 Mar;34(3):310-314. [doi: 10.1080/09540121.2021.1896661] [Medline: 33650447]
9. Reisner SL, Moore CS, Asquith A, et al. High risk and low uptake of pre-exposure prophylaxis to prevent HIV acquisition in a national online sample of transgender men who have sex with men in the United States. J Int AIDS Soc 2019 Sep;22(9):e25391. [doi: 10.1002/jia2.25391] [Medline: 31536171]
10. Yellin H, Levy ME, Magnus M, Kuo I, Siegel M. HIV risk perception, willingness to use PrEP, and PrEP uptake among young men who have sex with men in Washington, DC. AIDS Behav 2023 Sep;27(9):2844-2854. [doi: 10.1007/s10461-023-04008-0] [Medline: 36807246]
11. Wu Y, Xie L, Meng S, et al. Mapping potential pre-exposure prophylaxis users onto a motivational cascade: identifying targets to prepare for implementation in China. LGBT Health 2019 Jul;6(5):250-260. [doi: 10.1089/lgbt.2018.0256] [Medline: 31170020]

XSL•FO
RenderX

12. Fleming SM, Thomas CL, Dolan RJ. Overcoming status quo bias in the human brain. Proc Natl Acad Sci U S A 2010 Mar 30;107(13):6005-6009. [doi: 10.1073/pnas.0910380107] [Medline: 20231462]
13. Godefroid ME, Plattfaut R, Niehaves B. How to measure the status quo bias? A review of current literature. Manag Rev Q 2023 Dec;73(4):1667-1711. [doi: 10.1007/s11301-022-00283-8]
14. Li J, Liu M, Liu X. Why do employees resist knowledge management systems? An empirical study from the status quo bias and inertia perspectives. Comput Human Behav 2016 Dec;65:189-200. [doi: 10.1016/j.chb.2016.08.028]
15. Owens C, Hubach RD, Williams D, et al. Facilitators and barriers of pre-exposure prophylaxis (PrEP) uptake among rural men who have sex with men living in the Midwestern U.S. Arch Sex Behav 2020 Aug;49(6):2179-2191. [doi: 10.1007/s10508-020-01654-6] [Medline: 32219687]
16. Arnold T, Giorlando KK, Barnett AP, et al. Social, structural, behavioral, and clinical barriers influencing pre-exposure prophylaxis (PrEP) use among young black men who have sex with men in the South: a qualitative update to a 2016 Study. Arch Sex Behav 2024 Feb;53(2):785-797. [doi: 10.1007/s10508-023-02721-4] [Medline: 37891436]
17. Koppe U, Marcus U, Albrecht S, et al. Barriers to using HIV pre-exposure prophylaxis (PrEP) and sexual behaviour after stopping PrEP: a cross-sectional study in Germany. BMC Public Health 2021 Jan 19;21(1):159. [doi: 10.1186/s12889-021-10174-4] [Medline: 33468089]
18. Katz IT, Ngure K, Kamolloh K, et al. Multi-level factors driving pre-exposure prophylaxis non-initiation among young women at high risk for HIV in Kenya. AIDS Behav 2023 Jan;27(1):106-118. [doi: 10.1007/s10461-022-03748-9] [Medline: 35930203]
19. Samuelson W, Zeckhauser R. Status quo bias in decision making. J Risk Uncertainty 1988 Mar;1(1):7-59. [doi: 10.1007/BF00055564]
20. Johnson LM, Green HD Jr, Koch B, et al. Role of social networks and social norms in future PrEP use in a racially diverse sample of at-risk women and members of their social networks. J Acquir Immune Defic Syndr 2021 Apr 1;86(4):422-429. [doi: 10.1097/QAI.0000000000002571] [Medline: 33196549]
21. Dillon FR, Ertl MM, Eklund AC, et al. Sexual identity development and social ecological facilitators and barriers of PrEP uptake and adherence among Latinx men who have sex with men. Arch Sex Behav 2024 Mar;53(3):1197-1211. [doi: 10.1007/s10508-023-02776-3] [Medline: 38212437]
22. Rivera DB, Brady JP, Blashill AJ. Traditional machismo, caballerismo, and the pre-exposure prophylaxis (PrEP) cascade among a sample of Latino sexual minority men. J Sex Res 2021 Jan;58(1):21-28. [doi: 10.1080/00224499.2020.1743961] [Medline: 32243199]
23. Onwubiko UN, Murray SM, Rao A, et al. Individual & joint associations of sexual stigma and mental distress with PrEP uptake, adherence and persistence among US gay and bisexual men. Soc Sci Med 2024 Dec;363:117493. [doi: 10.1016/j.socscimed.2024.117493] [Medline: 39541830]
24. Agaku I, Nkosi L, Gwar J, Tsafa T. Gender norms, HIV risk, and attitudes towards pre-exposure prophylaxis and other HIV preventive interventions among South African adolescents. Pan Afr Med J 2022;41(136):136. [doi: 10.11604/pamj.2022.41.136.32881] [Medline: 35519158]
25. Irie WC, Mahone A, Nakka R, Ghebremichael M. Factors associated with comfort discussing PrEP with healthcare providers among black cisgender women. Trop Med Infect Dis 2023 Sep 7;8(9):436. [doi: 10.3390/tropicalmed8090436] [Medline: 37755897]
26. Wood W. Habits, goals, and effective behavior change. Curr Dir Psychol Sci 2024 Aug;33(4):226-232. [doi: 10.1177/09637214241246480]
27. Wood W, Mazar A, Neal DT. Habits and goals in human behavior: separate but interacting systems. Perspect Psychol Sci 2022 Mar;17(2):590-605. [doi: 10.1177/1745691621994226] [Medline: 34283681]
28. Leng T, Kessou L, Heitner J, et al. Potential impact and cost-effectiveness of oral HIV pre-exposure prophylaxis for men who have sex with men in Cotonou, Benin: a mathematical modelling study. Lancet Glob Health 2025 Jun;13(6):e1111-e1121. [doi: 10.1016/S2214-109X(25)00098-1] [Medline: 40412400]
29. Han J, Bouey JZ, Wang L, et al. PrEP uptake preferences among men who have sex with men in China: results from a National Internet Survey. J Int AIDS Soc 2019 Feb;22(2):e25242. [doi: 10.1002/jia2.25242] [Medline: 30724470]
30. Jaramillo J, Pagkas-Bather J, Waters K, et al. Perceptions of sexual risk, PrEP services, and peer navigation support among HIV-negative Latinx and black men who have sex with men (MSM) residing in western Washington. Sex Res Social Policy 2022 Sep;19(3):1058-1068. [doi: 10.1007/s13178-021-00595-6] [Medline: 36387840]
31. Farley JE, Beuchamp G, Bergman A, et al. The impact of stigma and sexual identity on PrEP awareness and use among at-risk men who have sex with men in four U.S. cities (HPTN 078). Stigma Health 2024 Aug;9(3):400-410. [doi: 10.1037/sah0000413] [Medline: 39148912]
32. Zhao M, Zhao X, Zhang Y, et al. How status quo bias shapes willingness, uptake, and adherence to PrEP among Chinese MSM: a behavioral economics perspective. Infect Dis Ther 2026 Mar;15(3):749-767. [doi: 10.1007/s40121-025-01296-y] [Medline: 41575704]
33. Storholm ED, Klein DJ, Pedersen ER, et al. Sociodemographic and behavioral risk correlates of PrEP interest and use among young adults experiencing homelessness in Los Angeles. AIDS Behav 2024 Apr;28(4):1216-1226. [doi: 10.1007/s10461-023-04144-7] [Medline: 37698638]

XSL•FO
RenderX

34. Tu Z, He S, Zhao R, et al. Preferences for HIV pre-exposure prophylaxis among gay, bisexual, and men who have sex with men in China: a discrete choice experiment. AIDS Behav 2025 Nov;29(11):3486-3497. [doi: 10.1007/s10461-025-04793-w] [Medline: 40465041]
35. Lu F, She B, Zhao R, et al. Identifying high-risk populations for sexually transmitted infections in Chinese men who have sex with men: a cluster analysis. Open Forum Infect Dis 2025 Jan;12(1):ofae754. [doi: 10.1093/ofid/ofae754] [Medline: 39829637]
36. Wasserstein RL, Lazar NA. The ASA statement on p-values: context, process, and purpose. Am Stat 2016 Apr 2;70(2):129-133. [doi: 10.1080/00031305.2016.1154108]
37. Wasserstein RL, Schirm AL, Lazar NA. Moving to a world beyond “p < 0.05”. Am Stat 2019;73:1-19. [doi: 10.1080/00031305.2019.1583913]
38. Acquired Immunodeficiency Syndrome Professional Group, Society of Infectious Diseases, Chinese Medical Association; Chinese Center for Disease Control and Prevention. Chinese guidelines for the diagnosis and treatment of human immunodeficiency virus infection/acquired immunodeficiency syndrome (2024 edition). Chin Med J (Engl) 2024 Nov 20;137(22):2654-2680. [doi: 10.1097/CM9.0000000000003383] [Medline: 39602327]
39. Li C, Xiong Y, Liu H, et al. Multifaceted barriers to rapid roll-out of HIV pre-exposure prophylaxis in China: a qualitative study among men who have sex with men. Int J Behav Med 2024 Apr;31(2):252-262. [doi: 10.1007/s12529-023-10177-y] [Medline: 37156941]
40. Hong C, Wang L, Wang Y, et al. Barriers and facilitators to engagement with the PrEP care continuum among sexual minority men in China: a systematic review. Sex Health 2025 Aug;22(4):SH25021. [doi: 10.1071/SH25021] [Medline: 40977211]
41. Case KK, Gomez GB, Hallett TB. The impact, cost and cost-effectiveness of oral pre-exposure prophylaxis in sub-Saharan Africa: a scoping review of modelling contributions and way forward. J Int AIDS Soc 2019 Sep;22(9):e25390. [doi: 10.1002/jia2.25390] [Medline: 31538407]
42. Badawy SM, Shah R, Beg U, Heneghan MB. Habit strength, medication adherence, and habit-based mobile health interventions across chronic medical conditions: systematic review. J Med Internet Res 2020 Apr 28;22(4):e17883. [doi: 10.2196/17883] [Medline: 32343250]
43. Drake C, Ryan C, Dowd B. Sources of inertia in the individual health insurance market. J Public Econ 2022 Apr;208:104622. [doi: 10.1016/j.jpubeco.2022.104622]
44. Di Ciaccio M, Delabre RM, Vilotitch A, et al. PrEP in the context of other HIV risk reduction strategies among men who have sex with men: results from the Flash! PrEP in Europe Survey. Arch Sex Behav 2022 Jul;51(5):2451-2464. [doi: 10.1007/s10508-022-02322-7] [Medline: 35725849]
45. Rotsaert A, Nöstlinger C, Van Landeghem E, et al. Pre-exposure prophylaxis users’ attitudes about sexually transmitted infections and its influence on condom use: a mixed-method study in Belgium. AIDS Patient Care STDS 2022 Dec;36(12):483-492. [doi: 10.1089/apc.2022.0172] [Medline: 36484763]
46. Rotsaert A, Smekens T, Vuylsteke B, et al. Patterns of PrEP and condom use among PrEP users in Belgium: a web-based longitudinal study. BMC Public Health 2023 May 26;23(1):970. [doi: 10.1186/s12889-023-15786-6] [Medline: 37237375]
47. Bunting SR, Feinstein BA, Calabrese SK, et al. The role of social biases, race, and condom use in willingness to prescribe HIV pre-exposure prophylaxis to MSM: an experimental, vignette-based study. J Acquir Immune Defic Syndr 2022 Dec 1;91(4):353-363. [doi: 10.1097/QAI.0000000000003072] [Medline: 36288542]
48. Liu W, Chen X, Wang S, Wang Y, Lin H, Yao Y. Trends of sexual behaviors and related factors in high-risk MSM from 2010 to 2023: a repeated cross-sectional study in Zhejiang, China. BMC Public Health 2025;25(1):4264. [doi: 10.1186/s12889-025-25517-8]
49. Andrawis A, Tapa J, Vlaev I, et al. Applying behavioural insights to HIV prevention and management: a scoping review. Curr HIV/AIDS Rep 2022 Oct;19(5):358-374. [doi: 10.1007/s11904-022-00615-z] [Medline: 35930186]
50. Mantell JE, Bauman LJ, Bonett S, et al. Innovation in providing equitable pre‐exposure prophylaxis services in the United States: expanding access in nontraditional settings. JAIDS 2025;98(5S):e156-e169. [doi: 10.1097/QAI.0000000000003610]
51. Merle JL, Benbow N, Li DH, et al. Improving delivery and use of HIV pre-exposure prophylaxis in the US: a systematic review of implementation strategies and adjunctive interventions. AIDS Behav 2024 Jul;28(7):2321-2339. [doi: 10.1007/s10461-024-04331-0] [Medline: 38564136]
52. Bonett S, Li Q, Sweeney A, Gaither-Hardy D, Safa H. Telehealth models for PrEP delivery: a systematic review of acceptability, implementation, and impact on the PrEP care continuum in the United States. AIDS Behav 2024 Sep;28(9):2875-2886. [doi: 10.1007/s10461-024-04366-3] [Medline: 38856846]
53. Shrader CH, Duncan DT, Driver R, et al. Investigating social network peer effects on HIV care engagement using a fuzzy-like matching approach: cross-sectional secondary analysis of the N2 Cohort Study. JMIR Public Health Surveill 2025 May 8;11:e64497. [doi: 10.2196/64497] [Medline: 40338201]
54. Zhao M, Hai L, Zhao X, Ji F, Zhang L, Liu H. How do egocentric sexual networks affect male migrants’ HIV sexual behavior patterns in gender-imbalanced China? The moderating role of migration and neighborhoods. Arch Sex Behav 2026 Feb 12. [doi: 10.1007/s10508-025-03347-4] [Medline: 41678118]

XSL•FO
RenderX

55. Zhao M, Liu H, Chen X, Feldman MW. Sexual network and condom use among male migrants in the context of China's gender imbalance. AIDS Care 2022 Aug;34(8):1048-1052. [doi: 10.1080/09540121.2021.1938965] [Medline: 34115571]
56. Rogers BG, Whiteley L, Haubrick KK, Mena LA, Brown LK. Intervention messaging about pre-exposure prophylaxis use among young, black sexual minority men. AIDS Patient Care STDS 2019 Nov;33(11):473-481. [doi: 10.1089/apc.2019.0139] [Medline: 31682168]
57. Wirawan GBS, Schimdt HM, Chan C, et al. PrEP use and willingness cascades among GBMSM in 15 Asian countries/territories: an analysis of the PrEP APPEAL survey. J Int AIDS Soc 2025 Apr;28(4):e26438. [doi: 10.1002/jia2.26438] [Medline: 40156168]

Abbreviations

AVE: average variance extracted
CBO: community-based organization
CFA: confirmatory factor analysis
CHERRIES: Checklist for Reporting Results of Internet E-Surveys
MINDSPACE: Messenger, Incentives, Norms, Defaults, Salience, Priming, Affect, Commitments, Ego
MSM: men who have sex with men
PrEP: pre-exposure prophylaxis
SQB: status quo bias

Edited by GJB Sousa; submitted 02.Dec.2025; peer-reviewed by K Jiao, X Chen; revised version received 07.Mar.2026; accepted 07.Mar.2026; published 10.Apr.2026.

Please cite as:
Zhao M, Li S, Zhao X, Hu J, Yin Z, Ji F, Tang W, Zhang L
Moderating Role of Condom-Use Inertia on the Association Between Status Quo Bias and Pre-Exposure Prophylaxis Resistance Intention Among Chinese Men Who Have Sex With Men: Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e88806
URL: https://publichealth.jmir.org/2026/1/e88806
doi:10.2196/88806

© Min Zhao, Shanshan Li, Xiang Zhao, Jincheng Hu, Zhuoheng Yin, Fanpu Ji, Weiming Tang, Lei Zhang. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 10.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Stigma, Food Insecurity, and Limited Social Support as Psychosocial Correlates of Antiretroviral Therapy Adherence Among Pregnant Women Living With HIV: Cross-Sectional Study

Setor K Sorkpor[1*], RN, MSN, MPH, PhD; Ibrahim Yigit[2*], PhD; Rachel G A Thompson[3*], PhD; Brittany L Lane[2*], MPH, PhD; Geoffrey Anguyo[4], MD; Jerry John Ouner[5*], RN, PhD

[1]College of Nursing, Brain Science and Symptom Management Center, Florida State University, Tallahassee, FL, United States

[2]College of Nursing, Florida State University, Tallahassee, FL, United States

[3]University of Ghana, Africa Interdisciplinary Research Institute, University of California San Francisco, 1975 4th St, San Francisco, CA, United States

[4]Kigezi Healthcare Foundation, Mbarara University of Science and Technology, Kabale, Uganda

[5]Department of Family Health Care Nursing, University of California San Francisco, San Francisco, CA, United States

[*]these authors contributed equally

Corresponding Author:
Rachel G A Thompson, PhD
University of Ghana, Africa Interdisciplinary Research Institute, University of California San Francisco, 1975 4th St, San Francisco, CA, United States

Abstract

Background: Adherence to antiretroviral therapy (ART) during pregnancy is critical for maternal health and the prevention of vertical HIV transmission. In Uganda, where HIV prevalence remains high, pregnant women living with HIV face intersecting structural and psychosocial challenges, including stigma, food insecurity, and limited social support. Although each factor has been linked to ART nonadherence, less is known about how these factors co-occur within individuals and jointly shape vulnerability to nonadherence during pregnancy.

Objective: This study used latent profile analysis to identify empirically derived psychosocial vulnerability profiles reflecting the co-occurrence of stigma, food insecurity, and limited social support among pregnant women living with HIV in Uganda and examine whether profile membership is associated with ART adherence.

Methods: We conducted a cross-sectional analysis of 167 pregnant women living with HIV recruited from 6 health facilities in Uganda between June and December 2020. Measures included experienced HIV stigma, internalized HIV stigma, household food insecurity, and interpersonal social support. Latent profile analysis identified psychosocial profiles, and linear regression models assessed associations between profile membership and ART adherence, adjusting for sociodemographic covariates.

Results: A 2-class solution provided the best model fit (entropy=0.93). The higher-risk profile (75/167, 45.1%) was characterized by increased experienced stigma (mean score 1.97, SE 0.05), higher internalized stigma (mean score 2.66, SE 0.07), severe food insecurity (mean score 7.53, SE 0.20), and lower social support (mean score 2.06, SE 0.04). The lower-risk profile (92/167, 54.9%) showed significantly lower internalized stigma (mean score 2.32, SE 0.04; $P<.001$), lower experienced stigma (mean score 1.72, SE 0.05; $P<.001$), minimal food insecurity (mean score 0.82, SE 0.15; $P<.001$), and slightly higher social support (mean score 2.14, SE 0.04; $P<.001$). Membership in the higher-risk profile was associated with significantly lower ART adherence compared with membership in the lower-risk profile (B=0.88; β=0.40, 95% CI 0.02-0.78; $P=.04$).

Conclusions: Distinct psychosocial profiles are meaningfully associated with ART adherence among pregnant women living with HIV in Uganda. By moving beyond single-risk models, these findings demonstrate the added value of person-centered analytic approaches for identifying subgroups of pregnant women living with HIV exposed to clustered psychosocial and structural vulnerabilities. The findings underscore the need for multicomponent, person-centered interventions that simultaneously address stigma, food insecurity, and limited social support rather than treating these challenges in isolation. Identifying empirically derived vulnerability profiles provides a targeted framework for prioritizing high-risk subgroups and informing contextually appropriate intervention design. Addressing these intersecting vulnerabilities is essential for improving maternal ART adherence and supporting efforts to prevent vertical HIV transmission in high-burden settings.

(JMIR Public Health Surveill 2026;12:e84914) doi:10.2196/84914

XSL•FO
RenderX

KEYWORDS

prenatal women; sub-Saharan Africa; discrimination; internalized stigma; interpersonal support; food scarcity; psychosocial factors

Introduction

Although HIV prevention and treatment coverage have expanded across sub-Saharan Africa (SSA), Uganda continues to shoulder a high HIV burden [1,2]. An estimated 1.5 million Ugandans live with HIV, 57% of whom are women [1]. Against this backdrop, pregnancy and the early postpartum period are particularly consequential for maternal health and prevention of vertical transmission [3,4]. Accordingly, consistent antiretroviral therapy (ART) adherence during this window is essential [4-6]. Uganda has adopted HIV treatment for all, a World Health Organization strategy that provides lifelong ART to all people living with HIV, including pregnant and breastfeeding women living with HIV regardless of clinical stage or CD4 T-lymphocyte count [6-9]. However, nonadherence remains common, particularly during the postpartum period. Evidence from Uganda and SSA indicates that up to one-third of women miss doses of ART, thereby increasing the risk of virologic failure and mother-to-child transmission [6,10-12].

In this context, a complex array of structural and psychosocial factors undermines ART adherence among pregnant and postpartum women living with HIV in Uganda [10]. HIV-related stigma remains one of the most pervasive barriers to sustained engagement in care [1,13,14]. Stigma operates at multiple levels: externalized stigma (discrimination, rejection, and social exclusion by others) and internalized stigma (self-directed shame, guilt, and diminished self-worth) [15,16]. Women living with HIV frequently report experiencing stigma from family members, community figures, and health care providers, which discourages disclosure and contributes to social withdrawal and disengagement from care [17,18]. Internalized stigma is also associated with increased depressive symptoms and reduced treatment self-efficacy among people entering HIV care, factors that can undermine ART adherence during pregnancy and the early postpartum period [19-21].

Beyond stigma, household food insecurity is another critical, often underappreciated barrier to optimal ART adherence among pregnant and postpartum women living with HIV. Recent phone survey data show that 48% of Ugandan households experience moderate food insecurity and 11% experience severe food insecurity, with the highest burden in rural areas [22]. For women managing HIV during pregnancy, food insecurity magnifies the challenges of chronic illness by intensifying common ART side effects such as nausea, dizziness, and gastrointestinal discomfort when medications are taken without food [23,24]. These amplified side effects discourage adherence, increase depressive symptom severity, and compromise viral suppression [25,26]. Food insecurity also heightens psychological stress, limits care seeking, and undermines social participation and economic independence, thereby reinforcing HIV-related stigma [25,27]. Collectively, these pathways elevate the risk of suboptimal adherence during pregnancy and the early postpartum period.

Importantly, interpersonal social support can buffer the harmful effects of stigma and food insecurity on HIV-related outcomes. Longitudinal studies in rural Uganda and western Kenya demonstrate that emotional and instrumental support reduce the impact of food insecurity on depression and improve ART adherence [28,29]. Peer-based interventions, including peer mother counseling and online support groups, have also shown promise in improving adherence and disclosure among pregnant and postpartum women living with HIV in Uganda [30,31]. However, such support is not universally accessible. Qualitative research has documented that many women experience relationship conflict, lack of disclosure, and rejection following an HIV diagnosis, which limits their access to emotional or practical assistance [17,32].

Despite this evidence, while stigma, food insecurity, and social support have each been studied individually as predictors of ART adherence [33-35], their co-occurrence and interaction remain poorly understood among pregnant and postpartum women in SSA [36,37]. Traditional variable-centered approaches such as early regression models [38] often examine these predictors in isolation and may overlook how multiple adversities interact to shape adherence outcomes [39]. Syndemic frameworks emphasize the synergistic clustering of social and structural challenges, which may amplify risk beyond the sum of individual effects [39,40]. Person-centered methods such as latent profile analysis (LPA) offer a rigorous approach to identifying subgroups of individuals who share similar constellations of psychosocial characteristics [41].

Accordingly, LPA has been increasingly applied in HIV research to detect syndemic vulnerability profiles and adherence-related behavioral patterns among adults in diverse contexts, including SSA [42-44]. However, its use among pregnant and postpartum women remains limited, particularly in high-burden, resource-constrained settings. Applying LPA in this context may uncover subgroups of women disproportionately affected by overlapping stressors and inform more targeted, tailored interventions.

This study addresses this gap by applying LPA to identify vulnerability subgroups among pregnant women living with HIV in Uganda based on 4 theoretically and empirically grounded indicators: experienced HIV stigma, internalized HIV stigma, household food insecurity, and perceived interpersonal social support. We then examined whether profile membership was associated with ART adherence adjusting for sociodemographic covariates. We hypothesized that members of higher-risk profiles, characterized by greater stigma and food insecurity combined with lower levels of support, would report significantly lower ART adherence compared with members of lower-risk profiles. By identifying distinct constellations of co-occurring risk factors, this study aimed to inform more targeted and contextually appropriate interventions to improve maternal adherence and advance Uganda's progress toward achieving its 95-95-95 targets.

XSL•FO
RenderX

Methods

Data, Participants, and Procedures

Data were drawn from a larger longitudinal study examining factors that influence retention in HIV care among pregnant and postpartum women in Uganda. Details of the parent study design and data collection procedures have been reported elsewhere [45]. Participants were recruited from 6 participating health care facilities through a clinic-based recruitment process. Midwives and nurses involved in routine antenatal and HIV care served as initial points of contact and systematically identified potentially eligible women during routine clinic visits. Eligibility criteria included living with HIV, being in the third trimester of pregnancy, and enrollment in an ART or prevention of mother-to-child transmission program.

Women who met the eligibility criteria and expressed interest were referred to trained study staff, who provided a detailed explanation of the study, confirmed eligibility, obtained written informed consent, and administered baseline questionnaires. Baseline data collection took place between June and August 2020. Approximately 3 months later, participants were recontacted to complete follow-up assessments, including additional sociodemographic and ART adherence–related measures, between October and December 2020. Recruitment was intentionally anchored in late pregnancy, when women are approaching delivery and facing heightened clinical, psychosocial, and structural demands, to facilitate follow-up into the early postpartum period in the parent longitudinal study.

This analysis used baseline data only and was conducted as a cross-sectional study. Baseline assessment during the third trimester provided a uniform and clinically salient time point, capturing psychosocial and structural vulnerabilities during a period of heightened relevance for ART adherence and prevention of vertical HIV transmission. Restricting analyses to baseline data minimized heterogeneity related to postpartum transitions and allowed for the identification of contemporaneous psychosocial vulnerability profiles prior to delivery. Longitudinal findings from the parent study examining postpartum outcomes have been reported previously [45].

Ethical Considerations

This study was conducted in accordance with the principles of the Declaration of Helsinki and was reviewed and approved by the Institutional Review Board of the University of California, San Francisco (approval 20-30261), and the Mbarara University of Science and Technology Research Ethics Committee (approval MUREC 1/7). Written informed consent was obtained from all participants prior to enrollment. Participants did not receive compensation for their participation. Participant confidentiality was maintained throughout the study, and all data were deidentified prior to analysis and stored securely, with access restricted to authorized study personnel. This analysis used data from an approved parent study, and all ethics approvals and consent procedures were conducted as part of the original study protocol [45].

Measures

Overview

All study measures were administered by trained research staff with experience working with pregnant women living with HIV in Uganda. Instruments were selected based on prior use in HIV research and conceptual relevance to the study objectives. The measures were reviewed and piloted prior to data collection to ensure clarity and contextual appropriateness for the study population. Participants were provided with assistance as needed during questionnaire administration to support comprehension.

Internalized HIV Stigma

Internalized HIV stigma was assessed using the Internalized Stigma of HIV/AIDS Tool [46], a 10-item measure previously used in HIV research, including studies conducted in SSA, in which respondents rate each statement on a 5-point Likert scale ranging from “strongly disagree” to “strongly agree.” Items prompt participants to reflect on their self-perceptions and feelings since receiving an HIV diagnosis. Example items include “I feel blemished” and “I feel ashamed about having HIV/AIDS.” We calculated mean scores, with higher scores indicating greater internalized HIV stigma. In this sample, the scale demonstrated good internal consistency (Cronbach α=0.81).

Experienced HIV Stigma

Experienced HIV stigma was assessed using the Externalized Stigma of HIV/AIDS Tool (ESAT), a 10-item measure capturing participants’ experiences of stigma and discrimination from others in response to their HIV status. The ESAT was adapted from the Internalized Stigma of AIDS Tool to capture enacted or experienced HIV-related stigma, which is conceptually distinct from internalized stigma. Participants rate each statement on a 5-point Likert scale ranging from “strongly disagree” to “strongly agree.” Example items include “I have been excluded from social events” and “I have been verbally abused or ridiculed.” Average scores were calculated, with higher scores indicating greater experienced HIV stigma. In this sample, the scale demonstrated good internal consistency (Cronbach α=0.74). The full list of ESAT items is provided in Multimedia Appendix 1.

Household Food Insecurity

Household food insecurity was measured using items from the Household Food Insecurity Access Scale (HFIAS) [47], which assesses household food access over the previous 30 days. The HFIAS includes paired items that first assess the presence of specific food insecurity conditions (“yes” or “no”) and, if endorsed, follow with questions assessing frequency. In this analysis, only the dichotomous HFIAS items assessing the presence of food insecurity conditions were used. The corresponding frequency follow-up items were not included because participants who responded “no” to the initial presence questions did not provide frequency data, resulting in substantial missingness on the frequency items. Dichotomous items were coded as 0 (“no”) or 1 (“yes”) and summed to create a composite score reflecting the number of food insecurity conditions endorsed, with higher scores indicating greater household food

XSL•FO
RenderX

insecurity. Internal consistency for this composite was high (Cronbach α=0.96)

Interpersonal Social Support

Perceived interpersonal social support was assessed using the Interpersonal Support Evaluation List–12 [48], a widely used and psychometrically validated measure of perceived availability of social resources. The Interpersonal Support Evaluation List–12 consists of 12 items rated on a 4-point Likert scale (0="definitely false"; 3="definitely true"). Sample items include "If I were sick, I could easily find someone to help me with my daily chores" and "There is someone I can turn to for advice about handling problems with my family." Item scores were averaged to yield a composite score, with higher scores indicating greater perceived interpersonal social support. In this study, the Cronbach α was 0.71, suggesting acceptable reliability.

ART Adherence

ART adherence was assessed using 3 items from the Center for Adherence Support Evaluation Adherence Index [49], which capture common adherence challenges. Participants responded to the following three items: (1) "How often do you feel that you have difficulty taking your HIV medications on time?" (response options: "never," "rarely," "most of the time," and "all of the time"); (2) "On average, how many days per week would you say that you missed at least one dose of your HIV medications?" (response options: "every day," "4 - 6 days/week," "2-3 days/week," "once a week," "less than once a week," and "never"); and (3) "When was the last time you missed at least one dose of your HIV medications?" (response options: "within the past week," "1 - 2 weeks ago," "3 - 4 weeks ago," "between 1 and 3 months ago," "more than 3 months ago," and "never"). Responses across these 3 items were coded and summed to create an ART adherence composite score ranging from 3 to 16, with higher scores indicating more optimal adherence to ART. The Cronbach α was calculated as 0.76, suggesting good internal reliability.

Data Analysis

Descriptive statistics were calculated for all study variables. LPA was conducted to identify unobserved subgroups (latent classes) of participants based on 4 continuous indicators: experienced HIV stigma, internalized HIV stigma, food insecurity, and interpersonal support. Model selection (the number of profiles) was guided by multiple fit indexes: the Akaike information criterion (AIC), Bayesian information criterion (BIC), consistent AIC, sample size–adjusted BIC, and classification likelihood criterion. Higher model fit was indicated by lower values of these indexes. Model interpretability and entropy values (ranging from 0 to 1) were also used to assess classification accuracy, with higher values indicating more precise assignment of individuals to latent classes. In determining the optimal number of classes, parsimony and fit indexes were balanced with conceptual clarity and class separation. Parameter estimates (means, variances, and SEs) for each class were extracted and compared across latent profiles to describe group characteristics.

Next, we conducted a linear regression analysis to examine the association between the latent classes and optimal ART adherence adjusting for age, educational status of both women living with HIV and their partners, religion, place of residence, partners' HIV status, and perceived financial status. All independent variables, including covariates, were entered simultaneously into the model. All analyses were cross-sectional and performed using Jamovi (version 2.3) [50].

Because participants were recruited through clinic-based sampling rather than a probability-based survey design, no survey weights were applied in the analyses. The study sample was not intended to be nationally representative, and therefore, weighting procedures designed to correct for unequal selection probabilities or population-level inference were not appropriate. All analyses were conducted using unweighted data, consistent with prior clinic-based studies of pregnant women living with HIV in similar settings. The analytic focus was on identifying within-sample psychosocial vulnerability profiles rather than producing population-level prevalence estimates.

Results

Descriptive Statistics

A total of 167 pregnant women living with HIV were included in the analysis. As shown in Table 1, the mean age was 28.01 (SD 6.14; range 17-45) years. Most participants (124/167, 74.3%) reported being in a relationship and living with a partner. Regarding educational level, most (118/167, 70.7%) indicated primary school education as their highest level. When asked about their partners' educational level, 47.0% (77/164) reported primary school education, whereas 19.5% (32/164) indicated that they did not know. Most participants (96/166, 57.8%) identified as Protestant, and most (141/166, 84.9%) lived in rural areas. Slightly more than half (84/160, 52.5%) reported that their partner was HIV positive, whereas others reported that their partner was HIV negative (46/160, 28.7%) or that they did not know their partner's status (30/160, 18.8%). Finally, most participants (139/166, 83.7%) described their financial situation as poor or not enough. Mean scores were 1.84 (SD 0.47; range 1.00 - 4.20) for experienced stigma, 2.48 (SD 0.54; range 1.40 - 3.80) for internalized stigma, 3.88 (SD 3.65; range 0 - 9) for food insecurity, 2.11 (SD 0.38; range 0.75 - 3.00) for interpersonal social support, and 14.5 (SD 2.31; range 7 - 16) for ART adherence.

XSL•FO
RenderX

Table . Descriptive statistics.

Variable	Values
Current relationship status (n=167), n (%)	
Widowed	3 (1.8)
Divorced	8 (4.8)
In a relationship—living with partner	124 (74.3)
In a relationship—not living with partner	27 (16.2)
Single, never married, or no current partner	5 (3.0)
Educational level—participant (n=167), n (%)	
Never been to school	14 (8.4)
Primary school	118 (70.7)
Secondary school (O-Level)	28 (16.8)
Higher secondary school (A-Level)	4 (2.4)
University degree or higher	3 (1.8)
Educational level—partner (n=164), n (%)	
Never been to school	9 (5.5)
Primary school	77 (47.0)
Secondary school (O-Level)	31 (18.9)
Higher secondary school (A-Level)	4 (2.4)
Diploma	5 (3.0)
University degree or higher	6 (3.7)
Did not know	32 (19.5)
Religion (n=166), n (%)	
Protestant	96 (57.8)
Catholic	54 (32.5)
Born-again Christian	7 (4.2)
Muslim	9 (5.4)
Place of residence (n=166), n (%)	
Rural	141 (84.9)
Urban	25 (15.1)
Partner HIV status (n=160), n (%)	
Positive	84 (52.5)
Negative	46 (28.7)
Did not know	30 (18.8)
Financial situation (n=166), n (%)	
Poor or not enough	139 (83.7)
Enough	27 (16.3)
Experienced HIV stigma score, mean (SD; range: possible range)	1.84 (0.47; 1.00-4.20; 1-5)
Internalized HIV stigma score, mean (SD; range; possible range)	2.48 (0.54; 1.40-3.80; 1-5)
Food insecurity score, mean (SD; range; possible range)	3.88 (3.65; 0-9; 0-9)
Interpersonal social support score, mean (SD; range; possible range)	2.11 (0.38; 0.75-3.00; 0-3)
ART[a] adherence score mean (SD; range; possible range)	14.5 (2.31; 7-16; 3-16)

[a]ART: antiretroviral therapy.

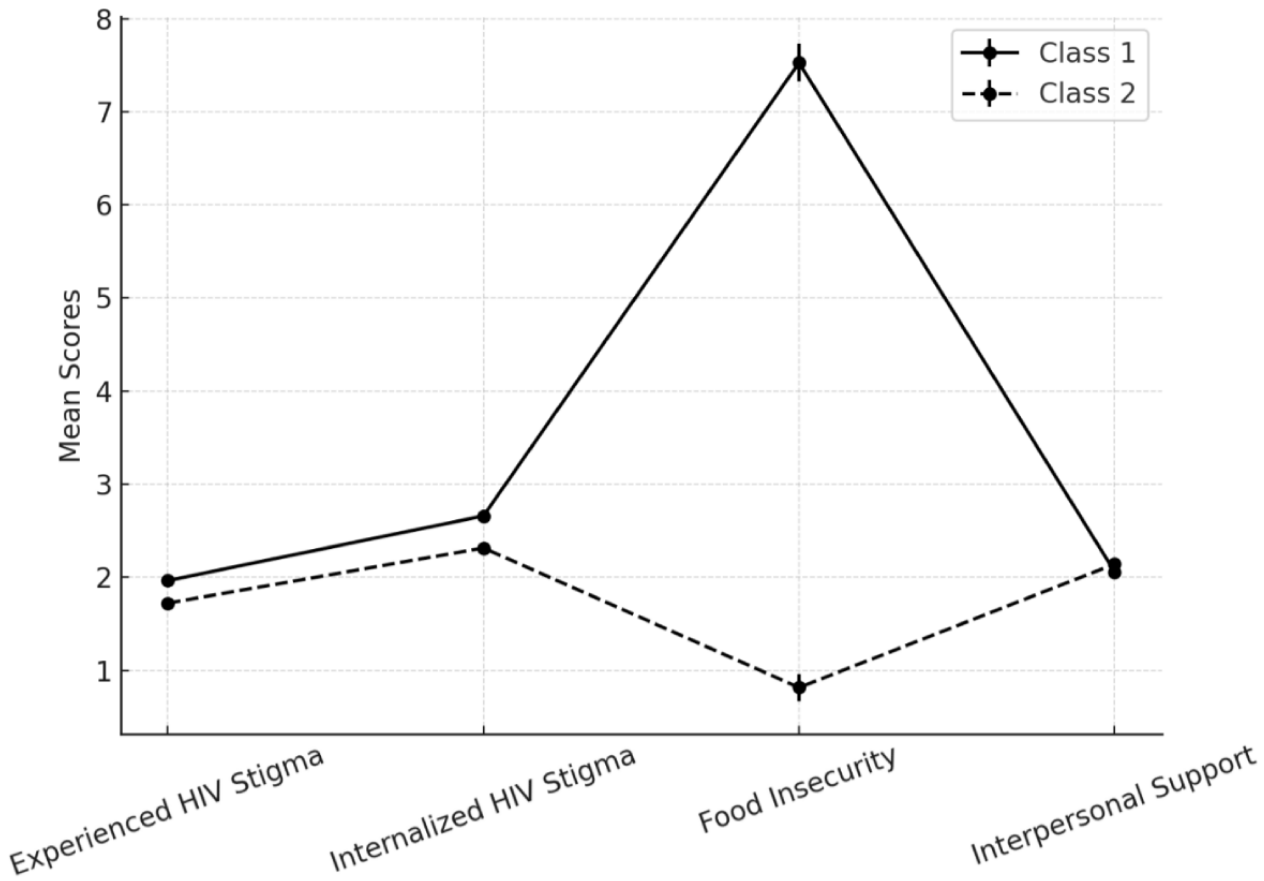
XSL•FO
RenderX

LPA Results

LPA was conducted to identify distinct subgroups of participants based on their scores on the selected indicators (ie, experienced HIV stigma, internalized HIV stigma, food insecurity, and interpersonal support). Competing models with 2 and 3 latent classes were compared. The results suggest that model fit indexes favored the 2-class solution over the 3-class solution. Specifically, the 2-class model had the lowest values for the AIC (1408), BIC (1449), consistent AIC (1462), and sample size–adjusted BIC (1407), all indicating better fit. The 2-class model also exhibited higher entropy (0.93) compared with the 3-class model (0.82), suggesting more accurate classification of individuals into latent classes. Although the 3-class model had slightly better classification likelihood (classification likelihood criterion=1411 vs 1384 for the 2-class model), this was not sufficient to outweigh the superior overall fit of the 2-class solution. Taken together, these criteria support the 2-class model as the most parsimonious and best-fitting solution.

As shown in Figure 1, there were 2 latent classes that emerged. Class 1 (higher-risk profile; 75/167, 45.1%) was characterized by greater experienced HIV stigma (mean 1.97; SE 0.05), higher internalized HIV stigma (mean 2.66; SE 0.07), high food insecurity (mean 7.53; SE 0.20), and lower interpersonal support (mean 2.06; SE 0.04). In contrast, class 2 (lower-risk profile; 92/167, 54.9%) showed slightly lower experienced HIV stigma (mean 1.72; SE 0.04), lower internalized HIV stigma (mean 2.32; SE 0.05), much lower food insecurity (mean 0.82; SE 0.15), and slightly higher interpersonal support (mean 2.14; SE 0.04). All differences between classes were statistically significant (P<.001). These findings indicate 2 interpretable and distinct profiles of pregnant women living with HIV in Uganda, differing substantially in structural and psychosocial vulnerability.

Figure 1. Profile plot of latent classes.



Linear Regression Analysis

A linear regression analysis was conducted to examine the association between latent profile membership (lower-risk profile vs higher-risk profile) and optimal ART adherence while adjusting for the covariates. Latent class membership emerged as a significant predictor of ART adherence. Specifically, compared with women living with HIV in the lower-risk profile (class 2), those in the higher-risk profile (class 1) reported significantly lower levels of optimal ART adherence (B=0.88; β=0.40, 95% CI 0.02-0.78; t_{133}=2.07; P=.04). These results suggest that latent profiles capturing structural and psychosocial vulnerabilities are meaningfully associated with ART adherence above and beyond key demographic and contextual factors.

Discussion

Principal Findings

This study aimed to identify psychosocial vulnerability profiles among pregnant women living with HIV in Uganda and examine their associations with ART adherence during pregnancy. Consistent with these objectives, we identified 2 empirically

XSL•FO
RenderX

distinct vulnerability profiles that differed significantly in levels of HIV stigma, food insecurity, and interpersonal social support and were significantly associated with ART adherence. Women living with HIV in the higher-risk profile, characterized by higher experienced and internalized HIV stigma, severe food insecurity, and lower interpersonal support, reported significantly lower ART adherence relative to their lower-risk counterparts. The converging of HIV stigma and food insecurity within the high-risk group suggests a pattern of intersecting stressors that may amplify psychological distress, impair medication adherence, and disrupt care engagement during the perinatal period.

These results align with syndemic theory, which posits that co-occurring adversities can interact synergistically to undermine health behaviors [39,40]. Unlike prior variable-centered studies, our LPA demonstrates how stigma, food insecurity, and social support co-occur to shape distinct vulnerability profiles. Our findings are consistent with those of prior Ugandan studies showing that stigma undermines ART adherence [51], food insecurity disrupts treatment consistency [52], and social support facilitates adherence [53]. Similar associations have been reported in neighboring countries, where stigma [54], food insecurity [55], and family support [56] have been identified as critical determinants of adherence. Our study extends this literature by highlighting how these factors co-occur within subgroups of pregnant women living with HIV, thereby highlighting heterogeneity in vulnerability that is obscured when determinants are examined individually. By situating our results within this evidence, we not only show alignment with prior work but also advance understanding by demonstrating the buffering role of interpersonal support, a dimension often overlooked in previous analyses. This person-centered perspective provides a more nuanced understanding of vulnerability among pregnant women living with HIV and complements prior studies that have examined these factors in isolation.

Our findings also underscore the need for interventions that address intersecting structural and psychosocial vulnerabilities identified in the higher-risk profile. Multicomponent approaches that respond to both material needs and social stressors are likely to be most effective in this context. In Uganda, the Friends for Life Circles peer group program combined peer mother mentoring, community ART distribution, and income-generating support, resulting in improved postpartum retention and engagement in care [57]. The program's design allowed peer mothers to provide psychosocial support and role modeling, whereas community-based ART delivery reduced barriers such as transport costs and clinic waiting times. The inclusion of income-generating activities further addressed underlying economic vulnerability, helping women sustain care engagement [57]. In Kenya, the Shamba Maisha [58,59] livelihood intervention provided small loans, farming tools, and agricultural and financial training to people living with HIV. Participants reported reduced internalized stigma and improved psychosocial well-being, attributed to increased economic productivity and a regained sense of social value [58,59]. By improving household food security and restoring participants' economic agency, the intervention not only strengthened material conditions but also enhanced social standing within the community [58,59].

Together, these examples highlight the role of economic empowerment in creating conditions that enable ART adherence and demonstrate the value of integrated, context-sensitive strategies that simultaneously address food insecurity, HIV stigma, and limited social support. Building on this evidence, future HIV interventions in Uganda should consider integrating community-based delivery models such as mentor mothers and community health workers with nutritional assistance. Potential strategies include food vouchers to meet immediate dietary needs and agricultural subsidies to support sustainable household food production, particularly in rural communities where subsistence farming is common. Conditional food support linked to engagement in antenatal HIV care may further help disrupt the cycle of poverty, stigma, and disengagement from care among pregnant women living with HIV.

Food insecurity may exacerbate stigma-related stress through multiple interacting pathways. Physiologically, taking ART without sufficient food intake has been linked to intensified side effects, which may undermine treatment adherence and persistence [60]. Psychologically, food insecurity contributes to depressive symptoms and reduced self-efficacy, particularly in rural Ugandan settings where poverty and HIV stigma are closely intertwined [25]. Interpersonally, social support serves a protective function by providing emotional reassurance and practical assistance, including help with medication adherence, clinic attendance, and disclosure [61]. Our findings align with these mechanisms: participants in the lower-risk latent profile reported higher levels of perceived support, suggesting that even modest increases in interpersonal support may buffer the adverse effects of structural and psychosocial stressors. This interpretation is further supported by meta-analytic evidence indicating that social support is a consistent facilitator of ART adherence among pregnant and breastfeeding women across SSA [62,63].

Although this study advances the HIV literature by applying a person-centered analytic approach during a critical perinatal period and integrating both structural and psychosocial determinants of ART adherence, several limitations should be noted. First, although the data were drawn from a longitudinal cohort, this analysis was cross-sectional. This limits causal inference, making it unclear whether stigma, food insecurity, and social support influence ART adherence or whether nonadherence may exacerbate these vulnerabilities. Future studies should use longitudinal or implementation science designs to clarify temporal and causal relationships. Second, ART adherence was assessed via self-report, which may be subject to recall and social desirability biases. Incorporating objective measures such as electronic dose monitoring or pharmacy refill records should strengthen adherence measurement in future research. Third, our analysis did not include other relevant determinants of adherence, such as depressive symptoms, intimate partner violence, or transportation barriers, which may further refine vulnerability profiles. Fourth, because the study sample was relatively small and recruited from rural health facilities, the findings may not be generalizable to urban populations or other regions.

XSL•FO
RenderX

Replication with larger and more diverse samples is warranted to strengthen external validity. Future research should also assess whether additional profiles emerge in larger and more diverse samples.

Although the data were collected in 2020, the structural and psychosocial challenges examined in this study, including HIV-related stigma, food insecurity, and limited social support, remain persistent barriers to ART adherence among pregnant women in Uganda and other high-burden settings. In addition, shifts in the global HIV funding landscape toward more constrained and efficiency-focused models heighten the importance of person-centered approaches that can identify and prioritize subgroups facing clustered vulnerabilities, thereby supporting more targeted and sustainable intervention strategies.

Conclusions

This study highlights the need to urgently address intersecting psychosocial and structural challenges that compromise ART adherence among pregnant women living with HIV. Identifying latent vulnerability profiles offers a targeted framework for directing limited resources to those at greatest risk of disengagement. Intervention strategies should be grounded in the lived experiences of pregnant women living with HIV and must address overlapping adversities rather than treating risks in isolation. Integrated approaches that combine HIV stigma reduction, nutritional support, and strengthened social networks are essential for efforts to promote maternal health and prevent vertical HIV transmission. By demonstrating how these determinants cluster within identifiable subgroups, our findings underscore the value of person-centered methods for guiding intervention design. Addressing these vulnerabilities is essential not only for advancing Uganda's progress toward the 95-95-95 targets but also for informing global strategies to improve maternal and child health outcomes in high-burden settings.

Acknowledgments

The authors are grateful to all the women who generously participated in this study.

Funding

This study was funded by the University of California, San Francisco (UCSF) School of Nursing Gaine Research Fund under grant number GRF-2020-03 and UCSF Population Health and Health Equity fellowship program under grant number 7504575.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

SKS conceptualized the analysis in collaboration with IY, drafted the manuscript, and supervised its development. IY conducted the data analysis and contributed to conceptualization and manuscript drafting. RGAT and BL contributed to manuscript drafting and editing. GA participated in data collection, drafting, and the review of the manuscript. JJO conceptualized the idea for the original study, secured funding, managed project administration, contributed to drafting, and reviewed the manuscript. All authors approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questionnaires used in the study.
[DOCX File, 63 KB - publichealth_v12i1e84914_app1.docx]

References

1. Uganda. UNAIDS. URL: https://www.unaids.org/en/regionscountries/countries/uganda [accessed 2026-03-19]
2. Global HIV & AIDS statistics — fact sheet. UNAIDS. URL: https://www.unaids.org/en/resources/fact-sheet [accessed 2026-03-19]
3. Townsend CL, Byrne L, Cortina-Borja M, et al. Earlier initiation of ART and further decline in mother-to-child HIV transmission rates, 2000-2011. AIDS 2014 Apr 24;28(7):1049-1057. [doi: 10.1097/QAD.0000000000000212] [Medline: 24566097]
4. Kiragga AN, Twinomuhwezi E, Banturaki G, et al. Outcomes of retained and disengaged pregnant women living with HIV in Uganda. PLoS One 2021 May 21;16(5):e0251413. [doi: 10.1371/journal.pone.0251413] [Medline: 34019568]
5. Drake AL, Wagner A, Richardson B, John-Stewart G. Incident HIV during pregnancy and postpartum and risk of mother-to-child HIV transmission: a systematic review and meta-analysis. PLoS Med 2014 Feb 25;11(2):e1001608. [doi: 10.1371/journal.pmed.1001608] [Medline: 24586123]

XSL•FO
RenderX

6. Consolidated guidelines for the prevention and treatment of HIV and AIDS in Uganda. : The Republic of Uganda Ministry of Health; 2022 URL: https://dsduganda.com/wp-content/uploads/2023/05/Consolidated-HIV-and-AIDS-Guidelines-20230516.pdf [accessed 2026-03-19]
7. Consolidated Guidelines on the Use of Antiretroviral Drugs for Treating and Preventing HIV Infection: Recommendations for a Public Health Approach: World Health Organization; 2016. URL: https://www.who.int/publications/i/item/9789241549684 [accessed 2026-03-31]
8. Kerschberger B, Boulle A, Kuwengwa R, Ciglenecki I, Schomaker M. The impact of same-day antiretroviral therapy initiation under the World Health Organization treat-all policy. Am J Epidemiol 2021 Aug 1;190(8):1519-1532. [doi: 10.1093/aje/kwab032] [Medline: 33576383]
9. Maximizing the Treatment and Prevention Potential of Antiretroviral Drugs: Early Country Experiences Towards Implementing a Treat-All Policy. : World Health Organization; 2015 URL: https://www.aidsdatahub.org/sites/default/files/resource/who-programmatic-update-maximizing-treatment-prevention-potential-antiretroviral-drugs.pdf [accessed 2026-03-19]
10. Mukose AD, Bastiaens H, Makumbi F, et al. What influences uptake and early adherence to Option B+ (lifelong antiretroviral therapy among HIV positive pregnant and breastfeeding women) in central Uganda? A mixed methods study. PLoS One 2021;16(5):e0251181. [doi: 10.1371/journal.pone.0251181] [Medline: 33951109]
11. Matthews LT, Orrell C, Bwana MB, et al. Adherence to HIV antiretroviral therapy among pregnant and postpartum women during the Option B+ era: 12-month cohort study in urban South Africa and rural Uganda. J Int AIDS Soc 2020 Aug;23(8):e25586. [doi: 10.1002/jia2.25586] [Medline: 32820622]
12. Haas AD, Msukwa MT, Egger M, et al. Adherence to antiretroviral therapy during and after pregnancy: cohort study on women receiving care in Malawi's Option B+ program. Clin Infect Dis 2016 Nov 1;63(9):1227-1235. [doi: 10.1093/cid/ciw500] [Medline: 27461920]
13. Turan JM, Bukusi EA, Onono M, Holzemer WL, Miller S, Cohen CR. HIV/AIDS stigma and refusal of HIV testing among pregnant women in rural Kenya: results from the MAMAS Study. AIDS Behav 2011 Aug;15(6):1111-1120. [doi: 10.1007/s10461-010-9798-5] [Medline: 20827573]
14. Nyblade L, Stockton MA, Giger K, et al. Stigma in health facilities: why it matters and how we can change it. BMC Med 2019 Feb 15;17(1):25. [doi: 10.1186/s12916-019-1256-2] [Medline: 30764806]
15. Turan B, Hatcher AM, Weiser SD, Johnson MO, Rice WS, Turan JM. Framing mechanisms linking HIV-related stigma, adherence to treatment, and health outcomes. Am J Public Health 2017 Jun;107(6):863-869. [doi: 10.2105/AJPH.2017.303744] [Medline: 28426316]
16. Earnshaw VA, Chaudoir SR. From conceptualizing to measuring HIV stigma: a review of HIV stigma mechanism measures. AIDS Behav 2009 Dec;13(6):1160-1177. [doi: 10.1007/s10461-009-9593-3] [Medline: 19636699]
17. Jolle J, Kabunga A, Okello TO, et al. HIV-related stigma experiences and coping strategies among pregnant women in rural Uganda: a qualitative descriptive study. PLoS One 2022 Oct 7;17(10):e0272931. [doi: 10.1371/journal.pone.0272931] [Medline: 36206276]
18. Wanyenze RK, Musinguzi G, Kiguli J, et al. "When they know that you are a sex worker, you will be the last person to be treated": perceptions and experiences of female sex workers in accessing HIV services in Uganda. BMC Int Health Hum Rights 2017 May 5;17(1):11. [doi: 10.1186/s12914-017-0119-1] [Medline: 28476153]
19. Yigit I, Bayramoglu Y, Weiser SD, et al. Changes in internalized stigma and HIV health outcomes in individuals new to HIV care: the mediating roles of depression and treatment self-efficacy. AIDS Patient Care STDS 2020 Nov;34(11):491-497. [doi: 10.1089/apc.2020.0114] [Medline: 33147084]
20. Brittain K, Mellins CA, Phillips T, et al. Social support, stigma and antenatal depression among HIV-infected pregnant women in South Africa. AIDS Behav 2017 Jan;21(1):274-282. [doi: 10.1007/s10461-016-1389-7] [Medline: 27052843]
21. Turan B, Smith W, Cohen MH, et al. Mechanisms for the negative effects of internalized HIV-related stigma on antiretroviral therapy adherence in women: the mediating roles of social isolation and depression. J Acquir Immune Defic Syndr 2016 Jun 1;72(2):198-205. [doi: 10.1097/QAI.0000000000000948] [Medline: 26885803]
22. Uganda - High-Frequency Phone Survey 2020-2024. : World Bank; 2024 URL: https://microdata.worldbank.org/index.php/catalog/3765/pdf-documentation [accessed 2026-03-19]
23. Omonaiye O, Kusljic S, Nicholson P, Manias E. Medication adherence in pregnant women with human immunodeficiency virus receiving antiretroviral therapy in sub-Saharan Africa: a systematic review. BMC Public Health 2018 Jun 27;18(1):805. [doi: 10.1186/s12889-018-5651-y] [Medline: 29945601]
24. Wiesner A, Skrońska M, Gawlik G, Marcinkowska M, Zagrodzki P, Paśko P. Interactions of antiretroviral drugs with food, beverages, dietary supplements, and alcohol: a systematic review and meta-analyses. AIDS Behav 2023 May;27(5):1441-1468. [doi: 10.1007/s10461-022-03880-6] [Medline: 36318429]
25. Tsai AC, Bangsberg DR, Emenyonu N, Senkungu JK, Martin JN, Weiser SD. The social context of food insecurity among persons living with HIV/AIDS in rural Uganda. Soc Sci Med 2011 Dec;73(12):1717-1724. [doi: 10.1016/j.socscimed.2011.09.026] [Medline: 22019367]
26. Chohan BH, Ronen K, Khasimwa B, et al. Food insecurity, drug resistance and non-disclosure are associated with virologic non-suppression among HIV pregnant women on antiretroviral treatment. PLoS One 2021 Aug 18;16(8):e0256249. [doi: 10.1371/journal.pone.0256249] [Medline: 34407133]

XSL•FO
RenderX

27. Aibibula W, Cox J, Hamelin AM, McLinden T, Klein MB, Brassard P. Association between food insecurity and HIV viral suppression: a systematic review and meta-analysis. AIDS Behav 2017 Mar;21(3):754-765. [doi: 10.1007/s10461-016-1605-5] [Medline: 27837425]
28. Tsai AC, Bangsberg DR, Frongillo EA, et al. Food insecurity, depression and the modifying role of social support among people living with HIV/AIDS in rural Uganda. Soc Sci Med 2012 Jun;74(12):2012-2019. [doi: 10.1016/j.socscimed.2012.02.033] [Medline: 22513248]
29. Wang M, Miller JD, Collins SM, et al. Social support mitigates negative impact of food insecurity on antiretroviral adherence among postpartum women in western Kenya. AIDS Behav 2020 Oct;24(10):2885-2894. [doi: 10.1007/s10461-020-02839-9] [Medline: 32212069]
30. Kabami J, Balzer LB, Kabageni S, et al. Peer-mother counseling improves HIV treatment adherence and status disclosure over time among pregnant and postpartum women in rural Uganda. AJE Adv 2025 May 15;1(1):uuaf002. [doi: 10.1093/ajeadv/uuaf002]
31. Kiirya Y, Kitaka S, Kalyango J, et al. Acceptability of an online peer support group as a strategy to improve antiretroviral therapy adherence among young people in Kampala district, Uganda: qualitative findings. BMC Infect Dis 2025 Apr 3;25(1):461. [doi: 10.1186/s12879-025-10831-8] [Medline: 40181265]
32. Ashaba S, Kaida A, Coleman JN, et al. Psychosocial challenges facing women living with HIV during the perinatal period in rural Uganda. PLoS One 2017 May 1;12(5):e0176256. [doi: 10.1371/journal.pone.0176256] [Medline: 28459866]
33. Weidle PJ, Wamai N, Solberg P, et al. Adherence to antiretroviral therapy in a home-based AIDS care programme in rural Uganda. Lancet 2006 Nov 4;368(9547):1587-1594. [doi: 10.1016/S0140-6736(06)69118-6] [Medline: 17084759]
34. Hardon AP, Akurut D, Comoro C, et al. Hunger, waiting time and transport costs: time to confront challenges to ART adherence in Africa. AIDS Care 2007 May;19(5):658-665. [doi: 10.1080/09540120701244943] [Medline: 17505927]
35. Weiser S, Wolfe W, Bangsberg D, et al. Barriers to antiretroviral adherence for patients living with HIV infection and AIDS in Botswana. J Acquir Immune Defic Syndr 2003 Nov 1;34(3):281-288. [doi: 10.1097/00126334-200311010-00004] [Medline: 14600572]
36. Sabin LL, Simmons E, Halim N, et al. Real-time feedback to improve HIV treatment adherence in pregnant and postpartum women in Uganda: a randomized controlled trial. AIDS Behav 2022 Dec;26(12):3834-3847. [doi: 10.1007/s10461-022-03712-7] [Medline: 35704124]
37. Tsai AC, Mendenhall E, Trostle JA, Kawachi I. Co-occurring epidemics, syndemics, and population health. Lancet 2017 Mar 4;389(10072):978-982. [doi: 10.1016/S0140-6736(17)30403-8] [Medline: 28271848]
38. Carlucci JG, Kamanga A, Sheneberger R, et al. Predictors of adherence to antiretroviral therapy in rural Zambia. J Acquir Immune Defic Syndr 2008 Apr 15;47(5):615-622. [doi: 10.1097/QAI.0b013e318165dc25] [Medline: 18209678]
39. Tsai AC. Syndemics: a theory in search of data or data in search of a theory? Soc Sci Med 2018 Jun;206:117-122. [doi: 10.1016/j.socscimed.2018.03.040] [Medline: 29628175]
40. Harkness A, Bainter SA, O'Cleirigh C, Mendez NA, Mayer KH, Safren SA. Longitudinal effects of syndemics on ART non-adherence among sexual minority men. AIDS Behav 2018 Aug;22(8):2564-2574. [doi: 10.1007/s10461-018-2180-8] [Medline: 29860556]
41. Collins LM, Lanza ST. Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences: John Wiley & Sons; 2013.
42. Badejo O, Wouters E, Van Belle S, et al. Latent class analysis of barriers to HIV testing services and associations with sexual behaviour and HIV status among adolescents and young adults in Nigeria. PLoS One 2024;19(4):e0300220. [doi: 10.1371/journal.pone.0300220] [Medline: 38635546]
43. Pantelic M, Boyes M, Cluver L, Meinck F. HIV, violence, blame and shame: pathways of risk to internalized HIV stigma among South African adolescents living with HIV. J Int AIDS Soc 2017 Aug 21;20(1):21771. [doi: 10.7448/IAS.20.1.21771] [Medline: 28853517]
44. McBain RK, Schuler MS, Rukundo T, Wanyenze RK, Wagner GJ. Trajectories of perinatal depression among women living with HIV in Uganda. J Glob Health 2024 Sep 20;14:04147. [doi: 10.7189/jogh.14.04147] [Medline: 39301593]
45. Nutor JJ, Agbadi P, Hoffmann TJ, Anguyo G, Camlin CS. Examining the relationship between interpersonal support and retention in HIV care among HIV+ nursing mothers in Uganda. BMC Res Notes 2021 Jun 3;14(1):224. [doi: 10.1186/s13104-021-05639-z] [Medline: 34082834]
46. Phillips KD, Moneyham L, Tavakoli A. Development of an instrument to measure internalized stigma in those with HIV/AIDS. Issues Ment Health Nurs 2011;32(6):359-366. [doi: 10.3109/01612840.2011.575533] [Medline: 21692574]
47. Coates J, Swindale A, Bilinsky P. Household Food Insecurity Access Scale (HFIAS) for measurement of food access: indicator guide (version 3). : Food and Nutrition Technical Assistance Project, Academy for Educational Development; 2007 URL: https://www.fao.org/fileadmin/user_upload/eufao-fsi4dm/doc-training/hfias.pdf [accessed 2026-03-19]
48. Merz EL, Roesch SC, Malcarne VL, et al. Validation of Interpersonal Support Evaluation List-12 (ISEL-12) scores among English- and Spanish-speaking Hispanics/Latinos from the HCHS/SOL Sociocultural Ancillary Study. Psychol Assess 2014 Jun;26(2):384-394. [doi: 10.1037/a0035248] [Medline: 24320763]
49. Mannheimer SB, Mukherjee R, Hirschhorn LR, et al. The CASE adherence index: a novel method for measuring adherence to antiretroviral therapy. AIDS Care 2006 Oct;18(7):853-861. [doi: 10.1080/09540120500465160] [Medline: 16971298]

XSL•FO
RenderX

50. Jamovi (version 23). URL: https://www.jamovi.org [accessed 2002-03-31]
51. Tsai AC, Bangsberg DR, Weiser SD. Harnessing poverty alleviation to reduce the stigma of HIV in sub-Saharan Africa. PLoS Med 2013 Nov;10(11):e1001557. [doi: 10.1371/journal.pmed.1001557] [Medline: 24319400]
52. Weiser SD, Tsai AC, Gupta R, et al. Food insecurity is associated with morbidity and patterns of healthcare utilization among HIV-infected individuals in a resource-poor setting. AIDS 2012 Jan 2;26(1):67-75. [doi: 10.1097/QAD.0b013e32834cad37] [Medline: 21904186]
53. Atukunda EC, Musiimenta A, Musinguzi N, et al. Understanding patterns of social support and their relationship to an ART adherence intervention among adults in rural southwestern Uganda. AIDS Behav 2017 Feb;21(2):428-440. [doi: 10.1007/s10461-016-1559-7] [Medline: 27671479]
54. Wanjala SW, Nyongesa MK, Mapenzi R, Luchters S, Abubakar A. A qualitative inquiry of experiences of HIV-related stigma and its effects among people living with HIV on treatment in rural Kilifi, Kenya. Front Public Health 2023 Jun 21;11:1188446. [doi: 10.3389/fpubh.2023.1188446] [Medline: 37427260]
55. Exavery A, Charles J, Kuhlik E, et al. Hunger and adherence to antiretroviral therapy: learning from HIV positive caregivers of orphans and vulnerable children in Tanzania. Front Public Health 2022 Feb 21;9:719485. [doi: 10.3389/fpubh.2021.719485] [Medline: 35265581]
56. Mshangila JD, Mwanga HH, Shao ME, Msesa DJ, Mankaba CN. Adherence to antiretroviral therapy among people living with HIV attending medication-assisted treatment clinics in Dar es Salaam, Tanzania: a cross-sectional study. Pan Afr Med J 2024 Apr 16;47:186. [doi: 10.11604/pamj.2024.47.186.40498] [Medline: 39092019]
57. Amone A, Gabagaya G, Wavamunno P, et al. Enhanced peer-group strategies to support the prevention of mother-to-child HIV transmission leads to increased retention in care in Uganda: a randomized controlled trial. PLoS One 2024 Apr 19;19(4):e0297652. [doi: 10.1371/journal.pone.0297652] [Medline: 38640123]
58. Tsai AC, Hatcher AM, Bukusi EA, et al. A livelihood intervention to reduce the stigma of HIV in rural Kenya: longitudinal qualitative study. AIDS Behav 2017 Jan;21(1):248-260. [doi: 10.1007/s10461-015-1285-6] [Medline: 26767535]
59. Weiser SD, Sheira LA, Weke E, et al. An agricultural livelihood intervention is associated with reduced HIV stigma among people with HIV. AIDS 2024 Aug 1;38(10):1553-1559. [doi: 10.1097/QAD.0000000000003928] [Medline: 38752558]
60. Anema A, Vogenthaler N, Frongillo EA, Kadiyala S, Weiser SD. Food insecurity and HIV/AIDS: current knowledge, gaps, and research priorities. Curr HIV/AIDS Rep 2009 Nov;6(4):224-231. [doi: 10.1007/s11904-009-0030-z] [Medline: 19849966]
61. Ashaba S, Kaida A, Burns BF, et al. Understanding coping strategies during pregnancy and the postpartum period: a qualitative study of women living with HIV in rural Uganda. BMC Pregnancy Childbirth 2017 May 8;17(1):138. [doi: 10.1186/s12884-017-1321-9] [Medline: 28482821]
62. Fassinou LC, Songwa Nkeunang D, Delvaux T, Nagot N, Kirakoya-Samadoulougou F. Adherence to option B + antiretroviral therapy and associated factors in pregnant and breastfeeding women in sub-Saharan Africa: a systematic review and meta-analysis. BMC Public Health 2024 Jan 5;24(1):94. [doi: 10.1186/s12889-023-17004-9] [Medline: 38183014]
63. Nachega JB, Uthman OA, Anderson J, et al. Adherence to antiretroviral therapy during and after pregnancy in low-income, middle-income, and high-income countries: a systematic review and meta-analysis. AIDS 2012 Oct 23;26(16):2039-2052. [doi: 10.1097/QAD.0b013e328359590f] [Medline: 22951634]

Abbreviations

AIC: Akaike information criterion
ART: antiretroviral therapy
BIC: Bayesian information criterion
ESAT: Externalized Stigma of HIV/AIDS Tool
HFIAS: Household Food Insecurity Access Scale
LPA: latent profile analysis
SSA: sub-Saharan Africa

Edited by A Mavragani, T Sanchez; submitted 26.Sep.2025; peer-reviewed by D Murphy, J Jesson; revised version received 03.Feb.2026; accepted 25.Feb.2026; published 27.Apr.2026.

Please cite as:
Sorkpor SK, Yigit I, Thompson RGA, Lane BL, Anguyo G, Ouner JJ
Stigma, Food Insecurity, and Limited Social Support as Psychosocial Correlates of Antiretroviral Therapy Adherence Among Pregnant Women Living With HIV: Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e84914
URL: https://publichealth.jmir.org/2026/1/e84914
doi:10.2196/84914

XSL•FO
RenderX

© Setor K Sorkpor, Ibrahim Yigit, Rachel G A Thompson, Brittany L Lane, Geoffrey Anguyo, Jerry John Ouner. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 27.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Preferences for mHealth Features to Support Engagement in the HIV Preexposure Prophylaxis Cascade Among Men Who Have Sex With Men in Peru: Cross-Sectional Online Survey

Jorge A Gallardo-Cartagena[1], MD, MPH; Dora Leidy Geraldine German-Quiñones[1], BS; Kelika Anne Konda[2], PhD, MHS; Susan Buchbinder[3], MD; Jorge Sanchez[1], MD, MPH; Frederick L Altice[4], MD

[1]Centro de Investigaciones Tecnológicas, Biomédicas y Medioambientales, Universidad Nacional Mayor de San Marcos, José Santos Chocano 199, Bellavista, Callao, Peru

[2]Keck School of Medicine, University of Southern California, Los Angeles, CA, United States

[3]San Francisco Department of Public Health, San Francisco, CA, United States

[4]Center for Interdisciplinary Research on AIDS, Yale School of Medicine, Yale University, New Haven, CT, United States

Corresponding Author:
Jorge A Gallardo-Cartagena, MD, MPH
Centro de Investigaciones Tecnológicas, Biomédicas y Medioambientales, Universidad Nacional Mayor de San Marcos, José Santos Chocano 199, Bellavista, Callao, Peru

Abstract

Background: Despite policy-level progress, implementation of oral HIV preexposure prophylaxis (PrEP) remains limited in Latin America. In Peru, men who have sex with men (MSM) account for most new HIV diagnoses, yet uptake remains low. Widespread smartphone ownership and the use of digital platforms present an opportunity to expand access through mobile health (mHealth) interventions. However, limited data exist on user preferences to guide the design of mHealth tools in Spanish-speaking Latin American settings.

Objective: This study aimed to assess preferences for mHealth features to support PrEP engagement among Peruvian MSM and their association with PrEP cascade stages.

Methods: We conducted a cross-sectional online survey (June-August 2023) among 600 HIV-negative MSM residing in Peru (median age 29, IQR 24-35 years), recruited via Facebook, Instagram, WhatsApp (Meta Platforms, Inc), and Grindr (Grindr LLC). The survey assessed communication platform use, interest in mHealth features measured on a 4-point Likert scale, and PrEP cascade stages. Exploratory factor analysis (principal axis factoring with Promax rotation) identified domains of mHealth preferences, from which median domain scores were calculated. Bivariate analyses used chi-square tests and Wilcoxon rank sum tests. Multivariable logistic regression models (α=.05), with covariates selected using stepwise procedures from candidate sociodemographic and behavioral variables, estimated associations between each domain score and PrEP cascade stages, each modeled as a separate binary outcome.

Results: Nearly all participants (589/600, 98.2%) reported owning a smartphone. WhatsApp was the most frequently used and preferred platform for PrEP support, with 547 (91.2%) reporting frequent use and 302 (50.3%) ranking it first. Exploratory factor analysis identified three mHealth preference domains: informational support (Cronbach α=0.94), self-management tools (Cronbach α=0.94), and interactive communication (Cronbach α=0.91). Among participants, 483 (80.5%) had decided to use PrEP, 190 (31.7%) had sought PrEP, and 109 (18.2%) had initiated PrEP. Higher informational support was associated with the decision to use PrEP (adjusted odds ratio [aOR] 4.54, 95% CI 3.36 - 6.28; P<.001), seeking PrEP (aOR 1.43, 95% CI 1.10 - 1.89; P=.001), and PrEP initiation (aOR 1.64, 95% CI 1.16 - 2.44; P=.009). Self-management tools showed similar associations with the decision to use PrEP (aOR 3.23, 95% CI 2.51 - 4.22; P<.001), seeking PrEP (aOR 1.34, 95% CI 1.06 - 1.70; P=.02), and PrEP initiation (aOR 1.49, 95% CI 1.11 - 2.05; P=.01). Interactive communication was associated with the decision to use PrEP (aOR 2.74, 95% CI 2.15 - 3.53; P<.001) but not with initiation.

Conclusions: Preferences for mHealth features were associated with engagement at multiple stages of the PrEP cascade among MSM in Peru. Informational support features demonstrated the most consistent associations with cascade engagement. These findings provide empirical evidence on user-prioritized digital functions that could support early engagement in HIV prevention services in a Latin American implementation context. Integrating culturally tailored mHealth tools within widely used platforms such as WhatsApp may strengthen early PrEP cascade engagement and support scalable digital strategies for HIV prevention in Peru and similar settings.

(JMIR Public Health Surveill 2026;12:e84982) doi:10.2196/84982

XSL•FO
RenderX

KEYWORDS

HIV prevention; preexposure prophylaxis; mHealth; digital health; smartphones; human-centered design; implementation science; men who have sex with men; Peru; Latin America; mobile health

Introduction

Despite important gains in HIV prevention, HIV remains a pressing public health concern across Latin America, with a disproportionately high incidence among gay, bisexual, and other men who have sex with men (MSM), as well as transgender women [1]. In Peru, these populations account for the vast majority of new HIV diagnoses, reflecting persistent gaps in prevention services and structural inequities [2].

Oral preexposure prophylaxis (PrEP) is an evidence-based intervention to prevent HIV acquisition [3]. Although many Latin American countries have begun integrating PrEP into their public health policies, uptake remains limited and uneven [4]. In Peru, fewer than 8000 individuals were using PrEP among an estimated 140,000 who could benefit by 2025 [5]. Gaps along the PrEP cascade, including seeking, initiation, and adherence, underscore the need for scalable, user-centered approaches to support delivery.

Digital communication technologies are widely used among MSM in Latin America [6]. Social media, mobile apps, and geosocial networking platforms are common sources of both social connection and health information [7-9]. This widespread engagement presents opportunities to integrate digital solutions, such as mobile health (mHealth) apps, messaging platforms, social media–based tools, and AI-enabled chatbots, into HIV prevention efforts.

Globally, mHealth interventions have shown promise in promoting HIV testing, linkage to care, and engagement with HIV prevention services, including PrEP uptake and adherence [10-12]. Randomized trials of mHealth apps in the United States have demonstrated measurable effects on HIV testing and progression along the PrEP cascade. For example, HealthMindr-PrEP (Emory University) [13] increased PrEP uptake overall, particularly among individuals with baseline PrEP indications. In a national randomized trial among young sexual minority men, MyChoices (Keymind) [14] increased HIV testing compared with standard prevention services, whereas LYNX (APT Mobility) [14] showed a nonsignificant increase in testing; neither intervention significantly increased PrEP initiation. Other trials have focused on adherence and retention among individuals using PrEP. SMS text messaging–based interventions such as PrEPmate (WelTel Inc) [15] and iTAB (University of California San Diego) [16] improved visit retention and some adherence outcomes, while app-based tools such as the Amsterdam PrEP app (Amsterdam PrEP Project/H-TEAM) [17], PrEP-iT! (San Diego State University) [18], P3 (Florida State University) [19], and DOT Diary (AiCure LLC) [20] demonstrated signals of improved adherence. Messaging platforms have also been leveraged for adherence support, including reminder systems delivered through WeChat (Tencent Mobile International Limited) [21] and weekly text messaging interventions implemented in Brazil [22].

Additional approaches target service navigation and demand generation. Clinic-integrated platforms such as JomPrEP (JomPrEP study team) [23,24] enable HIV self-testing, teleconsultation, and rapid PrEP initiation through mobile interfaces, while digital peer-navigation tools such as PrEPresent (Florida State University) [25] support engagement with PrEP services. Decision-support tools such as MyPrEP (University of Washington) [26] have also been developed to help individuals assess HIV risk and evaluate whether PrEP is appropriate before clinical consultation, and have been associated with improved early PrEP persistence in a randomized trial. AI-enabled chatbot–based conversational agents are also emerging as digital prevention tools. For example, Amanda Selfie (PrEP1519 study team) [27], MYHIV365 (Yale University) [28,29], and CHIA (Brown University) [30] deliver automated HIV prevention information and counseling through interactive chat interfaces and have demonstrated usability and acceptability in early testing phases. Formative work further highlights the importance of culturally tailored design; SaludFindr (Emory University) [31] was developed to address barriers to PrEP uptake among Latino sexual minority men in the United States, and DOT Diary and PrEPmate have been adapted to support adherence among linguistically diverse populations [32]. Ongoing trials such as Combine [33] are also evaluating mobile interventions designed to increase HIV testing and PrEP initiation among underserved populations, including rural MSM in the southern United States.

While this body of work reflects substantial innovation, most evaluated interventions have been implemented in high-income settings, and evidence specific to Latin American health care systems remains limited, particularly for Spanish-speaking populations. Although Spanish is widely spoken in Latin America, communication norms, health system navigation pathways, and colloquial vocabulary vary across countries, potentially limiting the direct transferability of digital tools developed elsewhere [34]. Because these contextual factors may influence how digital interventions are understood and used, assessing user preferences is an important first step before developing or adapting existing mHealth tools for Peru and similar Latin American settings.

To address this gap, we conducted a survey among MSM in Peru to characterize communication technology use and assess preferences for mHealth features to support PrEP engagement. We examined whether interest in distinct domains of mHealth support was associated with engagement in three PrEP cascade stages: (1) decision to start PrEP; (2) actively seeking PrEP services; and (3) PrEP initiation. We hypothesized that greater interest in specific mHealth domains would be associated with a higher likelihood of having reached these stages. Because these stages represent distinct behavioral transitions, digital supports may operate differently at each stage. While grounded in the Peruvian context, these findings may inform the design of culturally responsive, user-centered interventions to improve

XSL•FO
RenderX

PrEP engagement across Latin America, with appropriate adaptations for local settings.

Methods

Study Design

We conducted a cross-sectional online survey between June 16 and August 31, 2023, to explore the use of communication technologies and preferences for mHealth features to support HIV prevention among Peruvian MSM. The study followed the Journal Article Reporting Standards for quantitative research [35].

Participant Selection

Respondents were recruited using convenience, nonprobability sampling through paid advertisements on Instagram, Facebook, Grindr, and via unpaid standard posts ("organic content") shared on social media and WhatsApp groups managed by participating HIV and AIDS service organizations. Recruitment materials targeted MSM aged 18 years and older residing in Peru.

Participants accessed the survey through a secure link and completed it using their own internet-connected electronic devices; no data subsidies were provided. The survey platform required an active internet connection; offline completion was not supported. Electronic informed consent was required before initiating the survey. The survey instrument was administered in Spanish using standard Peruvian terminology reviewed for clarity and cultural appropriateness. Responses were anonymous, and participants were allowed to skip any question they did not wish to answer.

To ensure data integrity, responses were restricted to one submission per IP address, and IP addresses, timestamps, and completion times were reviewed to identify duplicate, automated, or implausibly rapid entries; none were identified.

Measures and Covariates

Survey Design

The 34-item survey included multiple-choice and Likert-scale questions covering sociodemographic and behavioral characteristics, achieved stages in the HIV prevention cascade (refer to Variable Definitions for operational definitions) [36,37], smartphone and internet access, frequency of use of communication channels (eg, phone calls, SMS text messaging, WhatsApp, social media, and internet browsing), and preferences for features of mHealth tools designed to support PrEP use.

mHealth Feature Preferences

To assess preferences for mHealth tools to support PrEP use, participants were presented with a series of 11 potential features and asked to rate their level of interest in each. The stem question read, "How interested would you be in receiving the following types of PrEP-related support through your mobile phone?" Responses were collected on a 4-point Likert scale, ranging from "Not interested" to "Very interested." The list of features was developed based on existing literature and used human-centered design by refining items through input from community partners to ensure local relevance and clarity [15,20,26,32].

The features covered both informational and behavioral support functions. Informational items included general PrEP education (eg, what PrEP is, how it works, and how to take it), more detailed information (eg, whether someone is a good candidate, side effects, and steps to start) [26], and logistical details (eg, where PrEP is offered, clinic hours, frequency of visits, and cost). Behavioral support features included reminders for PrEP appointments and pill-taking, interactive communication through a virtual assistant or a real person [15], and several app-based tools: push notifications to take PrEP, a daily pill tracker, a sexual activity log, and a feature that assesses whether the user is currently protected from HIV based on their reported pill use and sexual behavior [20].

Exploratory Factor Analysis

To identify latent constructs representing participant preferences for mHealth features, we conducted an exploratory factor analysis (EFA) on 11 items assessing interest in different functionalities of hypothetical mHealth PrEP support tools [38,39]. Before the EFA, data suitability was evaluated using the Kaiser-Meyer-Olkin measure of sampling adequacy and Bartlett test of sphericity. We used principal axis factoring with Promax (oblique) rotation, given the expected correlation among features. The number of factors retained was based on eigenvalues >1.0, scree plot inspection, parallel analysis, and conceptual clarity. Resulting factors were used to generate composite scores that represented core domains of mHealth tool interest, which were then used in regression analyses.

Variable Definitions

Preferred Communication Channel Category

The preferred communication channel category for PrEP support was defined according to the participant's most preferred platform for receiving PrEP-related communication. "Traditional channels" referred to phone calls or SMS text messaging, which do not require smartphone-specific apps. "Digital channels" included WhatsApp, a customized PrEP app, or other messaging or social media platforms such as Telegram (Telegram FZ-LLC), Messenger and Instagram (Meta Platforms, Inc), or TikTok (ByteDance Ltd), all of which require a smartphone and an internet connection.

PrEP Cascade Outcomes

PrEP cascade outcomes reflected 3 distinct stages of engagement in the HIV prevention cascade [36,37], selected for their proximity to PrEP uptake and their potential to be directly influenced by mHealth interventions. The decision to use PrEP was defined as responding "Yes" or "Probably yes" to the question "Have you decided to start using PrEP?" Sought PrEP referred to self-reported attempts to obtain PrEP from a clinic or provider. PrEP initiation was defined as having ever started PrEP, including both current and past use.

Sociodemographic and Behavioral Covariates

Sociodemographic covariates included age (<25 years vs ≥25 years; based on regional PrEP implementation evidence showing lower engagement among younger individuals [40]), educational

XSL•FO
RenderX

level (elementary and secondary vs post-secondary), monthly income (<1000 Peruvian soles [PEN; 1000 PEN=US $268 as of August 15, 2023], equivalent to the minimum monthly wage at the time of the study [41], vs ≥1000 PEN), and city of residence (Metropolitan Lima vs other cities). Behavioral variables referred to the past 6 months and included self-perceived HIV risk (yes or no), engagement in anal sex (yes or no), number of anal sex partners (≤5 vs>5), condom use (always vs not always), transactional sex (yes or no), and recent bacterial sexually transmitted infection (STI) diagnosis (chlamydia, gonorrhea, or syphilis; yes or no).

Statistical Analysis

Analytic Sample Determination

The analytic sample included participants who completed all exposure and outcome measures. Respondents who reported being assigned female at birth, identified as a gender other than cisgender male, or self-reported living with HIV were excluded from the analytic sample. No missing data were observed among the 600 participants included in the analysis; therefore, complete-case analysis was performed.

Descriptive Statistics

We summarized participant characteristics using frequencies and percentages for categorical variables and medians with IQRs for continuous variables given nonnormal distributions. Group comparisons used 2-tailed chi-square tests or Fisher exact tests when expected cell counts were <5. Continuous variables were compared using Wilcoxon rank sum tests. To compare preferences across digital communication platforms, we used pairwise 2-tailed Wilcoxon signed-rank tests on participant rankings, applying Bonferroni correction for multiple comparisons. Statistical significance was defined as α=0.05.

Logistic Regression Analyses

We conducted bivariate and multivariable logistic regression analyses with a logit link to examine associations between interest in mHealth tool features and engagement at 3 distinct stages of the HIV prevention cascade. Interest in mHealth features was operationalized using median composite scores derived from EFA. Each PrEP cascade outcome was modeled independently to reflect conceptually distinct dimensions of PrEP engagement with potentially different determinants. Because the cascade stages were assessed cross-sectionally and were not strictly sequential in all participants, we modeled each stage as a separate binary outcome rather than imposing hierarchical or conditional restrictions that would reduce sample size and could introduce selection bias. This approach allowed us to estimate associations with having reached each cascade stage without assuming temporal ordering or monotonic progression across stages.

For each outcome, stepwise selection based on the lowest Akaike Information Criterion was first applied to sociodemographic and behavioral variables to identify relevant covariates independently for each PrEP cascade stage. Akaike Information Criterion was used to balance model fit and parsimony while reducing the risk of overfitting. Variance inflation factors were examined to assess multicollinearity, with all retained covariates having variance inflation factor <2. The selected variables were then used to adjust separate multivariable logistic regression models for each mHealth composite factor score. Results are presented as odds ratios and adjusted odds ratios (aORs), with 95% CI. As a sensitivity analysis, multinomial logistic regression models were fitted, treating each cascade stage as a categorical outcome to assess whether associations observed in stage-specific models were consistent when modeled jointly. All analyses were conducted using R version 4.3.1 (R Foundation for Statistical Computing).

Ethical Considerations

The informed consent form and survey instrument (Multimedia Appendix 1) were reviewed and approved in Spanish by the Comité Institucional de Bioética de Vía Libre (approval 9008 - 2023 a). Electronic informed consent was obtained before survey access, and a waiver of documentation of written informed consent was issued to preserve participant anonymity. Participants indicated consent by selecting “I agree” before accessing survey items. The consent form clearly stated that the survey was intended for adult MSM living in Peru, and self-identification was confirmed upon accepting the consent. No personally identifying information was collected, and IP addresses were used solely for duplicate detection. All data were stored on secure, password-protected servers accessible only to the research team. Participation was voluntary, and respondents could skip any question or exit the survey at any time without penalty. No financial compensation was provided. Upon completion, participants received information on HIV prevention services and resources. No identifiable images or personal data are included in this paper.

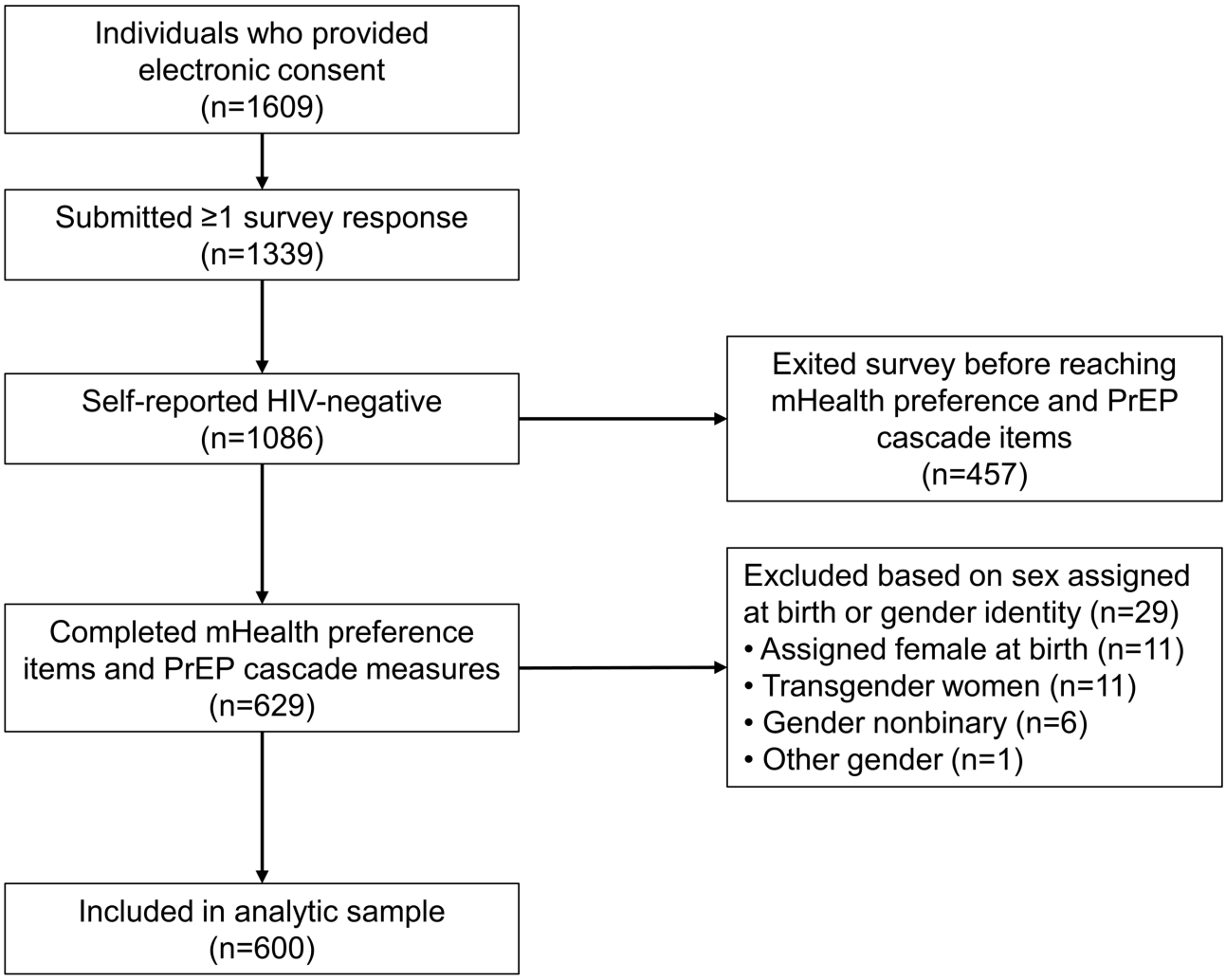
Results

Participant Flow

Of the 1609 individuals who provided informed consent, 1339 submitted at least one survey response. Among these, 1086 self-reported being HIV-negative and were retained for further screening. A total of 457 respondents were excluded because they did not complete the required exposure or outcome items due to survey dropout before reaching those sections. Compared to those with incomplete surveys, respondents who completed the full survey were older (median 29.0, IQR 24.0 - 35.0 years vs median 27, IQR 21.0 - 33.0 years; *P*<.001), more likely to reside in Metropolitan Lima (473/629, 75.2% vs 262/457, 57.3%; *P*<.001), have postsecondary education (494/629, 78.5% vs 328/457, 71.8%; *P*=.01), report higher perceived HIV risk (455/629, 72.3% vs 263/457, 60.9%; *P*=.001), and report more than 5 anal sex partners in the past 6 months (177/629, 35.3% vs 69/457, 25%; *P*=.005). There were no significant differences in sex, gender identity, income, or behavioral characteristics. We further excluded 29 respondents who did not identify as cisgender men, resulting in a final analytic sample of 600 HIV-negative MSM who completed the survey. Participant flow is summarized in Figure 1.

XSL•FO
RenderX

Figure 1. Participant flow diagram for a cross-sectional online survey assessing communication technology use and preferences for mobile health (mHealth) features to support HIV preexposure prophylaxis (PrEP) engagement among HIV-negative men who have sex with men in Peru (June 16-August 31, 2023).



Participant Characteristics

Participants were primarily recruited through organic content on social media and WhatsApp groups (305/600, 50.8%), followed by paid advertisements on Instagram or Facebook (216/600, 36%) and on Grindr (79/600, 13.2%) (Table S1 in Multimedia Appendix 2). The median age in the final analytic sample was 29 (IQR 24‐35) years, with 152 (25.3%) aged <25 years. Most participants had completed postsecondary education (479/600, 79.8%), and 403 (67.2%) reported a monthly income exceeding 1000 PEN. A total of 450 (75%) resided in Metropolitan Lima. In the past 6 months, 476 (79.3%) engaged in anal sex, 354 (59%) reported condomless sex, and 168 (28%) reported a recent bacterial STI diagnosis. Additionally, 95 (15.8%) engaged in transactional sex, and 440 (73.3%) perceived themselves to be at risk for HIV.

Engagement along the PrEP cascade varied. While 518 (86.3%) participants expressed interest in using PrEP and 483 (80.5%) had made the decision to use it, fewer had taken steps to access it. A total of 190 (31.7%) had sought PrEP, 109 (18.2%) had initiated it, and only 43 (7.2%) were currently using PrEP with self-reported good adherence.

Communication Behaviors

Smartphone ownership was nearly universal (589/600, 98.2%). Participants reported frequent use of multiple digital platforms, with WhatsApp emerging as the most widely used platform; 547 (91.2%) participants reported using it multiple times per day or throughout the day. Frequent use of internet browsers (498/600, 83%) and social media (405/600, 67.5%) was also common. Phone calls remained widely used, with 432 (72%) participants reporting several calls per day or all-day use. In contrast, SMS text messaging was less popular, with only 193 (32.2%) participants reporting similarly frequent use (Figure S1 in Multimedia Appendix 2).

Preferred Channels for PrEP Support

Participants expressed a strong preference for receiving PrEP-related support through digital channels. As shown in Figure 2, WhatsApp was the most preferred platform, with 302 (50.3%) participants ranking it as their top choice (Figure S2 in Multimedia Appendix 2). This preference was statistically stronger than for any other channel (all adjusted P<.001). Preferences across the remaining platforms also differed significantly in nearly all pairwise comparisons. The only exceptions were the comparisons between the customized PrEP

XSL•FO
RenderX

app and SMS text messaging (adjusted *P*=.24), and between other chat apps and phone calls (adjusted *P*=.07). When platforms were categorized as digital vs traditional, 450 (75%) participants selected digital channels (WhatsApp, other chat platforms, or a customized PrEP app) as their most preferred platform for receiving PrEP support, while only 150 (25%) participants ranked traditional channels (SMS text messaging or phone calls) as their top choice.

Figure 2. Mean rank scores of preferred communication channels for receiving HIV preexposure prophylaxis–related support among HIV-negative men who have sex with men in Peru (cross-sectional online survey, June 16-August 31, 2023; N=600). Lower mean rank scores indicate greater preference for each communication channel. mHealth: mobile health; PrEP: preexposure prophylaxis.

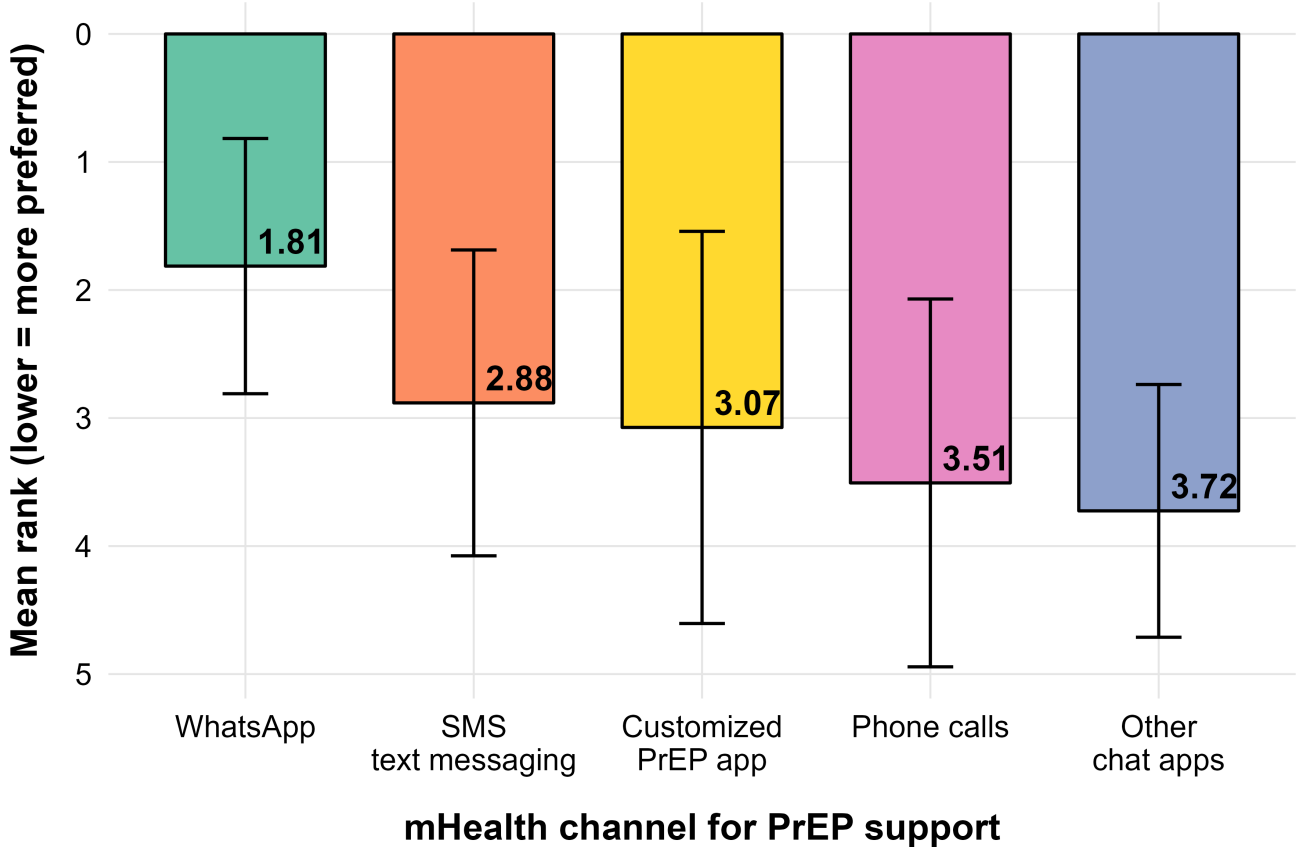


Table 1 provides sociodemographic and behavioral characteristics by preferred communication channel category for PrEP support. Compared to participants who preferred traditional channels, those who preferred digital channels were more likely to reside in Metropolitan Lima (352/450, 78.2% vs 98/150, 65.3%; *P*=.002) and to report having more than 5 anal sex partners in the past 6 months (137/450, 30.4% vs 32/150, 21.3%; *P*=.03). Participants preferring digital channels were also significantly more likely to own a smartphone (448/450, 99.6% vs 141/150, 94.0%; *P*<.001). Other characteristics, including age, income, education level, recruitment method, perceived HIV risk, STI history, and condom use, did not differ significantly between groups.

XSL•FO
RenderX

Table . Sociodemographic and behavioral characteristics of HIV-negative men who have sex with men participating in a cross-sectional online survey assessing communication technology use and preferences for mHealth[a] features to support HIV PrEP[b] engagement in Peru (June 16-August 31, 2023; N=600), stratified by preferred communication channel for PrEP support.

Characteristics	Overall (N=600), n (%)	Traditional channels[c] (n=150), n (%)	Digital channels[d] (n=450), n (%)	*P* value[e]
Age (years)				.12
<25	152 (25.3)	32 (21.3)	120 (26.7)	
≥25	448 (74.7)	118 (78.7)	330 (73.3)	
Monthly income (PEN)[f]				.65
<1000	197 (32.8)	47 (31.3)	150 (33.3)	
≥1000	403 (67.2)	103 (68.7)	300 (66.7)	
Educational level				.77
Elementary or secondary	121 (20.2)	29 (19.3)	92 (20.4)	
Postsecondary	479 (79.8)	121 (80.7)	358 (79.6)	
Location				.002
Metropolitan Lima	450 (75)	98 (65.3)	352 (78.2)	
All other cities	150 (25)	52 (34.7)	98 (21.8)	
Recruitment method				.06
Organic content[g] (unpaid)	305 (50.8)	64 (42.7)	241 (53.6)	
Instagram and Facebook ads	216 (36)	61 (40.7)	155 (34.4)	
Grindr ads	79 (13.2)	25 (16.7)	54 (12)	
Owns a smartphone				<.001
No	11 (1.8)	9 (6)	2 (0.4)	
Yes	589 (98.2)	141 (94.0)	448 (99.6)	
Perceived HIV risk[h]				.20
No	160 (26.7)	46 (30.7)	114 (25.3)	
Yes	440 (73.3)	104 (69.3)	336 (74.7)	
Engaged in anal sex[h]				.64
No	124 (20.7)	33 (22.0)	91 (20.2)	
Yes	476 (79.3)	117 (78.0)	359 (79.8)	
Anal sex partners[h]				.03
≤5 partners	431 (71.8)	118 (78.7)	313 (69.6)	
>5 partners	169 (28.2)	32 (21.3)	137 (30.4)	
Condom use[h]				.15
Always	246 (41)	69 (46)	177 (39.3)	
Not always	354 (59)	81 (54)	273 (60.7)	
Transactional sex[h]				.08
No	505 (84.2)	133 (88.7)	372 (82.7)	
Yes	95 (15.8)	17 (11.3)	78 (17.3)	
Recent STI[i] diagnosis[h]				.83
No	432 (72)	107 (71.3)	325 (72.2)	
Yes	168 (28)	43 (28.7)	125 (27.8)	

XSL•FO
RenderX

[a]mHealth: mobile health

[b]PrEP: preexposure prophylaxis.

[c]Includes phone calls and SMS text messaging.

[d]Includes WhatsApp, a customized PrEP app, and other messaging platforms (eg, Telegram, Messenger, Instagram, TikTok).

[e]*P* values correspond to chi-square tests, except where expected cell counts were <5, in which case Fisher exact test was used.

[f]PEN: Peruvian soles (1000 PEN approximately US $268 as of August 15, 2023).

[g]Organic content refers to standard posts shared via the participating HIV and AIDS service organizations' social media accounts and WhatsApp groups for study recruitment, without paid promotion.

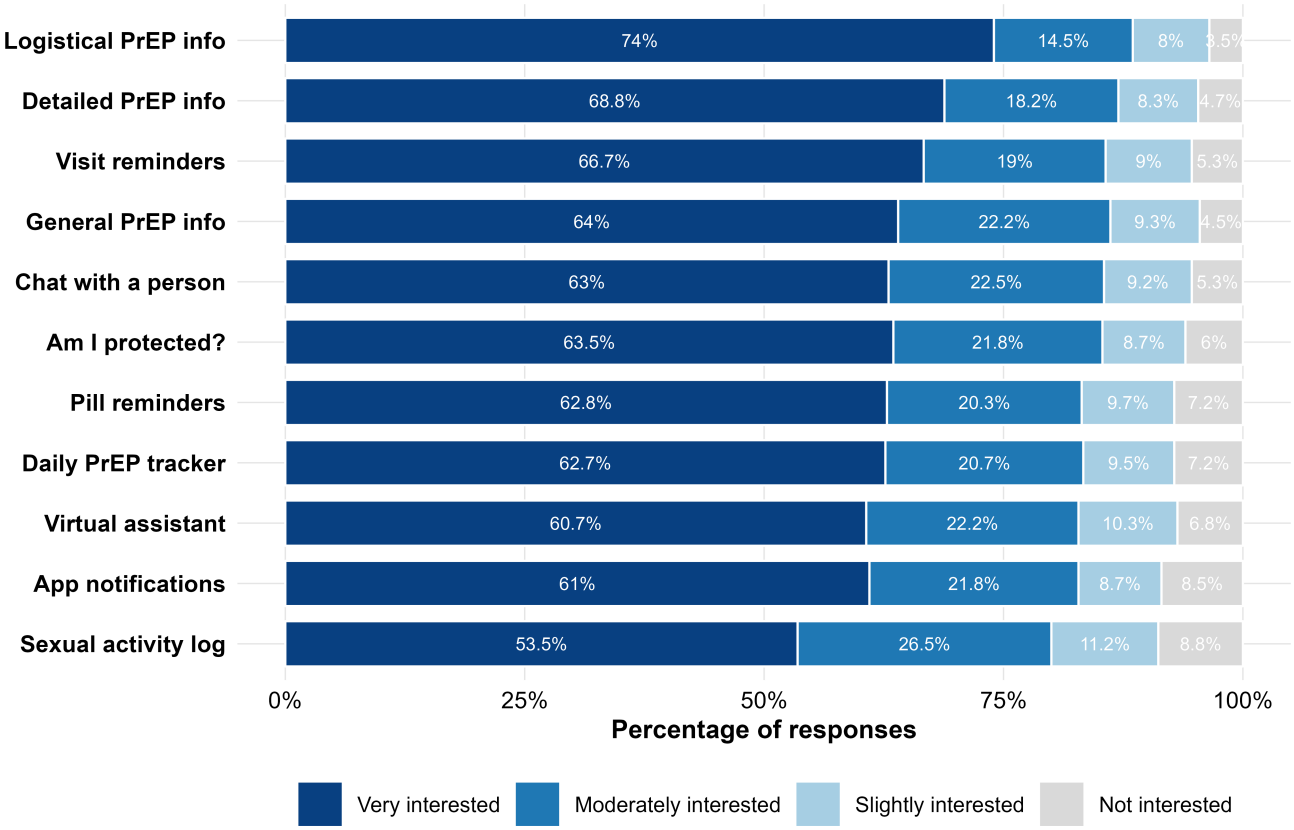
[h]Behavioral variables refer to the past 6 months.

[i]STI: sexually transmitted infection (chlamydia, gonorrhea, or syphilis).

Preferences for mHealth Tool Features for PrEP Support

Overall, interest was high across mHealth features for PrEP support (Figure 3 and Figures S4 and S5 in Multimedia Appendix 2). The most frequently endorsed feature was logistical information, such as the locations where PrEP is offered, clinic hours, frequency of visits, and cost, with 444 (74%) reporting they were "very interested." This was followed by detailed PrEP information, including whether someone is a good candidate, possible side effects, and the steps required to initiate PrEP (413/600, 68.8%), appointment reminders (400/600, 66.7%), and general PrEP information, such as what PrEP is, how it works, and how to take it (384/600, 64%). Features such as chatting with a person (378/600, 63%), pill reminders (377/600, 62.8%), a daily PrEP tracker (376/600, 62.7%), and a virtual assistant (364/600, 60.7%) also generated strong interest. App notifications (366/600, 61%) and the "Am I protected?" feature (381/600, 63.5%) were similarly well rated. The sexual activity log received the lowest endorsement, though over half of participants (321/600, 53.5%) still indicated they were "very interested."

Figure 3. Distribution of participant-reported interest in specific mobile health features to support HIV preexposure prophylaxis use among HIV-negative men who have sex with men in Peru (cross-sectional online survey, June 16-August 31, 2023; N=600). Bars represent the proportion of respondents selecting each Likert-scale category (not interested, slightly interested, moderately interested, and very interested), and segments sum to 100% per feature. PrEP: preexposure prophylaxis.



EFA of mHealth Tool Features for PrEP Support

The Kaiser-Meyer-Olkin measure indicated excellent sampling adequacy (0.93), and the Bartlett test of sphericity was significant (χ^2_{55}=7582.29; *P*<.001), supporting the suitability of the data for factor analysis. The EFA of the 11 items yielded a 3-factor solution explaining 81% of total variance and demonstrating high internal consistency for each factor (Table 2). The three domains were interpreted as:

1. Informational support (Cronbach α=0.94) included features providing educational content and logistical guidance related to PrEP. This factor encompassed general PrEP

XSL•FO
RenderX

information, more detailed content about candidacy and side effects, logistical details such as service locations and costs, and visit reminders.

2. Self-management tools (Cronbach α=0.94) captured interest in tools for daily adherence and self-monitoring. It included pill reminders, app notifications, a daily PrEP tracker, a sexual activity log, and the "Am I protected?" feature.
3. Interactive communication (Cronbach α=0.91) reflected a preference for real-time, 2-way support, whether through chatting with a person or a virtual assistant.

Table . Exploratory factor analysis[a] of participant-reported interest in mHealth[b] features to support HIV PrEP[c] use among HIV-negative men who have sex with men in Peru (cross-sectional online survey, June 16-August 31, 2023; N=600).

Items	Self-management tools (Cronbach α=0.94), factor loadings	Informational support (Cronbach α=0.94), factor loadings	Interactive communication (Cronbach α=0.91), factor loadings
General information about PrEP (eg, how it works, who it is intended for)	—[d]	0.80	—
Detailed medical information about PrEP (eg, side effects, lab tests, and dosing)	—	0.95	—
Logistical information on where and how to get PrEP (eg, nearby clinics, schedules, and costs)	—	0.90	—
Reminders for upcoming clinic appointments	—	0.63	—
Daily pill-taking reminders or dosing alerts	0.51	—	—
Push notifications from the app with updates or support	0.94	—	—
Daily tracker to monitor PrEP use	0.92	—	—
Log to record recent sexual activity (number of partners, condom use, etc)	0.74	—	—
Feature estimating HIV protection based on reported pill use and sexual activity	0.65	—	—
Virtual assistant (chatbot) providing automatic responses	—	—	0.54
Option to chat with a real person for support or questions	—	—	0.96

[a]Principal axis factoring with Promax (oblique) rotation. All retained factors had eigenvalues >1.0 and were supported by scree plot and parallel analysis. No substantial cross-loadings (>0.40 on more than one factor) were observed.

[b]mHealth: mobile health.

[c]PrEP: preexposure prophylaxis.

[d]Factor loadings <0.40 are not displayed for clarity.

Median scores were skewed toward the upper end of the scale, reflecting generally high levels of interest across all mHealth domains, with scores of 4.00 (IQR 3.00 - 4.00) for informational support and interactive communication, and 3.80 (IQR 3.00 - 4.00) for self-management tools. Even among participants who were unsure or not interested in PrEP at the time of the survey, interest in mHealth tools remained moderate to high, with median scores of 3.00 (IQR 2.00 - 3.50) for informational support, 2.70 (IQR 1.80 - 3.60) for self-management tools, and 3.00 (IQR 2.00 - 4.00) for interactive communication. As expected, interest was significantly higher among those who were currently interested in using PrEP, with median scores of 4.00 (IQR 3.50 - 4.00), 3.80 (IQR 3.00 - 4.00), and 4.00 (IQR 3.00 - 4.00) across the respective domains (all *P*<.001).

Association Between Interest in mHealth Tool Features for PrEP Support and PrEP Cascade Stages

Decided to Use PrEP

Participants who had decided to use PrEP (483/600, 80.5%) were more likely than those who were unsure or had not decided (117/600, 19.5%) to reside in Metropolitan Lima (375/483, 77.6% vs 75/117, 64.1%; *P*=.002), perceive themselves at risk for HIV (382/483, 79.1% vs 58/117, 49.6%; *P*<.001), report anal sex in the past 6 months (392/483, 81.2% vs 84/117, 71.8%; *P*=.03), and have more than 5 anal sex partners (148/483, 30.6% vs 21/117, 17.9%; *P*=.006). In adjusted analyses, greater interest in informational support (aOR 4.54, 95% CI 3.36 - 6.28; *P*<.001), self-management tools (aOR 3.23, 95% CI 2.51 - 4.22; *P*<.001), and interactive communication (aOR 2.74, 95% CI

XSL•FO
RenderX

2.15 - 3.53; *P*<.001) was associated with higher odds of having decided to use PrEP. Full results are provided in Table 3.

Table . Associations between interest in mHealth[a] feature domains and decision to use HIV PrEP[b] among HIV-negative men who have sex with men in Peru (cross-sectional online survey, June 16-August 31, 2023; N=600).

Characteristics	Decided to use PrEP[c]		Bivariate	Multivariable[d]
	Exposed, n/N (%)	Reference, n/N (%)	OR[e] (95% CI)	aOR[f] (95% CI)
Age ≥25 years	366/448 (81.7)	117/152 (77.0)	1.34 (0.85 - 2.07)	—[g]
Postsecondary education	389/479 (81.2)	94/121 (77.7)	1.24 (0.75 - 2.00)	—
Income >1000 PEN[h]	330/403 (81.9)	153/197 (77.7)	1.30 (0.85 - 1.97)	—
Resides in Metropolitan Lima	375/450 (83.3)	108/150 (72.0)	1.96 (1.25 - 3.03)[i]	1.92 (1.20 - 3.03)[i]
Self-perceived HIV risk[j]	382/440 (86.8)	101/160 (63.1)	3.85 (2.52 - 5.89)[k]	3.64 (2.36 - 5.63)[k]
Engaged in anal sex[j]	392/476 (82.4)	91/124 (73.4)	1.69 (1.06 - 2.67)[l]	—
Recent STI[m] diagnosis	138/168 (82.1)	345/432 (79.9)	1.16 (0.74 - 1.86)	—
>5 anal sex partners[j]	148/169 (87.6)	335/431 (77.7)	2.02 (1.23 - 3.44)[i]	1.54 (0.92 - 2.68)
Condomless sex[j]	293/354 (82.8)	190/246 (77.2)	1.42 (0.94 - 2.13)	—
Transactional sex[j]	82/95 (86.3)	401/505 (79.4)	1.64 (0.90 - 3.18)	—
mHealth features[n]				
Informational support[o]	2.75 (2.00 - 3.50)[p]	4.00 (3.75 - 4.00)[q]	4.99 (3.73 - 6.85)[k]	4.54 (3.36 - 6.28)[k]
Self-management tools[o]	2.80 (1.80 - 3.60)[p]	4.00 (3.20 - 4.00)[q]	3.53 (2.76 - 4.59)[k]	3.23 (2.51 - 4.22)[k]
Interactive communication[o]	3.00 (2.00 - 3.50)[p]	4.00 (3.00 - 4.00)[q]	2.97 (2.36 - 3.78)[k]	2.74 (2.15 - 3.53)[k]

[a]mHealth: mobile health.

[b]PrEP: preexposure prophylaxis.

[c]Exposed n/N (%) and reference n/N (%) indicate the number and proportion of participants with the outcome (decided to use PrEP) within each exposure category.

[d]Multivariable models were built using a two-step approach. First, stepwise selection was applied to all sociodemographic and behavioral variables to identify covariates independently associated with the decision to use PrEP. Second, each mHealth factor score was entered into a separate multivariable logistic regression model, adjusted for the covariates selected in step one. Variance inflation factors (VIFs) confirmed no multicollinearity (all <2).

[e]OR: odds ratio.

[f]aOR: adjusted odds ratio.

[g]Not available.

[h]PEN: Peruvian soles (1000 PEN approximately US $268 as of August 15, 2023).

[i]*P*<.01.

[j]Behavioral variables refer to the past 6 months.

[k]*P*<.001.

[l]*P*<.05.

[m]STI: sexually transmitted infection (chlamydia, gonorrhea, or syphilis).

[n]mHealth domain scores were derived from exploratory factor analysis using principal axis factoring with Promax rotation. Scores are presented as median (IQR) and were entered as continuous variables in regression models.

[o]Multivariable models adjusted for city of residence, self-perceived HIV risk, and number of anal sex partners.

[p]mHealth domain score among participants with the outcome (decided to use PrEP); median (IQR)

[q]mHealth domain score among participants without the outcome (decided to use PrEP); median (IQR)

Sought PrEP

Participants who had sought PrEP (190/600, 31.7%) were more likely than those who had not (410/600, 68.3%) to be aged 25 years or older (158/190, 83.2% vs 290/410, 70.7%; *P*=.001), have postsecondary education (161/190, 84.7% vs 318/410, 77.6%; *P*=.04), reside in Metropolitan Lima (161/190, 84.7% vs 289/410, 70.5%; *P*<.001), report condomless sex (124/190, 65.3% vs 230/410, 56.1%; *P*=.03), have more than 5 anal sex partners (72/190, 37.9% vs 97/410, 23.7%; *P*<.001), engage in transactional sex (41/190, 21.6% vs 54/410, 13.2%; *P*=.009), and report a recent STI diagnosis (68/190, 35.8% vs 100/410,

XSL•FO
RenderX

24.4%; *P*=.004). In adjusted analyses, higher interest in informational support (aOR 1.43, 95% CI 1.10 - 1.89; *P*=.001) and self-management tools (aOR 1.34, 95% CI 1.06 - 1.70; *P*=.02) was associated with having sought PrEP, whereas the association with interactive communication was marginal (aOR 1.25, 95% CI 1.00 - 1.57; *P*=.053). Full results are provided in Table 4.

Table . Associations between interest in mHealth[a] feature domains and having sought HIV PrEP[b] among HIV-negative men who have sex with men in Peru (cross-sectional online survey, June 16-August 31, 2023; N=600).

Characteristics	Sought PrEP[c]		Bivariate	Multivariable[d]
	Exposed, n/N (%)	Reference, n/N (%)	OR[e] (95% CI)	aOR[f] (95% CI)
Age ≥25 years	158/448 (35.3)	32/152 (21.1)	2.04 (1.34 - 3.20)[g]	1.74 (1.09 - 2.82)[h]
Postsecondary education	161/479 (33.6)	29/121 (24.0)	1.61 (1.03 - 2.58)[h]	1.49 (0.91 - 2.50)
Income >1000 PEN[i]	136/403 (33.7)	54/197 (27.4)	1.35 (0.93 - 1.97)	—[j]
Resides in Metropolitan Lima	161/450 (35.8)	29/150 (19.3)	2.33 (1.49 - 3.70)[k]	2.22 (1.41 - 3.57)[k]
Self-perceived HIV risk[l]	149/440 (33.9)	41/160 (25.6)	1.49 (1.00 - 2.25)	—
Engaged in anal sex[l]	153/476 (32.1)	37/124 (29.8)	1.11 (0.73 - 1.73)	—
Recent STI[m] diagnosis[l]	68/168 (40.5)	122/432 (28.2)	1.73 (1.19 - 2.51)[g]	1.47 (0.99 - 2.18)
>5 anal sex partners[l]	72/169 (42.6)	118/431 (27.4)	1.97 (1.36 - 2.85)[k]	1.62 (1.09 - 2.41)[h]
Condomless sex[l]	124/354 (35.0)	66/246 (26.8)	1.47 (1.03 - 2.11)[h]	—
Transactional sex[l]	41/95 (43.2)	149/505 (29.5)	1.81 (1.15 - 2.84)[g]	1.63 (1.01 - 2.62)[h]
mHealth features[n]				
Informational support[o]	4.00 (3.00 - 4.00)[p]	4.00 (3.50 - 4.00)[q]	1.54 (1.20 - 2.03)[g]	1.43 (1.10 - 1.89)[h]
Self-management tools[o]	3.60 (3.00 - 4.00)[p]	4.00 (3.20 - 4.00)[q]	1.42 (1.14 - 1.80)[g]	1.34 (1.06 - 1.70)[h]
Interactive communication[o]	4.00 (3.00 - 4.00)[p]	4.00 (3.00 - 4.00)[q]	1.32 (1.07 - 1.65)[h]	1.25 (1.00 - 1.57)

[a]mHealth: mobile health.

[b]PrEP: preexposure prophylaxis.

[c]Exposed n/N (%) and reference n/N (%) indicate the number and proportion of participants with the outcome (sought PrEP) within each exposure category.

[d]Multivariable models were built using a two-step approach. First, stepwise selection was applied to all sociodemographic and behavioral variables to identify covariates independently associated with having sought PrEP. Second, each mHealth factor score was entered into a separate multivariable logistic regression model, adjusted for the covariates selected in step one. Variance inflation factors (VIFs) confirmed no multicollinearity (all <2).

[e]OR: odds ratio.

[f]aOR: adjusted odds ratio.

[g]*P*<.01

[h]*P*<.05.

[i]PEN: Peruvian soles (1000 PEN: approximately US $268 as of August 15, 2023).

[j]Not available.

[k]*P*<.001

[l]Behavioral variables refer to the past 6 months.

[m]STI: sexually transmitted infection (chlamydia, gonorrhea, or syphilis).

[n]mHealth domain scores were derived from exploratory factor analysis using principal axis factoring with Promax rotation. Scores are presented as median (IQR) and were entered as continuous variables in regression models.

[o]Multivariable models adjusted for age, education, city of residence, recent STI diagnosis, number of anal sex partners, and transactional sex.

[p]mHealth domain score among participants with the outcome (sought PrEP); median (IQR).

[q]mHealth domain score among participants without the outcome (sought PrEP); median (IQR).

XSL•FO
RenderX

PrEP Initiation

Participants who had initiated PrEP (109/600, 18.2%) were more likely than those who had not (491/600, 81.8%) to report a monthly income above 1000 PEN (87/109, 79.8% vs 316/491, 64.4%; *P*=.002), have postsecondary education (96/109, 88.1% vs 383/491, 78%; *P*=.02), reside in Metropolitan Lima (94/109, 86.2% vs 356/491, 72.5%; *P*=.003), and report a recent STI diagnosis (42/109, 38.5% vs 126/491, 25.7%; *P*=.007). In adjusted analyses, higher interest in informational support (aOR 1.64, 95% CI 1.16 - 2.44; *P*=.009) and self-management tools (aOR 1.49, 95% CI 1.11 - 2.05; *P*=.01) was associated with PrEP initiation, whereas interactive communication was not significantly associated. Full results are provided in Table 5.

Table . Associations between interest in mHealth[a] feature domains and initiating HIV PrEP[b] among HIV-negative men who have sex with men in Peru (cross-sectional online survey, June 16-August 31, 2023; N=600).

Characteristics	Ever initiated PrEP[c]		Bivariate	Multivariable[d]
	Exposed, n/N (%)	Reference, n/N (%)	OR[e] (95% CI)	aOR[f] (95% CI)
Age ≥25 years	88/448 (19.6)	21/152 (13.8)	1.52 (0.93 - 2.61)	—[g]
Postsecondary education	96/479 (20.0)	13/121 (10.7)	2.08 (1.16 - 4.03)[h]	1.78 (0.95 - 3.55)
Income >1000 PEN[i]	87/403 (21.6)	22/197 (11.2)	2.19 (1.35 - 3.70)[j]	1.85 (1.10 - 3.22)[h]
Resides in Metropolitan Lima	94/450 (20.9)	15/150 (10.0)	2.38 (1.37 - 4.35)[j]	2.27 (1.28 - 4.17)[j]
Self-perceived HIV risk[k]	85/440 (19.3)	24/160 (15.0)	1.36 (0.84 - 2.26)	—
Engaged in anal sex[k]	83/476 (17.4)	26/124 (21.0)	0.80 (0.49 - 1.32)	—
Recent STI[l] diagnosis[k]	42/168 (25.0)	67/432 (15.5)	1.82 (1.17 - 2.80)[j]	1.83 (1.17 - 2.86)[j]
>5 anal sex partners[k]	39/169 (23.1)	70/431 (16.2)	1.55 (0.99 - 2.39)	—
Condomless sex[k]	67/354 (18.9)	42/246 (17.1)	1.13 (0.74 - 1.74)	—
Transactional sex[k]	18/95 (18.9)	91/505 (18.0)	1.06 (0.59 - 1.83)	—
mHealth features[m]				
Informational support[n]	4.00 (3.00 - 4.00)[o]	4.00 (3.75 - 4.00)[p]	1.76 (1.25 - 2.60)[j]	1.64 (1.16 - 2.44)[j]
Self-management tools[n]	3.60 (3.00 - 4.00)[o]	4.00 (3.40 - 4.00)[p]	1.57 (1.17 - 2.16)[j]	1.49 (1.11 - 2.05)[h]
Interactive communication[n]	4.00 (3.00 - 4.00)[o]	4.00 (3.00 - 4.00)[p]	1.25 (0.97 - 1.66)	1.20 (0.93 - 1.60)

[a]mHealth: mobile health.

[b]PrEP: preexposure prophylaxis.

[c]Exposed n/N (%) and reference n/N (%) indicate the number and proportion of participants with the outcome (initiated PrEP) within each exposure category.

[d]Multivariable models were built using a two-step approach. First, stepwise selection was applied to all sociodemographic and behavioral variables to identify covariates independently associated with PrEP initiation. Second, each mHealth factor score was entered into a separate multivariable logistic regression model, adjusted for the covariates selected in step one. Variance inflation factors (VIFs) confirmed no multicollinearity (all <2).

[e]OR: odds ratio.

[f]aOR: adjusted odds ratio.

[g]Not available.

[h]*P*<.05.

[i]PEN: Peruvian soles (1000 PEN: approximately US $268 as of August 15, 2023).

[j]*P*<.01.

[k]Behavioral variables refer to the past 6 months.

[l]STI: sexually transmitted infection (chlamydia, gonorrhea, or syphilis).

[m]mHealth domain scores were derived from exploratory factor analysis using principal axis factoring with Promax rotation. Scores are presented as median (IQR) and were entered as continuous variables in regression models.

[n]Multivariable models adjusted for education, income, city of residence and recent STI diagnosis.

[o]mHealth domain score among participants with the outcome (initiated PrEP); median (IQR).

[p]mHealth domain score among participants with the outcome (initiated PrEP); median (IQR).

XSL•FO
RenderX

Multinomial Sensitivity Analyses

As a sensitivity analysis, we fit multinomial logistic regression models treating PrEP cascade stage as a 4-level categorical outcome (not decided, decided only, sought PrEP, and initiated PrEP), using "not decided" as the reference category. The distribution of cascade stages was as follows: not decided (101/600, 16.8%), decided only (285/600, 47.5%), sought PrEP only (105/600, 17.5%), and initiated PrEP (109/600, 18.2%). Associations were broadly consistent with the stage-specific logistic regression models (Table 6).

Table . Multinomial analysis[a] of associations between interest in mHealth[b] feature domains and engagement across stages of the HIV PrEP[c] cascade among HIV-negative men who have sex with men in Peru (cross-sectional online survey, June 16-August 31, 2023; N=600).

Multinomial analysis	Decided only vs not decided, aRRR[d] (95% CI)	Sought PrEP vs not decided, aRRR (95% CI)	Initiated PrEP vs not decided, aRRR (95% CI)
PrEP cascade stage vs not decided (each mHealth domain entered separately)			
Score: informational support[e]	4.18 (2.97 - 5.87)[f]	3.02 (2.04 - 4.49)[f]	4.18 (2.65 - 6.57)[f]
Score: self-management tools[e]	3.25 (2.42 - 4.35)[f]	2.57 (1.82 - 3.64)[f]	3.31 (2.27 - 4.83)[f]
Score: interactive support[e]	2.93 (2.20 - 3.89)[f]	2.26 (1.62 - 3.17)[f]	2.45 (1.74 - 3.44)[f]
All mHealth scores in one model (mutually adjusted)			
Score: informational support[e]	2.99 (1.67 - 5.34)[f]	2.44 (1.22 - 4.89)[g]	3.37 (1.58 - 7.19)[h]
Score: self-management tools[e]	1.49 (0.84 - 2.64)	1.50 (0.76 - 2.95)	2.03 (0.99 - 4.16)
Score: interactive support[e]	0.98 (0.57 - 1.70)	0.85 (0.45 - 1.60)	0.63 (0.33 - 1.20)

[a]Multinomial models were adjusted for prespecified sociodemographic and behavioral covariates (age category, income, education, location, perceived HIV risk, number of anal sex partners, recent STI diagnosis, and transactional sex) included simultaneously in all models. Models in the upper panel include each mHealth domain score entered separately. Models in the lower panel include all three mHealth domain scores simultaneously (mutually adjusted).

[b]mHealth: mobile health.

[c]PrEP: preexposure prophylaxis.

[d]aRRR: adjusted relative risk ratio.

[e]mHealth domain scores were derived from exploratory factor analysis using principal axis factoring with Promax rotation and entered as continuous variables in regression models.

[f]$P<.001$.

[g]$P<.05$.

[h]$P<.01$.

In models including each mHealth domain separately, higher scores were strongly associated with engagement in more advanced cascade stages compared with not having decided to use PrEP. For example, for each 1-unit increase in the informational support score, the relative risk of being in the decided-only, sought, and initiated stages was approximately 3- to 4-fold higher compared with the not-decided stage, with similar magnitudes observed for the self-management tools and interactive support scores. When all 3 domain scores were included simultaneously, the informational support score remained independently associated with more advanced cascade stages relative to the not-decided stage, whereas associations for the self-management tools and interactive support scores were attenuated, suggesting shared variance across domains. When cascade stages were rereferenced to examine adjacent contrasts, domain scores were strongly associated with the contrast between decided-only and not decided but were not significantly associated with transitions from decided-only to sought PrEP or from sought PrEP to initiated PrEP.

Discussion

Principal Findings

In this cross-sectional survey of HIV-negative MSM in Peru, we found near-universal smartphone ownership and a strong preference for digital communication platforms, particularly WhatsApp, for receiving PrEP-related support. Interest in informational support, self-management tools, and interactive communication features was high overall. Greater interest in these mHealth domains was associated with engagement across the PrEP cascade. Informational support and self-management features demonstrated consistent associations with all 3 stages, whereas interactive communication features were associated only with the decision stage. Associations were strongest for the decision stage and more modest for later stages. Sensitivity analyses using multinomial models showed similar overall patterns.

These findings are consistent with prior evidence suggesting that mHealth interventions can support engagement with HIV prevention services. A recent meta-analysis of randomized controlled trials reported that digital interventions improve HIV testing and PrEP adherence and show emerging effects on PrEP

XSL•FO
RenderX

uptake [42]. Similarly, a systematic review among MSM concluded that digital tools are generally feasible and acceptable and may support engagement across the PrEP continuum when culturally tailored and user-centered [43]. Together, this body of evidence underscores the potential role of mHealth as an entry point into prevention services that can be integrated with offline initiation and follow-up within differentiated service delivery models [24,44]. Our findings extend this literature by identifying which types of digital support may be most salient at different stages of PrEP engagement. In particular, the stage-specific associations observed in our analyses provide insight into when digital supports may be most useful within the PrEP cascade.

The stronger associations at the decision stage, compared with the more modest associations at later stages, suggest that digital features may be particularly salient during early phases of PrEP engagement. At these stages, individuals may be evaluating their HIV risk, learning about prevention options, and determining whether and how to access services. Digital tools delivered through widely used communication platforms could therefore help translate risk awareness into concrete prevention decisions by providing accessible information about eligibility, service availability, and pathways to PrEP initiation. In contrast, progression through later cascade stages, such as seeking services and initiating PrEP, may depend more heavily on structural and health system factors, including service accessibility, appointment availability, and provider interactions.

Informational support demonstrated the strongest and most consistent associations across the PrEP cascade stages examined in this study, underscoring the continued importance of accurate, accessible, and actionable information about PrEP. Features that clarify where and how to obtain PrEP, including service locations, clinic hours, costs, and eligibility criteria, may be particularly relevant in contexts where program expansion is ongoing. Despite expansion of PrEP services in Peru, uncertainty about access pathways may persist. Digital informational tools may help address these barriers by centralizing prevention information and simplifying service navigation. For example, decision-support and navigation platforms such as MyPrEP [26], SaludFindr [31], and PrEPresent [25] have been designed to help users assess PrEP eligibility, locate services, and connect with prevention programs. Clinic-integrated platforms such as JomPrEP [23,24] extend these functions by combining informational resources with service navigation features, including HIV self-testing support, teleconsultation, and streamlined pathways for PrEP initiation, within a mobile interface. In the Peruvian context, similar tools could be adapted to incorporate location-specific service directories, appointment scheduling, and referral pathways. Emerging approaches, including the integration of generative artificial intelligence to dynamically update service directories, eligibility guidance, and prevention information, may further enhance scalability and timeliness of such tools [45-47], although their effectiveness in real-world PrEP delivery settings warrants evaluation.

Likewise, self-management tools including pill reminders, adherence trackers, protection status calculators, and sexual activity logs were also associated with greater engagement across the PrEP cascade, suggesting that users value tools that support adherence, routine formation, and ongoing self-monitoring. These functionalities are designed to support adherence and ongoing engagement after initiation, and evidence from mobile interventions in other settings suggests they can be feasible and acceptable to users. For example, tools such as PrEPmate [15], iTAB [16], the Amsterdam PrEP app [17], PrEP-iT! [18], P3 [19], and DOT Diary [20] incorporate adherence reminders, monitoring systems, or behavioral feedback features and have demonstrated high usability and feasibility, with some studies reporting improvements in PrEP adherence. In the Peruvian context, similar tools could be adapted with appropriate cultural and linguistic tailoring. Core features such as reminders and simplified adherence tracking could also be integrated into widely used messaging platforms. Emerging digital approaches may facilitate scalable adaptation; however, their effectiveness and implementation feasibility should be evaluated within local health system contexts.

By contrast, interactive communication features appeared to be most relevant earlier in the cascade, with the strongest association observed for the decision to use PrEP. This pattern suggests that personalized, bidirectional communication may support individuals as they explore PrEP and clarify uncertainties. Digital conversational tools, including AI-enabled chatbots, have been proposed as scalable mechanisms to deliver tailored prevention information and facilitate service navigation while operating continuously and triaging more complex inquiries to human providers [48]. Early implementations of chatbot-based HIV prevention tools such as Amanda Selfie [27] and CHIA [30] have demonstrated feasibility and user acceptability in pilot studies. MYHIV365, a chatbot-based digital intervention developed in Malaysia to support HIV prevention among MSM, integrates conversational guidance on HIV testing, PrEP awareness, and risk reduction with culturally tailored content designed to address stigma and informational barriers [28,29]. Early evaluations have reported high usability and acceptability, suggesting that conversational agents may provide a scalable mechanism to deliver prevention information and support decision-making in stigmatized or resource-constrained settings.

Because such conversational systems can be deployed through widely used messaging platforms such as WhatsApp, they may offer a practical approach to delivering interactive prevention support within existing communication ecosystems. The widespread and frequent use of WhatsApp observed in our sample further supports the feasibility of this approach. Nearly all participants owned smartphones and reported frequent use of WhatsApp for daily communication, consistent with prior reports documenting the central role of messaging platforms in digital health engagement across diverse settings [8,28,49-52]. In contexts such as Peru, where developing and maintaining dedicated health applications may be resource-intensive for public health systems, leveraging existing communication platforms may offer a pragmatic strategy for delivering scalable interactive prevention support. Earlier work among MSM in Lima reported substantially lower smartphone access in 2015 [7], suggesting that the rapid expansion of digital connectivity over the past decade may enable broader implementation of messaging-based prevention tools. As conversational

XSL•FO
RenderX

technologies evolve, integrating AI-enabled chatbots with clinical messaging workflows may provide scalable models for delivering personalized PrEP guidance and follow-up within routine prevention services.

The potential value of these digital supports will likely depend on how well they align with local communication habits, user preferences, social norms, health system capacity, and service pathways. Adapting mHealth tools for the Peruvian context by embedding local clinical pathways, incorporating culturally and linguistically appropriate content, and integrating with widely used platforms may enhance their relevance across different stages of the PrEP cascade. Prior Spanish-language adaptations of digital HIV interventions have demonstrated the feasibility of this approach and underscore the importance of aligning tools with local communication norms and service structures [31,32]. Together, these considerations support the development of contextually grounded digital strategies rather than the direct transplantation of tools developed for other health system environments.

In addition to linguistic and service-delivery adaptation, digital tools may also be relevant because they can address barriers related to privacy and anticipated stigma. Although stigma was not directly measured in this study, prior research in Peru indicates that discrimination and discomfort in health care settings remain important barriers to PrEP uptake [53-55]. Digital support tools have been proposed as an implementation strategy to reduce structural and interpersonal barriers by allowing individuals to access information and decision support more privately [12,56]. By offering discreet engagement with credible sources and supportive communication, mHealth approaches may lower some perceived barriers to initial PrEP exploration. Digital support strategies should be viewed as complementary to, rather than replacements for, broader efforts to address stigma within health care systems.

Beyond individual feature preferences, our findings also have implications for how digital supports might be embedded within broader PrEP delivery models. Two relevant approaches are online-to-offline linkage models and telePrEP. The high level of interest in digital support features observed in this study is consistent with the potential relevance of online-to-offline approaches for PrEP delivery. In these models, individuals are initially engaged through digital platforms such as targeted messaging, educational content, or virtual navigation, and subsequently linked to in-person services through facilitated referrals, electronic appointment systems, or digital vouchers [57]. Such models may be particularly pertinent in settings where the geographic concentration of services and perceived stigma create barriers to in-person engagement [58,59]. The Adam's Love online-to-offline model in Thailand, which increased HIV testing and PrEP uptake through integrated online counseling and streamlined clinic check-in processes, illustrates how digital engagement can be operationalized within routine service delivery [60]. Adaptation of similar models to the Peruvian context would require integration with local service pathways and evaluation of their feasibility and impact.

TelePrEP represents a complementary strategy within differentiated service delivery models [56]. By incorporating virtual consultations, remote prescribing, and home-based monitoring, telePrEP has been proposed as a mechanism to expand access for individuals who prefer remote care or who reside in areas with limited clinic availability [61]. Integration of mHealth features such as messaging-based reminders, adherence tracking, and location-specific informational support into telePrEP platforms may facilitate continuity between digital engagement and clinical services. In combination, online-to-offline approaches and telePrEP models illustrate how digital strategies could be configured to support more flexible and geographically inclusive PrEP delivery frameworks in Peru. Formal evaluation would be needed to determine their effectiveness and scalability within the national health system.

As access to long-acting PrEP expands, including twice-yearly lenacapavir, which has been approved in the United States and recommended by the World Health Organization as an additional PrEP option [62,63], priorities for digital support may shift toward earlier stages of the prevention cascade. Because twice-yearly injectable PrEP reduces daily adherence demands compared with oral regimens, digital tools may become increasingly relevant for awareness, risk appraisal, demand generation, and linkage to services, while still supporting safety monitoring and follow-up. Ongoing research is also evaluating lenacapavir formulations administered at extended intervals [64], which could further reduce adherence-related barriers and increase the relative importance of timely, localized information regarding eligibility, service availability, cost, and evolving guidance. Integration of these informational and navigation functions into widely used platforms such as WhatsApp may support engagement as long-acting options become available, while maintaining pathways to human support.

This study has several limitations. First, the cross-sectional online survey design likely oversampled individuals who are urban and digitally connected, which may limit generalizability to populations with more limited internet access and to other groups not included in this analysis. Second, all measures were self-reported, including behavioral indicators and interest in digital features, which may be subject to recall or reporting bias. Although anonymous online surveys may reduce social desirability bias [65], measurement error remains possible. Third, interest was assessed for hypothetical mHealth features rather than actual tool use or clinical outcomes and therefore observed associations with PrEP cascade stages should not be interpreted as evidence of effectiveness. Fourth, the cross-sectional design precludes causal inference regarding the directionality of associations. Finally, stigma was not directly measured and therefore could not be examined as a mediator or moderator of digital engagement preferences.

These limitations underscore the need for prospective implementation research to evaluate the feasibility, acceptability, and effectiveness of specific mHealth interventions within diverse Peruvian settings. Implementation science frameworks may provide structured guidance for adapting and evaluating digital tools in real-world contexts [51,66]. By identifying which domains of digital support are most strongly associated with engagement at different stages of the PrEP cascade, our findings offer practical guidance for designing user-centered digital

XSL•FO
RenderX

strategies to strengthen HIV prevention delivery in Peru and similar settings.

Conclusions

In summary, this study provides novel, stage-specific evidence on digital preferences that may support PrEP engagement in a Latin American implementation context. Informational support and self-management tools were consistently associated with engagement across multiple stages of the PrEP cascade, whereas interactive communication features appeared most relevant during early decision-making phases. Among these domains, informational support emerged as the most consistent correlate of reaching more advanced cascade stages, highlighting the importance of accessible and actionable information in facilitating progression along the prevention continuum. These findings extend the existing literature by identifying which digital functions may be most salient at different points in the prevention cascade.

From a public health perspective, the near-universal smartphone ownership and strong preference for WhatsApp suggest that culturally tailored, messaging-based digital strategies aligned with specific cascade stages could be feasibly integrated within differentiated service delivery models, including telePrEP and online-to-offline frameworks. As long-acting PrEP options expand and delivery models evolve, contextually adapted mHealth tools may help align prevention services with users' communication practices and complement efforts to address structural barriers to engagement. Rigorous implementation research is needed to determine their effectiveness, scalability, and sustainability within Peru's national PrEP program and similar health system environments.

Acknowledgments

We thank the Community Engagement Unit of the Peru Clinical Trials Unit for their contributions to language review and community validation of the survey instrument. Generative artificial intelligence tools were not used in any portion of the paper preparation.

Funding

This work was supported by the National Institute of Allergy and Infectious Diseases, National Institutes of Health (R01AI177082; FLA), and the Fogarty International Center of the National Institutes of Health (D43TW010540; JAG-C).

Data Availability

The anonymized dataset analyzed during this study is available from the corresponding author upon reasonable request.

Authors' Contributions

JAG-C led the conceptualization, data curation, formal analysis, investigation, methodology, project administration, visualization, and writing (original draft of the paper). KAK contributed to formal analysis and methodology and provided cosupervision alongside SB, with JS serving as the local supervisor and FLA as the senior supervisor. DLGG-Q supported investigation and project administration. All authors contributed to writing (review and editing of the paper), with JAG-C leading revisions.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Mobile health survey in Spanish.
[DOCX File, 45 KB - publichealth_v12i1e84982_app1.docx]

Multimedia Appendix 2
Additional tables and figures.
[DOCX File, 520 KB - publichealth_v12i1e84982_app2.docx]

References

1. Nachega JB, Musoke P, Kilmarx PH, et al. Global HIV control: is the glass half empty or half full? Lancet HIV 2023 Sep;10(9):e617-e622. [doi: 10.1016/S2352-3018(23)00150-9] [Medline: 37506723]
2. Situación epidemiológica del VIH-SIDA en el Perú [Webpage in Spanish]. Gobierno del Perú, Ministerio de Salud, Centro Nacional de Epidemiología, Prevención y Control de Enfermedades (CDC MINSA). 2024. URL: https://www.dge.gob.pe/vih/uploads/nacional_vih.html [accessed 2025-11-13]
3. Fonner VA, Dalglish SL, Kennedy CE, et al. Effectiveness and safety of oral HIV preexposure prophylaxis for all populations. AIDS 2016 Jul 31;30(12):1973-1983. [doi: 10.1097/QAD.0000000000001145] [Medline: 27149090]

XSL•FO
RenderX

4. Gonzales-Gavancho C, Araoz-Salinas JM, Ramon Tapia R, et al. Updates on HIV pre-exposure prophylaxis in Latin America: available drugs and implementation status. Infez Med 2025;33(1):29-49. [doi: 10.53854/liim-3301-4] [Medline: 40071257]
5. Menacho L, Konda KA, Lecca L, et al. Optimising PrEP uptake and use in Peru: no time to lose!. Lancet HIV 2024 Apr;11(4):e204-e206. [doi: 10.1016/S2352-3018(24)00038-9] [Medline: 38538157]
6. Reyes-Díaz M, Celly A, Folch C, et al. Latin American Internet Survey for Men Who Have Sex With Men (LAMIS-2018): design, methods and implementation. PLoS One 2022;17(11):e0277518. [doi: 10.1371/journal.pone.0277518] [Medline: 36395121]
7. Krishnan A, Ferro EG, Weikum D, et al. Communication technology use and mHealth acceptance among HIV-infected men who have sex with men in Peru: implications for HIV prevention and treatment. AIDS Care 2015;27(3):273-282. [doi: 10.1080/09540121.2014.963014] [Medline: 25285464]
8. Castanheira D, Pereira CR, Jalil CM, et al. Young men who have sex with men and digital technologies for HIV prevention and care: qualitative findings from the conectad@s study. Culture, Health & Sexuality 2025 Dec 2;27(12):1517-1532. [doi: 10.1080/13691058.2024.2428329]
9. Shrestha R, Lim SH, Altice FL, et al. Use of smartphone to seek sexual health information online among Malaysian men who have sex with men (MSM): implications for mHealth intervention to increase HIV testing and reduce HIV risks. J Community Health 2020 Feb;45(1):10-19. [doi: 10.1007/s10900-019-00713-x] [Medline: 31375976]
10. Mhando F, Nyankomo M, Hall C, et al. Digital intervention services to promote HIV self-testing and linkage to care services: a bibliometric and content analysis—global trends and future directions. Public Health Rev 2024;45:1606354. [doi: 10.3389/phrs.2024.1606354] [Medline: 38434540]
11. Muessig KE, Pike EC, Legrand S, Hightow-Weidman LB. Mobile phone applications for the care and prevention of HIV and other sexually transmitted diseases: a review. J Med Internet Res 2013 Jan 4;15(1):e1. [doi: 10.2196/jmir.2301] [Medline: 23291245]
12. Goldstein M, Archary M, Adong J, et al. Systematic review of mHealth interventions for adolescent and young adult HIV prevention and the adolescent HIV continuum of care in low to middle income countries. AIDS Behav 2023 May;27(Suppl 1):94-115. [doi: 10.1007/s10461-022-03840-0] [Medline: 36322217]
13. Jones J, Schneider I, Valencia R, et al. Randomized controlled trial of a HIV prevention app to increase pre-exposure prophylaxis uptake among gay and bisexual men who have sex with men in the United States: the HealthMindr PrEP study. J Acquir Immune Defic Syndr 2025 Dec 15;100(5):425-434. [doi: 10.1097/QAI.0000000000003747] [Medline: 40828543]
14. Biello KB, Mayer KH, Scott H, et al. The effects of MyChoices and LYNX mobile apps on HIV testing and pre-exposure prophylaxis use by young US sexual minority men: results from a national randomized controlled trial. JMIR Public Health Surveill 2025 Feb 5;11:e63428. [doi: 10.2196/63428] [Medline: 39908084]
15. Liu AY, Vittinghoff E, von Felten P, et al. Randomized controlled trial of a mobile health intervention to promote retention and adherence to preexposure prophylaxis among young people at risk for human immunodeficiency virus: the EPIC study. Clin Infect Dis 2019 May 30;68(12):2010-2017. [doi: 10.1093/cid/ciy810] [Medline: 30239620]
16. Moore DJ, Jain S, Dubé MP, et al. Randomized controlled trial of daily text messages to support adherence to preexposure prophylaxis in individuals at risk for human immunodeficiency virus: the TAPIR study. Clin Infect Dis 2018 May 2;66(10):1566-1572. [doi: 10.1093/cid/cix1055] [Medline: 29228144]
17. van den Elshout MAM, Hoornenborg E, Achterbergh RCA, et al. Improving adherence to daily preexposure prophylaxis among MSM in Amsterdam by providing feedback via a mobile application. AIDS 2021 Sep 1;35(11):1823-1834. [doi: 10.1097/QAD.0000000000002949] [Medline: 34001705]
18. Horvath KJ, Helm JL, Black A, et al. A pilot randomized controlled trial of an mhealth intervention to improve PrEP adherence among young sexual minority men. AIDS Behav 2024 Aug;28(8):2804-2820. [doi: 10.1007/s10461-024-04374-3] [Medline: 38816592]
19. Hightow-Weidman LB, Rainer C, Schader L, et al. Prepared, Protected, EmPowered (P3): primary results of a randomized controlled trial using a social networking, gamification, and coaching app to promote pre-exposure prophylaxis (PrEP) adherence for sexual and gender minority (SGM) youth living in the United States. AIDS Behav 2025 Feb;29(2):652-663. [doi: 10.1007/s10461-024-04547-0] [Medline: 39531118]
20. Liu AY, Laborde ND, Coleman K, et al. DOT Diary: developing a novel mobile app using artificial intelligence and an electronic sexual diary to measure and support PrEP adherence among young men who have sex with men. AIDS Behav 2021 Apr;25(4):1001-1012. [doi: 10.1007/s10461-020-03054-2] [Medline: 33044687]
21. Li C, Xiong Y, Maman S, et al. An instant messaging mobile phone application for promoting HIV pre-exposure prophylaxis uptake among Chinese gay, bisexual and other men who have sex with men: a mixed methods feasibility and piloting randomized controlled trial study. PLoS One 2023;18(11):e0285036. [doi: 10.1371/journal.pone.0285036]
22. Marins LMS, Torres TS, Moreira RI, et al. Weekly text messages to support adherence to oral pre-exposure prophylaxis (PrEP) among gay, bisexual, and other cisgender men who have sex with men (MSM) and transgender women: pilot randomized controlled trial nested in PrEP Brasil study. J Med Internet Res 2025 Jul 18;27:e72360. [doi: 10.2196/72360] [Medline: 40680163]

XSL•FO
RenderX

23. Shrestha R, Wickersham JA, Khati A, et al. Clinic-integrated mobile health intervention ("JomPrEP" App) to improve uptake of HIV testing and pre-exposure prophylaxis among men who have sex with men in Malaysia: protocol for an intervention development and multiphase trial. JMIR Res Protoc 2022 Dec 21;11(12):e43318. [doi: 10.2196/43318] [Medline: 36542425]
24. Shrestha R, Altice FL, Khati A, et al. Clinic-integrated smartphone app (JomPrEP) to improve uptake of HIV testing and pre-exposure prophylaxis among men who have sex with men in Malaysia: mixed methods evaluation of usability and acceptability. JMIR Mhealth Uhealth 2023 Feb 16;11:e44468. [doi: 10.2196/44468] [Medline: 36795465]
25. Calvetti S, Lei B, Stocks JB, et al. Feasibility and acceptability of a mHealth patient navigation intervention to increase pre-exposure prophylaxis uptake in racially and ethnically diverse sexual and gender minority youth in Los Angeles (PrEPresent): pilot randomized controlled trial. JMIR Form Res 2025 Oct 31;9:e69255. [doi: 10.2196/69255] [Medline: 41171997]
26. Celum C, Seidman D, Travill D, et al. A decision support tool has similar high PrEP uptake and increases early PrEP persistence in adolescent girls and young women in South Africa: results from a randomized controlled trial. J Int AIDS Soc 2023 Aug;26(8):e26154. [doi: 10.1002/jia2.26154] [Medline: 37634942]
27. Massa P, de Souza Ferraz DA, Magno L, et al. A transgender chatbot (Amanda Selfie) to create pre-exposure prophylaxis demand among adolescents in Brazil: assessment of acceptability, functionality, usability, and results. J Med Internet Res 2023 Jun 23;25:e41881. [doi: 10.2196/41881] [Medline: 37351920]
28. Cheah MH, Gan YN, Altice FL, et al. Testing the feasibility and acceptability of using an artificial intelligence chatbot to promote HIV testing and pre-exposure prophylaxis in Malaysia: mixed methods study. JMIR Hum Factors 2024 Jan 26;11:e52055. [doi: 10.2196/52055] [Medline: 38277206]
29. Ni Z, Oh S, Saifi R, Azwa I, Altice FL. Evaluating the usability of an HIV prevention artificial intelligence chatbot in Malaysia: national observational study. JMIR Hum Factors 2025 Jul 15;12:e70034. [doi: 10.2196/70034] [Medline: 40663792]
30. Tao J, Pavlick E, Grondin A, et al. Evaluation of an artificial intelligence conversational chatbot to enhance HIV preexposure prophylaxis uptake: development and usability internal testing. J Med Internet Res 2026 Feb 3;28:e79671. [doi: 10.2196/79671] [Medline: 41632955]
31. Cantos VD, Hagen K, Duarte AP, et al. Development of a mobile app to increase the uptake of HIV pre-exposure prophylaxis among Latino sexual minority men: qualitative needs assessment. JMIR Form Res 2023 Feb 8;7:e43844. [doi: 10.2196/43844] [Medline: 36625855]
32. Liu AY, Alleyne CD, Doblecki-Lewis S, et al. Adapting mHealth interventions (PrEPmate and DOT Diary) to support PrEP retention in care and adherence among English and Spanish-speaking men who have sex with men and transgender women in the United States: formative work and pilot randomized trial. JMIR Form Res 2024 Mar 27;8:e54073. [doi: 10.2196/54073] [Medline: 38536232]
33. Jones J, Manley G, Glynn TR, et al. Evaluating the effectiveness of a mobile HIV prevention app to increase HIV and sexually transmitted infection testing and pre-exposure prophylaxis initiation among rural men who have sex with men in the southern United States: protocol for a randomized controlled trial. JMIR Res Protoc 2025 Jul 23;14:e69540. [doi: 10.2196/69540] [Medline: 40699887]
34. Sevilla-Gonzalez MDR, Moreno Loaeza L, Lazaro-Carrera LS, et al. Spanish version of the System Usability Scale for the assessment of electronic tools: development and validation. JMIR Hum Factors 2020 Dec 16;7(4):e21161. [doi: 10.2196/21161] [Medline: 33325828]
35. Appelbaum M, Cooper H, Kline RB, Mayo-Wilson E, Nezu AM, Rao SM. Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. Am Psychol 2018 Jan;73(1):3-25. [doi: 10.1037/amp0000191] [Medline: 29345484]
36. Parsons JT, Rendina HJ, Lassiter JM, Whitfield THF, Starks TJ, Grov C. Uptake of HIV pre-exposure prophylaxis (PrEP) in a national cohort of gay and bisexual men in the United States. J Acquir Immune Defic Syndr 2017 Mar 1;74(3):285-292. [doi: 10.1097/QAI.0000000000001251] [Medline: 28187084]
37. Gallardo-Cartagena JA, Oliveros D, Cabello R, et al. Identifying implementation gaps in the HIV prevention cascade among Peruvian men who have sex with men. AIDS Behav 2026 May 5. [doi: 10.1007/s10461-026-05147-w] [Medline: 42084836]
38. Haig BD. Exploratory factor analysis, theory generation, and scientific method. Multivariate Behav Res 2005;40(3):303-329. [doi: 10.1207/s15327906mbr4003_2] [Medline: 26794686]
39. Joliffe IT, Morgan BJ. Principal component analysis and exploratory factor analysis. Stat Methods Med Res 1992;1(1):69-95. [doi: 10.1177/096228029200100105] [Medline: 1341653]
40. Veloso VG, Cáceres CF, Hoagland B, et al. Same-day initiation of oral pre-exposure prophylaxis among gay, bisexual, and other cisgender men who have sex with men and transgender women in Brazil, Mexico, and Peru (ImPrEP): a prospective, single-arm, open-label, multicentre implementation study. Lancet HIV 2023 Feb;10(2):e84-e96. [doi: 10.1016/S2352-3018(22)00331-9] [Medline: 36565708]
41. Decreto Supremo N° 003-2022-TR. Gobierno del Perú, Ministerio de Trabajo y Promoción del Empleo. 2022. URL: https://www.gob.pe/institucion/mtpe/normas-legales/2890811-003-2022-tr [accessed 2025-05-26]

XSL•FO
RenderX

42. Li F, Xie C, Xiang F. Impact of mHealth on enhancing pre-exposure prophylaxis adherence and strengthening the HIV prevention cascade among key populations: a systematic review and meta-analysis. Front Public Health 2025;13(1600773):1600773. [doi: 10.3389/fpubh.2025.1600773] [Medline: 40642253]
43. Du J, Jin G, Zhang H, et al. Effectiveness of digital health interventions in promoting the pre-exposure prophylaxis (PrEP) care continuum among men who have sex with men (MSM): a systematic review of randomized controlled trials. Curr HIV/AIDS Rep 2025 Mar 18;22(1):25. [doi: 10.1007/s11904-025-00733-4] [Medline: 40100510]
44. Khati A, Wickersham JA, Gautam K, et al. Integrated online-to-offline model of care for HIV prevention and treatment among men who have sex with men in Malaysia: protocol for an intervention development and a multiphase trial. JMIR Res Protoc 2024 Oct 23;13:e60962. [doi: 10.2196/60962] [Medline: 39441624]
45. Marcus JL, Sewell WC, Balzer LB, Krakower DS. Artificial intelligence and machine learning for HIV prevention: emerging approaches to ending the epidemic. Curr HIV/AIDS Rep 2020 Jun;17(3):171-179. [doi: 10.1007/s11904-020-00490-6] [Medline: 32347446]
46. Ratevosian J, Reid M, Ni Z, et al. Reimagining HIV prevention with artificial intelligence. Lancet HIV 2025 Oct;12(10):e670-e671. [doi: 10.1016/S2352-3018(25)00158-4]
47. Odone A, Barbati C, Amadasi S, Schultz T, Resnik DB. Artificial intelligence and infectious diseases: an evidence-driven conceptual framework for research, public health, and clinical practice. Lancet Infect Dis 2026 Mar;26(3):e152-e167. [doi: 10.1016/S1473-3099(25)00412-8] [Medline: 40972628]
48. Ni Z, Peng ML, Balakrishnan V, et al. Implementation of chatbot technology in health care: protocol for a bibliometric analysis. JMIR Res Protoc 2024 Feb 15;13:e54349. [doi: 10.2196/54349] [Medline: 38228575]
49. Fabens I, Makhele C, Igaba NK, et al. WhatsApp versus SMS for 2-way, text-based follow-up after voluntary medical male circumcision in South Africa: exploration of messaging platform choice. JMIR Form Res 2024 Oct 16;8:e62762. [doi: 10.2196/62762] [Medline: 39412842]
50. Marin-Gomez FX, Garcia Cuyas F, Reig-Bolano R, et al. Social networking app use among primary health care professionals: web-based cross-sectional survey. JMIR Mhealth Uhealth 2018 Dec 21;6(12):e11147. [doi: 10.2196/11147] [Medline: 30578175]
51. Peng ML, Wickersham JA, Altice FL, et al. Formative evaluation of the acceptance of HIV prevention artificial intelligence chatbots by men who have sex with men in Malaysia: focus group study. JMIR Form Res 2022 Oct 6;6(10):e42055. [doi: 10.2196/42055] [Medline: 36201390]
52. Manji K, Hanefeld J, Vearey J, Walls H, de Gruchy T. Using WhatsApp messenger for health systems research: a scoping review of available literature. Health Policy Plan 2021 Jun 1;36(5):594-605. [doi: 10.1093/heapol/czab024] [Medline: 33860314]
53. Calabrese SK. Understanding, contextualizing, and addressing PrEP stigma to enhance PrEP implementation. Curr HIV/AIDS Rep 2020 Dec;17(6):579-588. [doi: 10.1007/s11904-020-00533-y] [Medline: 32965576]
54. Dubov A, Galbo P, Altice FL, Fraenkel L. Stigma and shame experiences by MSM who take PrEP for HIV prevention: a qualitative study. Am J Mens Health 2018 Nov;12(6):1843-1854. [doi: 10.1177/1557988318797437] [Medline: 30160195]
55. Oliveros Gómez D, Machavariani E, Altice FL, et al. Influence of stigma on engagement in HIV care and adherence to antiretroviral therapy in specialized HIV clinics targeting men who have sex with men and transgender women in Lima, Peru. AIDS Behav 2024 Aug;28(8):2755-2768. [doi: 10.1007/s10461-024-04401-3] [Medline: 38878137]
56. Mantell JE, Bauman LJ, Bonett S, et al. Innovation in providing equitable pre-exposure prophylaxis services in the United States: expanding access in nontraditional settings. J Acquir Immune Defic Syndr 2025 Apr 15;98(5S):e156-e169. [doi: 10.1097/QAI.0000000000003610] [Medline: 40163068]
57. Anand T, Nitpolprasert C, Phanuphak N. Online-to-offline models in HIV service delivery. Curr Opin HIV AIDS 2017 Sep;12(5):447-457. [doi: 10.1097/COH.0000000000000403] [Medline: 28682799]
58. Galea JT, Kinsler JJ, Salazar X, et al. Acceptability of pre-exposure prophylaxis as an HIV prevention strategy: barriers and facilitators to pre-exposure prophylaxis uptake among at-risk Peruvian populations. Int J STD AIDS 2011 May;22(5):256-262. [doi: 10.1258/ijsa.2009.009255] [Medline: 21571973]
59. Naz-McLean S, Clark J, Huerta L, et al. Social, economic, and physical side effects impact PrEP uptake and persistence among transgender women in Peru. BMC Public Health 2024 Jul 25;24(1). [doi: 10.1186/s12889-024-19474-x]
60. Anand T, Nitpolprasert C, Trachunthong D, et al. A novel online-to-offline (O2O) model for pre-exposure prophylaxis and HIV testing scale up. J Int AIDS Soc 2017 Mar 13;20(1):21326. [doi: 10.7448/IAS.20.1.21326] [Medline: 28362062]
61. Bonett S, Li Q, Sweeney A, Gaither-Hardy D, Safa H. Telehealth models for PrEP delivery: a systematic review of acceptability, implementation, and impact on the PrEP care continuum in the United States. AIDS Behav 2024 Sep;28(9):2875-2886. [doi: 10.1007/s10461-024-04366-3] [Medline: 38856846]
62. Kelley CF, Acevedo-Quiñones M, Agwu AL, et al. Twice-yearly lenacapavir for HIV prevention in men and gender-diverse persons. N Engl J Med 2025 Apr 3;392(13):1261-1276. [doi: 10.1056/NEJMoa2411858] [Medline: 39602624]
63. Guidelines on lenacapavir for HIV prevention and testing strategies for long-acting injectable pre-exposure prophylaxis. World Health Organization. 2025. URL: https://www.who.int/publications/i/item/9789240111608?utm [accessed 2025-09-18]
64. Jogiraju V, Pawar P, Yager J, et al. Pharmacokinetics and safety of once-yearly lenacapavir: a phase 1, open-label study. The Lancet 2025 Apr;405(10485):1147-1154. [doi: 10.1016/S0140-6736(25)00405-2]

XSL•FO
RenderX

65. Gnambs T, Kaspar K. Socially desirable responding in web-based questionnaires: a meta-analytic review of the Candor hypothesis. Assessment 2017 Sep;24(6):746-762. [doi: 10.1177/1073191115624547] [Medline: 26739360]
66. Zhai S, Yuwen W, Zyniewicz TL, et al. Evaluating behavioral intervention technologies: integrating human-centered design and implementation science outcomes. Digit Health 2025;11:20552076251348579. [doi: 10.1177/20552076251348579] [Medline: 40520137]

Abbreviations

aOR : adjusted odds ratio
EFA: exploratory factor analysis
mHealth: mobile health
MSM: men who have sex with men
PEN: Peruvian soles
PrEP: preexposure prophylaxis
STI : sexually transmitted infection

Edited by S Brini; submitted 29.Sep.2025; peer-reviewed by K Mngadi, SA John; revised version received 09.Mar.2026; accepted 13.Apr.2026; published 14.Jul.2026.

Please cite as:
Gallardo-Cartagena JA, German-Quiñones DLG, Konda KA, Buchbinder S, Sanchez J, Altice FL
Preferences for mHealth Features to Support Engagement in the HIV Preexposure Prophylaxis Cascade Among Men Who Have Sex With Men in Peru: Cross-Sectional Online Survey
JMIR Public Health Surveill 2026;12:e84982
URL: https://publichealth.jmir.org/2026/1/e84982
doi:10.2196/84982

© Jorge A Gallardo-Cartagena, Dora Leidy Geraldine German-Quiñones, Kelika Anne Konda, Susan Buchbinder, Jorge Sanchez, Frederick L Altice. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 14.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Profiles of Willingness to Use Pre-Exposure Prophylaxis Modalities and an HIV Vaccine Among Sexual and Gender Minority Individuals in Brazil, Mexico, and Peru: Cross-Sectional Online Survey

Jazmin Qquellon[1], MSc; Kelika A Konda[1,2], PhD; Oliver Elorreaga[1], MSc; Hamid Vega-Ramirez[3], MD, PhD; Centli Guillén-Díaz-Barriga[4], PhD; Dulce Díaz-Sosa[5], PhD; Brenda Hoagland[6], MD, PhD; Juan V Guanira[1], MPH, MD; Marcos Benedetti[6], MSc; Cristina Pimenta[6], PhD; Beatriz Grinsztejn[6], MD, PhD; Carlos F Caceres[1], MD, PhD; Valdilea G Veloso[6], MD, PhD; Thiago S Torres[6], PhD

[1]Centro de Investigación Interdisciplinaria en Sexualidad, Sida y Sociedad, Universidad Peruana Cayetano Heredia, Lima, Peru

[2]Division Disease Prevention, Policy and Global Health, Department of Population and Public Health Sciences, Keck School of Medicine, University of Southern California, Los Angeles, California, CA, United States

[3]Division of Epidemiology and Psychosocial Research, Instituto Nacional de Psiquiatría Ramón de la Fuente Muñiz, Mexico City, Mexico

[4]Faculty of Psychology, National Autonomous University of Mexico, Mexico City, Mexico

[5]Department of Health Sciences, Metropolitan Autonomous University-Lerma (UAM-L), Lerma de Villada, Mexico

[6]Instituto Nacional de Infectologia Evandro Chagas, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil

Corresponding Author:
Jazmin Qquellon, MSc
Centro de Investigación Interdisciplinaria en Sexualidad, Sida y Sociedad
Universidad Peruana Cayetano Heredia
Av. Armendáriz 497, Miraflores
Lima, 15074
Peru
Phone: 51 203 3300
Email: luz.qquellon@upch.pe

Abstract

Background: HIV incidence continues to disproportionately affect sexual and gender minority (SGM) individuals in Latin America. Pre-exposure prophylaxis (PrEP), including long-acting products, urgently need scaling up in the region. Understanding PrEP modality preferences can help design effective implementation.

Objective: This study examined willingness to use different PrEP modalities and an HIV vaccine among SGM individuals aged 18 years or older from Brazil, Mexico, and Peru and factors associated with willingness to use 4 PrEP modalities.

Methods: We conducted a cross-sectional online survey in 2021; participants were recruited via apps (Grindr and Hornet) and social media (Facebook, Instagram, and WhatsApp). We used multivariate Poisson regression with robust variance (α=.05) to estimate prevalence ratios, identifying differences in willingness to use daily oral, event-driven oral, monthly oral, and bimonthly injectable PrEP. Models were constructed for each PrEP modality and adjusted for age, country, race, education, income, HIV risk score, HIV testing, and HIV risk perception. Variables were retained in the final adjusted models regardless of statistical significance.

Results: Among 16,951 respondents, 10,385 (61.3%) were Brazilian, 4996 (29.5%) Mexican, and 1570 (9.3%) Peruvian. Median age was 32 (IQR 26-39) years. Among the total respondents, 12,621 (74.4%) were willing to use monthly oral PrEP; 11,153 (65.8%) daily oral PrEP; 10,212 (60.2%) bimonthly injectable PrEP; and 14,044 (82.8%) an HIV vaccine. Only 6442 (38%) were willing to use event-driven oral PrEP. In Brazil, 6082 (66.8%) were willing to use daily oral, 3162 (36.1%) event-driven, 7640 (74.9%) monthly oral, and 6450 (63.3%) injectable PrEP; in Mexico, 3242 (67.8%) daily oral, 1958 (41%) event-driven, 3728 (76.5%) monthly oral, and 2805 (57.5%) injectable PrEP; in Peru, 799 (53.1%) daily oral, 584 (39.9%) event-driven, 830 (62.5%) monthly oral, and 610 (45.9%) injectable PrEP. In multivariable models, willingness to use each of the 4 PrEP modalities was positively associated with high self-perceived HIV risk (adjusted prevalence ratios [aPRs] 1.10-1.25) and higher HIV Incidence

XSL•FO
RenderX

Risk Index scores (aPRs 1.08-1.22). Having lower education was associated with lower willingness for monthly oral and bimonthly injectable PrEP (aPR=0.93, 95% CI 0.89-0.97 and aPR=0.94, 95% CI 0.90-0.99, respectively). Never having been tested for HIV and testing more than 6 months ago were associated with lower willingness for daily oral PrEP (aPR=0.89, 95% CI 0.83-0.95 and aPR=0.95, 95% CI 0.91-0.99, respectively) and bimonthly injectable PrEP (aPR=0.80, 95% CI 0.74-0.86 and aPR=0.90, 95% CI 0.86-0.94, respectively).

Conclusions: Our results suggest a strong preference for long-acting formulations, including monthly oral and bimonthly injectable, among SGM individuals in Latin America. Further research is needed to address gaps in the understanding of prevention modalities. As additional PrEP modalities are included in HIV prevention programs, the development of accessible tools and community-based strategies will be essential to support informed PrEP choices and ensure equitable implementation across the region.

(JMIR Public Health Surveill 2026;12:e75753) doi:10.2196/75753

KEYWORDS

HIV prevention; vaccine; men who have sex with men; transgender women; cabotegravir; lenacapavir

Introduction

In 2024, an estimated 2.5 million people were living with HIV in Latin America [1]. Latin America is one of the few regions where HIV incidence has risen, with a 13% increase from 2010 to 2024 [2]. The incidence of HIV remains disproportionately higher among sexual and gender minority (SGM) individuals, who accounted for 66% of cases in Latin America in 2022. HIV prevalence among gay, bisexual, and other men who have sex with men (MSM) and transgender persons aged 15 to 49 years (2019-2023) was 10% and 9.5%, respectively [3]. Additionally, a study conducted in Brazil and Peru found high annualized HIV incidence among SGM individuals not using pre-exposure prophylaxis (PrEP; 3.9%, 95% CI 2.9-4.9), with higher incidence among participants from Peru, those aged 18 to 30 years, and those reporting lower incomes [4].

Currently, the World Health Organization (WHO) recommends daily or on-demand oral emtricitabine 200 mg and tenofovir disoproxil fumarate 300 mg (FTC/TDF), long-acting injectable cabotegravir (CAB-LA), long-acting injectable lenacapavir (LEN), and the dapivirine vaginal ring as effective HIV PrEP modalities [5]. As of July 2024, a total of 152 countries worldwide had adopted WHO recommendations on PrEP into their national guidelines, and 12 countries had adopted policies for CAB-LA [6]. In Latin America, oral PrEP was implemented as part of public policy in most countries, with Brazil initiating the provision of PrEP through the public health system (Sistema Único de Saúde [SUS]) in 2017, Mexico in 2021, and Peru in 2023 [7-9]. In 2024, a total of 286,862 persons were using PrEP in the region, with Brazil accounting for 58% of them [10]. In addition to regional disparities, more is needed to reach those who are at higher vulnerability to HIV. During the ImPrEP study, the largest PrEP implementation study in Latin America, younger participants, those who were Black or of mixed race (*pardo* or *mestizo*), those with lower educational attainment, and transgender women had higher HIV incidence and increased odds of discontinuing PrEP, lower adherence, and reduced persistence [11-13]. Similarly, PrEP uptake and persistence are worse among younger, Indigenous, and Black Brazilians according to Brazilian national data [14]. In response to implementation concerns with existing oral PrEP regimens, long-acting products need to be offered and accessible to those who need them most.

In 2021, CAB-LA was shown to be a highly effective alternative to daily oral FTC/TDF for HIV prevention among cisgender women, cisgender MSM, and transgender women [15,16]. Recently, subcutaneous LEN, administered twice yearly, was superior to FTC/TDF among cisgender women and SGM individuals in the phase 3 PURPOSE 1 and PURPOSE 2 clinical trials [17,18]. The frequency of administration of CAB-LA and LEN, bimonthly or biannually, respectively, is attractive for people who have difficulty adhering to an oral pill regimen. However, barriers to implementing injectable PrEP must be considered, including cost-effectiveness, staffing and other implementation constraints, acceptability and preferences among potential users, factors influencing timely attendance at injection visits, the prolonged subtherapeutic pharmacokinetic effect, and the feasibility of alternative administration sites [19,20]. While CAB-LA and LEN are not widely available in Latin America, Brazil is conducting the ImPrEP CAB Brasil and the ImPrEP LEN Brasil studies to evaluate the implementation of CAB-LA and LEN for SGM individuals aged 18 to 30 years in the public health system [21-23].

Different PrEP technologies, such as monthly oral, infusion of monoclonal antibodies, topical microbicides, and other long-acting products, are being developed [20,24]. For instance, subdermal implants have several advantages, such as easy removal; biodegradability; minimal health care system interaction; lower daily dosage; and more consistent, predictable drug release [20,25]. Islatravir delivered through a subdermal implant demonstrated promise as a potentially effective and well-tolerated approach for HIV prevention [26].

PrEP urgently needs to be scaled-up among SGM individuals in Latin America, including the provision of long-acting products, such as CAB-LA and LEN [3]. In this sense, understanding the willingness to use PrEP modalities could provide information for future adaptations or implementation of HIV prevention policies. We aimed to describe willingness to use different PrEP modalities and an HIV vaccine among SGM individuals in Brazil, Mexico, and Peru. We also assessed factors associated with willingness to use 4 PrEP modalities

XSL•FO
RenderX

(daily oral, event-driven oral, monthly oral, and bimonthly injectable).

Methods

Study Design

A cross-sectional online survey was conducted between April and August 2021 among individuals who self-identified as SGM individuals and aged 18 years or older from Brazil, Peru, and Mexico. Respondents were recruited using a convenience sampling approach through advertisements on dating apps, such as Grindr (3236/16,943, 19.1%) and Hornet (6400/16,943, 37.8%), and social media platforms, such as Facebook (3374/16,943, 19.9%), Instagram (2994/16,943, 17.7%), and WhatsApp (431/16,943, 2.5%), as well as other online sources (508/16,943, 3%). Each advertisement included a link to the survey website, where participants were provided with detailed information about the study aims, procedures, and confidentiality measures. Those who provided their electronic informed consent were able to access and complete a computer-based questionnaire that included information on sociodemographics, behavior, HIV testing, and prevention. We excluded participants who reported living with HIV.

Anonymity was ensured by not collecting any personally identifiable information. The survey platform prevented multiple submissions from the same device using browser cookies and IP restrictions. Participants were permitted to select only 1 response option per question to ensure data consistency and avoid multiple selections.

Variables

Outcomes

Participants were asked to report their willingness to use each of the PrEP modalities using the following question: “Considering that all PrEP modalities are available, how likely would you be to use them?” with answers recorded using a 4-point Likert scale: very unlikely, unlikely, likely, and very likely. “Very likely” answers were classified as “willing,” following the definition from previous studies [27,28]. We evaluated willingness to use the following PrEP modalities: daily oral, event-driven oral, monthly oral, bimonthly injectable in the gluteal muscle, daily topical, event-driven topical, patch, implant, infusion of monoclonal antibodies, and subcutaneous monoclonal antibodies. We also inquired whether participants were willing to use a vaccine to prevent HIV. Only factors associated with 4 PrEP modalities (daily oral, event-driven oral, monthly oral, and bimonthly injectable) were reported, as these were currently available, implemented, or under investigation in Latin America when the survey was launched.

Sociodemographics

For gender identity, participants answered the question “What gender do you currently identify with?” with the following possible options: cisgender man, cisgender woman, transgender woman, transgender man, queer or nonbinary, and other. Age was categorized as 18 to 24, 25 to 30, and more than 30 years. Ethnicity or race was classified as White, mixed race (*pardo* or *mestizo*), and Black, Asian, Indigenous, or other. Level of education was dichotomized as complete secondary education or less (lower education) and more than secondary education (higher education). Regarding individual monthly income, we considered the minimum wage per country in 2021 (Brazil: US $213 per month, Mexico: US $215 per month, and Peru: US $257 per month) and dichotomized as minimum wage or less (lower income) and more than minimum wage (higher income).

Perceived Risk of HIV Acquisition, HIV Incidence Risk Index, and HIV Testing

The perceived risk for HIV was evaluated using the question “Considering your current sexual practices, in your opinion, what is your risk of acquiring HIV in the next 12 months?” which was recategorized as low (none, low, or moderate) and high (high, or certainty of infection). The HIV Incidence Risk Index (HIRI) was constructed by combining sexual behavior and substance use variables, as described previously [29]. Scores equal to or greater than 10 were considered as “engaging in high HIV sexual exposure.” Regarding HIV testing, participants answered the question “When was the last time you took an HIV test?” which was categorized as never, 6 months or less, and more than 6 months.

Ethical Considerations

Ethics approval for the study was obtained from the respective ethical review boards in each country. In Brazil, the study was approved by the Institutional Review Board of Instituto Nacional de Infectologia Evandro Chagas-FIOCRUZ (CAAE 82021918.0.0000.5262); in Mexico, by the Research Ethics Committee of the National Institute of Psychiatry Ramón de la Fuente Muñiz (CEI/C/038/2018); and in Peru, by the Ethical Committee for Research with Human Subjects at Universidad Peruana Cayetano Heredia (101460). All participants signed electronic informed consent form before initiating the study. No personally identifiable information was collected, except IP addresses, which were used exclusively for quality control purposes (duplicate entry prevention) and were not linked to survey responses. Data were stored on secure servers and analyzed in deidentified form. We provided no compensation to participants.

Statistical Analysis

Statistical analyses were conducted using Stata (version 18.0; StataCorp). Frequencies and percentages of sociodemographic and behavior characteristics and HIV testing were compared by country (Brazil, Mexico, and Peru) using the χ^2 test. We also described these variables according to the willingness to use 4 PrEP modalities (daily oral, event-driven oral, monthly oral, and bimonthly injectable) for each country. Descriptive analyses were based on observed data.

Missing data were assessed using Little's MCAR (missing completely at random) test, which indicated that the data were not missing completely at random (χ^2_{121}=5239; P<.001). Therefore, missing values in outcome and covariate variables were handled using multiple imputation by chained equations under the assumption of missing at random. Binary variables were imputed using logistic regression models and multinomial variables using multinomial logistic regression models. A total

XSL•FO
RenderX

of 20 imputed datasets were generated, and parameter estimates were combined using the Rubin rules.

We conducted multivariate Poisson regression models with robust variance to estimate adjusted prevalence ratios (aPRs) and 95% CIs for willingness to use each PrEP modality (daily oral, event-driven oral, monthly oral, and bimonthly injectable). Models were adjusted for covariates selected a priori (age, country, race, education, income, HIRI score, HIV testing, and HIV risk perception). Variables were retained in the final adjusted models regardless of statistical significance.

Results

A total of 35,541 individuals accessed the online questionnaire, of whom 7952 (22.4%) were ineligible for the following reasons: 1903 (23.9%) did not provide informed consent, 1523 (19.2%) had participated previously, 686 (8.6%) self-identified as cisgender women, 62 (0.8%) were younger than 18 years, 895 (11.3%) were men who did not self-identify as SGM individuals, and 2883 (36.2%) reported living with HIV (Figure 1). Among 27,589 eligible individuals, 10,638 (38.6%) were excluded for not answering questions about PrEP modalities. Overall, 16,951 (61.4%) respondents were included in this analysis; 10,385 (61.3%) Brazilians, 4996 (29.5%) Mexicans, and 1570 (9.3%) Peruvians. Among these, 16,183 (95.5%) self-identified as cisgender men, 99 (0.6%) as transgender women, 60 (0.4%) as transgender men, 423 (2.5%) as nonbinary or queer persons, and 165 (1%) as other. Median age was 32 (IQR 26-39) years; almost half of the respondents (7483/16,285, 46%) self-declared mixed race, and more than three-quarters (12,627/16,509, 76.5%) received more than 1 minimum wage per month (higher income). In addition, 10,379 (61.2%) respondents were identified as having high HIV sexual exposure according to the HIRI, 1961 (11.6%) had never been tested for HIV, and 6688 (39.6%) had their last HIV test more than 6 months ago.

Study population characteristics were compared by country (Table 1). More Peruvians (164/1570, 10.5%) self-identified as transgender or nonbinary persons compared with Brazilians (224/10,384, 2.2%) and Mexicans (379/4996, 7.6%) and were aged 18 to 24 years (Peru: 716/1570, 45.6%; Mexico: 1057/4996, 21.2%; Brazil: 1304/10,385, 12.6%). Regarding race, most Brazilians (5873/10,385, 56.5%) self-defined as White, while most Mexicans (3249/4330, 75%) and Peruvians (1109/1570, 70.6%) self-identified as mixed race. Most individuals in Brazil and Mexico received more than minimum wage monthly (8179/10,385, 78.8% and 3738/4729, 79%, respectively) compared with Peru (710/1395, 50.9%). In addition, more Peruvians (180/1194, 15.1%) perceived themselves to be at high risk of acquiring HIV compared with Brazilians (964/10,385, 9.3%) and Mexicans (439/4996, 8.8%); conversely, more Brazilians (6676/10,385, 64.3%) had high HIV sexual exposure according to HIRI scale compared with Mexicans (2965/4996, 59.4%) and Peruvians (738/1570, 47%). Additionally, a higher proportion of Peruvians (328/1550, 21.2%) had never been tested for HIV compared with Mexicans (757/4972, 15.2%) and Brazilians (876/10,385, 8.4%).

Regarding PrEP modalities, most of the respondents were willing to use monthly oral PrEP (12,621/16,951, 74.4%), daily oral PrEP (11,153/16,951, 65.8%), and bimonthly injectable PrEP (10,212/16,951, 60.2%), while only 38% (6442/16,951) were willing to use event-driven PrEP. Other PrEP modalities received lower willingness levels: 25.1% (4109/16,398) for daily topical, 25.6% (4202/16,398) for event-driven topical, 31.7% (5204/16,399) for a patch, 35.9% (5885/16,398) for an implant, 28.4% (4660/16,397) for infusion monoclonal antibodies, and 31.3% (5139/16,398) for subcutaneous monoclonal antibodies. Additionally, 82.8% (13,579/16,398) of participants reported willingness to use an HIV vaccine. Figure 2 shows the percentages of willingness to use these less common PrEP modalities and an HIV vaccine in each country.

We also found country-level differences in the willingness to use at least one of the 4 PrEP modalities evaluated (daily oral, event-driven oral, monthly oral, and bimonthly injectable). Mexicans reported the highest overall willingness (4449/4903, 90.7%), followed by Brazilians (8965/10,034, 89.4%) and Peruvians (1195/1423, 84%). Within each country, there were substantial differences in the willingness to use each of these 4 PrEP modalities. Peruvians were less willing to use daily oral (799/1506, 53.1%), monthly oral (830/1328, 62.5%), and bimonthly injectable (610/1328, 45.9%) compared with Brazilians (daily oral: 6082/9106, 66.8%; monthly oral: 7640/10,197, 74.9%; bimonthly injectable: 6450/10,197, 63.3%) and Mexicans (daily oral: 3242/4784, 67.8%; monthly oral: 3728/4875, 76.5%; bimonthly injectable: 2805/4875, 57.5%). Willingness to use event-driven PrEP was lower for Brazilians (3162/8761, 36.1%).

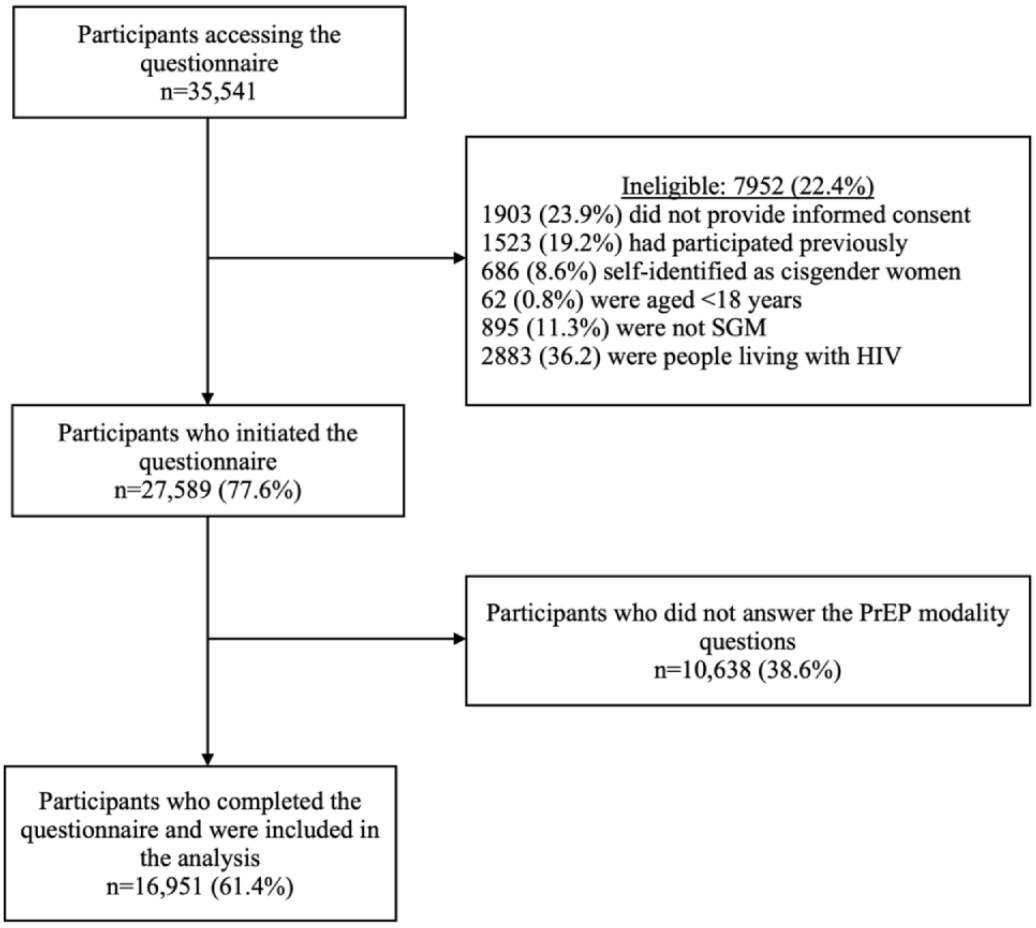
In all 3 countries, cisgender MSM were more willing to use monthly oral PrEP compared with transgender or nonbinary participants (Table 2). In Brazil (4376/5780, 75.7%) and Mexico (619/788, 78.6%), White individuals were more willing to use monthly oral PrEP, while in Peru (610/951, 64.1%), mixed-race individuals showed higher willingness. In Mexico (469/788, 59.5%) and Peru (80/168, 47.6%), White participants were more willing to use bimonthly injectable PrEP, whereas in Brazil (884/6450, 64.7%), Black, Asian, or Indigenous individuals showed greater willingness. In Brazil and Mexico, daily oral PrEP was more preferred among individuals with lower education (2165/6082, 68.7% and 983/1429, 68.8%, respectively) and lower income (1368/2032, 67.3% and 664/969, 68.5%), while in Peru, it was more preferred among participants with higher education (519/958, 54.2%) and income (378/671, 56.3%). In Brazil and Mexico, participants who had an HIV test within the past 6 months showed greater willingness to use monthly oral PrEP (4148/5462, 75.9% and 1576/2010, 78.4%, respectively), while in Peru, this modality was more preferred by people who had an HIV test more than 6 months ago (313/491, 63.8%).

In the multivariable analysis (Table 3), willingness to use each of the 4 evaluated PrEP modalities was positively associated with high HIV self-perceived risk and higher HIRI scores. Peruvian respondents showed lower willingness to use 3 modalities (daily oral: aPR=0.81, 95% CI 0.75-0.87; monthly oral: aPR=0.84, 95% CI 0.78-0.91; and injectable: aPR=0.76, 95% CI 0.69-0.83). Mexican and Peruvian respondents reported higher willingness to use event-driven oral PrEP (aPR=1.14, 95% CI 1.08-1.22 and aPR=1.11, 95% CI 1.01-1.22,

XSL•FO
RenderX

respectively) compared with Brazilian respondents. Having secondary education or less was associated with lower willingness to use monthly oral and bimonthly injectable PrEP (aPR=0.93, 95% CI 0.89-0.97 and aPR=0.94, 95% CI 0.90-0.99, respectively). Having a minimum wage or less was associated with lower willingness for bimonthly injectable PrEP (aPR=0.95, 95% CI 0.90-1.00). Never having been tested for HIV and testing for more than 6 months ago were associated with lower willingness to use daily oral PrEP (aPR=0.89, 95% CI 0.83-0.95 and aPR=0.95, 95% CI 0.91-0.99, respectively) and bimonthly injectable PrEP (aPR=0.80, 95% CI 0.74-0.86 and aPR=0.90, 95% CI 0.86-0.94, respectively).

Figure 1. Study flowchart. PrEP: pre-exposure prophylaxis.



XSL•FO
RenderX

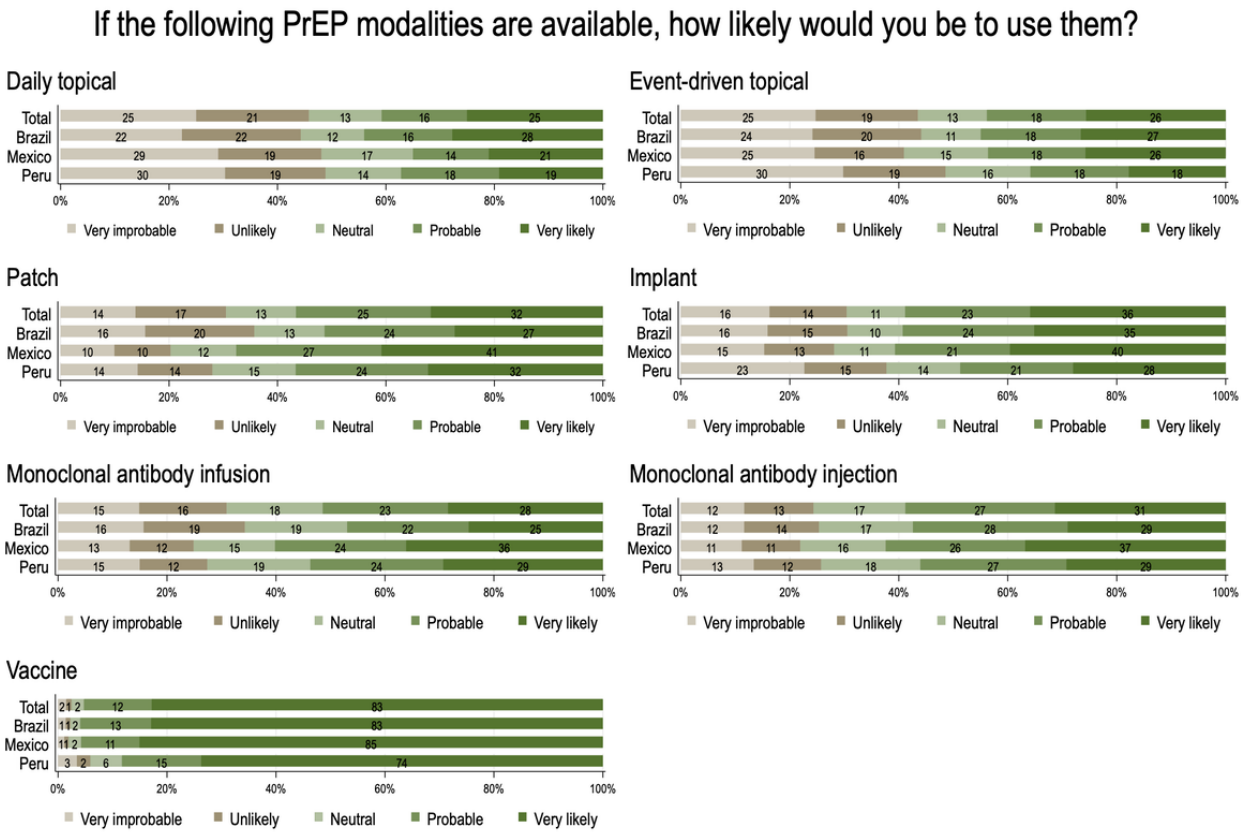
Table 1. Characteristics of an online sample of sexual and gender minority individuals (N=16,951) in Brazil (n=10,385, 61.3%), Mexico (n=4996, 29.5%), and Peru (n=1570, 9.3%) in 2021.

Characteristics	Total (n=16,951), n (%)	Brazil (n=10,385), n (%)	Mexico (n=4996), n (%)	Peru (n=1570), n (%)	*P* value
Gender identity (n=16,950)					<.001
Cisgender men	16,183 (95.5)	10,160 (97.8)	4617 (92.4)	1406 (89.5)	
Transgender or nonbinary persons	767 (4.5)	224 (2.2)	379 (7.6)	164 (10.5)	
Age (years; n=16,951)					<.001
18-24	3077 (18.1)	1304 (12.6)	1057 (21.2)	716 (45.6)	
25-30	4604 (27.2)	2712 (26.1)	1456 (29.1)	436 (27.8)	
>30	9270 (54.7)	6369 (61.3)	2483 (49.7)	418 (26.6)	
Race (n=16,285)					<.001
White	6879 (42.2)	5873 (56.5)	808 (18.7)	198 (12.6)	
Mixed race (*pardo* or *mestizo*)	7483 (46)	3125 (30.1)	3249 (75)	1109 (70.6)	
Black, Asian, or Indigenous	1923 (11.8)	1387 (13.4)	273 (6.3)	263 (16.8)	
Education level (n=16,951)					<.001
Less than or equal to secondary education	5455 (32.2)	3419 (32.9)	1474 (29.5)	562 (35.8)	
More than secondary education	11,496 (67.8)	6966 (67.1)	3522 (70.5)	1008 (64.2)	
Monthly income (n=16,509)					<.001
Less than or equal to 1 minimum wage	3882 (23.5)	2206 (21.2)	991 (21)	685 (49.1)	
More than 1 minimum wage	12,627 (76.5)	8179 (78.8)	3738 (79)	710 (50.9)	
HIV risk perception (n=16,575)					<.001
None, low, or moderate	14,992 (90.5)	9421 (90.7)	4557 (91.2)	1014 (84.9)	
High	1583 (9.5)	964 (9.3)	439 (8.8)	180 (15.1)	
HIRI[a] (n=16,951)					<.001
Low risk	6572 (38.8)	3709 (35.7)	2031 (40.6)	832 (53)	
High risk	10,379 (61.2)	6676 (64.3)	2965 (59.4)	738 (47)	
Last HIV test (n=16,907)					<.001
Never	1961 (11.6)	876 (8.4)	757 (15.2)	328 (21.1)	
Less than or equal to 6 months	8258 (48.8)	5543 (53.4)	2055 (41.3)	660 (42.6)	
More than 6 months	6688 (39.6)	3966 (38.2)	2160 (43.4)	562 (36.3)	

[a]HIRI: HIV Incidence Risk Index.

XSL•FO
RenderX

Figure 2. Willingness to use pre-exposure prophylaxis (PrEP) modalities and an HIV vaccine among sexual and gender minority individuals in Brazil, Mexico, and Peru.



XSL•FO
RenderX

Table 2. Characteristics of sexual and gender minority individuals willing to use each pre-exposure prophylaxis (PrEP) modality by country, 2021 (N=16,951).

Characteristics	Brazil (PrEP modality)				Mexico (PrEP modality)				Peru (PrEP modality)			
	Daily, n (%)	ED[a], n (%)	Monthly, n (%)	Inject n, (%)	Daily, n (%)	ED, n (%)	Monthly, n (%)	Inject, (%)	Daily, n (%)	ED, n (%)	Monthly, n (%)	Inject, n (%)
Willingness to use												
Yes	6082 (66.8)	3162 (36.1)	7640 (74.9)	6450 (63.3)	3242 (67.8)	1958 (41)	3728 (76.5)	2805 (57.5)	799 (53.1)	584 (39.9)	830 (62.5)	610 (45.9)
Gender identity												
Cisgender men	5953 (66.9)	3088 (36.1)	7477 (75)	6316 (63.3)	2995 (67.8)	1812 (41.1)	3473 (77.2)	2603 (57.8)	723 (53.6)	531 (40.1)	761 (63.3)	557 (46.3)
Transgender or nonbinary	129 (62.6)	74 (37.2)	163 (73.8)	134 (60.6)	247 (66.9)	146 (40)	255 (68)	202 (53.9)	76 (48.1)	53 (37.6)	69 (55.2)	53 (42.4)
Age (years)												
18-24	817 (66.8)	401 (34.1)	947 (74)	763 (59.6)	724 (70.2)	431 (43.3)	783 (76.2)	533 (51.9)	359 (51.4)	270 (40.1)	377 (62.2)	252 (41.6)
25-30	1689 (70.1)	876 (37.8)	2017 (75.9)	1706 (64.2)	1008 (70.9)	599 (42.5)	1098 (77.1)	855 (60)	217 (52.7)	161 (39.7)	240 (63)	194 (50.9)
>30	3576 (65.3)	1885 (35.8)	4676 (74.7)	3981 (63.6)	1510 (64.7)	928 (39.1)	1847 (76.2)	1417 (58.5)	223 (56.5)	153 (39.8)	213 (62.5)	164 (48.1)
Race												
White	3296 (64.6)	1755 (35.6)	4376 (75.7)	3609 (62.4)	510 (67.1)	327 (42.3)	619 (78.6)	469 (59.5)	92 (48.2)	78 (42.4)	98 (58.3)	80 (47.6)
Mixed race (*pardo or mestizo*)	1915 (69.1)	946 (35.8)	2275 (74.6)	1957 (64.2)	2107 (67.6)	1280 (41.1)	2456 (77.5)	1872 (59.1)	574 (54)	423 (40.4)	610 (64.1)	435 (45.7)
Black, Asian, or Indigenous	871 (70.9)	461 (38.7)	989 (72.4)	884 (64.7)	186 (71.5)	126 (47.6)	193 (71.8)	138 (51.3)	133 (53)	83 (35.6)	122 (58.4)	95 (45.5)
Education level												
Secondary education or less	2165 (68.7)	1057 (35.3)	2368 (71)	1976 (59.3)	983 (68.8)	587 (42)	1054 (73.2)	784 (54.4)	280 (51.1)	223 (43.1)	286 (59.3)	204 (42.3)
More than secondary education	3917 (65.8)	2105 (36.5)	5272 (76.8)	4474 (65.2)	2259 (67.3)	1371 (40.6)	2674 (77.9)	2021 (58.8)	519 (54.2)	361 (38.2)	544 (64.3)	406 (48)
Monthly income												
Minimum wage or less	1368 (67.3)	708 (36.7)	1543 (71.7)	1256 (58.3)	664 (68.5)	401 (42.6)	721 (74.6)	520 (53.8)	340 (51.1)	255 (39.9)	336 (58.7)	249 (43.5)
More than minimum wage	4714 (66.6)	2454 (35.9)	6097 (75.8)	5194 (64.6)	2428 (68.2)	1462 (40.8)	2810 (77)	2153 (59)	378 (56.3)	274 (41.1)	409 (66.8)	312 (51)
HIV risk perception												
None, low, or moderate	5345 (64.9)	2809 (35.4)	6851 (74.1)	5730 (62)	2887 (66)	1734 (39.8)	3364 (75.7)	2501 (56.3)	517 (52.8)	409 (40.3)	633 (62.4)	454 (44.8)
High	737 (84.9)	353 (42.3)	789 (83.2)	720 (76)	355 (87)	224 (53)	364 (84.9)	304 (70.9)	131 (79.4)	90 (50)	125 (69.4)	99 (55)
HIRI[b]												
Low risk	1972 (56.8)	1114 (33.7)	2523 (69.7)	1950 (53.9)	1190 (59.8)	746 (38.6)	1423 (72)	1018 (51.5)	367 (45.8)	277 (38.2)	355 (60.2)	249 (42.2)
High risk	4110 (72.9)	2048 (37.5)	5117 (77.8)	4500 (68.4)	2052 (73.4)	1212 (42.6)	2305 (79.5)	1787 (61.7)	432 (61.4)	307 (41.6)	475 (64.4)	361 (48.9)
Last HIV test												
Never	545 (62.7)	280 (34.7)	597 (70.8)	443 (52.6)	470 (62.7)	272 (38.6)	537 (74.1)	350 (48.3)	130 (40.6)	108 (37.5)	143 (58.6)	80 (32.8)

XSL•FO
RenderX

Characteristics	Brazil (PrEP modality)				Mexico (PrEP modality)				Peru (PrEP modality)			
	Daily, n (%)	ED[a], n (%)	Monthly, n (%)	Inject n, (%)	Daily, n (%)	ED, n (%)	Monthly, n (%)	Inject, (%)	Daily, n (%)	ED, n (%)	Monthly, n (%)	Inject, n (%)
6 months or less	2983 (69.2)	1544 (37.1)	4148 (75.9)	3684 (67.5)	1363 (73.1)	851 (43)	1576 (78.4)	1299 (64.6)	362 (59.2)	262 (41.7)	371 (63.6)	295 (50.6)
More than 6 months	2554 (65)	1338 (35.2)	2895 (74.4)	2323 (59.7)	1398 (65.1)	830 (40)	1600 (75.5)	1144 (54)	298 (53.8)	207 (38.9)	313 (63.8)	232 (47.3)

[a]ED: event-driven.

[b]HIRI: HIV Incidence Risk Index.

Table 3. Factors associated with willingness to use each pre-exposure prophylaxis modality among sexual and gender minorities in Brazil, Mexico, and Peru in 2021.

Characteristics	Daily oral, aPR[a] (95% CI)	Event-driven, aPR (95% CI)	Monthly oral, aPR (95% CI)	Injectable, aPR (95% CI)
Country				
Brazil	Reference	Reference	Reference	Reference
Mexico	1.02 (0.97-1.06)	*1.14 (1.08-1.22)* [b]	1.02 (0.98-1.07)	*0.93 (0.89-0.98)*
Peru	*0.81 (0.75-0.87)*	*1.11 (1.01-1.22)*	*0.84 (0.78-0.91)*	*0.76 (0.69-0.83)*
Age (years)				
18-24	1.01 (0.96-1.07)	1.02 (0.94-1.10)	1.04 (0.98-1.11)	0.95 (0.89-1.01)
25-30	1.03 (0.99-1.08)	1.04 (0.98-1.11)	1.01 (0.97-1.06)	1.00 (0.95-1.05)
>30	Reference	Reference	Reference	Reference
Race				
White	Reference	Reference	Reference	Reference
Mixed (Black and *pardo*)	1.04 (0.99-1.09)	0.99 (0.93-1.05)	0.99 (0.95-1.04)	1.02 (0.98-1.07)
Black, Asian, or Indigenous	1.07 (1.00-1.14)	1.05 (0.96-1.14)	0.95 (0.89-1.01)	1.01 (0.95-1.08)
Education level				
Secondary education or less	1.02 (0.98-1.07)	0.98 (0.92-1.04)	*0.93 (0.89-0.97)*	*0.94 (0.90-0.99)*
More than secondary education	Reference	Reference	Reference	Reference
Monthly income				
Minimum wage or less	0.98 (0.94-1.03)	1.02 (0.95-1.09)	0.96 (0.92-1.01)	*0.95 (0.90-1.00)*
More than minimum wage	Reference	Reference	Reference	Reference
HIV risk perception				
None, low, or moderate	Reference	Reference	Reference	Reference
High	*1.25 (1.18-1.32)*	*1.21 (1.12-1.32)*	*1.10 (1.04-1.17)*	*1.17 (1.10-1.25)*
HIRI[c]				
Low risk	Reference	Reference	Reference	Reference
High risk	*1.22 (1.17-1.27)*	*1.08 (1.02-1.14)*	*1.10 (1.06-1.14)*	*1.21 (1.16-1.26)*
Last HIV test				
Never	*0.89 (0.83-0.95)*	0.92 (0.85-1.00)	0.97 (0.91-1.03)	*0.80 (0.74-0.86)*
6 months or less	Reference	Reference	Reference	Reference
More than 6 months	*0.95 (0.91-0.99)*	0.95 (0.90-1.01)	0.99 (0.96-1.03)	*0.90 (0.86-0.94)*

[a]aPR: adjusted prevalence ratio.

[b]Italicization indicates statistical significance ($P<.05$).

[c]HIRI: HIV Incidence Risk Index.

XSL•FO
RenderX

Discussion

Principal Findings

In this paper, we provide information about willingness to use PrEP modalities and an HIV vaccine among SGM individuals from 3 countries in Latin America. Overall, participants were more interested in monthly oral PrEP, with daily oral and bimonthly injectable PrEP also being widely accepted and event-driven oral PrEP being the least preferred. Most participants reported willingness to use an HIV vaccine. Participants perceiving themselves at higher HIV risk and engaging in higher HIV sexual exposure were more willing to use all 4 PrEP modalities. Individuals with lower education were less willing to use long-acting PrEP (monthly oral and bimonthly injectable). Participants who had never tested for HIV or had been tested more than 6 months ago were less willing to use daily oral PrEP and bimonthly injectable. Our study provides insight into PrEP preferences among SGM individuals in Latin America and can guide the implementation of new PrEP modalities in prevention programs.

Higher willingness to use monthly oral and bimonthly injectable PrEP indicates a preference for long-acting formulations among SGM individuals in Latin America. In a discrete-choice experiment conducted in Brazil, SGM individuals preferred long-acting PrEP requiring less frequent dosing (monthly, bimonthly, or annually) as long as efficacy and side effects were similar or lower than those of oral PrEP [30]. Additionally, participants did not show a stronger preference for annual dosing compared with monthly or bimonthly dosing [30]. Previous studies identified injectable PrEP as the most preferred modality; however, they did not consider monthly oral PrEP as an option [31-33]. Currently, monthly oral PrEP is not available, although MK-8527, a novel investigational nucleoside reverse transcriptase translocation inhibitor, is under evaluation as an alternative to daily oral PrEP [34]. Two phase 3 clinical trials (EXPrESSIVE 10 and 11) are investigating the efficacy, safety, and tolerability of MK-8527 monthly compared with daily oral PrEP (FTC/TDF) among cisgender women and SGM individuals in several regions, including Latin America [35].

Long-acting injectables have been considered highly acceptable among SGM populations for HIV prevention [36]. This is corroborated by findings from a survey conducted among 3665 Brazilian SGM individuals in 2024 [37]. In the ImPrEP CAB Brasil study, most participants (83%) chose CAB-LA over daily oral PrEP [38]. Long-acting injectables offer potential advantages, such as reducing the frequency of clinic visits and improving adherence compared with daily oral PrEP [39]. Our study also revealed that individuals with lower education were less willing to use long-acting PrEP formulations, similar to findings related to injectable PrEP in Brazil [31]. This suggests that people using PrEP need accessible information to better understand the important features of these modalities.

Daily oral PrEP was also widely accepted; however, Peruvians were less likely to be willing to use this modality compared with Brazilians and Mexicans. This result is consistent with results from the ImPrEP study, which reported lower persistence and adherence to daily oral PrEP among MSM and transgender women from Peru compared with Brazil [11,12]. Possible contributing factors included low awareness of oral PrEP, HIV-related stigma during sexually transmitted infection clinic visits, and concerns about medication efficacy and side effects [27]. These concerns may be more noticeable with newer PrEP modalities, particularly long-acting formulations, due to the lack of information and perceived loss of autonomy in administration, which may contribute to a preference for the daily oral modality [33,40]. Additionally, familiarity with and routine use of a daily oral pill could also have contributed to this preference [40].

Differences among countries may also reflect structural and contextual barriers. Brazil and Mexico incorporated daily oral PrEP into their public health systems in 2017 and 2021, respectively, and have since implemented large-scale, community-supported programs that promote familiarity with and confidence in PrEP [8,41]. In contrast, the Peruvian Ministry of Health initiated offering oral PrEP to populations at high vulnerability for HIV in 2023 [9]. This delayed implementation, with limited-service availability, may have reduced visibility and accessibility among potential users. Strengthening PrEP implementation in Peru through decentralized service delivery, stigma reduction initiatives, and targeted communication campaigns could help improve uptake and promote more equitable access across the region.

Event-driven oral PrEP was the least acceptable among our participants. This finding aligns with previous studies [31,32], except for a study conducted in the United States, in which MSM reported that most of condomless anal sex events are either infrequent or can be anticipated [42]. In contrast, participants in the ImPrEP study showed low interest in switching from daily oral to event-driven oral PrEP [43]. Reasons included that fewer than 25% reported having sex less than 2 days per week (indicating infrequent sex), most were satisfied with the daily regimen, considered event-driven PrEP a difficult regimen to follow, and had concerns about its efficacy and anxiety about their own HIV risk [43].

Individuals who perceived themselves at high risk for HIV and who were engaging in high HIV sexual exposure were more likely to use any PrEP modality. A high perceived risk of HIV has been identified as a facilitator of PrEP acceptability and willingness to use it [44-46]. For example, in Mexico, transgender women reported a high willingness to use daily oral PrEP (95.5%) if they had high HIV risk perception [47]. Additionally, Torres et al [31] reported that higher HIRI scores increased the willingness of use injectable PrEP, while lower HIRI scores were associated with a preference for event-driven PrEP in Brazil and Mexico. In another Latin American study, Assaf et al [48] found that a higher risk for HIV was associated with PrEP awareness in Brazil but not among MSM in Mexico and Peru. Importantly, a study from Brazil [27] found that PrEP awareness was associated with willingness to use it, indicating the importance of continuous education campaigns about HIV prevention, including current and future PrEP modalities.

PrEP modalities involving topical agents, patches, implants, and monoclonal antibodies received lower willingness to use compared with the oral and injectable modalities. In contrast,

XSL•FO
RenderX

HIV vaccines were widely accepted by participants. Many of these alternative PrEP formulations are still in the clinical research phase [20,49], while no HIV vaccine has been efficacious to date. Limited awareness of the options still under study could have generated distrust. However, previous study from South Africa have reported high acceptability of topical agents among different populations, including SGM individuals [49]. Although there is currently no approved HIV vaccine, the MOSAICO study (HPX3002/HVTN706) enrolled Latin American participants before being discontinued due to lack of efficacy [50]. Nevertheless, previous studies from Brazil found a high willingness to use a hypothetical effective HIV vaccine, even if it was not free of charge [51,52].

We acknowledge that this study has certain limitations. Participants were recruited through dating apps and social media platforms and accessed the online survey using smartphones or other internet-connected devices. This recruitment strategy was appropriate for reaching SGM communities but may have introduced selection bias by excluding individuals with limited digital access, meaning that our sample may not be representative of all SGM individuals in Latin America, Brazil, Mexico, or Peru. Additionally, the cross-sectional design of our study limited our ability to establish causality. Notably, when the data were collected in 2021, only daily oral PrEP was available in these countries. However, oral PrEP information was more widely disseminated in Brazil, where it had been included in their public health care system since 2017 [41]. Finally, all responses were based on participants' self-reports, and social desirability bias may have occurred; however, online anonymous data collection may have reduced this bias. It is also important to note that "willingness" reflects hypothetical acceptability and may not directly translate into actual intention or behavior, which should be considered when interpreting these findings.

Conclusions

We found that SGM individuals from Latin America were more willing to use long-acting PrEP, including monthly oral and bimonthly injectable PrEP, but daily oral PrEP was also highly accepted. Further research and education are needed to better understand and address the gaps in knowledge about prevention modalities. The availability of additional choices to better address the prevention needs of SGM populations could empower individuals to use these methods. As additional PrEP modalities are included in HIV prevention programs, the development of accessible tools and community-based strategies will be essential to support informed PrEP choices and ensure equitable implementation across the region.

Acknowledgments

The authors express their gratitude to all study participants for their time and participation in this study. HV-R, CG-D-B, and DD-S thank the National Council of Humanities, Sciences and Technologies for acknowledging them as members of the National Researchers System. These results were presented in part as a poster presentation at the AIDS 2022 Conference in Montreal, QC. Artificial intelligence was not used in any portion of the manuscript writing.

Funding

This project was made possible through Unitaid's funding and support. Unitaid accelerates access to innovative health products and supports their scale-up by countries and partners. Unitaid is a hosted partnership of the World Health Organization. TST was supported by the National Council of Technological and Scientific Development (CNPq; 304417/2025-4 and 405558/2025-2) and the Carlos Chagas Filho Foundation for Research Support in the State of Rio de Janeiro (FAPERJ; E-26/201.270/2022). BG was supported by CNPq (313265/2023-2) and FAPERJ (E.26/200.946/2022).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

References

1. 2025 global AIDS update — AIDS, crisis and the power to transform. UNAIDS. URL: https://www.unaids.org/en/resources/documents/2025/2025-global-aids-update [accessed 2025-10-25]
2. AIDSinfo. UNAIDS. URL: https://aidsinfo.unaids.org/ [accessed 2025-10-27]
3. UNAIDS. The Urgency of Now: AIDS at a Crossroads. New York, NY: United Nations; 2024.
4. Torres TS, Teixeira SL, Hoagland B, Konda KA, Derrico M, Moreira RI, et al. Recent HIV infection and annualized HIV incidence rates among sexual and gender minorities in Brazil and Peru (ImPrEP seroincidence study): a cross-sectional, multicenter study. Lancet Reg Health Am 2023 Dec 2;28:100642 [FREE Full text] [doi: 10.1016/j.lana.2023.100642] [Medline: 38076411]
5. Pre-exposure prophylaxis (PrEP). World Health Organization. URL: https://www.who.int/teams/global-hiv-hepatitis-and-stis-programmes/hiv/prevention/pre-exposure-prophylaxis [accessed 2025-10-27]

XSL•FO
RenderX

6. Global state of PrEP. World Health Organization. URL: https://www.who.int/groups/global-prep-network/global-state-of-prep [accessed 2025-10-29]
7. Murphy L, Bowra A, Adams E, Cabello R, Clark JL, Konda K, et al. PrEP policy implementation gaps and opportunities in Latin America and the Caribbean: a scoping review. Ther Adv Infect Dis 2023 Apr 17;10:20499361231164030 [FREE Full text] [doi: 10.1177/20499361231164030] [Medline: 37114192]
8. Treviño-Pérez SC, Vega-Yáñez A, Martínez-Abarca CI, Estrada-Zarazúa G, Pérez-Camargo LA, Borrayo-Sánchez G. [Medical care of people living with HIV in the Instituto Mexicano del Seguro Social]. Rev Med Inst Mex Seguro Soc 2022 Dec 19;60(Suppl 2):96-102 [FREE Full text] [Medline: 36796006]
9. Ministerial resolution no. 576-2023-MINSA. Plataforma del Estado Peruano. 2023 Jun 19. URL: https://www.gob.pe/institucion/minsa/normas-legales/4327001-576-2023-minsa [accessed 2023-08-28]
10. HIV situation in the Americas. Pan American Health Organization. URL: https://www.paho.org/en/hiv-situation-americas [accessed 2025-10-24]
11. Veloso VG, Cáceres CF, Hoagland B, Moreira RI, Vega-Ramírez H, Konda KA, et al. Same-day initiation of oral pre-exposure prophylaxis among gay, bisexual, and other cisgender men who have sex with men and transgender women in Brazil, Mexico, and Peru (ImPrEP): a prospective, single-arm, open-label, multicentre implementation study. Lancet HIV 2023 Feb;10(2):e84-e96 [FREE Full text] [doi: 10.1016/S2352-3018(22)00331-9] [Medline: 36565708]
12. Konda KA, Torres TS, Mariño G, Ramos A, Moreira RI, Leite IC, et al. Factors associated with long-term HIV pre-exposure prophylaxis engagement and adherence among transgender women in Brazil, Mexico and Peru: results from the ImPrEP study. J Int AIDS Soc 2022 Oct;25 Suppl 5(Suppl 5):e25974 [FREE Full text] [doi: 10.1002/jia2.25974] [Medline: 36225148]
13. Caceres CF, Pines H, Konda KA, Borquez A, Moreira R, Leite I, et al. Baseline variables associated with subsequent HIV seroconversion among gay, bisexual and other men who have sex with men and transgender women: a prospective, multicenter PrEP demonstration study (ImPrEP). Lancet Reg Health Am 2025 Jun;46:101098 [FREE Full text] [doi: 10.1016/j.lana.2025.101098] [Medline: 40337320]
14. Painel PrEP. Ministério da Saúde do Brasil. 2022 Aug 23. URL: https://www.gov.br/aids/pt-br/indicadores-epidemiologicos/painel-de-monitoramento/painel-prep [accessed 2025-10-27]
15. Delany-Moretlwe S, Hughes JP, Bock P, Ouma SG, Hunidzarira P, Kalonji D, et al. Cabotegravir for the prevention of HIV-1 in women: results from HPTN 084, a phase 3, randomised clinical trial. Lancet 2022 May 07;399(10337):1779-1789 [FREE Full text] [doi: 10.1016/S0140-6736(22)00538-4] [Medline: 35378077]
16. Landovitz RJ, Donnell D, Clement ME, Hanscom B, Cottle L, Coelho L, et al. Cabotegravir for HIV prevention in cisgender men and transgender women. N Engl J Med 2021 Aug 12;385(7):595-608 [FREE Full text] [doi: 10.1056/NEJMoa2101016] [Medline: 34379922]
17. Bekker LG, Das M, Abdool Karim Q, Ahmed K, Batting J, Brumskine W, et al. Twice-yearly lenacapavir or daily F/TAF for HIV prevention in cisgender women. N Engl J Med 2024 Oct 03;391(13):1179-1192. [doi: 10.1056/NEJMoa2407001] [Medline: 39046157]
18. Gilead. Gilead's twice-yearly lenacapavir for HIV prevention reduced HIV infections by 96% and demonstrated superiority to daily Truvada® in second pivotal phase 3 trial. Gilead. 2024 Sep 12. URL: https://tinyurl.com/mvp89h7d [accessed 2024-10-27]
19. Guidelines on long-acting injectable cabotegravir for HIV prevention. World Health Organization. 2022 Jul 28. URL: https://www.who.int/publications-detail-redirect/9789240054097 [accessed 2024-06-11]
20. Coelho LE, Torres TS, Veloso VG, Landovitz RJ, Grinsztejn B. Pre-exposure prophylaxis 2.0: new drugs and technologies in the pipeline. Lancet HIV 2019 Nov;6(11):e788-e799. [doi: 10.1016/S2352-3018(19)30238-3] [Medline: 31558423]
21. Grinsztejn B, Torres TS, Hoagland B, Jalil EM, Moreira RI, O'Malley G, et al. Long-acting injectable cabotegravir for HIV preexposure prophylaxis among sexual and gender minorities: protocol for an implementation study. JMIR Public Health Surveill 2023 Apr 19;9:e44961 [FREE Full text] [doi: 10.2196/44961] [Medline: 37074775]
22. Pimenta MC, Torres TS, Hoagland B, Cohen M, Mann CG, Jalil CM, et al. Preparing for the implementation of long-acting injectable cabotegravir for HIV pre-exposure prophylaxis within the Brazilian public health system (ImPrEP CAB Brasil): qualitative study. JMIR Public Health Surveill 2024 Oct 24;10:e60961 [FREE Full text] [doi: 10.2196/60961] [Medline: 39446416]
23. Pimenta C, Mann CG, Hoagland B, Carvalheira E, Jalil C, Benedetti M, et al. Exploring perceptions and preferences for PrEP choice and of an mHealth intervention: insights from the ImPrEP CAB-Brasil study. J Int AIDS Soc 2025 Jul;28 Suppl 2(Suppl 2):e26493. [doi: 10.1002/jia2.26493] [Medline: 40600468]
24. Torres TS, Hoagland B. Emerging landscape for long-acting HIV prevention. Lancet HIV 2025 Aug;12(8):e538-e539. [doi: 10.1016/S2352-3018(25)00159-6] [Medline: 40587988]
25. Martinez O, Levine E, Munoz-Laboy M, Carballo-Diéguez A, Bauermeister JA, Chacon A, et al. More than just oral PrEP: exploring interest in rectal douche, dissolvable implant, removable implant and injection HIV prevention approaches among racially diverse men who have sex with men in the Northeast Corridor. BMJ Open 2022 Aug 18;12(8):e063474 [FREE Full text] [doi: 10.1136/bmjopen-2022-063474] [Medline: 35981775]
26. Matthews RP, Zang X, Barrett SE, Koynov A, Goodey A, Heimbach T, et al. A randomized, double-blind, placebo-controlled, phase 1 trial of radiopaque islatravir-eluting subdermal implants for pre-exposure prophylaxis against HIV-1 infection. J

XSL•FO
RenderX

Acquir Immune Defic Syndr 2023 Apr 01;92(4):310-316 [FREE Full text] [doi: 10.1097/QAI.0000000000003135] [Medline: 36450129]
27. Torres TS, Konda KA, Vega-Ramirez EH, Elorreaga OA, Diaz-Sosa D, Hoagland B, et al. Factors associated with willingness to use pre-exposure prophylaxis in Brazil, Mexico, and Peru: web-based survey among men who have sex with men. JMIR Public Health Surveill 2019 Jun 17;5(2):e13771 [FREE Full text] [doi: 10.2196/13771] [Medline: 31210143]
28. Torres TS, Marins LM, Veloso VG, Grinsztejn B, Luz PM. How heterogeneous are MSM from Brazilian cities? An analysis of sexual behavior and perceived risk and a description of trends in awareness and willingness to use pre-exposure prophylaxis. BMC Infect Dis 2019 Dec 19;19(1):1067 [FREE Full text] [doi: 10.1186/s12879-019-4704-x] [Medline: 31856746]
29. Konda KA, Qquellon J, Torres TS, Vega-Ramirez EH, Elorreaga O, Guillén-Díaz-Barriga C, et al. Awareness of U = U among sexual and gender minorities in Brazil, Mexico, and Peru: differences according to self-reported HIV status. AIDS Behav 2024 Jul;28(7):2391-2402 [FREE Full text] [doi: 10.1007/s10461-024-04336-9] [Medline: 38662277]
30. Pereira CC, Torres TS, Luz PM, Hoagland B, Farias A, Brito JD, et al. Preferences for pre-exposure prophylaxis (PrEP) among sexual and gender minorities: a discrete choice experiment in Brazil. Lancet Reg Health Am 2023 Mar;19:100432 [FREE Full text] [doi: 10.1016/j.lana.2023.100432] [Medline: 36950036]
31. Torres TS, Nascimento AR, Coelho LE, Konda KA, Vega-Ramirez EH, Elorreaga OA, et al. Preferences for PrEP modalities among gay, bisexual, and other men who have sex with men from Brazil, Mexico, and Peru: a cross-sectional study. Ther Adv Infect Dis 2023;10:20499361231153548 [FREE Full text] [doi: 10.1177/20499361231153548] [Medline: 36814515]
32. Chan C, Vaccher S, Fraser D, Grulich AE, Holt M, Zablotska-Manos I, et al. Preferences for current and future PrEP modalities among PrEP-experienced gay and bisexual men in Australia. AIDS Behav 2022 Jan;26(1):88-95. [doi: 10.1007/s10461-021-03344-3] [Medline: 34142274]
33. Parsons JT, Rendina HJ, Whitfield TH, Grov C. Familiarity with and preferences for oral and long-acting injectable HIV pre-exposure prophylaxis (PrEP) in a national sample of gay and bisexual men in the U.S. AIDS Behav 2016 Jul;20(7):1390-1399 [FREE Full text] [doi: 10.1007/s10461-016-1370-5] [Medline: 27000145]
34. Mayer K, Kotze P, Lombaard J, Caraco Y, Peer A, Poovan N, et al. Safety and pharmacokinetics of MK-8527 oral once-monthly: a phase 2 study in adults at low risk of HIV-1 exposure, oral. In: Proceedings of the 13th IAS Conference on HIV. 2025 Presented at: IAS 2025; July 13-17, 2025; Kigali, Rwanda URL: https://www.natap.org/2025/IAS/IAS_36.htm
35. MK-8527. PrEPWatch. URL: https://www.prepwatch.org/products/mk-8527/ [accessed 2025-10-24]
36. Tolley EE, Zangeneh SZ, Chau G, Eron J, Grinsztejn B, Humphries H, et al. Acceptability of long-acting injectable cabotegravir (CAB LA) in HIV-uninfected individuals: HPTN 077. AIDS Behav 2020 Sep;24(9):2520-2531 [FREE Full text] [doi: 10.1007/s10461-020-02808-2] [Medline: 32052214]
37. Torres T, Torres da Silva MS, Coutinho C, Hoagland B, Jalil EM, Cardoso SW, et al. Long-acting PrEP preference among brazilian sexual and gender minorities: key influencing factors. In: Proceedings of the Conference on Retroviruses and Opportunistic Infections. 2025 Presented at: CROI 2025; March 9-12, 2025; San Francisco, CA URL: https://www.croiconference.org/abstract/2473-2025/
38. Grinsztejn B, Coutinho C, Hoagland B, Farias AS, Madruga JV, Lima J, et al. PrEP choices among sexual and gender minorities in Brazil: the ImPrEP CAB-LA study. In: Proceedings of the Conference on Retroviruses and Opportunistic Infections. 2025 Presented at: CROI 2025; March 9-12, 2025; San Francisco, CA URL: https://www.croiconference.org/abstract/3411-2025/
39. Rogers BG, Chan PA, Sutten-Coats C, Zanowick-Marr A, Patel RR, Mena L, et al. Perspectives on long-acting formulations of pre-exposure prophylaxis (PrEP) among men who have sex with men who are non-adherent to daily oral PrEP in the United States. BMC Public Health 2023 Aug 28;23(1):1643 [FREE Full text] [doi: 10.1186/s12889-023-16382-4] [Medline: 37641018]
40. Jesus ML, Vargas MG, Iñiguez-Rueda L, Fernandez-Gonzalez E, Silva LA, Cabrera JM, et al. Users' expectations of the new pre-exposure prophylaxis for the HIV modalities. Cien Saude Colet 2024 Jun;29(6):e03042023 [FREE Full text] [doi: 10.1590/1413-81232024296.03042023] [Medline: 38896667]
41. Luz PM, Benzaken A, Alencar TM, Pimenta C, Veloso VG, Grinsztejn B. PrEP adopted by the Brazilian national health system: what is the size of the demand? Medicine (Baltimore) 2018 May;97(1S Suppl 1):S75-S77 [FREE Full text] [doi: 10.1097/MD.0000000000010602] [Medline: 29794605]
42. Hall EW, Heneine W, Sanchez T, Sineath RC, Sullivan P. Preexposure prophylaxis modality preferences among men who have sex with men and use social media in the United States. J Med Internet Res 2016 May 19;18(5):e111 [FREE Full text] [doi: 10.2196/jmir.5713] [Medline: 27199100]
43. Hoagland B, Torres T, Konda KA, Vega-Ramirez EH, Guanira JV, Vermandere H, et al. Awareness of ED-PrEP and interest in switching from daily oral PrEP to ED-PrEP in Brazil, Peru and Mexico ' the ImPrEP study. In: Proceedings of the IAS 2021. 2021 Presented at: IAS 2021; July 18-21, 2021; Berlin, Alemania URL: https://theprogramme.ias2021.org/Abstract/Abstract/1773
44. Sun Z, Gu Q, Dai Y, Zou H, Agins B, Chen Q, et al. Increasing awareness of HIV pre-exposure prophylaxis (PrEP) and willingness to use HIV PrEP among men who have sex with men: a systematic review and meta-analysis of global data. J Int AIDS Soc 2022 Mar;25(3):e25883 [FREE Full text] [doi: 10.1002/jia2.25883] [Medline: 35255193]

XSL•FO
RenderX

45. Soares F, Magno L, da Silva LA, Guimarães MD, Leal AF, Knauth D, et al. Perceived risk of HIV infection and acceptability of PrEP among men who have sex with men in Brazil. Arch Sex Behav 2023 Feb;52(2):773-782 [FREE Full text] [doi: 10.1007/s10508-022-02342-3] [Medline: 36169773]
46. Kesler MA, Kaul R, Myers T, Liu J, Loutfy M, Remis RS, et al. Perceived HIV risk, actual sexual HIV risk and willingness to take pre-exposure prophylaxis among men who have sex with men in Toronto, Canada. AIDS Care 2016 Nov;28(11):1378-1385. [doi: 10.1080/09540121.2016.1178703] [Medline: 27136725]
47. Guillen-Diaz-Barriga C, Diaz-Sosa D, Torres TS, Konda KA, Robles-Garcia R, Hoagland B, et al. HIV risk perception and pre-exposure prophylaxis (PrEP) awareness among transgender women from Mexico. AIDS Behav 2023 Mar;27(3):992-1002. [doi: 10.1007/s10461-022-03836-w] [Medline: 36121550]
48. Assaf RD, Konda KA, Torres TS, Vega-Ramirez EH, Elorreaga OA, Diaz-Sosa D, et al. Are men who have sex with men at higher risk for HIV in Latin America more aware of PrEP? PLoS One 2021;16(8):e0255557 [FREE Full text] [doi: 10.1371/journal.pone.0255557] [Medline: 34388155]
49. Atujuna M, Newman PA, Wallace M, Eluhu M, Rubincam C, Brown B, et al. Contexts of vulnerability and the acceptability of new biomedical HIV prevention technologies among key populations in South Africa: a qualitative study. PLoS One 2018;13(2):e0191251 [FREE Full text] [doi: 10.1371/journal.pone.0191251] [Medline: 29420549]
50. Crabtree Ramírez B, González Hernández LA, Cabrera C, Del Río C, González Rodríguez A, Sierra Madero J. Mexican perspective on the Mosaico HIV vaccine trial. Lancet HIV 2023 Jul;10(7):e426-e427. [doi: 10.1016/S2352-3018(23)00114-5] [Medline: 37216956]
51. Nogueira VG, Reis EA, Godman B, Martin AP, Godói IP. Acceptability and willingness to pay for a hypothetical HIV vaccine in Brazil and the implications: a cross-sectional study. Expert Rev Pharmacoecon Outcomes Res 2025 Jan;25(1):91-100 [FREE Full text] [doi: 10.1080/14737167.2024.2384543] [Medline: 39049463]
52. Torres TS, Luz PM, Giozza SP, Brito I, Pimenta C, Grinsztejn B, et al. Attitudes and beliefs regarding an HIV preventive vaccine among sexual and gender minorities in Brazil. In: Proceedings of the AIDS 2022. 2022 Presented at: AIDS 2022 Conference; July 29-August 2, 2022; Montreal, QC URL: https://programme.aids2022.org/Abstract/Abstract/?abstractid=5775

Abbreviations

aPR: adjusted prevalence ratio
CAB-LA: long-acting injectable cabotegravir
FTC/TDF: emtricitabine 200 mg and tenofovir disoproxil fumarate 300 mg
HIRI: HIV Incidence Risk Index
LEN: long-acting injectable lenacapavir
MSM: men who have sex with men
PrEP: pre-exposure prophylaxis
SGM: sexual and gender minority
WHO: World Health Organization

Edited by S Brini, A Mavragani; submitted 12.Apr.2025; peer-reviewed by DAR Silva, H Wang, L Garcia-Garcia; comments to author 03.Oct.2025; revised version received 24.Dec.2025; accepted 29.Dec.2025; published 09.Feb.2026.

Please cite as:
Qquellon J, Konda KA, Elorreaga O, Vega-Ramirez H, Guillén-Díaz-Barriga C, Díaz-Sosa D, Hoagland B, Guanira JV, Benedetti M, Pimenta C, Grinsztejn B, Caceres CF, Veloso VG, Torres TS
Profiles of Willingness to Use Pre-Exposure Prophylaxis Modalities and an HIV Vaccine Among Sexual and Gender Minority Individuals in Brazil, Mexico, and Peru: Cross-Sectional Online Survey
JMIR Public Health Surveill 2026;12:e75753
URL: https://publichealth.jmir.org/2026/1/e75753
doi:10.2196/75753
PMID:

©Jazmin Qquellon, Kelika A Konda, Oliver Elorreaga, Hamid Vega-Ramirez, Centli Guillén-Díaz-Barriga, Dulce Díaz-Sosa, Brenda Hoagland, Juan V Guanira, Marcos Benedetti, Cristina Pimenta, Beatriz Grinsztejn, Carlos F Caceres, Valdilea G Veloso, Thiago S Torres. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 09.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic

XSL•FO
RenderX

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

XSL•FO
RenderX

Original Paper

Impact of a Culturally Tailored Educational Intervention on Human Papillomavirus Knowledge and Vaccination Intentions Among Women in Saudi Arabia: Quasi-Experimental Study

Ebtesam A Elsayed[1,2], PhD; Fathia Ahmed Mersal[3], PhD; Heba Ahmed Osman Mohamed[4], PhD; Ala M Aljehani[5], PhD; Najla M Aljehani[2], PhD

[1]Family and Community Health Nursing Department, Faculty of Nursing, Ain Shams University, Cairo, Egypt
[2]Public Health Department, College of Health Sciences, Saudi Electronic University, Riyadh, Saudi Arabia
[3]Public Health Nursing Department, College of Nursing, Northern Border University, Arar, Saudi Arabia
[4]Maternal and Child Health Nursing Department, College of Nursing, Northern Border University, Arar, Saudi Arabia
[5]Department of Pathology, College of Medicine, Imam Mohammad Ibn Saud Islamic University, Riyadh, Saudi Arabia

Corresponding Author:
Ebtesam A Elsayed, PhD
Public Health Department
College of Health Sciences
Saudi Electronic University
3557, Mohammed Bin Abdulaziz Al Dughaithir St. Alsahafa District
Riyadh, 13321
Saudi Arabia
Phone: 966 505835391
Email: e.elsayed@seu.edu.sa

Abstract

Background: Human papillomavirus (HPV) is one of the most common sexually transmitted infections worldwide. Despite the availability of effective vaccines, uptake remains suboptimal in many settings, including Saudi Arabia, where cultural sensitivities, limited awareness, and vaccine hesitancy have been documented. Educational interventions grounded in behavioral theory, such as the information-motivation-behavioral skills (IMB) model, can help address informational and motivational barriers to HPV vaccination.

Objective: This study aimed to evaluate the impact of an educational intervention in improving HPV vaccination decision-making among women in Northern Saudi Arabia.

Methods: A quasi-experimental design (pretest-posttest with a control group) was used among women attending primary health care centers in the Northern Border Region of Saudi Arabia. A total of 380 participants, 190 in the control group and 190 in the intervention group, completed surveys at baseline (T1), 1-week post intervention (T2), and a 3-month follow-up (T3), assessing HPV-related knowledge, health awareness, attitudes, and vaccination intentions. The educational intervention consisted of a culturally adapted, evidence-based, 45-minute interactive group session delivered in primary health care centers.

Results: Study retention was excellent (380/389, 97.7%). At the 3-month follow-up, intervention participants demonstrated medium to large effect size improvements compared with controls: knowledge (adjusted β=0.84, 95% CI 0.53-1.15; $P<.001$; Cohen $d=0.45$, medium effect), health awareness (β=0.88, 95% CI 0.53-1.23; $P<.001$; $d=0.44$, medium effect), attitude (β=0.70, 95% CI 0.42-0.98; $P<.001$; $d=0.38$, medium effect), and vaccination intentions (β=0.86, 95% CI 0.60-1.12; $P<.001$; $d=0.82$, large effect). Exploratory subgroup analyses found no significant effect modification by age, education, or health literacy (all interaction $P>.10$), suggesting consistent intervention effects across demographic characteristics.

Conclusions: This culturally tailored educational intervention was effective in improving HPV-related knowledge, attitudes, health awareness, and vaccination intentions among women in Northern Saudi Arabia. Given the well-documented intention-behavior gap in vaccination uptake, future research should incorporate verified vaccine initiation and completion outcomes, along with longer follow-up, to assess whether enhanced intentions translate into actual vaccine uptake.

(JMIR Public Health Surveill 2026;12:e104529) doi:10.2196/104529

XSL•FO
RenderX

KEYWORDS

culturally adapted intervention; HPV vaccination; HPV; human papillomavirus; IMB; information-motivation-behavioral skills; Saudi Arabia; sexually transmitted infection

Introduction

The Global Challenge of Human Papillomavirus as a Non-HIV Sexually Transmitted Infection in Women

Human papillomavirus (HPV) is the most prevalent non-HIV sexually transmitted infection (STI) globally, and most sexually active individuals will acquire HPV during their lifetime [1-3]. Among women, persistent infection with high-risk oncogenic HPV types is a necessary cause of cervical cancer. It contributes to a broader spectrum of anogenital and oropharyngeal cancers, creating a substantial preventable burden across the life course [3-5]. HPV vaccination offers a critical opportunity for primary prevention of an STI-related cancer, with high efficacy against vaccine-type infection and intraepithelial neoplasia when administered before exposure [5-7]. Among young adult women beyond routine childhood immunization schedules, HPV vaccination is often positioned as a preference-sensitive decision requiring individualized counseling, yet vaccine coverage remains insufficient in many settings [8,9].

Incomplete vaccine uptake reflects both individual-level barriers (knowledge, attitudes, and motivation) and structural barriers (access, cost, and provider recommendation). While system-level changes are essential, individual-level educational interventions remain necessary to address modifiable cognitive and motivational barriers, particularly in contexts where structural access exists but is underused [9,10]. Individual barriers include limited awareness and misinformation, perceived stigma associated with STI prevention, low confidence in vaccine safety, concerns about fertility and long-term harms, and the "intention-behavior gap," the well-documented phenomenon in which individuals express willingness yet do not translate intention into action [11-13]. These factors are particularly relevant for women's health in culturally conservative contexts where sexual health discussions carry social stigma.

HPV Vaccination Challenges in Saudi Arabia: Epidemiological Context and Cultural Barriers

In Saudi Arabia, cervical cancer incidence is comparatively low but nonnegligible: the Information Centre on HPV and Cancer (ICO)/International Agency for Research on Cancer (IARC) country fact sheet [14] estimates approximately 358 new cases and approximately 179 deaths annually, with an age-standardized incidence rate (ASR) of approximately 2.8 per 100,000 women and ranking eighth among cancers in women overall and in those aged 15 to 44 years. Recent registry-based analyses also show regional variation, with Makkah consistently reporting the highest rates, followed by Riyadh and the Eastern Region, and a favorable shift toward earlier stage at diagnosis between 2005 and 2019 (increase in localized-stage diagnoses; decline in unknown stage), underscoring geographic and stage heterogeneity relevant to prevention strategies [15,16]. The Saudi Cancer Registry (under the National Cancer Center, Saudi Health Council) continues to publish periodic incidence reports and program updates, improving data quality for planning and evaluation [17].

Regarding vaccine policy and access, the Saudi Ministry of Health (MOH) health-education pages recommend HPV vaccination for adolescents (girls and boys) at ages 11 to 12 years (can start at age 9 years) and catch-up through age 26 years, aligning with international guidance; adults aged 27 to 45 years may benefit on a case-by-case basis. Since 2022, public communications from MOH and independent evaluations have described school-based campaigns introducing HPV vaccination for early adolescents, with media notices highlighting the importance of early vaccination and a multicohort school-based rollout reported in 2022 (with subsequent years' activity). Importantly, official, centralized MOH fee schedules for HPV vaccine are not published; doses delivered within school-based/public programs are provided at no charge to families, whereas private-sector prices are not standardized or publicly listed by MOH/Saudi Food and Drug Authority (SFDA) [18].

Contemporary studies in Saudi Arabia consistently demonstrate low HPV vaccine uptake despite rising awareness. For example, a 2023-2024 national survey found that although 60.85% of Saudi women had heard of HPV, only 8.25% had received the vaccine, and more than 70% expressed willingness, highlighting a clear intention-behavior gap [19]. Similarly, recent cross-sectional research shows that Saudi women continue to report fear of side effects, doubts about vaccine effectiveness, and low positive perceptions, indicating persistent informational and motivational barriers to HPV prevention efforts [20]. Among younger populations, such as nursing and medical students, studies report poor HPV knowledge, limited understanding of screening, and low vaccination rates, with over 70% unvaccinated, despite relatively high awareness [21,22].

Although overall awareness has significantly increased following Saudi MoH campaigns, HPV-related behaviors remain suboptimal. A national comparison from 2022 to 2024 documented a rise in HPV awareness from 38.8% to 71.1%, yet cervical cancer screening uptake remained low, with only 25.8% of women reporting ever having a Pap smear [23]. Cultural and social dynamics further shape preventive behaviors: barriers such as perceived lack of necessity, abstinence-based reasoning, safety concerns, and misconceptions about who the vaccine is "appropriate" for continue to impede uptake [24]. In family-centered decision contexts, parental acceptance of HPV vaccination is strongly influenced by education level, health care worker status, and especially receiving a recommendation from the MoH or clinicians, demonstrating how social and institutional cues shape women's access to preventive care [24-26].

Shared decision-making (SDM) principles may require adaptation in this context, as women's health care decisions often involve family consultation and are influenced by clinician paternalism [27-29]. Educational interventions can support SDM

XSL•FO
RenderX

by strengthening women's knowledge and confidence to engage in vaccine discussions, ask questions, and express preferences to clinicians [30,31]. Therefore, this study aims to improve determinants of informed vaccination decisions, a foundational step toward SDM-compatible counseling.

Educational Interventions, Theoretical Frameworks, and Evidence Gaps

Educational interventions have repeatedly demonstrated effectiveness in improving HPV-related knowledge, attitudes, and vaccine acceptance [11,32]. However, many interventions lack cultural tailoring and do not address context-specific concerns such as fertility misconceptions, stigma, and family influence, and they often fail to incorporate behavioral theory to explain how information converts to lasting motivation and action [11,33,34]. Ensuring equitable reach requires evaluating whether interventions maintain effectiveness across differing levels of health literacy, as literacy-related disparities substantially affect the ability to process health information and may reduce the intervention's public health impact [35-37].

The information-motivation-behavioral skills (IMB) model proposes that behavior change is driven by accurate, actionable information, motivation (personal and social), and the behavioral skills needed to execute decisions, an especially relevant structure for vaccine uptake behaviors that require navigating appointments, consent, and follow-through [38,39]. IMB-based educational interventions have improved HPV vaccination-related outcomes in other contexts, including web-based education among female college students [40]. Yet, within Saudi Arabia and the broader Middle Eastern region, many studies remain cross-sectional, focus on university settings, or assess knowledge and hesitancy without testing scalable, culturally adapted, primary care–compatible intervention models [20,41]. Additionally, registry and international sources emphasize that although overall cervical cancer incidence is low, regional disparities and ongoing cases justify prevention efforts, particularly in high-population regions such as Makkah and Riyadh.

Study Rationale, Objectives, and Research Questions

This study addresses these gaps directly by evaluating a culturally tailored educational intervention to improve HPV vaccination decision-making among women in Northern Saudi Arabia. This study makes 3 distinct contributions. First, it addresses HPV vaccination as both an informational challenge and an SDM implementation challenge, where women's concerns, stigma, and family/social context must be addressed to support informed and values-concordant decisions [27,29,30]. Second, it applies a behavioral theory lens (IMB) to structure how education targets the specific mechanisms most plausibly linked to vaccine initiation behaviors [38,39]. Third, it focuses on a Saudi setting where published evidence indicates persistent gaps in knowledge and uptake and where culturally tailored prevention strategies are needed to reduce HPV-related disease burden among women [4,19,20]; contemporaneous national and international data further justify the focus, citing current case counts, regional variation, and active public health efforts.

Research Questions

Research question 1 assesses whether a culturally tailored, IMB-based educational intervention improves HPV-related knowledge, health awareness, attitudes, and vaccination intentions among women in Northern Saudi Arabia compared to standard care. Research question 2 assesses whether intervention effects are sustained at 3-month follow-up. Research question 3 assesses whether intervention effects vary by age, education level, health literacy, or prior vaccination experience.

Hypotheses

Hypothesis 1 states that intervention participants will demonstrate significantly greater improvements in all 4 outcome domains compared to controls at post-intervention and 3-month follow-up. Hypothesis 2 states that effects will be sustained (no significant decay) between postintervention and 3-month assessments. Hypothesis 3 states that intervention effects will not significantly differ across demographic subgroups (consistent effectiveness).

Methods

Study Design

A quasi-experimental design (pretest-posttest with a control group), with a pragmatic, nonrandomized allocation approach, was used due to operational constraints within the primary health care (PHC) system that precluded individual-level randomization, combined with concerns about contamination between intervention and control participants within the same clinic sessions. Outcomes were assessed at 3 time points: baseline (T1), approximately 1 week after the intervention session (T2), and 3 months after baseline (T3). The study is reported in accordance with the Transparent Reporting of Evaluations with Nonrandomized Designs (TREND) statement for nonrandomized evaluations of behavioral and public health interventions; a completed TREND checklist is provided in Multimedia Appendix 1. Causal language is used cautiously given the nonrandomized design, and effect estimates are interpreted as associations consistent with an intervention effect.

Participants and Setting

Eligibility Criteria

Women who were currently pregnant were excluded because HPV vaccination is typically deferred until the postpartum period and because pregnancy-related clinical contexts could have introduced heterogeneity in decision urgency and confounded intervention effects. Additional exclusion criteria were (1) health care professionals or trainees likely to have specialized knowledge of HPV vaccination (eg, nursing, medical, or public health students or staff involved in health services), (2) current participation in another vaccine-related study, (3) cognitive impairment limiting informed consent or questionnaire completion, and (4) completion of the full HPV vaccination series. Numbers of women screened, excluded, and enrolled are reported in the Participant Flow and Retention section, in line with reporting standards for nonrandomized studies.

XSL•FO
RenderX

Recruitment and Enrollment

The study was implemented between December 2025 and May 2026 in 4 PHCs in the Northern Border Region, Saudi Arabia. Women attending the participating PHCs for routine services were approached by trained female research staff, screened for eligibility, and invited to participate; written informed consent was obtained from all enrolled women. Enrollment counts are reported in the Participant Flow and Retention section.

Allocation Procedure (Nonrandom)

Allocation was nonrandom and based on clinic scheduling/time of attendance to maximize feasibility and minimize disruption to routine PHC operations. Women attending designated morning clinic sessions at participating sites were invited to join the intervention arm, whereas women attending afternoon clinic sessions served as the control arm.

Threats to Internal Validity

Nonrandom, time-based allocation introduces potential selection bias, as timing of clinic attendance can correlate with employment status, caregiving responsibilities, health-seeking motivation, and socioeconomic factors. We prospectively collected and adjusted for key sociodemographic and health covariates (Outcome Measures section) to address measured confounding, but unmeasured factors associated with clinic attendance timing (eg, health motivation, urgency of care-seeking) could confound intervention effects. To quantify the magnitude of potential bias, standardized differences were calculated for all baseline covariates (Table 1), and sensitivity analyses were conducted (Statistical Analysis section). Although standardized differences <.10 suggested minimal observed imbalance across measured characteristics, causal inference remains limited, and estimates are interpreted as associations consistent with an intervention effect rather than as definitive causal effects. To minimize contamination, intervention and control sessions were scheduled on different days when feasible, and participants were asked not to share educational materials with other clinic attendees. Clinic staff provided routine care to both groups without knowledge of study arm allocation to the extent possible in this pragmatic trial. Participants and facilitators were not blinded to group allocation, which is inherent to the delivery of an educational intervention in a nonrandomized design. T3 telephone interviewers were partially blinded, as described in the Data Collection Procedures section.

XSL•FO
RenderX

Table 1. Baseline demographic characteristics and potential confounders. Data are presented as mean (SD) for continuous variables and n (%) for categorical variables.

Characteristic	Intervention group (n=190)	Control group (n=190)	*P* value[a]	Standardized difference[b]
Demographics				
Age (years), mean (SD)	28.4 (6.8)	28.9 (7.1)	.47	.07
Age categories, n (%)			.85	.03
18-26	82 (43.2)	79 (41.6)		
27-35	78 (41.1)	81 (42.6)		
36-45	30 (15.8)	30 (15.8)		
Marital status, n (%)			.76	.03
Single	94 (49.5)	91 (47.9)		
Married	81 (42.6)	85 (44.7)		
Divorced/widowed	15 (7.9)	14 (7.4)		
Socioeconomic Factors				
Education level, n (%)			.71	.05
High school or less	38 (20.0)	42 (22.1)		
University degree	126 (66.3)	121 (63.7)		
Postgraduate	26 (13.7)	27 (14.2)		
Employment status, n (%)			.68	.04
Employed	112 (58.9)	108 (56.8)		
Unemployed/student	78 (41.1)	82 (43.2)		
Monthly household income (SAR[c]), n (%)			.89	.04
<10,000	52 (27.4)	49 (25.8)		
10,000-20,000	89 (46.8)	92 (48.4)		
>20,000	49 (25.8)	49 (25.8)		
Health-related factors				
Health literacy score (0-50), mean (SD)	34.2 (8.6)	33.8 (8.9)	.64	.05
Previous vaccination experience, n (%)	142 (74.7)	138 (72.6)	.65	.05
HPV[d] vaccination history (partial), n (%)	8 (4.2)	6 (3.2)	.60	.06
Family history of cervical cancer, n (%)	18 (9.5)	21 (11.1)	.61	.05
Regular gynecological check-ups, n (%)	87 (45.8)	83 (43.7)	.69	.04
Baseline outcome measures				
Knowledge score (1-7; transformed)[e], mean (SD)	2.97 (1.40)	2.94 (1.42)	.86	.02
Health awareness score (1-7), mean (SD)	3.05 (1.43)	2.93 (1.41)	.41	.08
Attitude score (1-7, transformed)[e], mean (SD)	2.92 (1.43)	2.91 (1.40)	.92	.01
Vaccination intentions (1-7), mean (SD)	3.99 (1.04)	3.98 (1.08)	.94	.01

[a]*P* values were derived from chi-square tests for categorical variables and 2-tailed independent-samples *t* tests for continuous variables.

[b]Standardized differences <.10 indicate minimal baseline imbalance.

[c]SAR 1=US $0.27 as of September 5, 2026.

[d]HPV: human papillomavirus.

[e]Knowledge and attitudes scores are linearly transformed from the original scales (0-9 and 1-5, respectively) to a 1-7 scale for comparability across outcomes.

XSL•FO
RenderX

Retention and Follow-Up

Participant flow, retention, and completeness of follow-up are reported in the Participant Flow and Retention section. Baseline characteristics were compared between participants who completed all assessments and those lost to follow-up; results of this comparison between participants who completed follow-up vs those who did not are provided in Multimedia Appendix 2. All 380 participants included in the analytic sample completed follow-up within the prespecified window; the final 3-month interview was completed on May 30, 2026, and no further follow-up assessments were undertaken after that date.

Sample Size

Sample size was estimated using G*Power 3.1 for repeated-measures ANOVA with within-between interaction. The primary test was the F test of the group × time interaction in a 2 (group) × 3 (time) mixed design, for which the test statistic follows a noncentral F distribution with numerator $df = (2 - 1)(3 - 1) = 2$ and denominator $df = 2(N - 2)$, with sphericity assumed (ε=1). Assuming a moderate effect size (F=0.25, equivalent to Cohen $d \approx 0.50$), an α of .05, power of 0.80, a total of 3 measurement time points, and a within-subject correlation of 0.5, the sample size required was 82 participants per group. Anticipating 20% attrition, we targeted 103 per group (n=206). Actual enrollment (n=389) substantially exceeded this target, providing adequate power for primary analyses.

Although the achieved sample provides sufficient power for main effects, the exploratory subgroup analyses (Statistical Analysis section) were not formally powered and should be interpreted as hypothesis-generating: interaction tests require approximately 4 times the sample size of main-effect tests, and with 190 participants per group, statistical power for detecting subgroup interactions was estimated at 40% to 50%.

Intervention Development and Implementation

Theoretical Framework

The intervention was grounded in the IMB model [30], which proposes that health behavior adoption is driven by (1) accurate and accessible information, (2) motivation (personal and social), and (3) behavioral skills needed to translate decisions into action. The IMB model has been applied to sexual health behaviors and preventive decision-making, including STI prevention [42,43].

Patient and Public Involvement

Cultural Adaptation

A structured cultural adaptation framework consisting of 4 phases was used, developed with input from the target population and stakeholders (patient and public involvement): (1) information gathering through focus groups with 32 women from the target population and interviews with 15 stakeholders, including PHC clinicians (n=8), community health workers (n=4), and women's health advocates (n=3), to identify context-specific barriers (eg, STI stigma, HPV/HIV confusion, fertility concerns, and preference for gender-concordant delivery); (2) preliminary adaptation and development of culturally congruent messaging and delivery strategies; (3) pilot testing with 20 women to refine clarity, acceptability, and session length, after which the session was reduced from 60 minutes to 45 minutes and terminology was revised to replace "sexually transmitted" with "virus that can cause cancer" in several sections to reduce stigma (pilot acceptability rating 4.6/5.0 for clarity and 4.8/5.0 for cultural appropriateness); and (4) refinement and finalization of materials and delivery protocols.

All intervention materials (session scripts, facilitator guides, and the Arabic pamphlet) were developed in Arabic by the bilingual study team and back-translated into English to verify content fidelity; discrepancies identified during back-translation were resolved by consensus before pilot testing.

Intervention Content and Delivery

The intervention consisted of a 45-minute interactive group session delivered in private rooms within PHCs to groups of 10 to 15 participants, facilitated by 2 trained female health educators.

Content was standardized and organized into four modules aligned with IMB constructs: (1) understanding HPV and cervical cancer (information; 15 minutes); (2) HPV vaccines: safety, efficacy, and access (information; 15 minutes); (3) making an informed vaccination decision (motivation and behavioral skills; 10 minutes); and (4) addressing misconceptions (information and motivation; 5 minutes). Key misconceptions explicitly addressed included confusion between HPV and HIV; beliefs that vaccination causes infertility; concerns that vaccination is not permissible (Halal); and the misunderstanding that HPV only affects sexually promiscuous individuals. Key cultural adaptations included framing HPV primarily as a cancer-preventing vaccination to reduce STI-related stigma, ensuring female-only facilitation, clarifying differences between HPV and HIV, addressing common fertility concerns, and explaining HPV risk in married and monogamous relationships.

Participants received an 8-page Arabic pamphlet written at approximately a sixth-grade reading level (assessed using the Arabic adaptation of the Simple Measure of Gobbledygook [SMOG] readability formula) summarizing key points. Complete session scripts, facilitator guides, and educational materials are provided in Multimedia Appendix 3.

Facilitator Training and Fidelity Monitoring

Educators completed a standardized 2-day training program (16 hours) covering IMB principles, HPV vaccine content, communication skills for sensitive topics, and fidelity procedures. To assess fidelity, 20% (8/40) of sessions were audio-recorded with additional participant consent and independently coded by 2 trained raters using a 15-item fidelity checklist. Interrater reliability for fidelity coding was substantial (Cohen κ=0.78, 95% CI 0.69-0.87). Mean fidelity was 93% (range 87%-100%), indicating high adherence. Sessions with fidelity <90% (n=3) triggered facilitator feedback and retraining.

Control Condition

Control participants received standard PHC care without structured HPV education. Controls completed the same outcome assessments at T1 to T3. After completion of the

XSL•FO
RenderX

3-month follow-up, control participants were offered delayed access to the educational session (148/190, 78% accepted).

Outcome Measures

Instrument Development and Translation

A structured questionnaire was developed by adapting items from previously published HPV knowledge/attitude instruments [44,45] and tailoring language for cultural appropriateness. The instrument underwent forward-backward translation (English-Arabic), expert panel review (n=5 content experts), and cognitive interviewing (n=8 women from the target population). Pilot testing (n=20) provided preliminary support for internal consistency (Cronbach α≥0.78) and 2-week test-retest stability (intraclass correlation coefficient [ICC] ≥0.76). While these values suggest acceptable measurement properties, the small pilot sample and lack of comprehensive validation (eg, confirmatory factor analysis and convergent/discriminant validity testing) preclude definitive conclusions about scale validity. We therefore interpret the adapted instruments as research tools requiring further validation in Saudi populations.

HPV Knowledge (Coprimary Outcome)

HPV knowledge was assessed using 9 "true," "false," or "do not know" items covering HPV transmission, cancer risk, vaccine efficacy, and safety. Items were scored 1 for correct and 0 for incorrect or "do not know" responses (range 0-9). Because items were dichotomously scored, internal consistency was estimated using the Kuder-Richardson formula 20 (KR-20), which is mathematically equivalent to Cronbach α for binary items; the reported value (KR-20=0.78) refers to this estimate. "Do not know" responses were coded as incorrect by a priori specification, consistent with the interpretation that they do not reflect accurate knowledge. For comparability across outcomes, raw knowledge scores were linearly transformed to a 1 to 7 scale using the formula: transformed score = 1 + (raw score / 9) × 6; with higher scores indicating greater knowledge.

HPV-Related Health Awareness (Coprimary Outcome)

HPV-related health awareness was measured using 5 items assessing awareness of personal risk, disease burden, and cancer prevention on 7-point Likert scales (1=not at all aware to 7=extremely aware). The mean score was computed, with higher scores indicating greater awareness (Cronbach α=0.81).

Attitudes Toward HPV Vaccination (Coprimary Outcome)

Attitudes toward HPV vaccination were measured using 8 items assessing beliefs about vaccine safety, effectiveness, necessity, and appropriateness on Likert scales (1=strongly disagree to 7=strongly agree). The mean score was computed, with higher scores indicating more favorable attitudes (Cronbach α=0.84). Raw attitude scores (range 1-5) were linearly transformed to the common 1 to 7 metric using the formula: transformed score = 1 + (raw score / 5) × 6.

HPV Vaccination Intentions (Coprimary Outcome)

HPV vaccination intentions were measured using 5 items assessing likelihood of vaccine uptake behaviors (seeking information, discussing with family, and scheduling vaccination) on 7-point likelihood scales (1=extremely unlikely to 7=extremely likely). The mean score was computed, with higher scores indicating stronger intentions (Cronbach α=0.88).

Justification for Score Transformation

Because the 4 outcomes were measured on different raw scales, all scores were placed on a common 1 to 7 metric to permit direct comparability of effect sizes across outcomes. Linear transformation preserves within-outcome rank order and does not alter inferential statistics; all primary models were reestimated using original-scale scores, with identical conclusions**.** All outcome items and scoring procedures are reported in Multimedia Appendix 4.

Data Collection Procedures

At baseline (T1), participants provided written informed consent and completed the Arabic questionnaire in a private PHC room (approximately 20 to 25 minutes). Trained female research assistants were available to clarify procedural questions without guiding responses.

At T2, a mean of 6.8 (SD 1.2) days after the intervention session (approximately 1 week after baseline), both groups completed the same questionnaire using identical paper-based, in-person administration procedures.

At T3, approximately 3 months postbaseline (mean 91.4, SD 8.6 days), follow-up data were collected via standardized telephone interviews conducted by trained female staff. T3 telephone interviewers were partially blinded: they did not know participant group allocation at the time of the call but could potentially infer group status if participants spontaneously mentioned attending an education session. To minimize detection bias, interviewers used scripted questions, avoided unstructured conversation, and were trained to redirect discussion to assessment items if participants volunteered unsolicited information about study participation.

To assess potential mode effects (in-person paper vs telephone), we conducted sensitivity analyses comparing T2 (paper) vs T3 (telephone) response patterns within the control group, which received no intervention between these time points. No significant differences were observed (all P>.15), suggesting minimal mode-related bias (Multimedia Appendix 5).

Statistical Analysis

All statistical analyses were performed using R software (version 4.3.1; R Core Team), following a prespecified analysis plan.

Multiple Testing Considerations

We specified 4 coprimary outcomes to comprehensively assess intervention effects across the IMB model's theoretical domains (information, motivation, and behavioral intention). Because these outcomes are conceptually and empirically correlated (r=0.42-0.68 at baseline) and represent a unified construct (vaccine readiness), we did not apply multiple testing corrections for primary analyses of time × group effects, consistent with recommendations for evaluating conceptually linked outcomes. However, for exploratory subgroup analyses involving 4 moderators tested across 4 outcomes, we applied Bonferroni

XSL•FO
RenderX

correction (α=.05/4=.0125 per outcome domain, testing each moderator separately).

Descriptive and Baseline Balance Analyses

Baseline characteristics were summarized using mean (SD) for continuous variables and n (%) for categorical variables. Between-group balance was assessed using 2-tailed independent-samples *t* tests, chi-square tests, and standardized differences (Cohen *d* for continuous variables; Cramer V for categorical variables), with standardized differences <.10 interpreted as excellent balance and indicative of minimal confounding from measured covariates.

Primary Models

For each outcome, linear mixed-effects models with participant-level random intercepts were fitted to account for repeated measures, including fixed effects for time (T1, T2, and T3), group (intervention vs control), and the time × group interaction. Three nested models were estimated: model 0 (unadjusted); model 1 (adjusted for age, education, and marital status); and model 2, the primary model (additionally adjusted for health literacy, previous vaccination experience, and family history of cervical cancer). The time × group interaction at T2 and T3 represents the primary intervention effect. Effect sizes were calculated as Cohen *d* using pooled change-score standard deviations (T3 − T1) as the primary metric, appropriate for repeated-measures designs**;** for comparability with previous studies, Cohen *d* using pooled baseline SDs is also reported in the tables.

Secondary and Robustness Analyses

We computed within-group change scores (T3 − T1) and between-group differences in change with 95% CIs to provide interpretable estimates alongside model-based results.

Exploratory subgroup analyses tested effect modification by age (18-26 years vs 27-45 years; cutoff reflects typical transition from younger adult to family-building life stages), education (university vs below university level), health literacy (inadequate/marginal vs adequate) using validated Brief Health Literacy Screen cutoffs [46], and prior vaccination experience (any vs none) using 3-way interaction terms (time × group × moderator). Each moderator was tested separately with Bonferroni-adjusted α=.0125 per outcome domain.

Sensitivity analyses included intention-to-treat using last observation carried forward (LOCF), per-protocol analysis (participants attending ≥80% of session content), multiple imputation (50 datasets; multivariate imputation by chained equations [MICE]), complete-case analysis (primary), alternative covariance structures (unstructured and autoregressive), and assessment of potential administration mode effects (telephone vs paper). Because the intervention was delivered in groups across multiple clinics and facilitators, sensitivity models additionally included random intercepts for clinic (4 levels) and intervention session (40 levels), with cluster-robust (sandwich) SEs; ICCs are reported in Multimedia Appendix 6. These analyses addressed the potential clustering of responses within PHCs and intervention sessions**.**

Results from all sensitivity analyses are reported in Multimedia Appendix 6. Model assumptions were evaluated via residual diagnostics, quantile-quantile (Q-Q) plots, and variance inflation factors (VIF<2.5). Models were fitted using restricted maximum-likelihood with the lme4 package. Reporting followed the TREND statement for nonrandomized evaluations of behavioral and public health interventions [47]; a completed TREND checklist is provided in Multimedia Appendix 1.

Ethical Considerations

Ethical approval for this study was granted by the Local Committee of Bioethics at Northern Border University (HAP-09-A-043; decision number 90/25/H). All participants were informed about the study objectives and procedures, and written informed consent was obtained before participation. The research team adhered strictly to all bioethics guidelines stipulated by the committee throughout the study.

Consent procedures emphasized voluntariness, confidentiality, the right to withdraw, and the possibility of discomfort when discussing sensitive topics. Consent forms were written in Arabic at approximately a sixth-grade reading level (assessed using the Arabic SMOG formula) and reviewed orally by female staff to ensure comprehension.

Confidentiality was maintained using unique IDs, separate secure storage of identifiers and data, encrypted password-protected databases, and staff confidentiality training. Data will be retained for 5 years and then securely destroyed by institutional policy.

Risk mitigation strategies included female-only delivery teams, private spaces, flexible scheduling, permission to skip any question, and cancer-prevention framing to minimize stigma. Adverse events were monitored through participant self-report at each assessment and facilitator incident reports. No adverse events were reported. Control participants were offered delayed access to the intervention after completion of the 3-month follow-up (148/190, 78% accepted).

This nonrandomized interventional study was not registered in a public trials registry because trial registration is not a requirement for quasi-experimental evaluations under applicable journal policy; the analysis plan was prespecified in writing before data analysis. Study methods and procedures were reviewed and approved by the institutional bioethics committee; no protocol amendments were made after data collection began.

Results

Participant Flow and Retention

A total of 512 women were assessed for eligibility between December 2025 and May 2026 in the 4 participating PHCs. Of these, 123 (24%) were excluded: 69 did not meet inclusion criteria (18 were currently pregnant, 28 were health care professionals or trainees with specialized knowledge, 4 had cognitive impairment limiting informed consent, and 19 had previously completed the full HPV vaccination series), 32 declined to participate, and 22 were excluded for other reasons (eg, planned relocation or inability to attend scheduled sessions). The remaining 389 women (76% of those screened)

XSL•FO
RenderX

were enrolled and allocated to the educational intervention group (n=194) or the control group (n=195; Figure 1). All 380 participants included in the analytic sample completed the 3-month follow-up; the final follow-up assessment was completed on May 30, 2026.

In the intervention group, 99% (192/194) of enrolled participants received the educational intervention; 2 participants withdrew before the intervention session because of scheduling conflicts. In the control group, 97.9% (191/195) of enrolled participants completed the baseline assessment and remained in the study; 4 participants withdrew before follow-up assessments could begin (2 because of relocation and 2 because they could not be reached). During follow-up, an additional 2 participants from the intervention group and 1 from the control group were lost to follow-up (lost contact or declined further assessment). Consequently, 190 participants per group (380/389, 97.7%) completed all 3 assessments and constituted the analytic sample (Figure 1). Baseline characteristics did not differ significantly between completers and those lost to follow-up (all *P*>.40), suggesting minimal attrition bias.

Figure 1. Flow diagram of participant enrollment, allocation, follow-up, and analysis.

Baseline Characteristics and Balance Assessment

Baseline demographic, socioeconomic, and health-related characteristics were well-balanced between the study arms (Table 1). The mean age of participants was approximately 28.6 (SD 6.9) years, with the majority holding a university degree (n/N, 65%). Approximately 43% (n/N) of participants were married. Importantly, potential confounders such as health literacy, previous vaccination experience, and family history of cervical cancer showed no statistically significant differences between groups (all *P*>.05). Standardized differences for all covariates were less than 0.10, indicating minimal observed baseline imbalance despite nonrandom allocation and supporting comparability of groups for adjusted analyses.

Baseline outcome scores were also well-balanced: mean knowledge scores were 2.97 (intervention) vs 2.94 (control), health awareness 3.05 vs 2.93, attitude 2.92 vs 2.91, and vaccination intentions 3.99 vs 3.98 (all *P*>.40, standardized differences <.10). This balance across both baseline characteristics and outcome measures strengthens causal inference despite the nonrandomized design.

Primary Outcome Analyses

HPV Vaccine Knowledge

The intervention was associated with significant and sustained improvements in HPV vaccine knowledge at 3 months (Table 2). While baseline scores were comparable (intervention: 2.97, 95% CI 2.77-3.17; control: 2.94, 95% CI 2.74-3.14; *P*=.86), the intervention group increased to 3.47 (95% CI 3.29-3.65) at T2 and to 3.74 (95% CI 3.57-3.91) at T3; the control group remained essentially unchanged (T2: 3.09, 95% CI 2.90-3.28; T3: 2.87, 95% CI 2.69-3.05). Between-group comparisons were

XSL•FO
RenderX

statistically significant at T2 (*P*=.01) and T3 (*P*<.001). Effect sizes at T3 were Cohen *d*=0.58 using pooled baseline SD and Cohen *d*=0.45 using pooled change-score SD.

Table 2. Multivariable adjusted analysis of intervention effects on primary outcomes.

Outcome and time point	Unadjusted β (95% CI)	*P* value	Adjusted β[a] (95% CI)	*P* value	Fully adjusted β[b] (95% CI)	*P* value	R^2
Knowledge							
Postintervention	0.38 (0.08-0.68)	.01	0.36 (0.07-0.65)	.02	0.35 (0.06-0.64)	.02	0.14
3-month follow-up	0.87 (0.57-1.17)	<.001	0.85 (0.55-1.15)	<.001	0.84 (0.53-1.15)	<.001	0.21
Awareness							
Postintervention	0.52 (0.17-0.87)	.004	0.50 (0.15-0.85)	.005	0.49 (0.14-0.84)	.006	0.16
3-month follow-up	0.91 (0.56-1.26)	<.001	0.90 (0.55-1.25)	<.001	0.88 (0.53-1.23)	<.001	0.24
Attitudes							
Postintervention	0.40 (0.12-0.68)	.006	0.38 (0.10-0.66)	.008	0.38 (0.09-0.67)	.009	0.12
3-month follow-up	0.73 (0.45-1.01)	<.001	0.71 (0.43-0.99)	<.001	0.70 (0.42-0.98)	<.001	0.19
Intentions							
Postintervention	0.50 (0.24-0.76)	<.001	0.49 (0.23-0.75)	<.001	0.48 (0.21-0.75)	<.001	0.18
3-month follow-up	0.89 (0.63-1.15)	<.001	0.87 (0.61-1.13)	<.001	0.86 (0.60-1.12)	<.001	0.28

[a]Adjusted for age, education, and marital status.

[b]Additionally adjusted for health literacy, previous vaccination experience, and family history of cervical cancer.

HPV-Related Health Awareness

Awareness of HPV health risks increased significantly within the intervention arm (T1: 3.05, 95% CI 2.84-3.26; T2: 3.47, 95% CI 3.27-3.67; T3: 3.92, 95% CI 3.73-4.11), whereas the control group remained stable (T1: 2.93, 95% CI 2.73-3.13; T3: 3.01, 95% CI 2.81-3.21). Between-group differences were statistically significant at T2 (*P*=.004) and T3 (*P*<.001), with effect sizes of *d*=0.35 (baseline SD) at T2 and *d*=0.63 (baseline SD) at T3 (change-score *d*=0.44 in Table 3).

Table 3. Analysis of immediate postintervention effects: comparison of change scores.

Comparison and variable	Intervention Δ[a], mean (SD)	Control Δ[a], mean (SD)	Difference in change	*t* test (*df*)	*P* value	Cohen *d* (95% CI)[b]
Comparison of change scores (T2 – T1)						
Knowledge	0.50 (1.92)	0.15 (2.00)	0.36	1.77 (378)	.08	0.18 (–0.02 to 0.38)
Awareness	0.43 (1.86)	0.03 (2.04)	0.40	2.00 (378)	.05	0.21 (0.00 to 0.41)
Attitude	0.44 (1.88)	0.06 (2.02)	0.38	1.92 (378)	.06	0.20 (–0.01 to 0.40)
Intentions	0.66 (1.52)	0.17 (1.36)	0.49	3.31 (378)	.001	0.34 (0.14 to 0.54)
Comparison of change scores (T3 – T1)						
Knowledge	0.77 (1.72)	–0.07 (1.81)	0.84	4.69 (378)	<.001	0.45 (0.25 to 0.65)
Awareness	0.87 (1.69)	0.08 (1.88)	0.79	4.28 (378)	<.001	0.44 (0.24 to 0.64)
Attitude	0.82 (1.74)	0.10 (1.95)	0.72	3.80 (378)	<.001	0.38 (0.18 to 0.58)
Intentions	1.05 (1.28)	0.17 (1.24)	0.88	6.74 (378)	<.001	0.82 (0.62 to 1.02)

[a]Δ indicates the within-group change score.

[b]Cohen *d* was calculated using the pooled change score SD.

Attitudes Toward HPV Vaccination

Attitudes improved within the intervention arm (T1: 2.92, 95% CI 2.71-3.13; T3: 3.74, 95% CI 3.55-3.93), whereas control attitudes remained stable (T1: 2.91, 95% CI 2.71-3.11; T3: 3.01, 95% CI 2.81-3.21). Between-group analyses confirmed significantly more favorable attitudes in the intervention group at T2 (*P*=.006) and T3 (*P*<.001), with effect sizes of *d*=0.27 (baseline SD) and *d*=0.52 (baseline SD), respectively (change-score *d*=0.38 in Table 3).

Vaccination Intentions

The largest effects were observed for vaccination intentions. The intervention group showed a substantial increase from T1 (3.99, 95% CI 3.84-4.14) to T2 (4.65, 95% CI 4.49-4.81) and T3 (5.04, 95% CI 4.89-5.19). The control group showed small

XSL•FO
RenderX

increases (T1: 3.98, 95% CI 3.82-4.14; T3: 4.15, 95% CI 3.99-4.31). Between-group comparisons indicated significantly stronger vaccination intentions in the intervention group at both T2 and T3 (both $P<.001$), with effect sizes of $d=0.48$ and $d=0.71$ (baseline SD), respectively, and $d=0.82$ using change-score SD at T3 (Table 3).

Summary of Effect Sizes

Across outcomes, intervention effects using pooled baseline SD ranged from $d=0.27$ to $d=0.71$ at follow-up. Effect sizes calculated using the prespecified primary metric (pooled change-score SD) ranged from $d=0.38$ to $d=0.82$ at follow-up (Table 3), consistent with medium-to-large effects.

Adjusted Analyses Controlling for Confounders

Linear mixed-effects models were fitted for all primary outcomes to address potential baseline imbalance and confounding (Table 2). Three nested models were compared: model 0 (unadjusted), model 1 (adjusted for age, education, and marital status), and model 2, the primary model (additionally adjusted for health literacy, previous vaccination experience, and family history of cervical cancer).

The intervention effect (time × group interaction) was statistically significant across all adjusted models and across all 4 outcomes at T2 and T3. Adjustment for covariates produced minimal attenuation of effect estimates (β coefficients changed by <0.03 across models), indicating strong baseline balance and little measured confounding. Model R^2 values ranged from 0.12 to 0.28, with the fully adjusted models providing the best fit (Multimedia Appendix 6). Model diagnostics confirmed the adequacy of the linear mixed models: residual plots showed no systematic patterns; Q-Q plots indicated approximately normal residuals; and all VIFs were <2.5.

Cluster-Adjusted Sensitivity Analyses

Because the intervention was delivered in groups (10-15 participants) across multiple sessions (n=40), clinics (n=4), and facilitators (n=8), clustering of responses could have occurred. To address this concern (raised in BZ Major Comment 2), sensitivity models were refitted with random intercepts for clinic and intervention sessions, and with cluster-robust (sandwich) standard errors applied to the primary complete-case models. ICCs were modest at clinic and session levels (range 0.01-0.05). The cluster-adjusted models produced slightly larger point estimates and slightly wider confidence intervals than the primary models (maximum difference in β of ±0.04 for any outcome or time point), and all previously significant effects remained significant ($P<.01$). These analyses indicate that clustering does not materially alter inference for the primary outcomes.

Change Score Analysis and Effect Sizes

To quantify the specific magnitude of the intervention effect, within-group change scores were calculated and compared directly between groups. Table 3 presents immediate postintervention changes (T2 – T1) as well as sustained changes at 3-month follow-up (T3 – T1).

At the immediate postintervention assessment, the intervention group achieved significantly greater gains than controls for health awareness ($P=.05$; $d=0.21$) and vaccination intentions ($P=.001$; $d=0.34$). Knowledge and attitudes showed positive trends that did not reach statistical significance at this early time point. However, by the 3-month follow-up, all 4 outcomes demonstrated significant between-group differences in change scores, with effect sizes ranging from $d=0.45$ (knowledge) to $d=0.82$ (intentions) using the change score pooled SD. The progressive strengthening of effects from postintervention to follow-up suggests sustained engagement with intervention content and potential social diffusion or reinforcement of learning over time.

Subgroup Analyses and Effect Modification

Exploratory subgroup analyses were conducted to determine whether intervention effects varied by key demographic characteristics. We assessed interactions between the intervention and 4 prespecified moderators: age (18-26 years vs 27-45 years), education level (university vs below university level), health literacy (inadequate/marginal vs adequate using validated cutoffs), and previous vaccination experience (any vs none). Moderator effects were evaluated using 3-way interaction terms (time × group × moderator) in the mixed-effects models. Because time (2 levels), group (2 levels), and each moderator (2 levels) were binary, each interaction term has 1 numerator degree of freedom; denominator degrees of freedom were estimated using the Satterthwaite approximation. Using a Bonferroni-corrected significance level of $\alpha=.0125$ per outcome domain (to account for testing 4 moderators separately), no significant 3-way interactions (time × group × moderator) were detected for any outcome.

For vaccination intentions (the outcome with strongest main effects), the interaction tests were: age ($F_{1,424.31}=1.23$; $P=.28$), education ($F_{1,421.05}=0.87$; $P=.42$), health literacy ($F_{1,418.52}=1.41$; $P=.23$), and prior vaccination experience ($F_{1,425.14}=0.65$; $P=.52$). Similar nonsignificant patterns were observed for other outcomes (Multimedia Appendix 6 provides complete stratified results and interaction statistics).

Point estimates suggested numerically larger gains for younger women (age 18-26 years: mean 3-month change +1.10, 95% CI 0.89-1.31 vs age 27-45 years: mean 3-month change +0.98, 95% CI 0.72-1.24), university-educated participants (+1.15, 95% CI 0.96-1.34 vs below university level: +0.92, 95% CI 0.65-1.19), and those with adequate health literacy (+1.18, 95% CI 0.99-1.37 vs inadequate/marginal: +0.89, 95% CI 0.64-1.14), but these differences were not statistically significant after multiplicity correction. The study was not formally powered for interaction tests (estimated power approximately 40%-50%); therefore, the absence of significant interactions should not be interpreted as evidence of equal effect magnitudes across subgroups, and these analyses are best treated as hypothesis-generating.

Control Group Changes and Alternative Explanations

The control group also showed small increases across outcomes, particularly for vaccination intentions (from 3.98 at T1 to 4.15 at T3). Possible explanations discussed transparently include

XSL•FO
RenderX

secular trends from concurrent national HPV awareness campaigns in Saudi Arabia, repeated measurement effects (familiarity with questionnaire items), and nonspecific attention during routine clinic visits within the time period. Although these changes were small and did not reach statistical significance in isolation, the primary inference is based on adjusted between-group comparisons that account for any incremental control arm change.

Sensitivity Analyses

Comprehensive sensitivity analyses confirmed the robustness of primary findings across alternative analytic approaches (see Multimedia Appendix 4 for complete results). Intention-to-treat analysis using LOCF for the 9 participants lost to follow-up yielded effect estimates within 0.02 to 0.04 points of primary complete-case results, with all previously significant effects remaining significant (all $P<.01$).

Per-protocol analysis restricted to participants with 80% or greater session attendance (n=352; 92.7% of analytical sample) showed slightly stronger intervention effects, with β coefficients 5% to 8% larger than primary models (eg, vaccination intentions at 3 months: β=0.93, 95% CI 0.66-1.20 vs primary β=0.86). This dose-response pattern is consistent with greater engagement producing larger effects.

Multiple imputation with 50 datasets using MICE produced point estimates within 0.01 to 0.03 points of complete-case analysis, with nearly identical CIs and *P* values. Alternative covariance structures (unstructured and first-order autoregressive) yielded substantively identical conclusions (maximum difference in β=0.02; all previously significant effects remained $P<.001$).

Assessment of potential mode effects (in-person paper administration vs telephone) revealed no systematic bias. Within the control group (which received no intervention between T2 and T3), responses at T2 (paper) and T3 (telephone) showed no significant differences across all 4 outcomes (all $P>.15$; mean absolute difference <.08 points; Multimedia Appendix 5). This provides reassurance that the T3 telephone data collection mode did not introduce measurement artifacts. The consistency of findings across these diverse sensitivity analyses reinforces the internal validity and robustness of the trial results.

Discussion

Overview

This study demonstrates that a culturally tailored, IMB-based educational intervention significantly improved HPV-related knowledge, health awareness, attitudes, and vaccination intentions among women aged 18 to 45 years attending PHC centers in Northern Saudi Arabia. At 3-month follow-up, intervention participants showed medium-to-large effect-size improvements across all outcomes compared with controls, with effects sustained from the immediate postintervention assessments. These findings contribute to the limited evidence base on culturally tailored HPV education interventions in Middle Eastern contexts and provide the first controlled evaluation of an IMB-based approach delivered in Saudi primary care settings, but because allocation was nonrandom and based on clinic session timing, the present results estimate intervention association rather than a strict causal effect.

Principal Findings and International Contextualization

The intervention yielded a substantial rise in vaccination intention, from 3.99 at baseline to 5.04 at the 3-month follow-up, representing a 15% absolute increase on a 7-point scale and constituting the largest observed effect, a pattern consistent with prior educational interventions. Similar gains have been documented among pregnant women in Saudi Arabia, where structured educational sessions produced twofold to threefold increases in HPV-related awareness [48]. Comparable outcomes were also observed among Tunisian female students, who demonstrated significant knowledge improvement following brief video-based instruction, with mean scores rising from 4.23 preintervention to 6.16 postintervention, and vaccination acceptance reaching 68.4% [49]. These findings together suggest that theory-driven educational strategies addressing cultural concerns can produce reliably positive effects across Middle East and North Africa populations.

Our observed knowledge gains, rising from 2.97 at baseline to 3.74 at follow-up, mirror results from Chinese college populations who participated in web-based IMB model interventions, where significant increases in HPV knowledge and motivational determinants were consistently reported [40,50,51]. However, a key contrast emerges despite marked improvements in knowledge and intention: Chinese studies reported very limited translation into actual vaccine uptake, a gap attributed to systemic barriers such as high vaccine cost and restricted insurance coverage [52]. This intention-behavior discrepancy highlights that, although educational interventions can effectively enhance cognitive and motivational precursors to vaccination, their population-level impact ultimately depends on structural factors, including affordability, vaccine access, and provider recommendation practices.

Not all educational interventions achieve sustained behavioral impact. Systematic reviews of HPV vaccine education indicate that knowledge gains do not invariably translate into vaccination intentions or uptake, particularly when interventions fail to address context-specific motivational barriers such as perceived stigma, safety concerns, or family opposition [11]. Moreover, intention effects observed at short-term follow-up (1-3 months) often decay by 6 to 12 months absent booster sessions or reminder systems [13]. Our 3-month results should therefore be interpreted as promising but preliminary, requiring longer-term follow-up with objective vaccination outcomes to confirm durability and behavioral translation.

Theoretical Mechanisms and IMB Model Validation

The intervention's effectiveness supports IMB model predictions that targeting all 3 constructs, information, motivation, and behavioral skills, produces more robust effects than information-only approaches [38]. The information component addressed widespread misconceptions (HPV/HIV confusion, infertility concerns, and Halal permissibility) and provided evidence-based content on vaccine safety and efficacy, yielding significant knowledge gains. The motivational component emphasized cancer prevention framing rather than STI

XSL•FO
RenderX

prevention to reduce stigma, addressed personal susceptibility and severity perceptions, and normalized vaccination within cultural contexts, resulting in improved health awareness and attitudes. The behavioral skills component strengthened self-efficacy, communication competencies for discussing vaccination with providers and family members, and problem-solving for navigating appointment systems, contributing to the substantial intentions effect.

Notably, the progressive increase in intentions from postintervention (4.65) to 3-month follow-up (5.04), rather than decay, suggests that behavior change processes may require extended periods for consolidation as individuals integrate new information, negotiate social contexts, and build decision-making confidence. This temporal pattern aligns with IMB-based intervention studies reporting that behavioral skills acquisition requires longer maturation periods than informational or motivational gains [40].

The moderate-to-large effect sizes observed in this study compare favorably with meta-analytic estimates for health education interventions and exceed typical effects for single-session interventions [33]. Several design features likely contributed to effectiveness: (1) cultural adaptation through systematic formative research addressing local barriers (STI stigma, fertility concerns, and family decision-making dynamics), (2) theoretical grounding ensuring comprehensive coverage of behavioral determinants, (3) female-only facilitation enhancing comfort with sensitive topics, (4) interactive group format promoting social normative influences, and (5) high implementation fidelity (93%) maintaining intervention integrity.

Implications for Clinical Practice and Health Systems

Clinicians

These findings suggest that brief, structured education can be feasibly integrated into primary care encounters to prepare women for informed HPV vaccination discussions. Clinicians should prioritize addressing common misconceptions (fertility effects, HPV/HIV confusion, and Halal status) and emphasize cancer prevention rather than STI prevention to reduce stigma in conservative contexts. Providing culturally adapted written materials (pamphlets and decision aids) at appropriate literacy levels can supplement verbal counseling and facilitate family discussions outside clinical settings.

Health Systems

The intervention's standardized format (45 minutes, 2 facilitators, and 10-15 participants), modest resource requirements, and ability to provide high-fidelity support that scales within existing primary care infrastructure could facilitate integration into routine women's health services, for example, during well-woman visits, postpartum care, or family planning consultations, and could reach women at teachable moments when preventive health is salient. Group education formats offer additional benefits of peer modeling, normative influence, and cost-efficiency compared with individual counseling.

Policy

The findings support integrating HPV vaccination into Saudi Arabia's national immunization program for adult women, as current recommendations do not include this age group in the publicly funded schedule. This expansion should be paired with broad public education campaigns that address informational and motivational barriers at the population level. Policy planning must also account for persistent structural constraints, most notably vaccine cost and accessibility, given robust evidence that educational interventions alone are insufficient to overcome economic obstacles to uptake, as reflected in studies demonstrating pronounced intention-behavior gaps driven by resource limitations [52]. The Tunisian Ministry of Health's forthcoming school-based HPV vaccination rollout, anticipated for 2025 and informed by intervention research reporting 68.4% vaccine acceptance among female students [49], provides a regionally relevant policy model. Adaptation to Saudi Arabia, however, requires tailoring to local sociocultural dynamics, including family-centered decision-making processes and the gender-segregated organization of educational environments.

Shared Decision-Making

Educational interventions strengthen the informational foundation necessary for SDM by equipping women with knowledge, enhancing motivation, and building confidence to engage in vaccination discussions with providers. However, realizing SDM's full potential in Saudi health care contexts requires parallel efforts to train clinicians in patient-centered communication, modify directive counseling styles, and create structural supports for family involvement in decisions without undermining women's autonomous agency [27,28]. Future intervention research should evaluate whether enhanced patient readiness translates into higher-quality SDM processes (eg, measured via the Observing Patient Involvement in Decision Making [OPTION] scale) and whether SDM quality mediates relationships between education and vaccination uptake.

Notably, vaccination coverage is a common indicator of the effectiveness of vaccination strategies. In this context, the study by Guarducci et al [53] provides a valuable example of how an organizational and communication network can be used to implement HPV vaccination coverage for both males and females in the Central Tuscany Health Region. Their pilot model led to a significant increase in coverage, from an average of 49.2% in January 2022 to 63.9% in December 2022 [53]. These results support the idea that coordinated, multifaceted interventions can effectively address HPV vaccination uptake gaps and are consistent with the findings observed in our study. The success of these network-based approaches underscores the importance of integrating organizational strategies and community engagement to improve vaccination coverage.

Study Strengths

This study has several methodological strengths. First, the intervention was grounded in behavioral theory (IMB model) and developed through systematic cultural adaptation involving formative research with target populations and stakeholders, enhancing both internal validity and local relevance. Second, the study achieved excellent retention (97.7%) across 3 time

XSL•FO
RenderX

points, allowing assessment of both immediate and sustained effects. Third, despite nonrandomized allocation, baseline balance across measured covariates (standardized differences <.10) and comprehensive covariate adjustment strengthen confidence in effect estimates. Fourth, high implementation fidelity (93% adherence; substantial interrater reliability κ=0.78) and standardized delivery protocols support replicability. Fifth, the primary care setting and focus on adult women, an underserved population in HPV vaccination research, enhance ecological validity and policy relevance. Finally, comprehensive sensitivity analyses (intention-to-treat, per-protocol, multiple imputation, and alternative covariance structures) confirmed robustness of findings across analytic approaches.

Study Limitations

Several limitations qualify the interpretation and generalizability of findings.

Design Limitations

The nonrandomized, time-based allocation (morning vs afternoon clinic sessions) introduces potential selection bias, as clinic attendance timing may correlate with employment status, caregiving responsibilities, health motivation, and socioeconomic factors. While standardized differences less than 0.10 suggest minimal confounding from measured covariates, unmeasured factors (eg, health consciousness and urgency of care-seeking) could bias effect estimates. Future research should use randomized designs or, when randomization is infeasible, use propensity score matching and instrumental variable approaches to strengthen causal inference. Additionally, participants and facilitators were not blinded to group allocation, introducing potential performance and detection bias. T3 telephone interviewers were only partially blinded and could potentially infer allocation if participants mentioned attending education sessions.

Self-Reported Limitations

Outcomes relied exclusively on self-report, introducing potential social desirability bias, particularly for attitudes and intentions. The adapted outcome measures have not been comprehensively validated (eg, confirmatory factor analysis and convergent/discriminant validity) in Saudi populations; pilot testing (n=20) provided only preliminary psychometric support. Differential measurement modes (paper-based at T1/T2; telephone at T3) could introduce mode effects, though sensitivity analyses within the control group suggested minimal bias. Future studies could incorporate validated social desirability scales to statistically control for this bias.

Generalizability Limits

The study recruited predominantly educated (65% university degree), urban young women from a single region (Northern Border Region), limiting generalizability to rural populations, older women, those with lower educational attainment, and other Saudi regions with potentially different cultural norms and health care access patterns. Exclusion of pregnant women (a key vaccine-eligible group), health care workers, and those completing the full vaccination series further restricts generalizability. The Northern Border Region's geographic proximity to Jordan and exposure to cross-border cultural influences may have produced a sample with more liberal attitudes than central or southern Saudi regions.

Outcome and Follow-Up Limitations

This study assessed self-reported knowledge, awareness, attitudes, and vaccination intentions rather than verified vaccine initiation or series completion. While intention is a well-established predictor of health behavior, it does not guarantee actual vaccination. Therefore, our findings should be interpreted as preliminary evidence of potential impact, and claims regarding clinical meaningfulness, scalability, policy implementation, and broad applicability should be considered with caution. Future research should incorporate objective measures of vaccination behavior to confirm these findings.

The study measured vaccination intentions, a surrogate end point, rather than verified vaccine uptake. The 3-month follow-up is insufficient to assess vaccine uptake behaviors, which may require extended time frames for appointment scheduling, cost accumulation, and series completion (2-3 doses over 6 months). Unmeasured confounding remains a concern, as the study did not systematically assess male partner or family member attitudes toward HPV vaccination. Male partner and family attitudes likely represent meaningful, unmeasured determinants of HPV vaccination uptake. These factors are likely to exert substantial influence on women's preventive health behaviors in Saudi Arabia's collectivist cultural context. In this setting, future research should adopt dyadic or family-level designs evaluating concordance between individual intentions and family support and testing family-centered intervention approaches.

Process Evaluation Gaps

The study did not include a formal process evaluation to identify which intervention components were most effective, which women benefited most, or what barriers remained after education. Mediational analyses testing whether knowledge gains mediated attitude and intention improvements would strengthen theoretical understanding. Future studies should incorporate mixed methods process evaluations with qualitative interviews exploring participants' experiences, decision-making processes, and implementation barriers.

Future Research Directions

Future research should address the following priorities.

Objective Behavioral Outcomes

Conduct trials with 12-month follow-up measuring verified vaccine initiation and series completion through medical record linkage or immunization registry data. Include cost-effectiveness analysis to inform policy decisions about program implementation and scale-up.

Multisite Generalizability Studies

Use randomized designs spanning diverse Saudi regions (urban/rural, Northern, Central, and Southern) and demographic strata (age, education, and marital status) to establish intervention effectiveness across heterogeneous populations and identify moderators requiring adaptation.

XSL•FO
RenderX

Implementation Research

Conduct hybrid effectiveness-implementation trials evaluating both clinical outcomes and implementation processes (reach, adoption, fidelity, cost, and sustainability) under real-world conditions. Identify barriers and facilitators to routine integration into primary care workflows, training requirements for health care staff, and strategies for maintaining fidelity while allowing contextually appropriate adaptations.

Family-Centered Interventions

Develop and test interventions targeting male partners and family decision-makers, given evidence that family attitudes substantially influence women's vaccination decisions in collectivist contexts. Compare the effectiveness of individual-level, couple-level, and family-level intervention approaches.

Structural Barrier Reduction

Evaluate combined interventions addressing both informational/motivational barriers (education) and structural barriers (subsidized vaccines, integrated screening-vaccination services, provider training, and reminder/recall systems). Test whether addressing multiple levels simultaneously yields synergistic effects on uptake.

Comparative Effectiveness Research

Compare in-person group education (as tested here) with alternative delivery modalities, including individual counseling, web-based digital interventions, mobile health (mHealth) apps, and hybrid approaches combining digital content with personalized counseling. Identify optimal modality-population matches based on literacy levels, technology access, cultural preferences, and baseline motivation.

Conclusion

This study provides evidence that a culturally adapted, IMB-based educational intervention delivered in Saudi primary care settings significantly improves HPV-related knowledge, health awareness, attitudes, and vaccination intentions among women aged 18 to 45 years, with medium-to-large effect sizes sustained at 3-month follow-up. Effects were consistent across age, education, and health literacy subgroups, suggesting broad applicability. These findings demonstrate that educational interventions targeting cognitive and motivational barriers are feasible, acceptable, and effective in conservative Middle Eastern contexts where HPV vaccination faces cultural sensitivities.

Nevertheless, these results suggest that educational interventions represent a necessary, though insufficient, component of comprehensive HPV prevention strategies. HPV-related cancers are preventable, yet barriers persist at individual, interpersonal, health care system, and policy levels. Closing the gap between vaccine availability and uptake requires multilevel interventions that simultaneously address knowledge, motivation, structural barriers, and cultural realities. This study represents an important step toward evidence-based, culturally grounded HPV prevention programs that can contribute to cervical cancer reduction and women's health equity in Saudi Arabia and throughout the Middle Eastern region.

Acknowledgments

The authors acknowledge the contributions of community stakeholders, religious scholars, health care providers, and pilot study participants whose insights shaped this culturally adapted intervention. We also thank the facilitators who delivered the intervention with dedication and cultural sensitivity and the research assistants who ensured rigorous data collection and fidelity monitoring. The authors extend their sincere gratitude to all women who participated in this study. The authors also thank the Deanship of Scientific Research at Northern Border University, Arar, Saudi Arabia, for funding this research through project number "NBU-FFR-2026-3326-08."

Funding

This research received funding from the Deanship of Scientific Research at Northern Border University, Arar, Saudi Arabia, through project number "NBU-FFR-2026-3326-08."

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: FAM, HAOM, EAE, AMA, NMA
Data curation: HAOM, FAM, EAE
Formal analysis: FAM, HAOM
Methodology: EAE, HAOM, FAM
Project administration: FAM
Resources: NMA, AMA, EAE
Supervision: FAM
Writing—original draft: NMA, AMA
Writing—reviewing and editing: FAM, HAOM, EAE, AMA, NMA

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
TREND checklist.
[PDF File (Adobe PDF File), 166 KB - publichealth_v12i1e104529_app1.pdf]

Multimedia Appendix 2
Comparison of baseline characteristics: study completers vs noncompleters.
[DOCX File , 23 KB - publichealth_v12i1e104529_app2.docx]

Multimedia Appendix 3
Session scripts, facilitator guides, and educational materials.
[DOCX File , 48 KB - publichealth_v12i1e104529_app3.docx]

Multimedia Appendix 4
Complete subgroup analysis results: effect modification testing.
[DOCX File , 30 KB - publichealth_v12i1e104529_app4.docx]

Multimedia Appendix 5
Assessment of mode effects: telephone vs in-person paper administration.
[DOCX File , 22 KB - publichealth_v12i1e104529_app5.docx]

Multimedia Appendix 6
Model comparison statistics: unadjusted vs adjusted linear mixed-effects models.
[DOCX File , 27 KB - publichealth_v12i1e104529_app6.docx]

References

1. Chesson HW, Dunne EF, Hariri S, Markowitz LE. The estimated lifetime probability of acquiring human papillomavirus in the United States. Sex Transm Dis 2014;41(11):660-664 [FREE Full text] [doi: 10.1097/OLQ.0000000000000193] [Medline: 25299412]
2. Milano G, Guarducci G, Nante N, Montomoli E, Manini I. Human papillomavirus epidemiology and prevention: is there still a gender gap? Vaccines (Basel) 2023;11(6):1060-1076 [FREE Full text] [doi: 10.3390/vaccines11061060] [Medline: 37376449]
3. Human papillomavirus and cancer. World Health Organization. 2024 Mar 05. URL: https://www.who.int/news-room/fact-sheets/detail/human-papilloma-virus-and-cancer [accessed 2026-09-01]
4. Alsbeih G. HPV infection in cervical and other cancers in Saudi Arabia: implication for prevention and vaccination. Front Oncol 2014;4:65 [FREE Full text] [doi: 10.3389/fonc.2014.00065] [Medline: 24744990]
5. Joura EA, Giuliano AR, Iversen OE, Bouchard C, Mao C, Mehlsen J. A 9-valent HPV vaccine against infection and intraepithelial neoplasia in women. N Engl J Med 2015;372(8):711-723. [doi: 10.1056/NEJMoa1405044] [Medline: 25693011]
6. Petrosky E, Bocchini JJ, Hariri S, Chesson HW, Curtis C, Saraiya M, Centers for Disease ControlPrevention (CDC). Use of 9-valent human papillomavirus (HPV) vaccine: updated HPV vaccination recommendations of the Advisory Committee on Immunization Practices. MMWR Morb Mortal Wkly Rep 2015 Mar 27;64(11):300-304 [FREE Full text] [Medline: 25811679]
7. Seebauer C, Faluogy M, Sieg P, Olbrich H, Ludwig R. HPV vaccination and malignancy risks beyond cervical cancer: a retrospective global cohort study. Pharmacol Res 2025;218:107851 [FREE Full text] [doi: 10.1016/j.phrs.2025.107851] [Medline: 40653128]
8. Meites E, Szilagyi PG, Chesson HW, Unger ER, Romero JR, Markowitz LE. Human papillomavirus vaccination for adults: updated recommendations of the advisory committee on immunization practices. MMWR Morb Mortal Wkly Rep 2019;68(32):698-702 [FREE Full text] [doi: 10.15585/mmwr.mm6832a3] [Medline: 31415491]
9. Brumbaugh KQ, Gellert F, Mokdad AH. Understanding vaccine hesitancy: insights and improvement strategies drawn from a multi-study review. Vaccines (Basel) 2025;13(10):1003 [FREE Full text] [doi: 10.3390/vaccines13101003] [Medline: 41150390]
10. Kolobova I, Nyaku MK, Karakusevic A, Bridge D, Fotheringham I, O'Brien M. Vaccine uptake and barriers to vaccination among at-risk adult populations in the US. Hum Vaccin Immunother 2022;18(5):2055422 [FREE Full text] [doi: 10.1080/21645515.2022.2055422] [Medline: 35536017]

XSL•FO
RenderX

11. Fu LY, Bonhomme L, Cooper SC, Joseph JG, Zimet GD. Educational interventions to increase HPV vaccination acceptance: a systematic review. Vaccine 2014;32(17):1901-1920 [FREE Full text] [doi: 10.1016/j.vaccine.2014.01.091] [Medline: 24530401]
12. Hendry M, Lewis R, Clements A, Damery S, Wilkinson C. “HPV? Never heard of it!”: a systematic review of girls’ and parents’ information needs, views and preferences about human papillomavirus vaccination. Vaccine 2013;31(45):5152-5167. [doi: 10.1016/j.vaccine.2013.08.091] [Medline: 24029117]
13. Sheeran P, Webb TL. The intention-behavior gap. Soc Personal Psychol Compass 2016;10(9):503-518. [doi: 10.1111/spc3.12265]
14. Saudi Arabia: human papillomavirus and related cancers, fact sheet 2023. ICO/IARC Information Centre on HPV and Cancer. URL: https://hpvcentre.net/statistics/reports/SAU_FS.pdf [accessed 2026-09-01]
15. Cancer incidence report: Saudi Arabia. Saudi Health Council. URL: https://shc.gov.sa/Arabic/NewNCC/Activities/AnnualReports/Cancer%20Incidence%20Report%202022.pdf [accessed 2026-09-01]
16. Alshamsan B. Colorectal cancer incidence in Saudi Arabia, 2002-2023: age-specific trends and implications for screening and prevention. Prev Med Rep 2026;66:103505 [FREE Full text] [doi: 10.1016/j.pmedr.2026.103505] [Medline: 42199669]
17. Saudi Cancer Registry Annual Reports. Saudi Health Council. 2026. URL: https://shc.gov.sa/en/Centers/NCC/Pages/AnnualReports.aspx [accessed 2026-09-01]
18. Women’s health: human papillomavirus (HPV). Ministry of Health, Kingdom of Saudi Arabia. URL: https://www.moh.gov.sa/en/healthawareness/educationalcontent/wh/pages/human-papillomavirus.aspx [accessed 2026-09-01]
19. Alqarni SS, Alshehri SM, Alkhateeb MA, Alsudias LS. Assessing Saudi women's awareness about human papillomavirus (HPV) and their susceptibility to receive the vaccine. Hum Vaccin Immunother 2024;20(1):2395086 [FREE Full text] [doi: 10.1080/21645515.2024.2395086] [Medline: 39219415]
20. Rezq KA, Algamdi M, Alanazi R, Alanazi S, Alhujairy F, Albalawi R, et al. Knowledge, perception, and acceptance of HPV vaccination and screening for cervical cancer among Saudi females: a cross-sectional study. Vaccines (Basel) 2023;11(7):1188 [FREE Full text] [doi: 10.3390/vaccines11071188] [Medline: 37515004]
21. Alanazi AK, Kayal E, Alanzi S, Al Hodian H, Bin Rusayes A. Knowledge and awareness toward human papillomavirus vaccination among Saudi female nursing students. Front Glob Womens Health 2024;5:1470048 [FREE Full text] [doi: 10.3389/fgwh.2024.1470048] [Medline: 39512335]
22. Bakhsh H, Ali Altamimi SA, Aldosari FN, Hatim Aljohani L, Abdulrahman Alali S, Ibrahim Almutlaq N, et al. Barriers and predictors of HPV vaccine uptake among female medical students in Saudi Arabia: a multi-center cross-sectional study. Healthcare (Basel) 2025;13(19):2408 [FREE Full text] [doi: 10.3390/healthcare13192408] [Medline: 41095494]
23. Azzi A, Al-Dhelaan R, Al-Fahd A, Almukhalafi O, Al-Tayar N, Abutheeb A, et al. Assessment of awareness level among women about cervical smears, human papillomavirus (HPV), and HPV vaccine in Saudi Arabia. Sci Rep 2025;15(1):44171 [FREE Full text] [doi: 10.1038/s41598-025-27814-z] [Medline: 41419744]
24. Alshahrani MN, Almutairi D, Zahrani Y, Alsabaani A, Alraey Y, Alshahrani SM. Barriers associated with the parental acceptance of human papillomavirus (HPV) vaccination of minors in Asir, Saudi Arabia: a cross-sectional study. Hum Vaccin Immunother 2025;21(1):2513708 [FREE Full text] [doi: 10.1080/21645515.2025.2513708] [Medline: 40462525]
25. Othman S, Ghamri R, Alhamadah W, Alsatty A, Khesfaty D, Alghamdi O, et al. Prevalence and characteristics of HPV vaccine hesitancy among parents of adolescent females across Saudi Arabia. Front Public Health 2024;12:1501358 [FREE Full text] [doi: 10.3389/fpubh.2024.1501358] [Medline: 39917529]
26. Alghanmi R, Alkhalawi E, Albeladi R, Albeladi S, Alghamdi M, Fayoumi A, et al. Parental knowledge and acceptance of HPV vaccine in Rabigh’s school, Saudi Arabia. Ann Glob Health 2026;92(1):13 [FREE Full text] [doi: 10.5334/aogh.4866] [Medline: 41659211]
27. Alabdullah YY, Alzaid E, Alsaad S, Alamri T, Alolayan SW, Bah S, et al. Autonomy and paternalism in shared decision-making in a Saudi Arabian tertiary hospital: a cross-sectional study. Dev World Bioeth 2023;23(3):260-268. [doi: 10.1111/dewb.12355] [Medline: 35586963]
28. Charles C, Gafni A, Whelan T. Shared decision-making in the medical encounter: what does it mean? (or it takes at least two to tango). Soc Sci Med 1997;44(5):681-692. [doi: 10.1016/s0277-9536(96)00221-3] [Medline: 9032835]
29. Stiggelbout AM, Pieterse AH, De Haes JCJM. Shared decision making: concepts, evidence, and practice. Patient Educ Couns 2015;98(10):1172-1179. [doi: 10.1016/j.pec.2015.06.022] [Medline: 26215573]
30. Elwyn G, Frosch D, Thomson R, Joseph-Williams N, Lloyd A, Kinnersley P, et al. Shared decision making: a model for clinical practice. J Gen Intern Med 2012;27(10):1361-1367 [FREE Full text] [doi: 10.1007/s11606-012-2077-6] [Medline: 22618581]
31. Légaré F, Ratté S, Gravel K, Graham ID. Barriers and facilitators to implementing shared decision-making in clinical practice: update of a systematic review of health professionals' perceptions. Patient Educ Couns 2008;73(3):526-535. [doi: 10.1016/j.pec.2008.07.018] [Medline: 18752915]
32. Iova CF, Daina LG, Daina MD, Ghitea TC. The effectiveness of interventions targeting adolescents in HPV vaccination-a scoping review. Medicina (Kaunas) 2024;60(9):1550 [FREE Full text] [doi: 10.3390/medicina60091550] [Medline: 39336591]

XSL•FO
RenderX

33. Kreuter MW, Wray RJ. Tailored and targeted health communication: strategies for enhancing information relevance. Am J Health Behav 2003;27 Suppl 3:S227-S232. [doi: 10.5993/ajhb.27.1.s3.6] [Medline: 14672383]
34. Escoffery C, Petagna C, Agnone C, Perez S, Saber LB, Ryan G, et al. A systematic review of interventions to promote HPV vaccination globally. BMC Public Health 2023;23(1):1262 [FREE Full text] [doi: 10.1186/s12889-023-15876-5] [Medline: 37386430]
35. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. Ann Intern Med 2011;155(2):97-107. [doi: 10.7326/0003-4819-155-2-201107190-00005] [Medline: 21768583]
36. Chew LD, Bradley KA, Boyko EJ. Brief questions to identify patients with inadequate health literacy. Fam Med 2004 Sep;36(8):588-594. [Medline: 15343421]
37. National action plan to improve health literacy. Office of Disease Prevention and Health Promotion. 2010. URL: https://odphp.health.gov/our-work/national-health-initiatives/health-literacy/national-action-plan-improve-health-literacy [accessed 2026-09-01]
38. Fisher JD, Fisher WA. Changing AIDS-risk behavior. Psychol Bull 1992;111(3):455-474. [doi: 10.1037/0033-2909.111.3.455]
39. Perez GK, Cruess DG, Strauss NM. A brief information-motivation-behavioral skills intervention to promote human papillomavirus vaccination among college-aged women. Psychol Res Behav Manag 2016;9:285-296 [FREE Full text] [doi: 10.2147/PRBM.S112504] [Medline: 27799835]
40. Wang H, Wang X, Chen P, Xu H, Liu Y, Kang R, et al. Effect of health intervention web-based education on improving information-motivation-behavioral skills related to HPV vaccination among Chinese female college students. Int J Public Health 2023;68:1605596. [doi: 10.3389/ijph.2023.1605596] [Medline: 36816834]
41. Aldawood E, Dabbagh D, Alharbi S, Alzamil L, Faqih L, Alshurafa H, et al. HPV vaccine knowledge and hesitancy among health colleges' students at a Saudi university. J Multidiscip Healthc 2023;16:3465-3476 [FREE Full text] [doi: 10.2147/JMDH.S438633] [Medline: 38024130]
42. John S, Walsh J, Weinhardt L. The information-motivation-behavioral skills model revisited: a network-perspective structural equation model within a public sexually transmitted infection clinic sample of hazardous alcohol users. AIDS Behav 2017;21(4):1208-1218 [FREE Full text] [doi: 10.1007/s10461-016-1446-2] [Medline: 27260181]
43. Fullerton T, Rye B, Meaney G, Loomis C. Condom and hormonal contraceptive use by young women: an information-motivation-behavioral skills assessment. Can J Behav Sci 2013;45(3):196-209. [doi: 10.1037/a0033309]
44. Abdelaliem SMF, Kuaia AM, Hadadi AA, Alhujayri AK, Al Anazi AA, Hajar AA, et al. Knowledge and attitudes toward human papillomavirus and vaccination: a survey among nursing students in Saudi Arabia. Healthcare (Basel) 2023;11(12):1766 [FREE Full text] [doi: 10.3390/healthcare11121766] [Medline: 37372889]
45. Stout M, Christy S, Winger J, Vadaparampil S, Mosher C. Self-efficacy and HPV vaccine attitudes mediate the relationship between social norms and intentions to receive the HPV vaccine among college students. J Community Health 2020;45(6):1187-1195 [FREE Full text] [doi: 10.1007/s10900-020-00837-5] [Medline: 32418009]
46. Haun J, Luther S, Dodd V, Donaldson P. Measurement variation across health literacy assessments: implications for assessment selection in research and practice. J Health Commun 2012;17 Suppl 3:141-159 [FREE Full text] [doi: 10.1080/10810730.2012.712615] [Medline: 23030567]
47. Des Jarlais DC, Lyles C, Crepaz N, TREND Group. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. Am J Public Health 2004;94(3):361-366. [doi: 10.2105/ajph.94.3.361] [Medline: 14998794]
48. Almutairi SH, Alsedrah SA, Almutairi KS, Albayyahi YR, Almutairi OM, Alamri HM. The impact of education intervention on awareness of cervical cancer and barriers to HPV vaccination among Saudi pregnant women. Medical Science 2022;26(128):1-13. [doi: 10.54905/disssi/v26i128/ms404e2411]
49. Bouchhima N, Ammar M, Ksentini M, Zghal K, Hakim A, Mahmoud LB. Impact of an educational intervention video in HPV prevention among Tunisian female students: a before-and-after study. J Prev Med Hyg 2025;66(3):E279-E290. [doi: 10.15167/2421-4248/jpmh2025.66.3.3607] [Medline: 41292537]
50. Si M, Su X, Jiang Y, Wang W, Zhang X, Gu X, et al. An internet-based education program for human papillomavirus vaccination among female college students in Mainland China: application of the information-motivation-behavioral skills model in a cluster randomized trial. J Med Internet Res 2022;24(9):e37848 [FREE Full text] [doi: 10.2196/37848] [Medline: 36178723]
51. Si M, Su X, Jiang Y, Wang W, Zhang X, Gu X, et al. Effect of an IMB model-based education on the acceptability of HPV vaccination among college girls in mainland China: a cluster RCT. Cancer Control 2022;29:10732748211070719 [FREE Full text] [doi: 10.1177/10732748211070719] [Medline: 35088609]
52. Yi Y, Xiu S, Shi N, Huang Y, Zhang S, Wang Q, et al. Perceptions and acceptability of HPV vaccination among parents of female adolescents 9-14 in China: a cross-sectional survey based on the theory of planned behavior. Hum Vaccin Immunother 2023;19(2):2225994 [FREE Full text] [doi: 10.1080/21645515.2023.2225994] [Medline: 37340698]
53. Guarducci G, Chiti M, Fattore DC, Caldararo R, Messina G, Filidei P, et al. Development of a network model to implement the HPV vaccination coverage. Ann Ig 2024;36(6):636-643 [FREE Full text] [doi: 10.7416/ai.2024.2630] [Medline: 38647091]

XSL•FO
RenderX

Abbreviations

ASR: age-standardized incidence rate
HPV: human papillomavirus
IARC: International Agency for Research on Cancer
ICO: Information Centre on HPV and Cancer
IMB: information-motivation-behavioral skills
KR-20: Kuder-Richardson formula 20
LOCF: last observation carried forward
mHealth: mobile health
MICE: multivariate imputation by chained equations
MOH: Ministry of Health
OPTION: Observing Patient Involvement in Decision Making
PHC: primary health care
SDM: shared decision-making
SFDA: Saudi Food and Drug Authority
SMOG: Simple Measure of Gobbledygook
STI: sexually transmitted infection
TREND: Transparent Reporting of Evaluations with Nonrandomized Designs
VIF: variance inflation factor

Edited by Niranjani S; submitted 12.Jun.2026; peer-reviewed by G Guarducci, H Jangid, J Patel; comments to author 29.Jul.2026; revised version received 16.Aug.2026; accepted 29.Aug.2026; published 16.Sep.2026.

Please cite as:
Elsayed EA, Mersal FA, Mohamed HAO, Aljehani AM, Aljehani NM
Impact of a Culturally Tailored Educational Intervention on Human Papillomavirus Knowledge and Vaccination Intentions Among Women in Saudi Arabia: Quasi-Experimental Study
JMIR Public Health Surveill 2026;12:e104529
URL: https://publichealth.jmir.org/2026/1/e104529
doi:10.2196/104529
PMID:42748422

©Ebtesam A Elsayed, Fathia Ahmed Mersal, Heba Ahmed Osman Mohamed, Ala M Aljehani, Najla M Aljehani. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 16.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Prevalence, Trends, and Associated Factors of HIV, Syphilis, and Anti–Hepatitis C Virus Seropositivity Among Attendees at Voluntary Counseling and Testing Clinics in Zhejiang Province, China, From 2022 to 2024: Retrospective Quantitative Study

Jinlei Zheng[1], MM; Lin Chen[1], MM; Lin He[1], MM; Tingting Jiang[1], MM; Hui Wang[1], MD; Wanjun Chen[1], MM; Yapin Yao[1], MM

Department of HIV/AIDS and STDs Control and Prevention, Zhejiang Provincial Center for Disease Control and Prevention, Hangzhou, Zhejiang, China

Corresponding Author:
Lin Chen, MM
Department of HIV/AIDS and STDs Control and Prevention
Zhejiang Provincial Center for Disease Control and Prevention
3399th Binsheng Road
Binjiang District
Hangzhou, Zhejiang, 310051
China
Phone: 86 57187115149
Email: lchen@cdc.zj.cn

Abstract

Background: In the postpandemic era, trends in HIV, syphilis, and hepatitis C virus (HCV) infections, as well as the prevalence of coinfections among high-risk groups, may have shifted; however, relevant empirical studies are limited.

Objective: This study aimed to assess the seropositivity rates and coinfection rates for HIV, syphilis, and anti-HCV; and to explore the correlates of HIV, syphilis, and anti-HCV positivity among voluntary counseling and testing (VCT) clients at Centers for Disease Control and Prevention (CDC)–affiliated clinics in Zhejiang Province from 2022 to 2024.

Methods: This retrospective study analyzed routine data extracted in January 2025 from the National HIV Information Prevention and Control System. The Cochran-Armitage trend chi-square test was used to analyze trends in positivity rates over the years. Logistic regression was used to identify relevant indicators for positivity for HIV, syphilis, and anti-HCV in 2024.

Results: The HIV positivity rates decreased from 2.462% (743/30,182) in 2022 to 1.664% (681/40,916) in 2024 (χ^2_1 for trend =53.6; P<.001). The syphilis positivity rates were 1.378% (414/30,049), 1.502% (675/44,951), and 1.62% (647/39,929) in 2022, 2023, and 2024, respectively (χ^2_1 for trend=6.8; P=.03). The anti-HCV positivity rates were 0.309% (77/24,950), 0.387% (161/41,593), and 0.504% (192/38,098) in 2022, 2023, and 2024, respectively (χ^2_1 for trend=14.8; P<.001). In 2024, the positivity rates for dual HIV–anti-HCV coinfection, dual HIV-syphilis coinfection, and triple HIV-syphilis–anti-HCV coinfection were 0.011%, 0.203%, and 0.005%, respectively. Multivariable logistic regression analysis showed that age ≥35 years and HIV testing history were associated with HIV, syphilis, and anti-HCV positivity; compared with proactive consultation, outreach consultation was associated with lower odds of HIV positivity (adjusted odds ratio [aOR] 0.096, 95% CI 0.060-0.153), but higher odds of syphilis positivity (aOR 1.241, 95% CI 1.033-1.491) and anti-HCV positivity (aOR 2.837, 95% CI 1.882-4.276); compared with consulting for casual heterosexual behavior, consulting for homosexual behavior was associated with higher odds of HIV positivity (aOR 1.909, 95% CI 1.514-2.408) and lower odds of anti-HCV positivity (aOR 0.031, 95% CI 0.011-0.082).

Conclusions: From 2022 to 2024 in Zhejiang Province, the HIV and anti-HCV positivity rates among VCT clients at CDC-affiliated clinics declined, while the syphilis positivity rates increased. The coinfection positivity rates for these 3 infections remained at low levels. Individuals reporting homosexual behavior constituted a high-risk group for HIV positivity yet exhibited a lower positivity rate for anti-HCV. Among CDC-affiliated VCT clients, older age, lower educational attainment, and prior HIV testing history were associated with higher infection positivity rates, suggesting that these characteristics might help refine targeting of VCT and prevention efforts within this setting.

XSL•FO
RenderX

(JMIR Public Health Surveill 2026;12:e96584) doi:10.2196/96584

KEYWORDS

human immunodeficiency virus; HIV; hepatitis C virus; HCV; syphilis; human immunodeficiency virus–syphilis coinfection; HIV-syphilis coinfection; human immunodeficiency virus–hepatitis C virus coinfection; HIV-HCV coinfection

Introduction

To further strengthen the prevention and control of HIV-AIDS, the Joint United Nations Programme on HIV-AIDS set an ambitious 95-95-95 target for 2025: specifically, 95% of people within the subpopulation who are living with HIV know their HIV status, 95% of people living with HIV in key populations who know their HIV status are on treatment, and 95% of people within the subpopulations who are on treatment have suppressed viral loads [1]. Over the past decade, the treatment rates and viral suppression rates among people living with HIV in China have risen rapidly. Antiretroviral therapy (ART) coverage among diagnosed people living with HIV has approached 90%, whereas more than 94% of people living with HIV receiving treatment for at least 12 months have achieved viral suppression. However, the "detection rate of diagnosed cases" has remained <90% [2]. Consequently, HIV testing strategies have been continuously refined, with voluntary counseling and testing (VCT) serving as a crucial component.

HIV VCT refers to a process in which individuals who suspect they are at risk of HIV infection consult professionals and voluntarily undergo HIV testing along with related referrals and extended services with full informed consent and complete confidentiality [3]. Zhejiang Province has provided HIV VCT services since 2003 through facilities, including Centers for Disease Control and Prevention (CDC), general hospitals, and maternal and child health hospitals. As of 2024, there were 351 HIV VCT clinics in Zhejiang: 94 located in CDC facilities, 143 in general hospitals, 78 in maternal and child health hospitals, and 36 in community health service centers and township health centers, yielding a cumulative testing volume of approximately 90,000 person-times annually. In 2017, hepatitis C virus (HCV) antibody screening was incorporated into HIV VCT clinics, and in 2022, HCV testing was officially integrated into the HIV Testing and Counseling Information Management System.

HIV VCT plays a pivotal role in HIV-AIDS prevention and control [4-7]. According to an analysis by Chen et al [4], attendees of VCT and sexually transmitted disease clinics accounted for merely 4.5% of all individuals undergoing HIV testing, but they represented 23.7% of all HIV-positive cases identified in 2018. A study conducted in Turkey using an HIV transmission model projected that a 2-fold and 4-fold increase in the number of both VCT and other testing facilities would reduce the annual incidence by 71% and 91%, respectively, by 2033, whereas HIV prevalence was estimated to be 34% and 51% lower, respectively, than that in the scenario of taking no action [5].

With the release of guidelines by the World Health Organization and the rollout of national HCV elimination plans in China to eliminate the public health threat of HCV, coinfection with HIV, syphilis, and HCV has become a research focus [8-10]. Multiple studies have shown that the prognosis of patients with HIV coinfected with other diseases is poorer than that of patients with HIV monoinfection [11-14]. For instance, a study by Fan et al [12] demonstrated that patients with HIV-syphilis coinfection exhibited smaller increases in CD4+ T-cell counts and cluster of differentiation 4 (CD4)–cluster of differentiation 8 (CD8) ratios than patients with HIV monoinfection. HIV-HCV coinfection not only accelerates the progression of AIDS and increases the risk of liver-related end points but also affects the selection of ART regimens for AIDS, thereby compromising patient survival and prognosis [13,14]. Therefore, with the advancement of HIV-AIDS prevention and control efforts, the innovative strategy of "integrated prevention and control of multiple diseases" has been proposed globally [15].

To better characterize the prevalence of HIV, syphilis, and HCV monoinfection and coinfection among high-risk populations in Zhejiang Province, this study performed an in-depth analysis using data collected between 2022 and 2024 at CDC-affiliated VCT clinics. The dataset was large-scale and highly representative. These clinics were selected primarily because the data they generate are of higher quality than those obtained from other health care facilities. In addition, VCT clients at CDC-affiliated clinics are high-risk individuals who seek services proactively or are reached through outreach programs; these individuals exhibit multiple high-risk behaviors, making CDC-affiliated VCT clinics a critical data source for investigating the positivity rates of HIV, syphilis, and anti-HCV.

Methods

Study Population

Individuals who received HIV counseling and testing at CDC-affiliated VCT clinics in Zhejiang Province from 2022 to 2024 were included in this analysis. Individuals who received only counseling but did not undergo HIV testing were excluded. A total of 30,182, 46,060, and 40,916 individuals were enrolled in 2022, 2023, and 2024, respectively. Due to inconsistent implementation of multipathogen testing across VCT clinics, some VCT clinics did not offer syphilis or anti-HCV testing services during these 3 years: HIV test volumes were 30,182, 46,060, and 40,916; syphilis test volumes were 30,049, 44,951, and 39,929; and anti-HCV test volumes were 24,950, 41,593, and 38,098 in 2022, 2023, and 2024, respectively. The numbers of participants tested for HIV, syphilis, and anti-HCV during the same visit in 2022, 2023, and 2024 were 24,909, 41,563, and 38,065, respectively, accounting for 82.5%, 90.2%, and 93% of all VCT clients in each year.

Study Design and Data Collection

This was a retrospective analysis of repeated annual cross-sectional data from routine VCT surveillance. Data were obtained from the National HIV/AIDS Information Prevention and Control System. The VCT Clinics Database of Zhejiang

XSL•FO
RenderX

Province was downloaded in January 2025 and screened using the following criteria: (1) data submission years covered 2022 to 2024; (2) the reporting institution was a CDC facility; and (3) complete HIV test results were available. It was not until 2022 that syphilis and anti-HCV test results were added to VCT data collection nationwide. The dataset included submissions from 94 CDC sites in Zhejiang Province. Data were submitted by all 94 CDC-affiliated VCT clinics across the province, achieving full data coverage throughout the study period.

Measures

In this study, HIV positivity was defined as testing positive on a confirmed HIV test (Western blot confirmation); syphilis positivity was defined as testing positive for both specific and nonspecific antibodies (rapid plasma regain [RPR]–toluidine red unheated serum test [TRUST] and Treponema pallidum particle agglutination assay [TPPA]–enzyme-linked immunosorbent assay [ELISA]); and anti-HCV positivity was defined as testing positive for HCV antibodies (ELISA).

Relevant information on counseling region, gender, age, marital status, and educational background was analyzed to assess the sociodemographic characteristics of participants. The 11 counseling regions were divided into 3 groups based on the level of economic development: high-income regions included Hangzhou and Ningbo; middle-income regions included Wenzhou, Jiaxing, Shaoxing, and Taizhou; and low-income regions included Quzhou, Lishui, Zhoushan, and Huzhou.

Consultation types were divided into 3 categories: proactive, outreach, and referral consultation. Proactive consultation covers clients who voluntarily present at CDC-affiliated VCT clinics for HIV counseling and testing. Outreach consultation includes individuals who received counseling and testing during CDC on-site outreach activities. Referral consultation refers to clients who attend CDC-affiliated VCT clinics for HIV counseling and testing after being referred by other agencies or peers, such as medical institutions, family planning institutions, maternal and child health institutions, or community-based organizations (CBOs).

The primary reasons for consultation included a history of casual heterosexual behavior, homosexual behavior, blood exposure–related behaviors, having an HIV-positive intimate partner (spouse or regular partner), and injection drug use (IDU). Histories of plasma donation, surgery, and blood transfusion, or blood product use were grouped under blood exposure–related behaviors. The option “others” included a history of occupational exposure, maternal HIV positivity, no history of high-risk behavior, and other reasons not mentioned.

Statistical Analysis

Data were analyzed using SPSS (version 25.0; IBM Corp). Data from 2022 to 2024 were used to analyze the demographic characteristics and the positivity rates for HIV, syphilis, and anti-HCV; data from 2024 were used to analyze the predictors associated with these positivity rates. The calculation of positivity rates and influencing factors was based on the aforementioned data.

Descriptive analyses were used to describe the demographic characteristics of VCT clients at CDC-affiliated clinics, and statistical differences between groups were analyzed using the chi-square test. The Cochran-Armitage trend test or the chi-square test for trend was used for comparing proportions across years. Bar and line charts were used to display the test volumes, number of positive cases, and positivity rates for HIV, syphilis, and anti-HCV.

Missing values in marital status and education (approximately one-third of observations) were addressed using multiple imputation under the missing-at-random (MAR) assumption. Twenty imputed datasets were generated, incorporating the binary outcome variable and all covariates included in subsequent regression models. Univariable and multivariable binary logistic regression (LR) models (backward LR) were fitted separately on each imputed dataset. Pooled odds ratios (OR), adjusted odds ratios (aORs), 95% CIs, and *P* values were combined across imputed datasets using the Rubin rules, accounting for both within- and between-imputation variance. Complete-case analysis and an additional sensitivity analysis treating “unknown” as an independent category were performed to evaluate the robustness of the regression findings. $P<.05$ was considered to indicate statistical significance.

Ethical Considerations

The study was conducted in accordance with the Declaration of Helsinki, and all procedures were approved by the ethics committee of the Zhejiang Provincial Center for Disease Control and Prevention (2026-038-01). Ethics approval for the waiver of informed consent was granted because the data were routinely collected by province-wide VCT clinics for daily clinical service delivery, rather than collected specifically for a research project.

Results

A total of 30,182, 46,060, and 40,916 VCT clients were enrolled at CDC-affiliated clinics in Zhejiang Province in 2022, 2023, and 2024, respectively. In 2024, the proportions of individuals in high-income, middle-income, and low-income regions were 27.7% (n=11,335), 53% (n=21,671), and 19.3% (n=7910), respectively. Of these clients, 78.8% (n=32,228) were male, 52.1% (n=21,305) were aged 35 years, and 39.8% (n=16,267) were married. In total, 19.6% (n=8000) had a college education or higher. Of all respondents, 74.1% (n=30,307) received proactive consultation, 21.7% (n=8893) received outreach consultation, and 4.2% (n=1716) received referral consultation. For consultation reasons, 43% (n=17,601) consulted for homosexual behavior, followed by casual heterosexual behavior (n=11,972, 29.3%) and HIV-positive intimate partners (n=2077, 5.1%). Differences in these characteristics among the 3 years were statistically significant ($P<.05$; Table 1).

As shown in Figures 1 to 3, an analysis of the positivity rates for HIV, syphilis, and anti-HCV from 2022 to 2024 revealed the following: the HIV positivity rates decreased from 2.462% (743/30,182) in 2022 to 1.664% (681/40,916) in 2024, with a significant difference according to the Cochran-Armitage trend test (χ^2_1 for trend=53.6; $P<.001$; Figure 1). Between 2022 and 2024, the syphilis positivity rates were 1.378% (414/30,049),

XSL•FO
RenderX

1.502% (675/44,951), and 1.62% (647/39,929), respectively, with a significant difference according to the chi-square test for trend (χ^2_1 for trend=6.8; *P*=.03; Figure 2). The anti-HCV positivity rates were 0.309% (77/24,950), 0.387% (161/41,593), and 0.504% (192/38,098) in 2022, 2023, and 2024, respectively, with a significant difference according to the chi-square test for trend (χ^2_1 for trend=14.8; *P*<.001; Figure 3).

From 2022 to 2024, the positivity rates for dual HIV-syphilis coinfection were 0.3% (90/30,049), 0.249% (112/44,951), and 0.203% (93/39,929), respectively; those for dual HIV–anti-HCV coinfection were 0.012% (3/24,950), 0.002% (1/41,593), and 0.011% (4/38,098), respectively; and the positivity rates for triple HIV-syphilis–anti-HCV coinfection were 0.008% (2/24,909), 0% (0/41,563), and 0.005% (2/38,065), respectively (Table 2).

The results of univariable and multivariable analyses of factors associated with HIV, syphilis, and anti-HCV positivity are shown in Table 3.

Compared with high-income regions, middle-income regions had higher odds of HIV positivity (aOR 1.388, 95% CI 1.161-1.660) and anti-HCV positivity (aOR 3.568, 95% CI 1.924-6.615), while low-income regions had lower odds of HIV positivity (aOR 0.411, 95% CI 0.301-0.561). Female participants had a lower likelihood of HIV infection (aOR 0.511, 95% CI 0.385-0.676). The odds of HIV, syphilis, and anti-HCV positivity all rose with advancing age. Compared with unmarried clients, being married was associated with lower odds of HIV positivity (aOR 0.565, 95% CI 0.432-0.738) and higher odds of anti-HCV positivity (aOR 5.109, 95% CI 2.904-8.990).

Compared with participants with junior high school education or below, those with senior high school (aOR 0.455, 95% CI 0.356-0.583) or college education or above (aOR 0.646, 95% CI 0.515-0.811) had lower odds of HIV positivity; those with senior high school (aOR=0.665, 95% CI 0.541-0.818) or college education or above (aOR=0.734, 95% CI 0.584-0.922) had lower odds of syphilis positivity; and those with senior high school (aOR=0.448, 95% CI 0.254-0.792) or college education or above (aOR=0.392, 95% CI 0.228-0.672) had lower odds of anti-HCV positivity. Compared with proactive consultation, outreach consultation was associated with lower odds of HIV positivity (aOR 0.096, 95% CI 0.060-0.153), but higher odds of syphilis positivity (aOR 1.241, 95% CI 1.033-1.491) and anti-HCV positivity (aOR 2.837, 95% CI 1.882-4.276). Referral consultation was associated with higher odds of HIV positivity (aOR 1.632, 95% CI 1.249-2.131) and anti-HCV positivity (aOR 31.869, 95% CI 20.557-49.406).

Compared with participants seeking consultation for casual heterosexual behavior, those consulting for homosexual behavior had higher odds of HIV positivity (aOR 1.909, 95% CI 1.514-2.408) and lower odds of anti-HCV positivity (aOR 0.031, 95% CI 0.01-0.082). Participants consulting as intimate partners of HIV-positive individuals had higher odds of HIV positivity (aOR 3.526, 95% CI 2.546-4.882). Those consulting for IDU had higher odds of anti-HCV positivity (aOR 30.495, 95% CI 15.584-59.672).

Participants who had never been tested for HIV had higher odds of HIV positivity (aOR 1.455, 95% CI 1.208-1.754), syphilis positivity (aOR 1.294, 95% CI 1.105-1.516), and anti-HCV positivity (aOR 2.943, 95% CI 2.069-4.187).

XSL•FO
RenderX

Table 1. Sociodemographic characteristics of voluntary counseling and testing clients at Centers for Disease Control and Prevention–affiliated clinics from 2022 to 2024.

Variables	2022 (n=30,182), n (%)	2023 (n=46,060), n (%)	2024 (n=40,916), n (%)	χ^2 (df)	*P* value
Regions				201.7 (4)	<.001
High-income	9201 (30.5)	14,314 (31.1)	11,335 (27.7)		
Middle-income	14,769 (48.9)	23,364 (50.7)	21,671 (53)		
Low-income	6212 (20.6)	8382 (18.2)	7910 (19.3)		
Sex				328.9 (2)	<.001
Male	22,053 (73.1)	34,631 (75.2)	32,228 (78.8)		
Female	8129 (26.9)	11,429 (24.8)	8688 (21.2)		
Age (years)				252.1 (8)	<.001
16-24	5321 (17.6)	7788 (16.9)	6843 (16.7)		
25-34	10,273 (34)	15,914 (34.6)	12,768 (31.2)		
35-44	7061 (23.4)	10,506 (22.8)	9583 (23.4)		
45-54	4532 (15)	6638 (14.4)	6472 (15.8)		
≥55	2995 (9.9)	5214 (11.3)	5250 (12.8)		
Marital status				345.0 (6)	<.001
Unmarried	9196 (30.5)	13,635 (29.6)	12,501 (30.6)		
Married	12,304 (40.8)	16,716 (36.3)	16,267 (39.8)		
Divorced or widowed	1232 (4.1)	2016 (4.4)	1683 (4.1)		
Unknown	7450 (24.7)	13,693 (29.7)	10,465 (25.6)		
Educational background				334.5 (6)	<.001
Junior high school and under	10,301 (34.1)	13,487 (29.3)	12,910 (31.6)		
Senior high school	6466 (21.4)	9196 (20)	8367 (20.4)		
College and above	5320 (17.6)	8961 (19.4)	8000 (19.6)		
Unknown	8095 (26.8)	14,416 (31.3)	11,639 (28.4)		
Consultation types				775.6 (4)	<.001
Proactive	22,993 (76.2)	31,233 (67.8)	30,307 (74.1)		
Outreach	5882 (19.5)	12,531 (27.2)	8893 (21.7)		
Referral	1307 (4.3)	2296 (5)	1716 (4.2)		
Consultation reasons				2047.4 (10)	<.001
Casual heterosexual behavior	12,575 (41.7)	16,237 (35.2)	11,972 (29.2)		
Homosexual behavior	10,024 (33.3)	17,447 (37.9)	17,601 (43)		
Blood exposure–related behaviors	267 (0.9)	859 (1.9)	241 (0.6)		
Having an HIV-positive intimate partner	2008 (6.7)	2063 (4.5)	2077 (5.1)		
Injection drug use	194 (0.6)	307 (0.7)	74 (0.2)		
Others	5114 (16.9)	9147 (19.8)	8951 (21.9)		
HIV testing history				211.1 (2)	<.001
No	17,149 (56.8)	26,452 (57.4)	21,601 (52.8)		
Yes	13,033 (43.2)	19,608 (42.6)	19,315 (47.2)		

XSL•FO
RenderX

Figure 1. Trends in HIV test volume, HIV-positive cases, and HIV positivity rate among voluntary counseling and testing clients at Centers for Disease Control and Prevention–affiliated clinics in Zhejiang Province from 2022 to 2024, together with the results of the chi-square test for trend analysis.

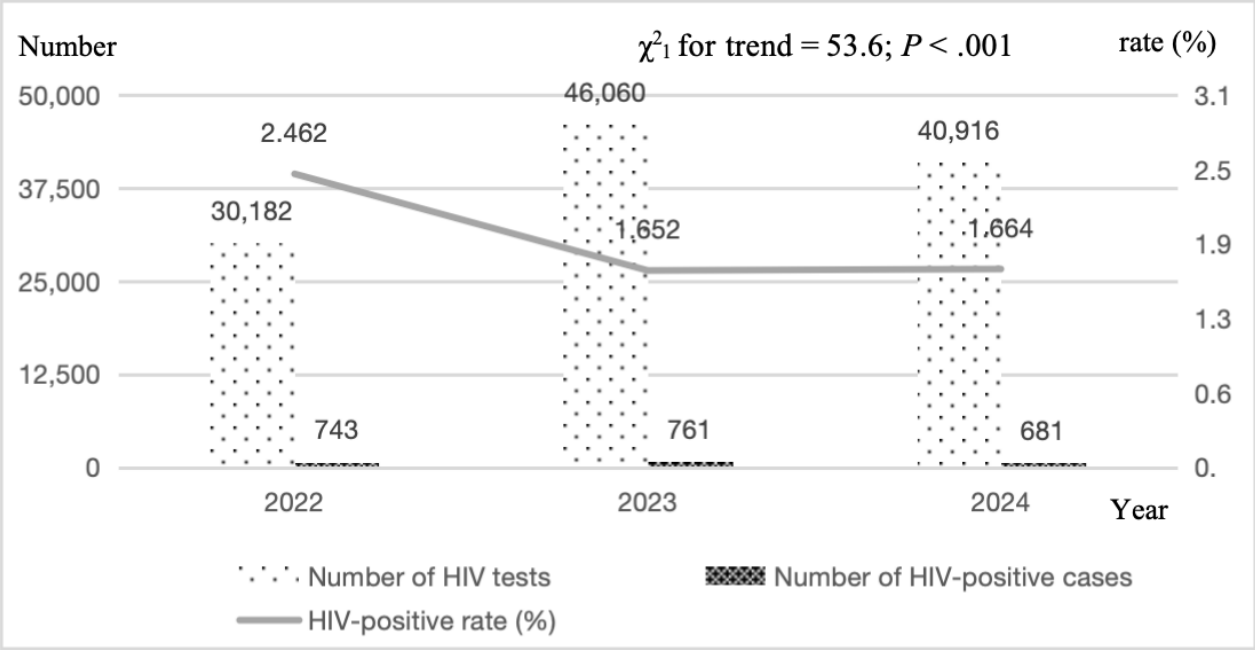
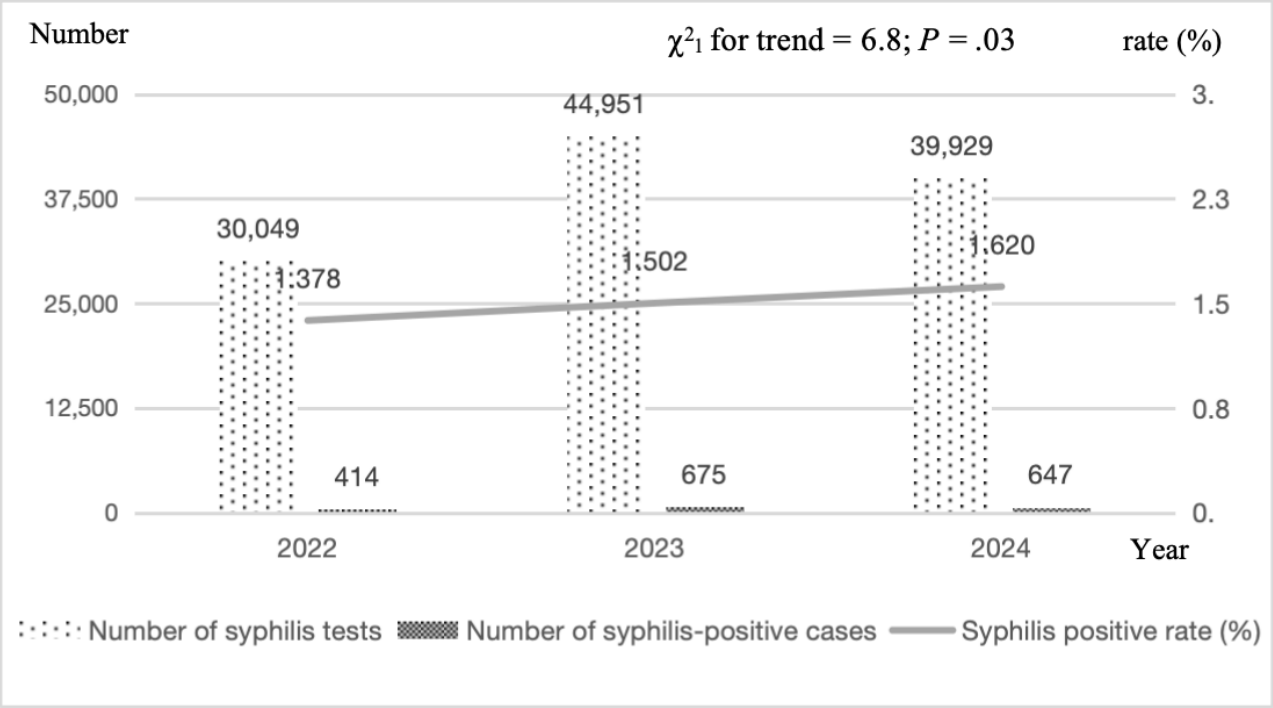


Figure 2. Trends in syphilis testing volume, syphilis-positive cases, and syphilis positivity rate among voluntary counseling and testing clients at Centers for Disease Control and Prevention–affiliated clinics in Zhejiang Province from 2022 to 2024. The statistical significance of the temporal trends was examined using the chi-square test for trend.



XSL•FO
RenderX

Figure 3. Trends in anti–hepatitis C virus (anti-HCV) test volume, anti-HCV–positive cases, and anti-HCV positivity rates among voluntary counseling and testing clients at Centers for Disease Control and Prevention–affiliated clinics in Zhejiang Province from 2022 to 2024. The statistical significance of the temporal trends was examined using the chi-square test for trend.

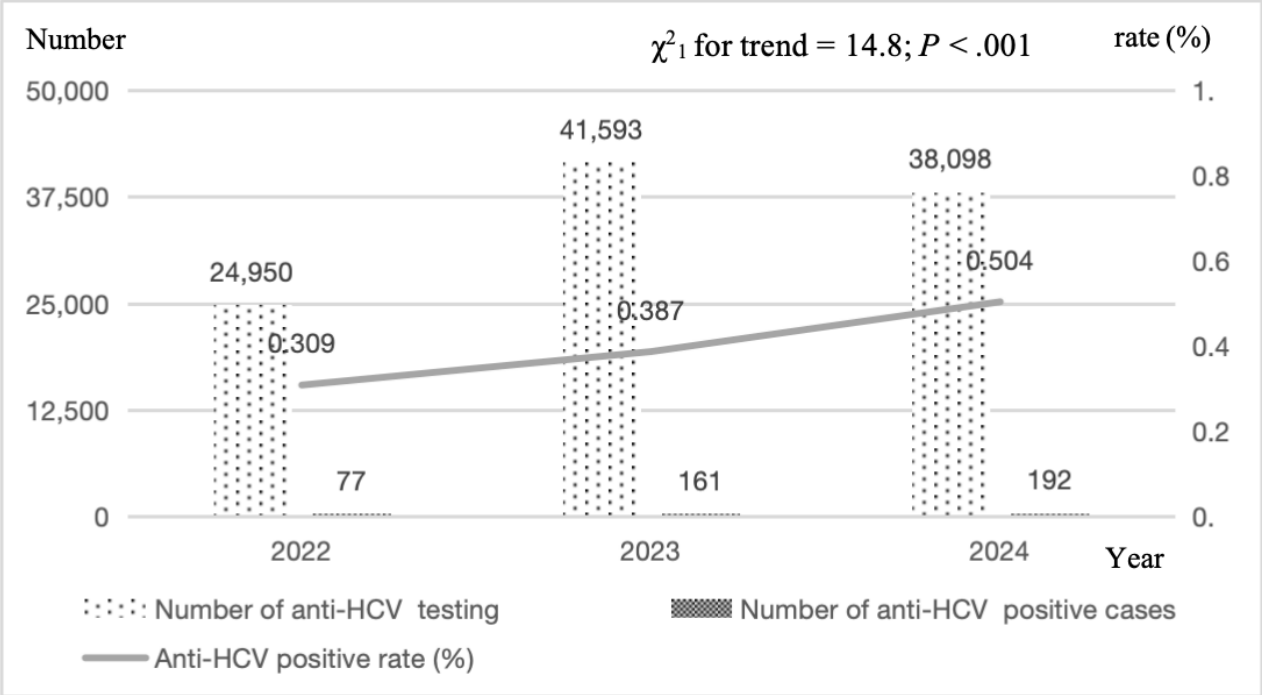


Table 2. Positivity rates for HIV, syphilis, and anti–hepatitis C virus (anti-HCV) coinfections among voluntary counseling and testing clients at Centers for Disease Control and Prevention–affiliated clinics from 2022 to 2024.

Year	HIV and anti-HCV			HIV and syphilis			HIV, syphilis, and anti-HCV		
	Testing number	Positive number	Positivity rate (95% CI)	Testing number	Positive number	Positivity rate (95% CI)	Testing number	Positive number	Positivity rate (95% CI)
2022	24,950	3	0.012 (0.003-0.035)	30,049	90	0.300 (0.240-0.360)	24,909	2	0.008 (0.001-0.029)
2023	41,593	1	0.002 (0.000-0.013)	44,951	112	0.249 (0.206-0.301)	41,563	0	0 (0)
2024	38,098	4	0.011 (0.003-0.027)	39,929	93	0.203 (0.189-0.295)	38,065	2	0.005 (0.001-0.019)

XSL•FO
RenderX

Table 3. Univariate and multivariable logistic regression analyses of the predictors associated with HIV, syphilis, and anti–hepatitis C virus (anti-HCV) positivity among voluntary counseling and testing clients at Centers for Disease Control and Prevention–affiliated clinics in 2024. Italicized values indicate 95% CIs that do not include 1.

Variables	HIV positivity			Syphilis positivity			Anti-HCV positivity		
	Positivity rate, n/N (%)	OR[a] (95% CI)	aOR[b] (95% CI)	Positivity rate, n/N (%)	OR (95% CI)	aOR (95% CI)	Positivity rate, n/N (%)	OR (95% CI)	aOR (95% CI)
Counseling regions									
High-income	233/11,335 (2.1)	1	1	181/11,311 (1.6)	1	—[c]	12/9,892 (0.1)	1	1
Middle-income	397/21,671 (1.8)	0.889 (0.755-1.047)	*1.388 (1.161-1.660)*	341/20,853 (1.6)	1.022 (0.852-1.226)	—	174/20,486 (0.8)	*7.053 (3.927-12.666)*	*3.568 (1.924-6.615)*
Low-income	51/7,910 (0.6)	*0.309 (0.228-0.419)*	*0.411 (0.301-0.561)*	125/7,765 (1.6)	1.006 (0.800-1.266)	—	6/7,720 (0.1)	0.640 (0.240-1.707)	0.648 (0.237-1.771)
Sex									
Male	601/32,228 (1.9)	1	1	490/31,459 (1.6)	1	—	136/29,895 (0.5)	1	1
Female	80/8,688 (0.9)	*0.489 (0.387-0.618)*	*0.511 (0.385-0.676)*	157/8,470 (1.9)	1.194 (0.996-1.431)	—	56/8,203 (0.7)	1.504 (0.101-2.055)	*0.677 (0.476-0.963)*
Age (years)									
16-34	290/19,611 (0.7)	1	1	234/19,086 (1.2)	1	1	12/18,199 (0.1)	1	1
35-54	272/16,055 (1.7)	1.148 (0.972-1.357)	*1.329 (1.038-1.703)*	290/15,676 (1.8)	*1.518 (1.277-1.806)*	*1.541 (1.224-1.940)*	140/14,883 (0.9)	*14.392 (7.978-25.962)*	*10.360 (5.629 - 19.069)*
≥55	119/5,250 (2.3)	*1.545 (1.246-1.917)*	*1.690 (1.235-2.312)*	123/5,167 (2.4)	*1.965 (1.576-2.449)*	*1.966 (1.473-2.624)*	40/5,016 (0.8)	*12.183 (6.387-23.241)*	*6.224 (3.154-12.281)*
Marital status									
Unmarried	254/12,501 (2)	1	1	172/12,332 (1.4)	1	1	11/11,767 (0.1)	1	1
Married	270/16,267 (1.7)	*0.807 (0.682-0.955)*	*0.565 (0.432-0.738)*	262/16,014 (1.6)	*1.219 (1.021-1.454)*	0.803 (0.638-1.011)	39/14,922 (0.3)	*5.109 (2.904-8.990)*	—
Divorced or widowed	56/1,683 (3.3)	*1.562 (1.167-2.089)*	1.153 (0.803-1.657)	41/1,677 (2.4)	*1.615 (1.161-2.245)*	1.047 (0.584-1.473)	9/1,642 (0.6)	*5.943 (2.688-13.140)*	—
Educational background									
Junior high school and under	307/12,910 (2.4)	1	1	268/12,697 (2.1)	1	1	48/11,973 (0.4)	1	—
Senior high school	93/8,367 (1.1)	*0.480 (0.381-0.606)*	*0.455 (0.356-0.583)*	95/8,257 (1.2)	*0.624 (0.508-0.767)*	*0.665 (0.541-0.818)*	8/7,824 (0.1)	*0.448 (0.254-0.792)*	—
College and above	165/8,000 (2.1)	0.909 (0.756-1.092)	*0.646 (0.515-0.811)*	93/7,910 (1.2)	*0.594 (0.477-0.739)*	*0.734 (0.584-0.922)*	3/7,387 (0.04)	*0.392 (0.228-0.672)*	—
Consultation types									
Proactive	594/30,307 (2)	1	1	459/29,607 (1.6)	1	1	56/27,815 (0.2)	1	1
Outreach	19/8,893 (0.2)	*0.107 (0.068-0.169)*	*0.096 (0.060-0.153)*	163/8,618 (1.9)	*1.224 (1.022-1.466)*	*1.241 (1.033-1.491)*	54/8,579 (90.6)	*3.140 (2.159-4.567)*	*2.837 (1.882-4.276)*

XSL•FO
RenderX

Variables	HIV positivity			Syphilis positivity			Anti-HCV positivity		
	Positivity rate, n/N (%)	OR[a] (95% CI)	aOR[b] (95% CI)	Positivity rate, n/N (%)	OR (95% CI)	aOR (95% CI)	Positivity rate, n/N (%)	OR (95% CI)	aOR (95% CI)
Referral	68/1,716 (4)	*2.064 (1.598-2.666)*	*1.632 (1.249-2.131)*	25/1,704 (1.5)	0.946 (0.630-1.418)	0.952 (0.634-1.430)	82/1,704 (4.8)	*25.060 (17.776-35.329)*	*31.869 (20.557-49.406)*
Consultation reasons									
Casual heterosexual behavior	126/11,972 (1.1)	1	1	189/11,823 (1.6)	1	—	33/11,570 (0.3)	1	1
Homosexual behavior	385/17,601 (2.2)	*2.102 (1.717-2.574)*	*1.909 (1.514-2.408)*	264/17,024 (1.6)	0.970 (0.803-1.170)	—	5/15,713 (0)	*0.111 (0.043-0.285)*	*0.031 (0.011-0.082)*
Having an HIV-positive intimate partner	89/2,077 (4.3)	*4.209 (3.195-5.544)*	*3.526 (2.546-4.882)*	45/2,034 (2.2)	1.393 (1.003-1.934)	—	7/1,980 (0.4)	1.240 (0.548-2.808)	0.826 (0.350-1.950)
Blood exposure–related behaviors	3/241 (1.2)	1.183 (0.374-3.751)	1.615 (0.505-5.167)	1/235 (0.4)	0.263 (0.037-1.885)	—	0/184 (0)	N/A[d]	N/A
Injection drug use	0/74 (0)	—	—	3/73 (4.1)	2.638 (0.823-8.453)	—	27/73 (37)	*205.204 (114.291-368.434)*	*30.495 (15.584-59.672)*
Others	78/8,951 (0.9)	0.826 (0.622-1.098)	1.005 (0.748-1.351)	145/8,740 (1.7)	1.038 (0.835-1.292)	—	120/8,578 (1.4)	*4.960 (3.371-7.299)*	*2.591 (1.675-4.006)*
HIV testing history									
No	253/21,601 (1.2)	1	1	311/21,162 (1.5)	1	1	69/20,679 (0.3)	1	1
Yes	428/19,315 (2.2)	*1.912 (1.635-2.236)*	*1.455 (1.208-1.754)*	336/18,767 (1.8)	*1.222 (1.046-1.428)*	*1.294 (1.105-1.516)*	123/17,41 (90.7)	*2.124 (1.581-2.854)*	*2.943 (2.069-4.187)*

[a]OR: odds ratio.

[b]aOR: adjusted odds ratio.

[c]These variables were not included in the model.

[d]N/A: not available. These data were not available because the odds ratio could not be estimated due to 0 events or model convergence issues.

Discussion

Prevalence of HIV, Syphilis, and Anti-HCV and Coinfection

From 2022 to 2024, the HIV positivity rates among VCT clients at CDC-affiliated clinics in Zhejiang Province decreased from 2.462% to 1.664%. This positivity rate was higher than the national average of 0.82% for VCT clinics in 2022 [16]. A study conducted in Chiang Mai Province, Thailand, found that the HIV positivity rates among individuals attending fixed-site and mobile VCT services were 5.6% and 2.6%, respectively [17]. The findings of this study indicated an overall decreasing trend in the newly reported HIV positivity rates, which were consistent with the downward trend in the overall HIV epidemic in Zhejiang Province.

This study showed that the syphilis positivity rates among VCT clients at CDC-affiliated clinics increased from 1.378% in 2022 to 1.62% in 2024. The reported incidence of syphilis in Zhejiang Province was 35.23 per 100,000 in 2022, with a decreasing trend from 2016 to 2022 [18]. The positivity rates observed in this study were higher than those reported in other regions of China and in studies conducted in the United States (2016-2023) and the United Kingdom [19-21].

The results of this study demonstrated that the anti-HCV positivity rates among VCT clients at CDC-affiliated clinics increased from 0.309% in 2022 to 0.504% in 2024. This rate was lower than the 1.8% reported for VCT clinics in Beijing [22] and the 1.5% reported by Conway [23] among CDC clients at community-based clinics across multiple European countries in 2018. The positivity rates for anti-HCV in Zhejiang Province showed a decreasing trend, with rates of 5.13 per 100,000, 3.97 per 100,000, and 3.61 per 100,000 in 2022, 2023, and 2024, respectively.

A study conducted among female sex workers in South China reported that the positivity rates for dual syphilis–anti-HCV, HIV-syphilis, HIV–anti-HCV, and triple HIV-syphilis–anti-HCV coinfection were 30 per 100,000, 0%, 0%, and 0%, respectively [24]. The probability of coinfection varied across different populations. In contrast, the prevalence of dual HIV-HCV coinfection was 39.1% among people who

XSL•FO
RenderX

inject drugs in Vietnam, 8.5% among prison inmates in India, and 4% among drug users in the United States [25-27]. In Mexico, the national positivity rates for syphilis, HIV, and their dual coinfection among men who have sex with men were 15.2%, 10.2%, and 5.7%, respectively [28]. In the present study, the positivity rates for triple HIV-syphilis–anti-HCV coinfection among VCT clients at CDC-affiliated clinics were extremely low, with wide 95% Clopper-Pearson CIs, indicating substantial statistical imprecision. Therefore, the results should be interpreted with caution.

Possible Explanations for Prevalence and Coinfection

The variation in the overall testing volume during the study period was significantly influenced by the COVID-19 pandemic. The peak testing volume observed in 2023 might be attributable to COVID-19 prevention and control restrictions implemented throughout 2022, which potentially biased the observed temporal trends in positivity rates for HIV and syphilis [29].

The increase in the number of VCT clients who engaged in male-male sexual behaviors at CDC-affiliated clinics in this study could be attributed to the integration of counseling and testing services for men who have sex with men provided by CBOs into the VCT information system in 2024. Although the number of men who have sex with men undergoing testing increased, the HIV positivity rates did not rise, which may be attributable to the generally stable HIV prevalence among key populations.

The rising trend on the syphilis epidemic may be associated with the economic prosperity of eastern China. The impact of HIV postexposure prophylaxis (PEP) and preexposure prophylaxis (PrEP) on syphilis infection warrants consideration. Recently, the coverage of PEP and PrEP services has expanded continuously. Currently, all 11 prefecture-level cities in Zhejiang Province have established 24-hour PEP clinics, and advancements in internet technology have provided robust support for PrEP service delivery [30]. However, evidence from available data is insufficient and warrants further research.

The generally low positivity rates for anti-HCV and concurrent HIV and anti-HCV infection could be attributed to the relatively small proportion of people who inject drugs among VCT clients at CDC-affiliated clinics, along with the notable effectiveness of harm reduction programs, such as clean needle exchange programs and methadone maintenance therapy in Zhejiang Province [31]. Conclusions regarding coinfection patterns should be interpreted cautiously and restricted solely to VCT attendees.

Correlates of HIV, Syphilis, and Anti-HCV Positivity Among VCT Clients at CDC-Affiliated Clinics

In the present study, an inverse association was observed between educational background and the risk of HIV, syphilis, and anti-HCV positivity. The association between education and infectious disease outcomes has been extensively examined in prior research, with a consistent protective effect reported across most studies [32]. Individuals with higher levels of education exhibited greater health consciousness, more access to health information, and higher use of health services and resources, which may contribute to a reduced risk of infection with the aforementioned infectious diseases.

Compared with individuals who actively sought services at CDC-affiliated VCT clinics, those who received referral-based testing had a higher risk of positivity for HIV–anti-HCV coinfection. In contrast, individuals receiving counseling and testing through outreach services had a lower risk of positivity for HIV-syphilis coinfection but a higher risk of anti-HCV positivity. Individuals who voluntarily attended CDC-affiliated VCT clinics for testing generally exhibited stronger risk awareness and self-protection capabilities, as well as higher levels of awareness regarding condom use [33]. On the other hand, referred individuals typically had lower risk awareness than those who actively sought services yet still engaged in high-risk behaviors, and this led to a higher infection rate. Outreach services are generally targeted toward hard-to-reach populations, such as community-based drug abstainers and those under supervision in correctional facilities, such as compulsory drug detoxification centers. This may explain the relatively high HCV infection rate observed in this subgroup.

Male homosexual behavior and having HIV-positive intimate partners were associated with HIV infection among VCT clients at CDC-affiliated clinics in our province. As our province had low HIV prevalence, sentinel surveillance data revealed that the HIV positivity rate among this population reached 4%, making it the high-risk group with the highest infection rate in the province [34]. Spousal notification and testing constituted a core component of HIV prevention and control efforts in the province. The HIV positivity rate was as high as 30% among spouses, which was attributed to delayed diagnosis of index patients [35].

Our study showed that individuals with an HIV testing history had a higher risk of HIV, syphilis, and HCV infection. A history of HIV testing was associated with risky behaviors and psychological traits; for example, people with high health awareness and those engaging in more high-risk behaviors are more likely to undergo HIV testing [36]. Moreover, a high frequency of high-risk behaviors was a significant risk factor for these 3 diseases, which may explain this result.

Programmatic Implications for VCT and Integrated Disease Control

The findings of this study yield several implications for intervention practices. First, testing for HIV, syphilis, and HCV in VCT clinics offers important data for epidemic surveillance. However, some VCT attendees at CDC clinics were not tested for syphilis and HCV, and anti-HCV–positive individuals failed to receive nucleic acid testing, which hinders epidemic monitoring. Screening strategies for syphilis and anti-HCV, as well as HCV nucleic acid testing, need to be improved. Second, this study found that VCT clients recruited through outreach and referral had higher positivity rates for HIV, syphilis, and anti-HCV compared with walk-in clients. Therefore, VCT clinics should scale up outreach and referral testing services. CBOs and high-risk behavior intervention teams, which have access to hard-to-reach high-risk populations, should be engaged to promote VCT services. Third, although the findings of this study cannot draw definitive conclusions, we still suggest delivering more targeted and individualized interventions during counseling sessions for the older adults, male participants, unmarried

XSL•FO
RenderX

individuals, people with low educational attainment, and populations identified via referral and outreach services, to reduce their risk of seroconversion.

Limitations

There were many limitations in this study. First, follow-up nucleic acid testing results for individuals who screened positive for HCV were unavailable, which limited accurate estimation of the actual HCV infection rate among VCT clients at CDC-affiliated clinics. Second, from the perspective of factor exploration, this study was based on an analysis of existing data without a rigorous scientific theoretical hypothesis. This may have led to the omission of key confounding factors. Third, the participants were not recruited via random sampling, which limited the generalizability of the study findings to the general population or to all key populations. Fourth, testing services provided by VCT clinics are anonymous, so duplicate individual records may exist in theory. Although the large sample size can mitigate their impact on analytical results, we still recommend excluding duplicate cases in future studies. Finally, the MAR assumption could not be empirically verified, which was noted as a study limitation. Thus, future research with scientifically rigorous study designs is warranted, building on the key findings of the present study.

Conclusions

From 2022 to 2024 in Zhejiang Province, the HIV and anti-HCV positivity rates among VCT clients at CDC-affiliated clinics declined, while the syphilis positivity rates increased. The coinfection positivity rates for these 3 infections remained at low levels. Individuals reporting homosexual behavior constituted a high-risk group for HIV yet exhibited a lower anti-HCV positivity rate. Among CDC-affiliated VCT clients, older age, lower educational attainment, and prior HIV testing history were associated with higher infection positivity rates, suggesting that these characteristics might help refine targeting of VCT and prevention efforts within this setting. It is suggested that VCT clinics enhance syphilis and HCV testing, expand outreach and referral services, and deliver targeted counseling and interventions for populations at high risk of infection.

Acknowledgments

The authors acknowledge the contributions of the staff of Centers for Disease Control and Prevention–affiliated voluntary counseling and testing clinics across the province, as well as the personnel from community-based organizations who contributed to data collection.

The generative AI tool used in the manuscript was Doubao (version 2.69.9_mac; ByteDance Ltd), which was used for manuscript proofreading and editing.

Funding

This research was supported by the Zhejiang Science and Technology Plan for Disease Prevention and Control (grants 2026JKZ009 and 2025JK154).

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' Contributions

JZ designed the study, performed the statistical analysis, and drafted the manuscript. LC performed the statistical analysis and reviewed and revised the manuscript. LH, TJ, HW, WC, and YY were responsible for on-site verification and data quality control. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Understanding measures of progress towards the 95–95–95 HIV testing, treatment and viral suppression targets. UNAIDS. 2024 Mar 11. URL: https://www.unaids.org/en/resources/documents/2024/progress-towards-95-95-95 [accessed 2024-05-10]
2. Zhao Y, Han M, Ma Y, Li D. Progress towards the 90-90-90 targets for controlling HIV - China, 2018. China CDC Wkly 2019 Nov;1(1):5-7 [FREE Full text] [Medline: 34594590]
3. Coovadia HM. Access to voluntary counseling and testing for HIV in developing countries. Ann N Y Acad Sci 2000 Nov;918:57-63. [doi: 10.1111/j.1749-6632.2000.tb05474.x] [Medline: 11131735]
4. Chen L, Luo M, Xu Y, Xia Y, Zhou X, Chen W, et al. The first 90: progress in HIV detection in Zhejiang Province, 2008-2018. PLoS One 2021 Apr 8;16(4):e0249517 [FREE Full text] [doi: 10.1371/journal.pone.0249517] [Medline: 33831067]
5. Erdogan ZM, Yaylali E. Expanded voluntary counselling and testing services in Türkiye: insights from an HIV model. BMC Public Health 2025 Oct 08;25(1):3402 [FREE Full text] [doi: 10.1186/s12889-025-24350-3] [Medline: 41063025]

XSL•FO
RenderX

6. Yigezu A, Alemayehu S, Hamusse SD, Ergeta GT, Hailemariam D, Hailu A. Cost-effectiveness of facility-based, stand-alone and mobile-based voluntary counseling and testing for HIV in Addis Ababa, Ethiopia. Cost Eff Resour Alloc 2020 Sep 11;18:34 [FREE Full text] [doi: 10.1186/s12962-020-00231-x] [Medline: 32944006]
7. Mlughu TS, Anaeli A, Joseph R, Sirili N. Voluntary HIV counseling and testing among commercial motorcyclist youths: an exploration of ethical challenges and coping mechanisms in Dar es Salaam. HIV AIDS (Auckl) 2020 Nov 03;12:675-685 [FREE Full text] [doi: 10.2147/HIV.S259997] [Medline: 33177884]
8. Li J, Pang L, Liu Z. Interpretation of the National Action Plan for Eliminating Hepatitis C as a Public Health Threat (2021-2030). China CDC Wkly 2022 Jul 22;4(29):627-630 [FREE Full text] [doi: 10.46234/ccdcw2022.139] [Medline: 35919827]
9. Aydoğan S, Kirca F, Gozalan A, Aslan FG, Gürler M, Toyran A, et al. Epidemiology of hepatitis B, hepatitis C, and syphilis co-infections in HIV-1 patients: a retrospective cross-sectional study of prevalence and viral load correlates. Ann Saudi Med 2025;45(5):345-352 [FREE Full text] [doi: 10.5144/0256-4947.2025.28.08.1204] [Medline: 40856990]
10. Doğan S, Mohamud SM, Ali AM, Orhan Z, Doğan A, Mohamud RY. Seroprevalence of hepatitis C, hepatitis B, HIV and syphilis among blood donors at a tertiary care hospital in Mogadishu-Somalia in 2020-2022: a retrospective study. BMC Public Health 2024 Sep 10;24(1):2463 [FREE Full text] [doi: 10.1186/s12889-024-19837-4] [Medline: 39256725]
11. Du M, Wang R, Yuan J, Lv X, Yan W, Liu Q, et al. Trends and disparities in 44 national notifiable infectious diseases in China: an analysis of national surveillance data from 2010 to 2019. J Med Virol 2023 Jan;95(1):e28353 [FREE Full text] [doi: 10.1002/jmv.28353] [Medline: 36443103]
12. Fan L, Yu A, Zhang D, Wang Z, Ma P. Consequences of HIV/syphilis co-infection on HIV viral load and immune response to antiretroviral therapy. Infect Drug Resist 2021 Jul 24;14:2851-2862 [FREE Full text] [doi: 10.2147/IDR.S320648] [Medline: 34335031]
13. Yakubu A, Hali B, Maiyaki AS. Prevalence and risk factors for hepatitis C virus co-infection among human immunodeficiency virus-infected patients and effect of hepatitis C virus infection on acquired immunodeficiency syndrome cases at baseline. Ann Afr Med 2021;20(4):297-301 [FREE Full text] [doi: 10.4103/aam.aam_65_20] [Medline: 34893569]
14. Zhou K, Hu F, Wang C, Xu M, Lan Y, Morano JP, et al. Genotypic distribution and hepatic fibrosis among HIV/HCV co-infected individuals in Southern China: a retrospective cross-sectional study. BMC Infect Dis 2015 Sep 30;15:401 [FREE Full text] [doi: 10.1186/s12879-015-1135-1] [Medline: 26424404]
15. Saayman E, Hechter VB, Sishwili NS, Sonderup MW. A decentralized point of service triple prevention and treatment model for hepatitis B, C and HIV in people who inject drugs in South Africa. Int J Drug Policy 2025 Sep;143:104894 [FREE Full text] [doi: 10.1016/j.drugpo.2025.104894] [Medline: 40578045]
16. Liu Y, Su R, Li D, Wang S, Han M. Temporal and spatial trends in HIV positivity rate for VCT clinics - China, 2015-2022. China CDC Wkly 2024 Jan 05;6(1):6-11 [FREE Full text] [doi: 10.46234/ccdcw2024.002] [Medline: 38223658]
17. Nanthaprut P, Manojai N, Chanlearn P, Mattawanon N, Chiawkhun P, Homkham N, et al. Comparison of HIV-positive incidence among transgender women and men who have sex with men at stand-alone and mobile voluntary counseling and testing facilities in Chiang Mai Province, Thailand. AIDS Patient Care STDS 2021 Apr;35(4):116-125. [doi: 10.1089/apc.2020.0258] [Medline: 33835852]
18. Wu H, Xue M, Wu C, Lu Q, Ding Z, Wang X, et al. Scaling law characteristics and spatiotemporal multicomponent analysis of syphilis from 2016 to 2022 in Zhejiang Province, China. Front Public Health 2023 Oct 25;11:1275551 [FREE Full text] [doi: 10.3389/fpubh.2023.1275551] [Medline: 37965512]
19. Slurink IA, Götz HM, van Aar F, van Benthem BH. Educational level and risk of sexually transmitted infections among clients of Dutch sexual health centres. Int J STD AIDS 2021 Oct;32(11):1004-1013. [doi: 10.1177/09564624211013670] [Medline: 33993803]
20. Liang Y, Menzies NA, Rönn MM. Trends in syphilis prevalence by race and ethnicity among people who are pregnant in the United States, 2016 to 2023. Sex Transm Dis 2025 Sep 01;52(9):535-541. [doi: 10.1097/OLQ.0000000000002187] [Medline: 40417976]
21. Richardson D, Pickering A, Nichols K, Buss Z, Trotman D, Fitzpatrick C, et al. Factors associated with testing positive for syphilis among MSM who present as sexual contacts of syphilis from a clinic-based population. Sex Transm Infect 2022 Mar;98(2):136-138. [doi: 10.1136/sextrans-2020-054878] [Medline: 33875566]
22. Xia YH, Chen W, Tucker JD, Wang C, Ling L. HIV and hepatitis C virus test uptake at methadone clinics in Southern China: opportunities for expanding detection of bloodborne infections. BMC Public Health 2013 Sep 30;13:899 [FREE Full text] [doi: 10.1186/1471-2458-13-899] [Medline: 24079351]
23. Conway A, Fernàndez-López L, Reyes-Urueña J, Casabona J, COBATEST Study Group. Hepatitis C screening in community-based voluntary counselling and testing services in Europe: an observational study from the COBATEST Network 2014-2018. J Community Health 2020 Jun;45(3):606-614 [FREE Full text] [doi: 10.1007/s10900-019-00780-0] [Medline: 31749115]
24. Xia M, Zhang XB, Zhang ZY, Guo Y, Liu CT, Chen HC, et al. Co-incidence of sexually transmitted diseases in female sex workers in Yunnan Province, 2023 [Article in Chinese]. Zhonghua Liu Xing Bing Xue Za Zhi 2025 Nov 10;46(11):1981-1987. [doi: 10.3760/cma.j.cn112338-20250422-00270] [Medline: 41443712]

XSL•FO
RenderX

25. Thinh VT, Li L, Matthieu D, Hoa VD, Anh NH, Giang LM. HCV and HIV co-infection among people who inject drugs in Vietnam. J Health Soc Sci 2020 Dec;5(4):573-586 [FREE Full text] [Medline: 34109283]
26. Singh GP, Lata S, Swu AK, Virk NS, Singh J, Thakkar S. A retrospective study to find the prevalence of HIV, HCV and dual HIV-HCV infection in the prison inmates. J Family Med Prim Care 2022 Oct;11(10):6250-6254 [FREE Full text] [doi: 10.4103/jfmpc.jfmpc_788_22] [Medline: 36618149]
27. Chapin-Bardales J, Asher A, Broz D, Teshale E, Mixson-Hayden T, Poe A, et al. Hepatitis C virus infection and co-infection with HIV among persons who inject drugs in 10 U.S. cities-National HIV Behavioral Surveillance, 2018. Int J Drug Policy 2025 Oct;144(Pt 1):104387. [doi: 10.1016/j.drugpo.2024.104387] [Medline: 38531730]
28. Vergara-Ortega DN, Tapia-Maltos A, Herrera-Ortíz A, García-Cisneros S, Olamendi-Portugal M, Sánchez-Alemán M. High prevalence of syphilis and syphilis/HIV coinfection among men who have sex with men who attend meeting places in Mexico. Pathogens 2023 Feb 21;12(3):356 [FREE Full text] [doi: 10.3390/pathogens12030356] [Medline: 36986278]
29. Weng R, Kwon JA, Hammoud M, Clifton B, Scott N, McGregor S, et al. Evaluating the impact of COVID-19 on the HIV epidemic among MSM in Australia. AIDS 2025 Dec 01;39(15):2286-2294. [doi: 10.1097/QAD.0000000000004344] [Medline: 40948424]
30. Ge R, Chai C, Zhu G, He L, Chen Z. Low pre-exposure prophylaxis (PrEP) uptake and high prevalence of sexually transmitted infections among men who have sex with men in Zhejiang Province, China: a large cross-sectional survey. BMC Public Health 2025 Nov 29;26(1):54 [FREE Full text] [doi: 10.1186/s12889-025-25807-1] [Medline: 41318494]
31. Liu Y, Li L, Zhang Y, Zhang L, Shen W, Xü H, et al. Assessment of attitudes towards methadone maintenance treatment between heroin users at a compulsory detoxification centre and methadone maintenance clinic in Ningbo, China. Subst Abuse Treat Prev Policy 2013 Aug 04;8:29 [FREE Full text] [doi: 10.1186/1747-597X-8-29] [Medline: 23915360]
32. Odimegwu CO, Imo CK, Amoo EO. HIV voluntary counselling and testing and behaviour changes among youths in Nigeria. J Biosoc Sci 2020 May;52(3):366-381. [doi: 10.1017/S0021932019000506] [Medline: 31409439]
33. Derebew B, Mola M, Kalyankar VB, Padwal ND, Chauhan NM, Humbe AS, et al. Determination of knowledge, attitude and practice of voluntary counseling testing on HIV among youths from Tepi Town, Ethiopia. PEC Innov 2022 Nov 15;2:100102. [doi: 10.1016/j.pecinn.2022.100102] [Medline: 37214491]
34. Liu W, Chen X, Wang S, Wang Y, Lin H, Yao Y. Trends of sexual behaviors and related factors in high-risk MSM from 2010 to 2023: a repeated cross-sectional study in Zhejiang, China. BMC Public Health 2025 Dec 18;25(1):4264 [FREE Full text] [doi: 10.1186/s12889-025-25517-8] [Medline: 41413811]
35. Jiang J, Chen L, Cheng W, Chen W, Yang J, Xu Y, et al. Characteristics of and factors associated with partner service uptake cascade among people with newly reported HIV/AIDS diagnoses in Southeastern China in 2022: cross-sectional survey. JMIR Public Health Surveill 2024 Sep 09;10:e59095 [FREE Full text] [doi: 10.2196/59095] [Medline: 39250196]
36. Awopegba OE, Ologunowa TO, Ajayi AI. HIV testing and self-testing coverage among men and women in South Africa: an exploration of related factors. Trop Med Int Health 2021 Feb;26(2):214-227 [FREE Full text] [doi: 10.1111/tmi.13514] [Medline: 33159363]

Abbreviations

aOR: adjusted odds ratio
ART: antiretroviral therapy
CBO: community-based organization
CD4: cluster of differentiation 4
CD8: cluster of differentiation 8
CDC: Centers for Disease Control and Prevention
ELISA: enzyme-linked immunosorbent assay
HCV: hepatitis C virus
IDU: injection drug use
LR: logistic regression
MAR: missing-at-random
OR: odds ratio
PEP: postexposure prophylaxis
PrEP: preexposure prophylaxis
RPR: rapid plasma regain
TPPA: Treponema pallidum particle agglutination assay
TRUST: toluidine red unheated serum test
VCT: voluntary counseling and testing

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 31.Mar.2026; peer-reviewed by D Pilalas, H Kahsay; comments to author 16.Jun.2026; revised version received 25.Aug.2026; accepted 26.Aug.2026; published 29.Sep.2026.

Please cite as:
Zheng J, Chen L, He L, Jiang T, Wang H, Chen W, Yao Y
Prevalence, Trends, and Associated Factors of HIV, Syphilis, and Anti–Hepatitis C Virus Seropositivity Among Attendees at Voluntary Counseling and Testing Clinics in Zhejiang Province, China, From 2022 to 2024: Retrospective Quantitative Study
JMIR Public Health Surveill 2026;12:e96584
URL: https://publichealth.jmir.org/2026/1/e96584
doi:10.2196/96584
PMID:

©Jinlei Zheng, Lin Chen, Lin He, Tingting Jiang, Hui Wang, Wanjun Chen, Yapin Yao. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 29.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Critical Health Literacy Podcast to Counter Health Misinformation at Scale: Randomized Controlled Trial

Vanesa Mora Ringle[1], PhD; Amanda Jensen-Doss[2], PhD

[1]Department of Education and Human Services, College of Education, Lehigh University, 111 Research Drive, Bethlehem, PA, United States
[2]Department of Psychology, College of Arts and Sciences, University of Miami, Coral Gables, FL, United States

Corresponding Author:
Vanesa Mora Ringle, PhD
Department of Education and Human Services, College of Education, Lehigh University, 111 Research Drive, Bethlehem, PA, United States

Abstract

Background: Widespread misinformation and low critical health literacy pose major barriers to public health worldwide. Rapid, scalable, and evidence-informed digital interventions are urgently needed to strengthen the public's ability to make informed health decisions.

Objective: Informed by critical health literacy frameworks, we developed and tested a brief, story-based critical thinking podcast, *Parents Making Informed Health Choices*, that was designed to improve critical health literacy and decision-making among US parents.

Methods: We conducted a 2-phase study. First, 5 parents participated in the user testing of the prototype podcast and provided qualitative feedback to refine content and delivery. The final podcast delivered 9 evidence-based practice principles through relatable scenarios about mental and physical health. In the second phase, we conducted a 2-arm randomized controlled trial (N=250) with a national online sample of US parents. Participants were randomly assigned to listen to either the critical thinking podcast (n=128, 51.2%) or a control podcast (n=122, 48.8%). There were no significant preintervention group differences except for age, which was controlled for in all analyses. Primary outcomes included critical thinking about health claims, intended health behaviors, attitudes toward evidence-based mental health practices, and treatment preferences.

Results: On average, parents were aged 35 (SD 7.8) years; 49% (121/247) were female, 75% (185/248) were White; and 60.0% (148/248) had a bachelor's degree or higher. Parents who listened to the critical thinking podcast demonstrated significantly improved critical thinking about health information compared to the control group (B=2.56; $P<.001$; $\Delta R^2=0.06$). They also reported stronger critical thinking–aligned intended behaviors (B=0.252; $P=.001$; $\Delta R^2=0.015$), and more evidence-informed treatment preferences (B=4.89; $P=.038$; $\Delta R^2=0.02$). The effect sizes were small to moderate across outcomes.

Conclusions: Findings suggest that a brief, story-based digital podcast can meaningfully improve critical thinking about health information, intended behaviors, and evidence-based practice attitudes. Podcasts represent a promising, low-cost, and scalable strategy for promoting critical health literacy and countering health misinformation in the general public.

(JMIR Public Health Surveill 2026;12:e78003) doi:10.2196/78003

KEYWORDS

critical health literacy; evidence-based practice; EBP; health misinformation; digital health intervention; public health education; randomized controlled trial; RCT; parents

Introduction

Background

Health misinformation is now widely recognized as a public health threat [1]. During the COVID-19 pandemic and beyond, inaccurate health claims about vaccines and treatments have spread rapidly across digital platforms, undermining public trust and distorting health-related decision-making. Parents, in particular, face a deluge of conflicting health messages as they seek information about their children's physical and mental health needs. However, few public health interventions directly aim to strengthen parents' ability to critically assess health claims and make evidence-informed choices. Furthermore, scalable, low-intensity interventions that leverage popular media formats, such as podcasts, represent an underexplored avenue for enhancing critical health literacy at the population level. Accordingly, we developed and tested the effectiveness of a brief, story-based podcast intervention aimed at improving critical health literacy among US parents.

Many individuals lack the critical thinking skills necessary to make empowered and well-informed health choices [2,3]. Low

XSL•FO
RenderX

critical thinking about health can lead to believing and acting on unreliable health information that is easily accessible by word of mouth, through mass media, and the internet [4-6]. Basing mental and physical health treatment decisions on unreliable information may lead to pursuing ineffective or potentially harmful services [7], delaying access to effective services (evidence-based practices [EBPs]), increasing risks, and wasting resources.

Unfortunately, few studies have focused on addressing the health-related critical thinking needs of the public, especially in the United States. Grounded in the Informed Health Choices Key Concepts framework, critical thinking about health extends beyond functional health literacy (ie, basic knowledge of health conditions and services) and involves applying EBP principles to access, understand, and evaluate sources of health information [8-10]. Studies across the world, even among high-income and highly educated populations, have found that the public engages in low levels of critical thinking about health [2,11-13].

Effective decision-making behavior in health care is complex, comprising not only critical thinking skills but also related attitudes and beliefs, among other characteristics and factors. Existing research has found that US adults and adolescents value scientific evidence when making health decisions; however, they have a limited understanding of what it means for practices to be evidence-based [14-16]. Additionally, US adults report that there are situations in which they would trust their doctor's opinion over scientific research [17]. Issues of public trust in medicine and medical research are particularly prominent in relation to vaccines [18]. Indeed, despite their demonstrated efficacy, safety, and benefits, antivaccination or vaccine hesitancy attitudes persist and have resulted in adverse public health consequences (eg, increase in outbreaks of diseases that were under control or eradicated, such as measles) [18,19]. As such, examining critical thinking about health alongside attitudinal factors such as vaccine hesitancy is an important next step toward deepening our understanding of public health literacy and health decision-making.

Considering the potential impact of increasing the public's ability to engage in critical thinking about health, it is crucial to develop and test brief, targeted, and scalable critical thinking interventions. This is especially important for parents, since they are responsible for health care decisions for both themselves and their children. Information and communication technologies play a crucial role in promoting critical health literacy by expanding access to information, fostering interactive learning environments, and enabling individuals to engage more actively with health content. Podcasts, in particular, offer a promising delivery method for addressing these challenges. They are widely accessible, cost-effective, and increasingly used for health communication and education [20-22]. As defined by the World Health Organization, digital health interventions are discrete uses of digital technologies, such as podcasts, to achieve health objectives [23]. Prior research and communication theory suggest that storytelling and narrative formats can improve engagement and promote behavior change, especially when combined with principles of EBP and tailored to target audiences [24]. However, few studies have tested the use of podcasts as a critical thinking intervention for health, and even fewer have done so using rigorous randomized designs.

Only one prior study has evaluated the effectiveness of such an intervention and focused on parents in Uganda [25]. Semakula et al [25] developed a 10-episode critical thinking podcast based on the needs of the Ugandan families. Via a randomized controlled trial (RCT) of 675 participants, the educational podcast was found to significantly improve parents' critical thinking compared to those who only listened to a public service announcement. Moreover, improvements in critical thinking were observed across varying levels of educational attainment, including those who had only completed elementary education. Unfortunately, this intervention has yet to be adapted for and tested in the United States, and it solely focused on physical health issues, without any mental health examples.

While educating the public is essential across all areas of health care, it is especially urgent in the context of mental health, where treatment decisions are often complex and nuanced. Public discourse in this area is frequently shaped by unscientific or pseudoscientific claims, which can contribute to misinformation and reduce public trust in evidence-based care. This environment poses a significant barrier to individuals' ability to make informed decisions and access appropriate mental health services [26-30]. Despite this, no existing critical health literacy intervention has specifically addressed decision-making in mental health care.

This Study

The ability to critically think about health claims and treatment effectiveness enables the public to make well-informed health care decisions. By evaluating the effectiveness of treatments, the public can avoid potentially harmful or ineffective care and be empowered to take an active role in their health management. Thus, the primary aim of this study was to develop and test whether a brief critical thinking podcast increased critical thinking about health and related intended behaviors in comparison to an active control. The current intervention was adapted from the previously validated podcast used in Uganda [25] with cultural and contextual tailoring for a US-based audience. Additionally, because the original Ugandan podcast focused solely on physical health conditions, the current podcast expands its scope to include decision-making scenarios related to mental health. Given that this was the first evaluation of a critical health literacy podcast in the United States, the study also served as a proof-of-concept assessment to determine whether the approach would translate effectively in a different cultural and health communication context. A secondary aim was to examine the effects of the podcast on EBP attitudes, vaccine safety concerns, and mental health treatment preferences. We hypothesized that the critical thinking podcast would have a positive effect on all measured outcomes.

Methods

Design and Procedures

The study was conducted in the United States and included both community-based and online components. It consisted of two

XSL•FO
RenderX

phases: (1) podcast development and user testing, and (2) podcast evaluation through an online RCT.

Podcast development included reviewing the scripts and storyboards of the Ugandan podcast by Semakula et al [25], and determining the needed adaptations (eg, changing examples about malaria). Communication theory on learning through entertainment informed the podcast design to ensure that messaging was not merely didactic, and there was a sense of relatability in the storylines and characters [24]. The first author and undergraduate research assistants scanned US mass media websites, such as Health News Review, for examples of mental health care claims relevant to US parents. For example, one of the episodes covers a claim regarding vaccines causing autism. Once the podcast script was finalized, it was produced by the Orange Umbrella, a production company housed within the university's School of Communication.

We then conducted user testing where 5 parents from South Florida listened to the podcast prototype and provided feedback. A sample size of 5 is common in usability research where small samples are sufficient to identify most usability issues. For user testing, we recruited parents through the local chapter of the National Alliance on Mental Illness, advertisements in public community spaces, and Facebook posts. Eligible participants were aged at least 18 years, had at least 1 child aged less than 18 years, and were fluent in English. Participants were compensated with a US $15.00 gift card for their time and feedback. User-testing sessions lasted approximately 60 to 90 minutes, including 32 minutes of estimated individual podcast listening, a 20 to 30-minute semistructured interview, and a 5-minute online satisfaction survey. Their feedback was then reviewed and applied, and the podcast was finalized for testing through an online RCT. The podcast is available in SoundCloud [31].

A total of 250 parents participated in the RCT in May 2019 through Amazon Mechanical Turk (MTurk) [32,33]. A priori power analysis indicated that this sample size would provide adequate power to detect a large effect. As previously demonstrated by Jensen-Doss et al [33], MTurk is an appropriate space for recruiting parents online, especially for mental health care–related studies such as this one where capturing the perspectives of potential treatment seekers may be pertinent. The study was advertised to MTurk workers in the United States who had an MTurk approval rating of 98% or higher and were not participating in any other related study [3]. In the RCT, participants first provided consent and completed a baseline questionnaire of demographic and other characteristics. As in the original trial of the *Informed Health Choices* podcast by Semakula et al [25], outcome measures were completed after random assignment and after listening to the podcasts. Exposing participants to questionnaires prior to listening to the podcasts would have affected podcast listening, and it was expected that randomization would ensure that participants were equivalent at baseline (see the Randomization Success section). Following completion of baseline measures, participants were randomly assigned on a 1:1 ratio to either the critical thinking podcast or a control podcast. Participants in both conditions could not move on from the podcast listening page until 32 minutes (length of podcasts) had lapsed and they also could not fast forward. We also built in several "listening fidelity" checks (described in the Measures section). After they listened to the entire podcast, participants completed postintervention measures. Participants received a payment of US $7.25, in line with minimum wage at the time of the study. This study followed the CONSORT (Consolidated Standards of Reporting Trials) guidelines (Checklist 1).

Podcast Intervention and Control Condition

The *Parents Making Informed Health Choices* podcast [31] consisted of nine 2 to 4-minute episodes, including an introduction episode and a conclusion episode. The entire podcast was 32 minutes long. The 7 main episodes covered 9 different EBP principles using medical and mental health conditions (eg, depression). Following the storylines developed for the Uganda *Informed Health Choices* podcast, the US podcast included 2 main characters who engaged in back-and-forth conversations while explaining and applying EBP principles to different medical and mental health conditions. Table 1 lists the EBP principles covered and provides examples of health care claims. The control podcast (a 33-minute-long mindfulness meditation) served as an inert condition that was similar in length to the critical thinking podcast, but included no content related to critical thinking.

XSL•FO
RenderX

Table . Evidence-based practice (EBP) principles [10] addressed in the *Parents Making Informed Health Choices* podcast. Examples are compared to those in the original *Informed Health Choices* podcast used in Uganda [25].

EBP principle	US claim examples	Uganda claim examples
Treatments should be compared.	Elderberry is an effective treatment for child flu.	Quail eggs make you very strong.
Treatments should be compared fairly.	Cognitive-mental therapy for youth anxiety and depression works.	Group support is helpful for someone who is depressed.
Findings from small studies can be misleading.	We can know that vaccines cause autism based on information from one small study.	Washing hands with soap does not stop children from getting diarrhea.
Association is not the same as causation[a].	Contraceptive pills cause women to gain weight[a].	A lot of women gain weight when they take contraceptive pills.
Expert opinion is not always right[a].	Contraceptive pills cause women to gain weight[a].	Eating some hot pepper will heal ulcers.
Anecdotes are unreliable evidence.	Butter can heal burns.	Putting cooking oil on a burn will heal it.
Treatments might be harmful[a].	An herbal treatment for ADHD[b] with no side effects exists[a].	Quinine can cure malaria. It can also give you nausea and make you vomit.
Treatments have benefits and harms.[a]	Herbal treatment for ADHD with no side effects exists[a].	Herbal medicines exist for malaria treatment that cure malaria and do not have any bad effects.
Common practice does not mean it is beneficial or safe.	Physical discipline is the best strategy for managing child behavior problems.	An herbal treatment called kyogero stops babies from getting infections.

[a]Covered within the same episode of the *Parents Making Informed Health Choices Podcast*.

[b]ADHD: attention-deficit hyperactivity disorder.

Measures

Demographics and Other Characteristics

Participants' age, gender, ethnicity, educational attainment, employment status, podcast listening habits or consumption (listening time in min), and other characteristics data (eg, treatment-seeking history) were collected.

Critical Thinking About Health

We assessed critical thinking about health using 18 items from the Claim Evaluation Tools, which consists of over 100 multiple choice questions that can be used with people aged 10 years and older [34,35]. We selected 18 questions that tapped into 9 different key concepts based on their cultural relevance to the United States and how well they performed in a validation study [34]. We also included 3 new questions that were worded exactly like Claim Evaluation Tools' questions but replaced physical health conditions with mental health conditions. The overall critical thinking score was the number of correct responses out of 21 items. The internal consistency of the final 21-item critical thinking measure was good (α=.86).

Intended Behavior

Intended behavior related to critical thinking was assessed through 3 items created and used by Semakula et al [25]. The questions asked about the likelihood that someone will (1) find out what a treatment claim is based on, (2) find out if a claim is based on a research study comparing the treatment to no treatment (a fair comparison), and (3) saying "yes" if asked to participate in a research study comparing 2 treatments for an illness they have. Response options were on a 4-point Likert scale from very unlikely to very likely, and also included an "I don't know" option.

Attitudes Toward EBPs in Mental Health

The Consumer Attitudes Towards Evidence Based Services Scale (CAEBS) [36] consists of 29 items that load onto 4 factors: (1) beliefs regarding therapists' practices, (2) attitudes about mental health policy, (3) negative personal-level attitudes toward EBPs, and (4) negative societal-level attitudes towards EBPs. Items are rated on a 5-point scale from "strongly disagree" to "strongly agree." Internal consistencies in this sample were as follows: factor 1, α=.77; factor 2, α=.59; factor 3, α=.88; and factor 4, α=.73. To increase reliability, we removed item 13 from the factor 2 scale (new α=.73), and item 24 from the factor 4 scale (new α=.80).

Attitudes Regarding Empiricism in Mental Health Treatment

We used 5 items created by Kirk et al [37] that assessed attitudes regarding empiricism in mental health treatment on a 5-point Likert scale, with higher scores indicating more agreement. The internal consistency of this 5-item scale in this sample was adequate (α=.77).

Treatment Preferences

We assessed treatment preferences using a measure developed by Kirk et al [37] that asked participants to rate how important various mental treatment components were to them by allocating 99 points across the following: (1) treatment being evidence-based, (2) therapeutic alliance, (3) therapist experience, (4) empathic qualities of the therapist, and (5) client speaking for the majority of sessions [37]. Participants were instructed that more points signify higher preference for that treatment component.

XSL•FO
RenderX

Vaccine Safety Concerns

Following prior studies [38,39], we examined vaccine safety concerns through three items: (1) "vaccines are unsafe," (2) "vaccines have long-term negative side effects," and (3) "If I had another infant today, I do not want him/her to get all the recommended vaccinations." Participants rated these on a 5-point scale with higher scores indicating more negative vaccine attitudes. The internal consistency of this scale in this sample was α=.93.

Podcast Listening Fidelity

We assessed participant attention to the podcast through 7 questions about podcast content details that anyone who listened to the entire podcast should be able to answer, with specific questions for each study condition. Questions 1 to 6 were true or false, and question 7 required an open-ended response to a question about content at the end of the podcast. Responses to question 7 were assigned 0 to 2 points by the principal investigator; thus, participants could earn up to 8 points total.

Podcast Satisfaction

Participants in both conditions also rated their satisfaction with the podcast through 5 items that asked about overall satisfaction, likelihood of recommending to others, and relevance.

Ethical Considerations

All procedures were approved by the University of Miami Institutional Review Board (20180596). The submitted protocol document was the one submitted to the institutional review board and received approval. There were no deviations from the protocol. Participants provided informed consent electronically prior to participating in the study. Data were collected anonymously through secure online surveys (Qualtrics) and analyzed in deidentified form. All data were stored on password-protected, encrypted servers to ensure participant privacy and confidentiality. The 5 user-testing participants were compensated with a US $15.00 gift card for their time and feedback. RCT participants received a payment of US $7.25, in line with minimum wage.

Analytic Plan

Analyses were run using SPSS (version 28; IBM Corp). Descriptive statistics examined participant demographics, podcast satisfaction, and measures of data quality. We used multiple measures of data quality, including length of study participation (anticipating it would take at least 60 minutes), attention check questions, and 2 questions measuring consistency in the report of demographic information (age and youngest child age) strategically placed at the beginning and at the end of the study. Analyses checking randomization success examined differences in demographic variables between the 2 study conditions. We tested hypotheses about podcast effects using multiple linear regression for continuous variables and binomial logistic regression for categorical variables. The Cohen *d* values of 0.20, 0.50, and 0.80 indicated small, medium, and large effects, respectively; and R^2 values of 0.01, 0.09, and 0.25 indicated small, medium, and large effects, respectively [40].

Results

Participants

Five parents participated in the user-testing phase of the study, and 250 parents participated in the online RCT. On average, parents in the RCT were aged 35 (SD 7.8) years with 49% (121/147) being female, 75% (185/248) being White, and 60.0% (148/248) having a bachelor's degree or higher. Table 2 provides additional sample characteristics.

XSL•FO
RenderX

Table . Demographic characteristics of US parents who participated in the study's user-testing and randomized controlled trial phases (N=250). Values in italics in the same row differed significantly (*P*<.05).

Characteristic	User-testing (n=5)	Online RCT[a] (N=250)	Critical thinking podcast (n=128)	Control podcast (n=122)
Age (y), mean (SD; range)	36.40 (6.31; 30 - 43)	34.98 (7.8; 20 - 77)	*33.81 (7.5; 20-59)*	*36.20 (8.0; 23 - 77)*
Sex (female)[c], n (%)	5 (100)	121 (49.0)	59 (48.8)	62 (51.2)
Race or ethnicity[bc], n (%)				
African American	0 (0)	41 (16.5)	19 (14.8)	22 (18.3)
American Indian or Alaska Native	0 (0)	2 (0.8)	1 (0.8)	1 (0.8)
Asian	1 (20)	13 (5.2)	4 (3.1)	9 (7.5)
Hispanic	2 (40)	35 (14.1)	14 (10.9)	21 (17.5)
White	3 (60)	185 (74.6)	101 (78.9)	84 (70)
Highest level of education achieved[c], n (%)				
Some high school, no diploma	0 (0)	1 (0.4)	1 (0.8)	0 (0)
High school	0 (0)	29 (11.7)	14 (10.9)	15 (12.5)
Some college, no degree	0 (0)	44 (17.7)	22 (17.2)	22 (18.3)
Associate's or technical degree	0 (0)	26 (10.5)	12 (9.4)	14 (11.7)
Bachelor's degree	2 (40)	123 (49.6)	69 (56.1)	54 (43.9)
Master's degree	2 (40)	23 (9.32)	10 (7.8)	13 (10.8)
Doctoral or other graduate degree	0 (0)	2 (0.8)	1 (0.8)	1 (0.8)
Employment status[c], n (%)				
Currently working	5 (100)	213 (85.9)	115 (89.8)	98 (81.7)
Unemployed	0 (0)	9 (3.6)	3 (2.4)	6 (5)
Retired	0 (0)	2 (0.8)	1 (0.8)	1 (0.8)
Homemaker	0 (0)	22 (8.9)	8 (6.3)	14 (11.7)
Student or other	0 (0)	2 (0.8)	1 (0.8)	1 (0.8)
Annual household income[c] (US $), n (%)				
Less than 19,999	—[d]	20 (8.0)	11 (8.6)	9 (7.5)
20,000 to 39,999	—	56 (22.4)	34 (26.6)	22 (18.3)
40,000 to 59,999	—	64 (25.6)	33 (25.8)	31 (25.9)
60,000 to 79,999	—	55 (22)	21 (16.4)	34 (28.4)
80,000 to 99,999	—	24 (9.6)	12 (9.4)	12 (10)
More than 100,000	—	29 (11.6)	17 (13.3)	12 (10)
Insured[c], n (%)	—	220 (88)	114 (89.1)	106 (86.9)
Children (n), mean (SD; range)	2.20 (0.84; 1 - 3)	1.73 (0.94; 1 - 6)	1.69 (0.96; 1 - 6)	1.77 (0.91; 1 - 6)
Mental health history and service use[c], n (%)				
Mental health diagnosis	—	109 (43.6)	52 (40.6)	57 (47.5)
Ever received psychotherapy for a psychological problem	—	114 (45.6)	54 (42.2)	60 (50.0)

XSL•FO
RenderX

Characteristic	User-testing (n=5)	Online RCT[a] (N=250)	Critical thinking podcast (n=128)	Control podcast (n=122)
Ever taken medication for a psychological problem	—	91 (36.4)	42 (32.8)	49 (40.8)
Podcast listening habits[c] (min), n (%)				
<30	—	71 (28.4)	33 (25.8)	38 (31.4)
30 - 60	—	84 (33.6)	43 (33.6)	41 (33.9)
61 - 90	—	42 (16.8)	26 (20.3)	16 (13.2)
>90	—	52 (20.8)	26 (20.3)	26 (21.5)

[a]RCT: randomized controlled trial.

[b]These percentages do not sum to 100 due to overlap across categories.

[c]Data missing (n=1-3).

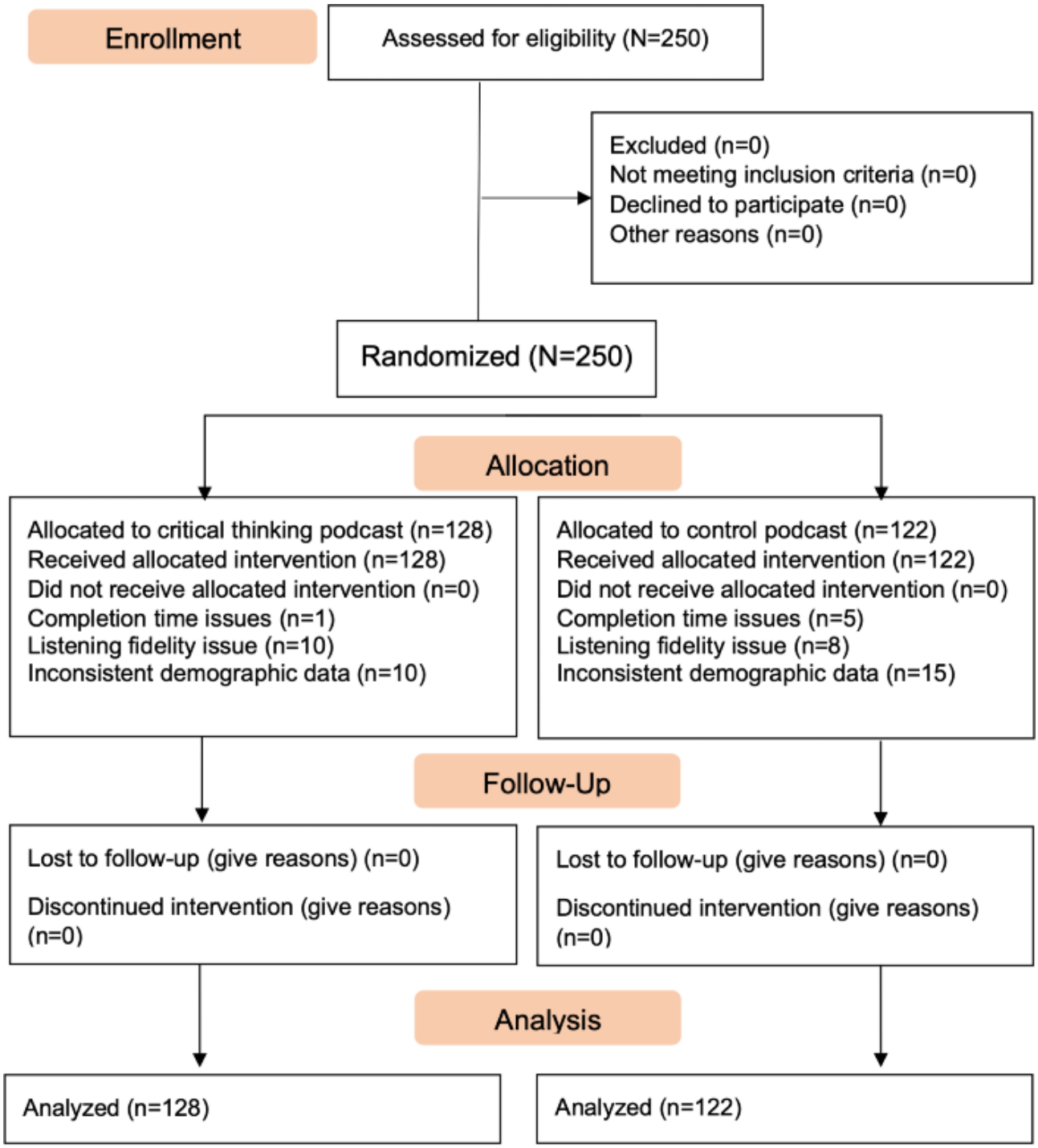
[d]Not applicable.

Analysis of Data Quality

Out of 250 participants, 6 (2.4%) completed the study under 42 minutes (70% of the projected completion time); 18 (7.2%) failed attention check questions or obtained a score of 4 or lower on the listening fidelity measure; and 25 (10%) provided inconsistent demographic data. Consistent with an intent-to-treat approach, all randomized participants who completed study procedures were included in the primary study analyses (Figure 1). For completeness, we report exploratory post hoc analyses excluding the 49 participants with data quality issues at the end of this section. Parents in the critical thinking podcast condition obtained an average score of 6.73 (SD 1.68; range 1 - 8) on the listening fidelity test and parents in the control condition earned an average score of 6.76 (SD 0.84; range 1 - 8).

XSL•FO
RenderX

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) diagram of participant flow through the *Parents Making Informed Health Choices Podcast* randomized controlled trial. This figure depicts participant flow for main study analyses where all randomized participants were retained in accordance with an intent-to-treat approach; exploratory post hoc analyses excluding 49 participants with data quality issues (indicated in the allocation section of the figure) are reported separately.



Randomization Success

There were no significant differences between the critical thinking podcast condition and control condition at preintervention based on gender (χ^2_1=0.679; *P*=.41), education level (χ^2_1=0.88; *P*=.35) annual household income (t_{248}=0.720; *P*=.47), podcast listening habits (t_{248}=–0.744; *P*=.46), and podcast listening fidelity scores (t_{248}=0.155; *P*=.88). However, the control group was significantly older, by 2 years (t_{248}=2.43; *P*=.02); therefore, we controlled for age in all analyses. Given some variability and to reduce the possibility that significant relationships were due to third variables, we also included education and listening fidelity as control variables in outcome analyses. Table 3 includes descriptive statistics for study measures, including means and SDs for this sample.

XSL•FO
RenderX

Table . Descriptive data for study variables by condition and overall sample (N=250). Values in italics in the same row differed significantly at *P*<.05 based on regression analyses controlling for age, education, and listening fidelity.

Variable	Overall (N=250)	Critical thinking podcast (n=128)	Control podcast (n=122)
Critical thinking—continuous, mean (SD; range)	14.36 (5.41; 2 - 21)	*15.48 (5.19; 2-21)*	*13.19 (5.42; 2 - 21)*
Critical thinking—dichotomous, n (%)			
High score (1=18 or more correct out of 21)	99 (39.6)	*67 (52.3)*	*32 (26.2)*
Low score (0=17 or less correct out of 21)	151 (60.4)	*61 (47.7)*	*90 (73.8)*
Intended behavior (1-4), mean (SD; range)			
Find out what a claim was based on	3.53 (0.71; 1 - 4)	3.54 (0.73; 1 - 4)	3.52 (0.70; 1 - 4)
Find out if a claim was based on a fair comparison study	3.61 (0.63; 1 - 4)	*3.72 (0.55; 1-4)*	*3.49 (0.69; 1 - 4)*
How likely are you to participate in a fair comparison study	3.08 (0.87; 1 - 4)	3.13 (0.85; 1 - 4)	3.03 (0.88; 1 - 4)
EBP[a] attitudes (1-5), mean (SD; range)			
Beliefs regarding therapists' practices	24.74 (3.54; 10 - 30)	24.85 (3.57; 13 - 30)	24.61 (3.51; 10 - 30)
Attitudes about mental health policy	13.40 (3.34; 5 - 20)	13.65 (3.19; 6 - 20)	13.14 (3.48; 5 - 20)
Negative personal-level attitudes towards EBPs	23.97 (7.82; 9 - 45)	*23.01 (7.43; 9-41)*	*24.98 (8.11; 9 - 45)*
Negative societal-level attitudes toward EBPs	14.48 (3.47; 5 - 20)	14.36 (3.51; 5 - 20)	14.60 (3.44; 4 - 20)
Empiricism attitudes (1-5)	4.08 (0.63; 2.2 - 5)	4.15 (0.60; 2.4 - 5)	4.01 (0.66; 2.2 - 5)
Vaccine safety concerns (1-5)	6.1 (3.58; 3 - 15)	6.03 (3.57; 3 - 15)	6.12 (3.60; 3 - 15)
Treatment preferences, mean (SD; range)			
Therapeutic alliance	24.66 (16.14; 0 - 99)	23.57 (16.51; 0 - 99)	25.80 (15.73; 0 - 99)
Scientific studies show therapy is highly effective	28.20 (18.82; 0 - 99)	*30.02 (20.55; 0-99)*	*26.29 (16.70; 0 - 89)*
Therapist experience	16.33 (10.63; 0 - 49)	16.02 (11.04; 0 - 49)	16.66 (10.21; 0 - 48)
Empathic therapist	18.39 (12.74; 0 - 71)	17.77 (12.53; 0 - 70)	19.05 (12.97; 0 - 71)
Client speaking majority of session	11.42 (10.04; 0 - 52)	11.63 (10.90; 0 - 52)	11.20 (9.09; 0 - 42)
Podcast satisfaction, mean (SD; range)			
Overall satisfaction (1-5)	4.18 (0.89; 1 - 5)	4.16 (0.93; 1 - 5)	4.20 (0.85; 1 - 5)
Continue listening (1-5)	3.61 (1.16; 1 - 5)	3.56 (1.19; 1 - 5)	3.65 (1.14; 1 - 5)
Recommend to others (0 - 10)	6.71 (2.89; 0 - 10)	6.76 (2.79; 0 - 10)	6.66 (3.01; 0 - 10)
Relevance to mental health questions (1-4)	2.97 (0.94; 1 - 4)	3.05 (0.95; 1 - 4)	2.88 (0.92; 1 - 4)
Relevance to physical health questions (1-4)	2.80 (1.03; 1 - 4)	3.14 (0.89; 1 - 4)	2.43 (1.04; 1 - 4)

[a]EBP: evidence-based practice.

XSL•FO
RenderX

Effects on Critical Thinking

As hypothesized, parents who listened to the critical thinking podcast performed significantly better on the critical thinking measure than those who listened to the control podcast (B=2.56; P<.001), with a small-medium effect size (ΔR^2=0.06). Table 4 presents all results from linear regression analyses.

Table . Results from linear regression analyses examining critical thinking podcast effect on critical thinking, intended behavior, attitudes, treatment preferences, and overall podcast satisfaction (control variables: age, education, and listening fidelity).

Dependent variable	Critical thinking		
	B	SE	β
Critical thinking	2.56	0.515	.237[a]
Intended behavior			
Find out what a claim was based on	0.042	0.092	.029
Find out if a claim was based on a fair comparison study	0.252	0.078	.201[b]
How likely are you to participate in a fair comparison study	0.109	0.116	.063
EBP[c] attitudes			
Beliefs Regarding Therapists' Practice	0.073	0.093	.050
Attitudes About Mental Health Policy	0.495	0.416	.074
Negative Personal-Level Attitudes towards EBPs	–1.96	0.893	–.127[d]
Negative Societal-Level Attitudes towards EBPs	–0.160	0.446	–.023
Attitudes regarding empiricism	0.156	0.080	.124
Vaccine safety concerns	–0.185	0.400	–.026
Treatment preferences			
Therapeutic alliance	–1.82	2.08	–.056
Scientific studies show therapy is highly effective	4.89	2.34	.131[d]
Therapist experience	–1.50	1.33	–.071
Empathic therapist	–1.59	1.64	–.063
Client speaking majority of session	0.024	1.26	.001
Overall podcast satisfaction	–0.050	0.113	–.028

[a]P<.001.

[b]P<.01.

[c]EBP: evidence-based practice.

[d]P<.05.

Effects on Intended Behavior

Compared to control parents, parents in the critical thinking podcast condition reported a significantly higher likelihood of "finding out if a claim was based on a fair comparison study" in the future (B=0.252; P=<.001; ΔR^2=.04). There were no significant differences between the conditions regarding the likelihood of engaging in the other 2 intended behaviors assessed: "finding out what a claim was based on," and "participating in fair comparison study."

Effects on Attitudes

Between-group differences in EBP attitudes were examined across the 4 attitudes factors of the CAEBS. Parents in the critical thinking condition had lower scores on the negative personal-level attitudes towards EBPs scale (B=−1.96; P<.05; ΔR^2=0.015). There were no significant differences between podcast groups in the 3 other CAEBS EBP attitudes scales, attitudes regarding empiricism in mental health, or vaccine safety concerns.

XSL•FO
RenderX

Effects on Treatment Preferences

Parents who listened to the critical thinking podcast expressed a significantly higher preference for "receiving effective therapies backed by scientific studies" compared to the control group (B=4.89; P<.05; ΔR^2=0.02). All other mental health treatment preferences did not vary significantly based on condition.

Post Hoc Analyses Without Intent to Treat Sample

We also ran analyses excluding the 49 participants who failed data-quality checks. When we excluded these individuals in the comparison analysis of parents' critical thinking postintervention, we found that the magnitude of the intervention effect on critical thinking increased from 0.06 to 0.11 (B=0.525; P<.001). In regard to other outcomes, when we removed the intent to treat sample, we found that parents who listened to the critical thinking podcast reported significantly more positive attitudes regarding empiricism in mental health treatment (B=0.210; P<.05; ΔR^2=0.03), and no longer found a significant difference between podcast groups in the negative personal-level attitudes towards EBPs scale.

Discussion

Principal Findings

To address the public health misinformation crisis at scale, we developed and tested the efficacy of a story-based, educational podcast—the *Parents Making Informed Health Choices Podcast*—a brief, scalable, low-intensity intervention designed to increase critical thinking about health among US parents. Through an online RCT, we found that listening to the podcast improved parent critical thinking and had an effect on intended behaviors, attitudes, and treatment preferences. To our knowledge, this is the first online RCT of a mass media, brief critical thinking intervention for a US audience.

A previous study's findings of low levels of critical thinking among US parents and young adult college students demonstrated a need for a scalable critical thinking intervention for the US public [3]. The impacts on critical thinking demonstrated in this study add to the growing body of empirical evidence demonstrating the efficacy of critical thinking learning resources designed for the lay public [41-43]. The critical thinking effects observed in this US sample are similar to those observed in the original Ugandan podcast, where they also found a significant difference in postintervention critical thinking performance favoring parents in the critical thinking condition [25].

Out of the 3 intended behaviors assessed, parents who listened to the critical thinking podcast were more likely to report intending to find out if a treatment claim is based on a fair comparison study with control parents, but contrary to the hypotheses, these 2 other intended behaviors were not significantly different between the 2 study conditions. A possible explanation for this may be that the podcast directly and repeatedly mentions "fair comparisons," whereas the other intended behaviors are less frequently mentioned. These findings also differ from the Ugandan podcast RCT, where they did not find any significant behavioral intention differences between study conditions.

In terms of attitudinal measures related to critical thinking about health, parents who listened to the critical thinking podcast had less negative personal-level attitudes toward EBPs (as measured by 1 of the 4 factors of CAEBS) [36], and reported a higher preference for receiving effective therapies backed by scientific studies. Contrary to expectations, conditions did not differ on any other EBP attitudes or vaccine safety concerns. Notably, this study was conducted prior to the COVID-19 pandemic; as such, public attitudes regarding vaccines may have changed.

Of the 4 CAEBS factors regarding consumer EBP attitudes (beliefs regarding therapists' practices, attitudes about mental health policy, negative personal-level attitudes toward EBPs, and negative societal-level attitudes towards EBPs), the personal-level attitudes toward EBP factor is most in line with the content of the critical thinking podcast, with items such as "I don't feel comfortable making treatment decisions." Nevertheless, the null findings regarding 3 out of the 4 CAEBS scales were surprising given the general similarity between the concepts covered in the critical thinking podcast and the factors assessed by the CAEBS [36]. Upon closer scrutiny of the CAEBS' scales and items, possible explanations become apparent. First, while the *Parents Making Informed Health Choices Podcast* centers on EBP principles and making them accessible to the lay public, it rarely mentions "evidence-based" terms because the focus is on teaching a skill rather than *telling* the audience they should seek out EBPs. On the other hand, the CAEBS frequently and explicitly mentions the term "evidence-based" and defines it in the instructions. Perhaps this suggests that evidence-based health care proponents and researchers have some middle ground to reach in regards to balancing efforts to teach *science and health literacy–based skills* (ie, critical thinking about health) and increasing consumer knowledge of *health care terms*. It would be interesting for future studies to examine whether explicitly and repeatedly mentioning EBPs in the podcast has a different effect on EBP attitudes as measured by the CAEBS. Another potential explanation for the null findings is that CAEBS items have a strong emphasis on society and policy themes, whereas the critical thinking podcast solely focused on individual actions around EBP principles.

It should be noted that the *Parents Making Informed Health Choices Podcast* achieved these effects through an online, one-time, brief, audio intervention whereas the original Ugandan podcast RCT used a lengthier podcast listening procedure. Thus, our choosing to prioritize ecological validity based on how the public interacts with health information on the internet did not diminish the intervention's effects. Notably, parents in this study were tested on their critical thinking abilities and other factors a few minutes after listening to the podcasts. As such, a follow-up study is necessary to examine long-term effects.

Future studies should also determine who is most likely to benefit from these critical thinking interventions. For example, including a sample with a broader range of levels of educational attainment could help determine if education is a moderator. Another important future direction for critical thinking research

XSL•FO
RenderX

is how to incorporate pragmatic tips on how parents can carry out critical thinking in conversations with health care providers, especially in situations where the evidence base may not be strong, but where a treatment is still recommended either because it is the only option or for other valid reasons. There are EBP principles that address these realities (eg, EBP principle "how certain is the evidence?") [44], and future intervention or podcast development should focus on creating additional short episodes that capture these EBP principles in decision-making (eg,[45].

Furthermore, critical thinking interventions could complement a shared-decision making framework [46,47] in that critical thinking interventions encourage patient activation. Thus, future studies might consider combining or embedding critical thinking interventions within shared decision-making frameworks. Indeed, although this is not directly communicated in the *Parents Making Informed Health Choices Podcast*, characters in the stories often do model appropriate interactions with health care providers; however, the storylines do not conclude in a shared decision about health care services.

Limitations

As with any investigation, this study has several limitations that suggest additional directions for future research. First, although online-convenience samples of parents generally provide reliable data [33] and this study sample was relatively diverse, especially in regard to socioeconomic status, they may not always be representative of all US parents, especially of parents in low-resource, community settings. Thus, future studies should focus on testing the efficacy of the podcast with more representative samples of parents from community settings that could greatly benefit from interventions such as this one. Along those lines, the *Parents Making Informed Health Choices Podcast* is currently only available to English-speaking parents, creating a disparity for parents who speak other languages, especially predominantly Spanish-speaking parents, who make up the fastest growing linguistic population in the United States. Future efforts should focus on translating these materials and testing their efficacy in other languages, especially Spanish. Some of this work is already starting with children in Spain [48]. Second, we have found evidence of only the very short-term efficacy of the *Parents Making Informed Health Choices Podcast*. Therefore, an important future direction is to conduct a long-term follow-up assessment of critical thinking with parents who participated in this online RCT, especially given that a 1-year follow-up study on the effects of the Ugandan critical thinking podcast found a decline in critical thinking abilities [49]. Finally, all outcome variables were about parent attitudes and preferences regarding mental health services for themselves, but we did not specifically ask about their attitudes and preferences in regards to mental health services for their children, which may be an important addition to future studies.

Conclusions

This pilot trial provided empirical evidence regarding the efficacy of a brief, story-based critical health literacy podcast for US parents—a population frequently targeted by health misinformation. We found that listening to a brief podcast significantly improved parents' critical thinking about health, as well as increased self-reported intended behavior, positive attitudes toward EBPs, and preference for EBPs. By delivering accessible, engaging content in a familiar digital format, podcasts may represent a scalable public health strategy to promote evidence-based decision-making and strengthen critical health literacy. Future research should examine long-term effects, implementation in diverse communities, and potential integration into broader digital health and health education ecosystems.

Funding

This work was supported by a departmental research grant from University of Miami Psychology Department. The first author gratefully acknowledges support from the University of Miami's Dean's Dissertation Fellowship, which also provided funding to conduct this research.

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Checklist 1
CONSORT checklist.
[PDF File, 137 KB - publichealth_v12i1e78003_app1.pdf]

References

1. Wilhelm E, Ballalai I, Belanger ME, et al. Measuring the burden of infodemics: summary of the methods and results of the Fifth WHO Infodemic Management Conference. JMIR Infodemiology 2023;3:e44207. [doi: 10.2196/44207] [Medline: 37012998]

XSL•FO
RenderX

2. Dahlgren A, Furuseth-Olsen K, Rose CJ, Oxman AD. The Norwegian public's ability to assess treatment claims: results of a cross-sectional study of critical health literacy. F1000Res 2021;9:179. [doi: 10.12688/f1000research.21902.2] [Medline: 38585673]
3. Mora Ringle VA, Dahlgren A, Rosenbaum S, Jensen-Doss A. Critical thinking about health and treatments in the United States: a cross-sectional assessment of parents and undergraduate college students. BMC Public Health 2025 Jan 27;25(1):336. [doi: 10.1186/s12889-025-21291-9] [Medline: 39871266]
4. Iyengar S, Massey DS. Scientific communication in a post-truth society. Proc Natl Acad Sci U S A 2019 Apr 16;116(16):7656-7661. [doi: 10.1073/pnas.1805868115] [Medline: 30478050]
5. Oxman AD, Paulsen EJ. Who can you trust? A review of free online sources of “trustworthy” information about treatment effects for patients and the public. BMC Med Inform Decis Mak 2019 Feb 20;19(1):35. [doi: 10.1186/s12911-019-0772-5] [Medline: 30786889]
6. Swire-Thompson B, Lazer D. Public health and online misinformation: challenges and recommendations. Annu Rev Public Health 2020 Apr 2;41(1):433-451. [doi: 10.1146/annurev-publhealth-040119-094127] [Medline: 31874069]
7. McKay D, Jensen-Doss A. Harmful treatments in psychotherapy. Clin Psychol: Sci Pract 2021 Mar;28(1):2-4. [doi: 10.1037/cps0000023]
8. Nutbeam D. The evolving concept of health literacy. Soc Sci Med 2008 Dec;67(12):2072-2078. [doi: 10.1016/j.socscimed.2008.09.050] [Medline: 18952344]
9. Nutbeam D, Lloyd JE. Understanding and responding to health literacy as a social determinant of health. Annu Rev Public Health 2021 Apr 1;42:159-173. [doi: 10.1146/annurev-publhealth-090419-102529] [Medline: 33035427]
10. Oxman AD, García LM. Comparison of the informed health choices key concepts framework to other frameworks relevant to teaching and learning how to think critically about health claims and choices: a systematic review. F1000Res 2020;9:164. [doi: 10.12688/f1000research.21858.1] [Medline: 33224475]
11. Aranza D, Milavić B, Marusic A, Buzov M, Poklepović Peričić T. A cross-sectional study on adaptation and initial validation of a test to evaluate health claims among high school students: croatian version. BMJ Open 2021 Aug 10;11(8):e048754. [doi: 10.1136/bmjopen-2021-048754] [Medline: 34376450]
12. Aronson JK, Barends E, Boruch R, et al. Key concepts for making informed choices. Nature New Biol 2019 Aug;572(7769):303-306. [doi: 10.1038/d41586-019-02407-9] [Medline: 31406318]
13. Roozenbeek J, Schneider CR, Dryhurst S, et al. Susceptibility to misinformation about COVID-19 around the world. R Soc Open Sci 2020 Oct;7(10):201199. [doi: 10.1098/rsos.201199] [Medline: 33204475]
14. Becker SJ, Spirito A, Vanmali R. Perceptions of 'evidence-based practice' among the consumers of adolescent substance use treatment. Health Educ J 2016 Apr;75(3):358-369. [doi: 10.1177/0017896915581061] [Medline: 27087698]
15. Mora Ringle VA, Walsh LM, Maxwell CA, et al. Understanding of evidence-based mental health care and the perceived importance of scientific information in a sample of U.S. adults. J Clin Psychol 2020 Jan;76(1):161-175. [doi: 10.1002/jclp.22856] [Medline: 31491053]
16. Scheyett A, McCarthy E, Rausch C. Consumer and family views on evidence-based practices and adult mental health services. Community Ment Health J 2006 Jun;42(3):243-257. [doi: 10.1007/s10597-005-9027-2] [Medline: 16532382]
17. Carman KL, Maurer M, Mangrum R, et al. Understanding an informed public's views on the role of evidence in making health care decisions. Health Aff (Millwood) 2016 Apr;35(4):566-574. [doi: 10.1377/hlthaff.2015.1112] [Medline: 27044953]
18. Badur S, Ota M, Öztürk S, Adegbola R, Dutta A. Vaccine confidence: the keys to restoring trust. Hum Vaccin Immunother 2020 May 3;16(5):1007-1017. [doi: 10.1080/21645515.2020.1740559] [Medline: 32298198]
19. Lo NC, Hotez PJ. Public health and economic consequences of vaccine hesitancy for measles in the United States. JAMA Pediatr 2017 Sep 1;171(9):887-892. [doi: 10.1001/jamapediatrics.2017.1695] [Medline: 28738137]
20. Newman J, Liew A, Bowles J, Soady K, Inglis S. Podcasts for the delivery of medical education and remote learning. J Med Internet Res 2021 Aug 27;23(8):e29168. [doi: 10.2196/29168] [Medline: 34448719]
21. Carrotte E, Hopgood F, Blanchard M, Groot C, Phillips L. A new podcast for reducing stigma against people living with complex mental health issues: co-design study. JMIR Form Res 2023 May 5;7(1):e44412. [doi: 10.2196/44412] [Medline: 37145860]
22. Carrotte ER, Groot C, Blanchard M, Hopgood F, Phillips L. The wavelengths study: a randomized controlled trial assessing the effectiveness of a codesigned podcast on mental illness stigma. Stigma Health 2024;10(4):599-609. [doi: 10.1037/sah0000537]
23. WHO Guideline Recommendations on Digital Interventions for Health System Strengthening: World Health Organization; 2019. URL: https://www.ncbi.nlm.nih.gov/books/NBK541904 [accessed 2025-10-16]
24. Moyer-Gusé E. Toward a theory of entertainment persuasion: explaining the persuasive effects of entertainment-education messages. Commun Theory 2008 Aug;18(3):407-425. [doi: 10.1111/j.1468-2885.2008.00328.x]
25. Semakula D, Nsangi A, Oxman AD, et al. Effects of the Informed Health Choices podcast on the ability of parents of primary school children in Uganda to assess claims about treatment effects: a randomised controlled trial. The Lancet 2017 Jul;390(10092):389-398. [doi: 10.1016/S0140-6736(17)31225-4]
26. Freiheit SR, Vye C, Swan R, Cady M. Cognitive-behavioral therapy for anxiety: is dissemination working? Behav Ther 2004;27(2):25-32.

XSL•FO
RenderX

27. Hipol LJ, Deacon BJ. Dissemination of evidence-based practices for anxiety disorders in Wyoming: a survey of practicing psychotherapists. Behav Modif 2013 Mar;37(2):170-188. [doi: 10.1177/0145445512458794] [Medline: 23012685]
28. Lilienfeld SO. Psychological treatments that cause harm. Perspect Psychol Sci 2007 Mar;2(1):53-70. [doi: 10.1111/j.1745-6916.2007.00029.x] [Medline: 26151919]
29. Lilienfeld SO. Public skepticism of psychology: why many people perceive the study of human behavior as unscientific. Am Psychol 2012;67(2):111-129. [doi: 10.1037/a0023963] [Medline: 21668088]
30. Sayer NA, Rosen CS, Bernardy NC, et al. Context matters: team and organizational factors associated with reach of evidence-based psychotherapies for PTSD in the Veterans Health Administration. Adm Policy Ment Health 2017 Nov;44(6):904-918. [doi: 10.1007/s10488-017-0809-y] [Medline: 28597238]
31. Parents making informed health choices. SoundCloud. URL: https://soundcloud.com/pmihcpodcast [accessed 2025-12-16]
32. Bohannon J. Psychologists grow increasingly dependent on online research subjects. Science. 2016. URL: https://www.science.org/content/article/psychologists-grow-increasingly-dependent-online-research-subjects [accessed 2022-09-25]
33. Jensen-Doss A, Patel ZS, Casline E, Mora Ringle VA, Timpano KR. Using Mechanical Turk to study parents and children: an examination of data quality and representativeness. J Clin Child Adolesc Psychol 2022;51(4):428-442. [doi: 10.1080/15374416.2020.1815205] [Medline: 33448875]
34. Austvoll-Dahlgren A, Guttersrud Ø, Nsangi A, Semakula D, Oxman AD, IHC Group. Measuring ability to assess claims about treatment effects: a latent trait analysis of items from the "Claim Evaluation Tools" database using rasch modelling. BMJ Open 2017 May 25;7(5):e013185. [doi: 10.1136/bmjopen-2016-013185] [Medline: 28550019]
35. Austvoll-Dahlgren A, Semakula D, Nsangi A, et al. Measuring ability to assess claims about treatment effects: the development of the "Claim Evaluation Tools". BMJ Open 2017 May 17;7(5):e013184. [doi: 10.1136/bmjopen-2016-013184] [Medline: 28515181]
36. Teh LB, Hayashi K, Latner J, Mueller CW. Consumer attitudes towards evidence based mental health services among American mental health consumers. Int J Ment Health Nurs 2016 Oct;25(5):462-471. [doi: 10.1111/inm.12234] [Medline: 27291072]
37. Kirk A, Broman-Fulks J, Bergquist J. National attitudes towards mental health treatment: the importance of research evidence. Cogn Behav Ther 2016 Nov;45(6):458-472. [doi: 10.1080/16506073.2016.1201847] [Medline: 27352993]
38. Gust DA, Strine TW, Maurice E, et al. Underimmunization among children: effects of vaccine safety concerns on immunization status. Pediatrics 2004 Jul;114(1):e16-e22. [doi: 10.1542/peds.114.1.e16] [Medline: 15231968]
39. Moran MB, Frank LB, Chatterjee JS, Murphy ST, Baezconde-Garbanati L. Information scanning and vaccine safety concerns among African American, Mexican American, and non-Hispanic White women. Patient Educ Couns 2016 Jan;99(1):147-153. [doi: 10.1016/j.pec.2015.08.016] [Medline: 26321294]
40. Cohen J. Statistical Power Analysis for the Behavioral Sciences, 2nd edition: Academic Press; 1988. [doi: 10.4324/9780203771587]
41. Castle JC, Chalmers I, Atkinson P, et al. Establishing a library of resources to help people understand key concepts in assessing treatment claims—the "Critical thinking and Appraisal Resource Library" (CARL). PLoS ONE 2017;12(7):e0178666. [doi: 10.1371/journal.pone.0178666] [Medline: 28738058]
42. Chalmers I, Atkinson P, Badenoch D, et al. The James Lind initiative: books, websites and databases to promote critical thinking about treatment claims, 2003 to 2018. Res Involv Engagem 2019;5(1):6. [doi: 10.1186/s40900-019-0138-2] [Medline: 30766728]
43. Cusack L, Del Mar CB, Chalmers I, Gibson E, Hoffmann TC. Educational interventions to improve people's understanding of key concepts in assessing the effects of health interventions: a systematic review. Syst Rev 2018 May 2;7(1):68. [doi: 10.1186/s13643-018-0719-4] [Medline: 29716639]
44. Chalmers I, Oxman AD, Austvoll-Dahlgren A, et al. Key concepts for informed health choices: a framework for helping people learn how to assess treatment claims and make informed choices. BMJ EBM 2018 Feb;23(1):29-33. [doi: 10.1136/ebmed-2017-110829]
45. Oxman M, Habib L, Jamtvedt G, Kalsnes B, Molin M. Using claims in the media to teach essential concepts for evidence-based healthcare. BMJ Evid Based Med 2021 Oct;26(5):234-236. [doi: 10.1136/bmjebm-2020-111390] [Medline: 33158855]
46. Cheng H, Hayes D, Edbrooke-Childs J, Martin K, Chapman L, Wolpert M. What approaches for promoting shared decision-making are used in child mental health? A scoping review. Clin Psychol Psychother 2017 Nov;24(6):O1495-O1511. [doi: 10.1002/cpp.2106] [Medline: 28752631]
47. Muscat DM, Shepherd HL, Nutbeam D, Trevena L, McCaffery KJ. Health literacy and shared decision-making: exploring the relationship to enable meaningful patient engagement in healthcare. J Gen Intern Med 2021 Feb;36(2):521-524. [doi: 10.1007/s11606-020-05912-0] [Medline: 32472490]
48. Bayruns Garcia E. Expression-style exclusion. Soc Epistemol 2019 May 4;33(3):245-261. [doi: 10.1080/02691728.2019.1621403]
49. Semakula D, Nsangi A, Oxman AD, et al. Effects of the Informed Health Choices podcast on the ability of parents of primary school children in Uganda to assess the trustworthiness of claims about treatment effects: one-year follow up of a randomised trial. Trials 2020 Feb 14;21(1):187. [doi: 10.1186/s13063-020-4093-x] [Medline: 32059694]

XSL•FO
RenderX

Abbreviations

CAEBS: Consumer Attitudes Towards Evidence Based Services Scale
CONSORT: Consolidated Standards of Reporting Trials
EBP: evidence-based practice
MTurk: Mechanical Turk
RCT: randomized controlled trial

Edited by A Mavragani; submitted 23.May.2025; peer-reviewed by D Cunha, F Yeh; accepted 17.Oct.2025; published 28.Jan.2026.

Please cite as:
Mora Ringle V, Jensen-Doss A
A Critical Health Literacy Podcast to Counter Health Misinformation at Scale: Randomized Controlled Trial
JMIR Public Health Surveill 2026;12:e78003
URL: https://publichealth.jmir.org/2026/1/e78003
doi:10.2196/78003

© Vanesa Mora Ringle, Amanda Jensen-Doss. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 28.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Legal Infoveillance of Unlicensed Medical Practices in South Korea Through Criminal Court Decisions Using Machine Learning: Retrospective Observational Study

Min Ji Kim[1], BSN, MPH, PhD; Min Choi[2], BS, MS, PhD

[1]Department of Medical Humanities and Ethics, College of Medicine, Hanyang University, Seoul, Republic of Korea
[2]LBox Corporation, Seoul, Republic of Korea

Corresponding Author:
Min Ji Kim, BSN, MPH, PhD
Department of Medical Humanities and Ethics
College of Medicine
Hanyang University
222 Wangsimni-ro, Seongdong-gu
Seoul, 04763
Republic of Korea
Phone: 82 2 2220 2461
Email: drphkim@gmail.com

Abstract

Background: Unlicensed medical practices (UMPs) pose a substantial threat to patient safety and public health, but their clandestine nature makes them difficult to monitor through conventional surveillance systems. Legal epidemiology offers a framework for using judicial data to study hidden health-related misconduct, and machine learning (ML) may help convert unstructured legal texts into analyzable public health information.

Objective: This study aimed to characterize prosecuted UMP cases in South Korea using a legal infoveillance framework and evaluate the utility of ML-assisted extraction from criminal court decisions for public health surveillance.

Methods: We conducted a retrospective observational study of 1532 criminal court decisions involving UMP-related convictions in South Korea between 2005 and 2023. Using an ML-assisted extraction pipeline with human-in-the-loop verification, we transformed unstructured judicial texts into structured legal and medical variables. Analyses were conducted at the case, charge, defendant category, and legal ruling levels. In addition to descriptive analyses, we performed exploratory inferential analyses to examine factors associated with legal rulings and professionals' involvement.

Results: Of 1718 charge entries, 987 (57.5%) were related to Article 5 of the Act on Special Measures for the Control of Public Health Crimes, and 731 (42.5%) were related to Article 27(1) of the Medical Service Act. Profit motive was coded in 91.6% (1404/1532) of the cases. At the legal ruling level (n=2004 entries), suspended sentences, meaning sentences whose execution was conditionally suspended under Korean criminal law, were the most common outcome (1261/2004, 62.9%), followed by fines (421/2004, 21%) and imprisonment without suspension (209/2004, 10.4%). Of 1716 defendant category entries, ordinary persons accounted for 1294 (75.4%), health care professionals accounted for 264 (15.4%), and health care providers accounted for 158 (9.2%). Physicians were the largest subgroup among health care professionals. In exploratory multinomial models, licensed personnel–only and mixed ordinary person and licensed personnel cases were more likely than ordinary person–only cases to result in fines or imprisonment without suspension rather than suspended sentences. Secondary exploratory analyses also suggested distinctive patterns of professional involvement and possible scope-of-practice or delegation-related boundary violations.

Conclusions: ML-assisted analysis of criminal court decisions can serve as a useful supplementary surveillance method for hidden UMPs. In South Korea, prosecuted UMPs were predominantly profit driven and involved both ordinary persons and licensed personnel. The findings support closer monitoring of scope-of-practice and delegation-related violations and demonstrate the value of judicial records as a source of public health intelligence.

(JMIR Public Health Surveill 2026;12:e92591) doi:10.2196/92591

XSL•FO
RenderX

KEYWORDS

unlicensed medical practice; machine learning; natural language processing; patient safety; scope of practice; legal infoveillance; artificial intelligence; AI; Republic of Korea

Introduction

The World Health Organization identifies unsafe care as a major global health challenge, with millions of patients harmed annually due to inadequate safety standards and unqualified practitioners [1]. One persistent but difficult-to-measure contributor is medical or quasi-medical practices by persons acting outside legally defined professional boundaries. Unlicensed medical practices (UMPs) include both practices performed by persons with no health-related license and practices performed by licensed personnel who exceed, improperly delegate, or otherwise violate their lawful scope of practice [2].

UMPs are difficult to study using conventional surveillance tools because they are inherently concealed. Administrative health data, claims systems, and routine public health reporting rarely capture such activity directly. In South Korea, as universal health coverage coexists with the strict licensing regime, UMPs may persist not because care is unavailable [3] but because commercially motivated noncompliance and legal ambiguity regarding the boundaries of "medical practice" create opportunities for underground or improperly delegated services [4].

In South Korea, the Medical Service Act strictly prohibits UMPs under Article 27(1) to protect patient safety [5]. Additionally, Article 5 of the Act on Special Measures for the Control of Public Health Crimes imposes severe penalties, including imprisonment, on nonprofessionals engaging in UMPs for profit [6]. However, the statutory provisions do not explicitly define "medical practice," leaving courts with wide discretion in interpreting the boundaries given potential risks to public health. Furthermore, these practices often lead to clinical negligence, compromising both patient safety and the integrity of the health care system. Therefore, establishing a stringent regulatory regime based on accurate situational awareness is crucial [7,8].

Despite its significant threat to public health, the inherently clandestine nature of UMPs makes them difficult to quantify [9]. Previous studies have been limited to regulatory reviews [10,11] or localized cross-sectional surveys [12], leaving a critical empirical gap regarding the actual prevalence and distinct characteristics of these offenses. Therefore, investigating the current status of UMPs using objective, large-scale data is essential to provide foundational evidence for regulatory policymaking.

South Korea provides a strategically informative setting for this approach because court decisions are highly digitized [13,14] and can be assembled at scale for computational analysis [15]. Recent work in legal natural language processing has made it increasingly feasible to extract structured variables from unstructured judicial text and analyze legal outcomes systematically [16,17]. Within the framework of legal epidemiology, which treats law and legal processes as measurable public health phenomena [18], we propose the term "legal infoveillance" to denote the systematic use of legal and judicial records as a source of surveillance for population health questions. This differs from conventional digital infoveillance, which relies primarily on internet-based or digitally generated information streams [19,20], because legal infoveillance relies on formal case records reflecting detected, investigated, prosecuted, and adjudicated events. Furthermore, while legal epidemiology examines the law as a factor in the distribution and prevention of disease and digital infoveillance monitors public health trends using internet-based behavioral data, our framework synthesizes and extends these domains. By integrating principles of legal epidemiology with recent advancements in artificial intelligence [20], we propose a scalable solution to this data bottleneck. Machine learning (ML) techniques can now process massive volumes of legal texts to predict outcomes and understand the factors influencing judicial decisions [16,17]. Applying these technologies to UMP precedents allows for a novel form of public health surveillance, turning retrospective legal data into actionable public health intelligence [18].

This study aimed to assess the status and characteristics of UMPs by analyzing a comprehensive dataset of criminal precedents using ML techniques. To our knowledge, this is the first study to use automated legal text analysis to generate evidence-based knowledge for global UMP policymaking and regulatory design.

Methods

Study Design and Data Collection

This study was a retrospective observational study of criminal court decisions involving UMP-related convictions in South Korea between 2005 and 2023. We identified closed criminal cases in which at least one defendant was convicted of an offense related to UMPs. The dataset was filtered from a comprehensive corpus of over 3 million court precedents collected by LBox, a legal technology company in South Korea [21]. We selected 2005 as the starting year because systemic indexing in the source corpus began in that year, allowing for consistent retrieval and longitudinal comparison across the study period [13,15]. This dataset serves as a proxy for the detected portion of UMPs, offering a rare window into the types of practices and defendants capture through regulatory oversight. This manuscript was reported with reference to the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) statement [22], and the completed checklist is provided in Multimedia Appendix 1.

Data Processing

Processing raw legal data posed significant challenges because court decisions are often stored as unstructured images or PDFs. To address this, we used a specialized data processing pipeline [15]. First, a rule-based layout classifier distinguished text-only pages from those containing tables or figures. Then, texts were

XSL•FO
RenderX

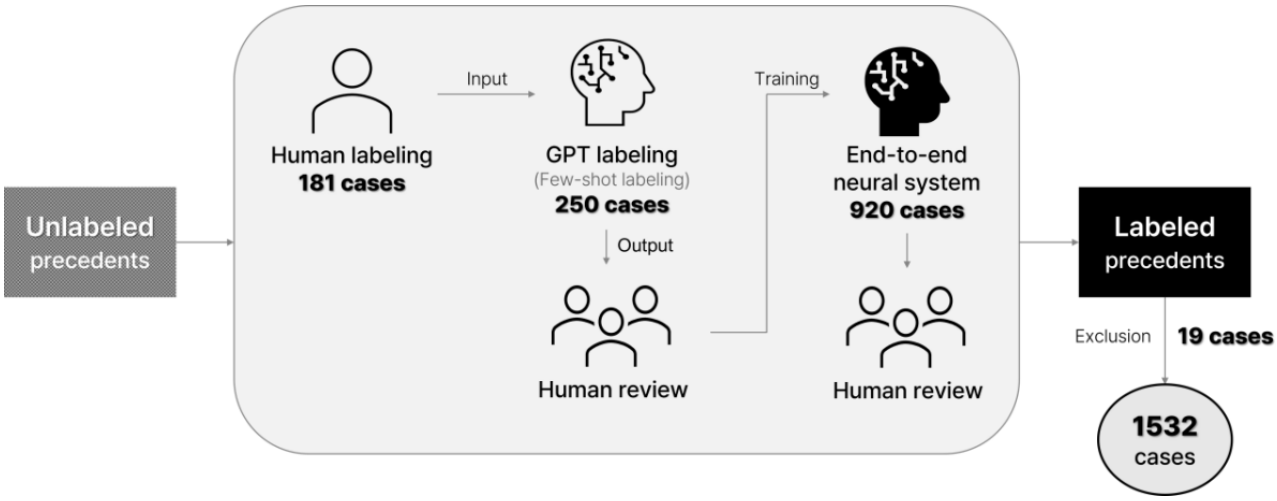
extracted using LBox's custom rule-based parser for PDFs and an optical character recognition engine for image files [23]. Third, to extract variables from the unstructured text bundles, a custom-made parser based on bidirectional encoder representations from transformers [24] segmented the texts into functional components, such as facts, legal rulings, reasons for sentencing, and conclusions. This methodology aligns with emerging standards in legal text mining for public health, enabling the conversion of complex narrative data into structured epidemiological variables.

ML-Assisted Extraction and Human Verification

We used an enhanced proprietary version of the end-to-end Information Extractor for Statistical Legal Analysis [25], an end-to-end neural system based on a domain-adapted mT5-Large architecture [26], for automated information extraction. The foundational model has been validated, achieving an average F_1-score of 93.1 across Korean legal information extraction tasks [25].

Given the high-stakes nature of legal and health data, we addressed the "black box" problem of artificial intelligence by implementing a rigorous human-in-the-loop validation process. Initially, 181 cases were labeled by the researchers to create the gold-standard set, which provided the few-shot examples for GPT-4 (OpenAI) [27] to generate labels for an additional 250 cases [28]. Second, these model-generated labels were thoroughly reviewed and corrected by 3 professional labelers, yielding a verified dataset of 631 cases. Third, this corrected dataset was used to fine-tune the Information Extractor for Statistical Legal Analysis–based system, which subsequently generated variable values autoregressively for the remaining cases. In total, 920 cases were labeled by the fine-tuned system and verified by the labelers. Cases involving complex compound charges (eg, UMPs combined with drug-related charges) were excluded (n=19) to ensure reliability. Finally, 1532 unique labeled cases were included in the final analysis (Figure 1).

Figure 1. Machine learning–assisted information extraction and human-in-the-loop labeling workflow. GPT: generative pretrained transformer.



Analytic Units and Variable Definitions

As a single judicial decision could involve multiple charges, defendants, and legal rulings, counts across analytic levels were not directly comparable. Therefore, all percentages are reported with their explicit denominators. Analyses were conducted at 4 levels: cases (n=1532), charges (n=1718), defendants (n=1716), and legal rulings (n=2004).

Variables were defined through close collaboration between experts in computer science and medical law. For practitioner classification, we used 3 operational categories: health care professionals, health care providers, and ordinary persons. Health care professionals were defined as physicians, dentists, doctors of Korean medicine, midwives, and registered nurses under the Medical Service Act of South Korea [5]. Health care providers were defined as other licensed or certified health care–related personnel, such as nurse aides, medical technologists, and physical therapists. Finally, we defined ordinary persons as defendants without a health care–related professional license. Legal rulings were classified as imprisonment without suspension, fines, suspended sentences, suspension of imposition of sentence, acquittal, or other. Sentencing factors (eg, settlement with victims and self-reflection) were categorized according to the guidelines of the Sentencing Commission of South Korea [29].

Data Analysis

The variables were labeled and classified into legal and medical sections for each case. Before labeling the variables, ontologies were built for the individual sections, and a labeling page was set up using Label Studio (HumanSignal) [30]. To ensure the validity of the results, the researchers independently labeled the variables for each case (181 samples) and then cross-checked them. The section on legal facts and rulings was labeled by a PhD in Engineering at LBox (MC), which provides a cloud-based service to support the digitization of legal practices by processing court documents and records in South Korea [21]. A PhD in Medical Law and Ethics (MJK) verified the medical sections to ensure accuracy.

Statistical analyses were performed using SPSS (version 27.0; IBM Corp) and SAS (version 9.4; SAS Institute Inc). We first conducted descriptive analyses at the case, charge, defendant category, and ruling entry levels. We then examined annual case counts from 2005 to 2023 to visualize the temporal

XSL•FO
RenderX

distribution of UMP cases prosecuted. Because the dataset was conditioned on prosecuted cases with at least one UMP-related conviction, inferential analyses were framed as exploratory models of sentencing patterns among detected, prosecuted, and adjudicated cases rather than as estimates of underlying UMP prevalence, probability of offending, probability of conviction, or guilt. For the primary inferential analysis, we fit a multinomial logistic regression using the principal legal ruling recorded in each case as the dependent variable. Suspended sentence was used as the reference, and suspension of imposition of sentence, acquittal, and other rare outcomes were collapsed into a single "others" category due to small cell counts. Predictors included case composition (ordinary person only, licensed personnel only, mixed ordinary person and licensed personnel, or uncoded), involvement in the Act on Special Measures for the Control of Public Health Crimes charges, profit motive, and practice type. As secondary analyses, we fit logistic models of professionals' involvement among category-coded cases and of health care providers' coinvolvement among licensed personnel–involved cases as an exploratory proxy for delegation- or boundary-related scenarios.

Ethical Considerations

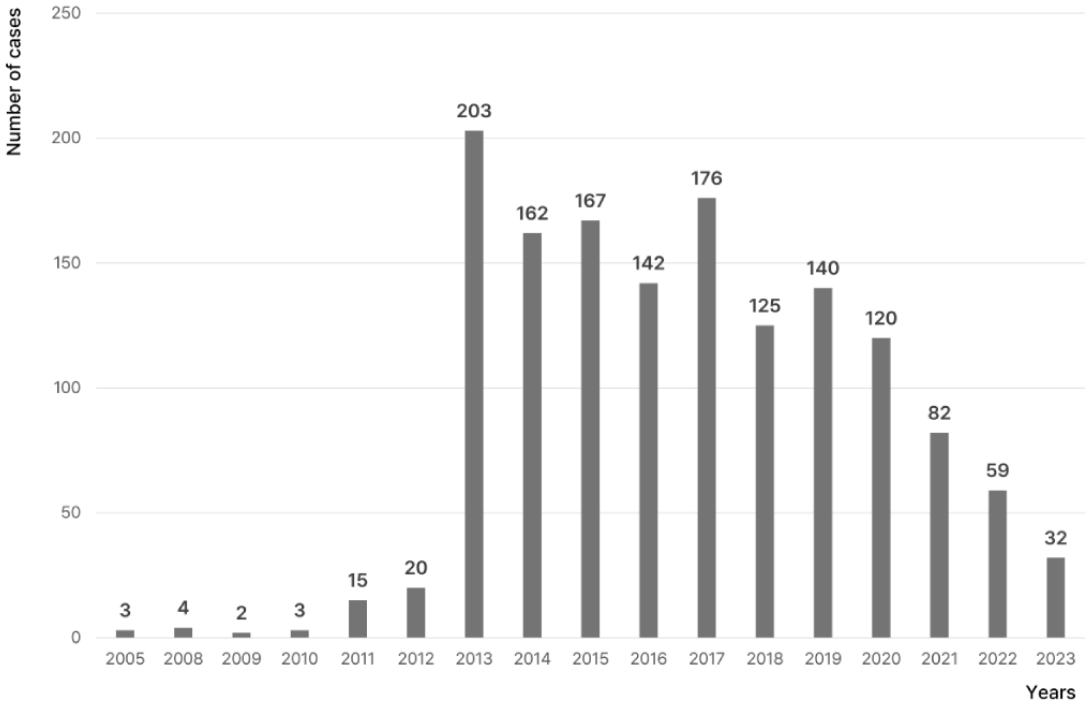
The joint institutional review board designated by the Ministry of Health and Welfare of South Korea confirmed that the study protocol met the criteria for exemption from regular review because this research was based on literature that was publicly accessible by any individual and owing to the absence of direct or indirect involvement of human materials, animal experiments, or pathogens (P01-202406-01-025).

Results

Overview of UMP Cases

A total of 1532 criminal cases were analyzed. The annual distribution of these cases from 2005 to 2023 is illustrated in Figure 2, showing the longitudinal trends in UMP prosecutions. Annual case counts remained low until 2011, increased sharply from 2012, peaked at 203 cases in 2013, remained high until 2019, and steadily declined to 32 cases by 2023.

Figure 2. Annual distribution of 1532 prosecuted unlicensed medical practice cases identified from Korean criminal court decisions (2005-2023).



Legal Characteristics and Rulings

All cases contained 1718 charge entries, of which 987 (57.5%) were related to violations of Article 5 of the Act on Special Measures for the Control of Public Health Crimes, whereas 731 (42.5%) were related to violations of Article 27(1) of the Medical Service Act. Profit motive was coded in 91.6% (1404/1532) of cases, indicating that the vast majority of prosecuted UMP cases were commercially motivated. At the legal ruling level (n=2004), suspended sentences were most common (1261/2004, 62.9%), followed by fines (421/2004, 21%), imprisonment without suspension (209/2004, 10.4%), and acquittal (57/2004, 2.8%). Analysis of sentencing factors revealed that payment or deposit for victims was recorded in 72.4% (63/87) of cases, settlements with victims were recorded in 75.4% (101/134) of cases, and self-reflection was recorded in 95.6% (586/613) of cases. Prior similar-offense counts were coded for 26.9% (412/1532) of cases, with a mean of 0.8 prior similar offenses per coded case (SD 1.2; range 0-12), and prior different-offense counts were coded for 11.3% (173/1532) of cases, with a mean of 0.3 prior different offenses per coded case (SD 0.6; range 0-5; Table 1).

XSL•FO
RenderX

Table 1. Legal characteristics and sentencing factors in prosecuted unlicensed medical practice cases identified from Korean criminal court decisions (2005-2023) in a retrospective observational study.

Category and variable	Values
Charges (n=1718), n (%)	
Violation of Article 5 of the Act on Special Measures for the Control of Public Health Crimes	987 (57.5)
Violation of Article 27(1) of the Medical Service Act	731 (42.5)
For-profit motive (n=1532), n (%)	
Yes	1404 (91.6)
No	128 (8.4)
Legal ruling (n=2004), n (%)	
Suspended sentences	1261 (62.9)
Fine	421 (21)
Imprisonment without suspension	209 (10.4)
Acquittal	57 (2.8)
Suspension of imposition of sentence	51 (2.5)
Others	5 (0.2)
Sentencing factors	
Payment or deposit for victims (n=87), n (%)	
Done	63 (72.4)
Not done	24 (27.6)
Settlement with victims (n=134), n (%)	
Done	101 (75.4)
Not done	33 (24.6)
Self-reflection (n=613), n (%)	
Yes	586 (95.6)
No	27 (4.4)
Criminal records, mean (SD; range)	
Prior similar offenses per coded cases (n=412)	0.8 (1.2; 0-12)
Prior different offenses per coded cases (n=173)	0.3 (0.6; 0-5)

Medical Characteristics and Defendant Profiles

Defendant category coding was available for 97.8% (1499/1532) of cases, yielding 1716 defendant category entries as some cases involved more than one defendant category. Ordinary persons accounted for 75.4% (1294/1716); whereas health care professionals accounted for 15.4% (264/1716); and health care providers accounted for 9.2% (158/1716), including nurse aides, medical technologists, or physical therapists. Among health care professionals, physicians were the most frequent offenders (166/1716, 9.7%), followed by doctors of Korean medicine (41/1716, 2.4%) and dentists (35/1716, 2%). At the case level (n=1532), general UMP was the most common practice type (1056/1532, 68.9%), followed by unlicensed Korean medical practice (302/1532, 19.7%) and unlicensed dental practice (175/1532, 11.4%). Notably, one case was double coded for both general unlicensed medical and Korean medical practice (Table 2).

XSL•FO
RenderX

Table 2. Medical characteristics, defendant categories, and practice types in prosecuted unlicensed medical practice cases identified from Korean criminal court decisions (2005-2023) reported by analytic unit.

Category and variable	Frequency, n (%)
Defendant category (defendant entries; n=1716)	
Ordinary person	1294 (75.4)
Health care provider	158 (9.2)
Health care professional	264 (15.4)
Physician	166 (9.7)
Doctor of Korean medicine	41 (2.4)
Dentist	35 (2)
Registered nurse	22 (1.3)
Type of practice (cases; n=1532)	
Unlicensed medical practice[a]	1056 (68.9)
Unlicensed Korean medical practice	302 (19.7)
Unlicensed dental practice	175 (11.4)

[a]The general unlicensed medical practice category included tattoo-related cases when the conduct was adjudicated as unlicensed medical practice under Korean law; however, these cases were not analyzed as a separate practice type category in the revised analyses.

Sanctions by Defendant Category

The distribution of legal rulings varied by defendant category. In the cross-tabulation of legal rulings and defendant categories among 1499 category-coded cases, suspended sentences (1129/1294, 87.2%) were the most frequently reported among ordinary person–involved cases. In contrast, fines were most frequently observed among physician-involved cases (130/166, 78.3%), followed by dentist-involved cases (26/35, 74.3%) and doctor of Korean medicine–involved cases (31/41, 75.6%; Table 3).

Table 3. Case-level multiple-response cross-tabulation of legal rulings and defendant categories in prosecuted unlicensed medical practice cases identified from Korean criminal court decisions (2005-2023)[a].

Defendant category	Cases, n/N (%)[b]	Imprisonment without suspension, n/N (%)[c]	Fine, n/N (%)[c]	Suspended sentences, n/N (%)[c]	Suspension of imposition of sentence, n/N (%)[c]	Acquittal, n/N (%)[c]	Others, n/N (%)[c]
Health care professionals							
Physician	166/1499 (11.1)	36/166 (21.7)	130/166 (78.3)	90/166 (54.2)	24/166 (14.5)	17/166 (10.2)	1/166 (0.6)
Dentist	35/1499 (2.3)	4/35 (11.4)	26/35 (74.3)	15/35 (42.9)	2/35 (5.7)	4/35 (11.4)	0/35 (0)
Doctor of Korean medicine	41/1499 (2.7)	7/41 (17.1)	31/41 (75.6)	26/41 (63.4)	6/41 (14.6)	3/41 (7.3)	0/41 (0)
Registered nurse	22/1499 (1.5)	4/22 (18.2)	11/22 (50)	9/22 (40.9)	2/22 (9.1)	5/22 (22.7)	0/22 (0)
Health care providers	158/1499 (10.5)	31/158 (19.6)	110/158 (69.6)	99/158 (62.7)	24/158 (15.2)	14/158 (8.9)	0/158(0)
Ordinary persons	1294/1499 (86.3)	167/1294 (12.9)	249/1294 (19.2)	1129/1294 (87.2)	20/1294 (1.5)	34/1294 (2.6)	5/1294 (0.4)
Total	1499/1499 (100)	201/1499 (13.4)	397/1499 (26.5)	1232/1499 (82.2)	49/1499 (3.3)	53/1499 (3.5)	5/1499 (0.3)

[a]Because a single case could involve multiple defendant categories and multiple legal rulings, counts across rows are non–mutually exclusive; therefore, row percentages may add up to more than 100.

[b]The column shows the number and percentage of total analyzed cases (n=1499) that involved at least one defendant in the specified category.

[c]Values constitute the number and percentage of cases in each defendant category.

Exploratory Inferential Analysis of Legal Rulings

To address the determinants of legal rulings, we fit an exploratory multinomial logistic regression model. Compared with ordinary person–only cases, licensed personnel–only cases were more likely to result in fines (adjusted relative risk ratio [aRRR] 2.13, 95% CI 1.36-3.32) and imprisonment without suspension rather than suspended sentences (aRRR 1.79, 95%

XSL•FO
RenderX

CI 1.02-3.16). Mixed ordinary person and licensed personnel cases showed even stronger associations with fines (aRRR 3.56, 95% CI 1.62-7.81) and imprisonment without suspension (aRRR 2.75, 95% CI 1.18-6.40). Act on Special Measures for the Control of Public Health Crimes–related cases were comparatively more likely to end in suspended sentences than in fines (aRRR 0.02, 95% CI 0.01-0.03) or imprisonment without suspension (aRRR 0.55, 95% CI 0.37-0.83; Table 4).

Table 4. Exploratory multinomial logistic regression of principal judicial outcomes among prosecuted unlicensed medical practice cases identified from Korean criminal court decisions (2005-2023; n=1532)[a].

Predictor	Fine, aRRR[b] (95% CI)[c]	*P* value	Imprisonment without suspension, aRRR (95% CI)[c]	*P* value	Others, aRRR (95% CI)[c]	*P* value
Licensed personnel–only cases (vs ordinary person only)	2.13 (1.36-3.32)	<.001	1.79 (1.02-3.16)	.04	5.14 (2.65-9.99)	<.001
Mixed ordinary person and licensed personnel cases (vs ordinary person only)	3.56 (1.62-7.81)	.002	2.75 (1.18-6.40)	.02	7.43 (2.73-20.21)	<.001
Uncoded (vs ordinary person only)	1.41 (0.53-3.78)	.50	0.74 (0.19-2.93)	.67	2.32 (0.59-9.07)	.23
Act on Special Measures for the Control of Public Health Crimes–related charge involvement (vs not involved)	0.02 (0.01-0.03)	<.001	0.55 (0.37-0.83)	.004	0.09 (0.04-0.18)	<.001
Profit motive (vs no profit)	0.98 (0.57-1.69)	.96	0.57 (0.28-1.13)	.11	0.66 (0.32-1.39)	.28
Korean medical practice (vs general)	0.86 (0.55-1.35)	.51	1.16 (0.77-1.74)	.49	0.93 (0.43-2.02)	.86
Dental practice (vs general)	0.58 (0.30-1.11)	.10	0.49 (0.26-0.92)	.03	0.29 (0.09-1.01)	.05

[a]As the dataset was restricted to detected, prosecuted, and adjudicated unlicensed medical practice cases with at least one conviction, the values estimate associations with ruling type within the prosecuted case set, not the underlying prevalence of unlicensed medical practices, probability of offending, probability of conviction, or guilt. The principal judicial outcome recorded for each case was used as the dependent variable.

[b]aRRR: adjusted relative risk ratio.

[c]The suspended-sentence category was the reference category.

Secondary Exploratory Analyses of Professional-Related Case Patterns

Table 5 illustrates the secondary exploratory logistic model among 1499 category-coded cases. Professional involvement was less common in Act on Special Measures for the Control of Public Health Crimes–related cases (adjusted odds ratio [aOR] 0.10, 95% CI 0.06-0.14), profit-driven cases (aOR 0.47, 95% CI 0.28-0.77), and Korean medical practice cases (aOR 0.09, 95% CI 0.05-0.16).

In a further exploratory model restricted to 248 cases involving at least one health care professional, physician involvement, compared with involvement of other health care professionals, was associated with health care provider coinvolvement (aOR 4.67, 95% CI 2.40-9.10), a pattern consistent with possible delegation-related or boundary violation scenarios.

Table 5. Exploratory logistic regression of health care professional involvement among category-coded prosecuted unlicensed medical practice cases identified from Korean criminal court decisions (2005-2023; n=1499)[a].

Predictor	aOR[b] (95% CI)	*P* value
Act on Special Measures for the Control of Public Health Crimes–related charge involvement (vs not involved)	0.10 (0.06-0.14)	<.001
Profit motive (vs no profit)	0.47 (0.28-0.77)	.003
Korean medical practice (vs general)	0.09 (0.05-0.16)	<.001
Dental practice (vs general)	0.55 (0.33-0.93)	.03

[a]This model examined whether a case involved at least one health care professional.

[b]aOR: adjusted odds ratio.

Discussion

Principal Findings

This retrospective study used end-to-end neural systems to analyze 1532 criminal precedents and demonstrate the public health utility of legal infoveillance for prosecuted UMPs in South Korea. Three major findings emerged. First, commercially motivated offending predominated, with 91.6% (1404/1532) of cases coded as profit driven. Second, suspended sentences were the most common judicial outcome. Third, UMPs were not confined to ordinary persons without licenses; licensed health care professionals and providers were also substantially represented in the prosecuted case set. Taken together, these findings suggest that UMPs in South Korea should be understood not only as lay “quackery” but also as a problem of

XSL•FO
RenderX

regulatory boundary violation within and around formal health care settings.

Health care professionals, particularly physicians, constituted the largest subgroup among them, and fines were common among professional-involved cases. The exploratory inferential analyses further suggested that licensed personnel–only and mixed ordinary person and licensed personnel cases were more likely than ordinary person–only cases to result in fines or imprisonment without suspension rather than suspended sentences. Although judicial records cannot reveal underlying institutional arrangements with certainty, this pattern is consistent with broader concerns about scope-of-practice breaches and commercially driven delegation or task shifting [4,31,32]. In this context, the coinvolvement of physicians and health care providers should be interpreted cautiously as a signal of possible boundary violation scenarios rather than as direct proof of unlawful delegation.

Our findings also help explain why UMPs may persist even in a setting with universal health coverage. In South Korea, the persistence of prosecuted UMPs is unlikely to be explained solely by lack of access to formal care [3]. Rather, the concentration of profit-driven cases and the involvement of licensed actors suggest that legal ambiguity, commercial incentives, and inadequately deterred noncompliance may all contribute to the continuing presence of UMPs [4]. The predominance of suspended sentences and fines may further indicate that many courts viewed these offenses as serious but still compatible with conditional leniency, which may weaken deterrence when the underlying conduct is economically rewarding.

From a methodological standpoint, this study supports legal infoveillance as a supplementary surveillance approach for hidden health-related misconduct. Because UMPs are clandestine and rarely captured directly in routine public health or administrative datasets [19], digitized court decisions provide a rare observational window into detected, investigated, prosecuted, and adjudicated events [18,33]. The framework is potentially transferable to other jurisdictions, but its usefulness will depend on the degree of judicial digitization [14], public access to court records [13], and availability of natural language processing tools capable of handling jurisdiction-specific legal language [34].

The findings need to be carefully interpreted in light of several limitations. Most importantly, the dataset represents prosecuted cases rather than the true prevalence of UMPs. This creates selection and detection bias [33] and makes the observed patterns sensitive to enforcement practices, prosecutorial priorities, and reporting behavior [35]. Accordingly, the results describe the enforcement-visible portion of UMPs rather than the absolute population burden. Therefore, these models should be read only as exploratory descriptions of sentencing patterns among detected, prosecuted, and adjudicated cases; they do not estimate the underlying prevalence of UMPs, the probability that an individual will offend, or the probability that a reported or prosecuted case will result in conviction. Formal interannotator agreement metrics were not calculated, and some complex judicial features had to be simplified into structured variables. Nevertheless, this study narrows an important empirical gap by showing that large-scale legal text analysis can generate population-relevant evidence from otherwise difficult-to-study judicial corpora [24,36].

Conclusions

This study demonstrates that ML-assisted analysis of criminal court precedents can serve as a useful supplementary method for monitoring hidden UMPs. In South Korea, prosecuted UMPs were predominantly profit driven and involved both ordinary persons and licensed personnel, with exploratory analyses suggesting distinct judicial outcome patterns when health care professionals were involved. These findings support stronger oversight of scope-of-practice and delegation-related violations and reinforce the value of judicial records as a practical source of public health intelligence. Future research should build on this descriptive foundation with richer inferential modeling, improved annotation reliability metrics, and cross-jurisdictional comparisons of how legal design and enforcement shape the detected profile of UMPs.

Acknowledgments

First, the authors appreciate LBox for providing data and cooperating on this study. Second, the authors declare the use of generative artificial intelligence in the research and writing process. According to the Generative Artificial Intelligence Delegation Taxonomy (2025), proofreading and editing were delegated to Gemini 3 (Google) under full human supervision. Responsibility for the final manuscript lies entirely with the authors. Generative artificial intelligence tools are not listed as authors and do not bear responsibility for the final outcomes. This declaration is made under the collective responsibility of all authors.

Funding

The authors declared no financial support was received for this work.

Data Availability

The structured analytic dataset generated during this study is available from the corresponding author on reasonable written request subject to ethical, legal, and data use restrictions. The original court decision corpus obtained through LBox cannot be redistributed by the authors.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: MJK
Data curation: MC
Formal analysis: MJK, MC
Investigation: MJK
Methodology: MJK, MC
Project administration: MJK
Software: MC
Validation: MC
Visualization: MJK, MC
Writing—original draft: MJK
Writing—review and editing: MJK, MC

Conflicts of Interest

MC was affiliated with LBox until August 2024, which provided data and technical cooperation for this study. The authors declare that LBox's commercial interests or products did not influence the study design, interpretation, or conclusions. The other author declares no other conflicts of interest.

Multimedia Appendix 1
STROBE checklist.
[DOCX File , 94 KB - publichealth_v12i1e92591_app1.docx]

References

1. Slawomirski L, Klazinga N. The economics of patient safety: from analysis to action. Organisation for Economic Co-operation and Development. 2022. URL: https://www.oecd.org/content/dam/oecd/en/publications/reports/2022/08/the-economics-of-patient-safety_dda2a072/761f2da8-en.pdf [accessed 2026-02-01]
2. Singh MP, Chand SK, Saha KB, Singh N, Dhiman RC, Sabin LL. Unlicensed medical practitioners in tribal dominated rural areas of central India: bottleneck in malaria elimination. Malar J 2020 Jan 14;19(1):18 [FREE Full text] [doi: 10.1186/s12936-020-3109-z] [Medline: 31937329]
3. Kwon S. Thirty years of national health insurance in South Korea: lessons for achieving universal health care coverage. Health Policy Plan 2009 Jan;24(1):63-71. [doi: 10.1093/heapol/czn037] [Medline: 19004861]
4. Cho SJ, Yang SR, Ryu S, Park SM. Ghost surgeries in South Korea: variations in practice and the need for clear guidelines in resident surgical training. Korean J Med Ethics 2025 Jun;28(2):103-123. [doi: 10.35301/ksme.2025.28.2.103]
5. Medical Service Act, Act no. 19421. Korea Law Information Center. 2023. URL: https://law.go.kr/LSW/eng/engLsSc.do?menuId=2&section=lawNm&query=Act+on+Special+Measures+for+the+Control&x=0&y=0#AJAX [accessed 2024-06-25]
6. Act on Special Measures for the Control of Public Health Crimes (South Korea), Act No. 17793, Dec 29, 2020. Korean Law Information Center. 2020. URL: https://law.go.kr/LSW/eng/engLsSc.do?menuId=2&section=lawNm&query=Act+on+Special+Measures+for+the+Control&x=0&y=0#liBgcolor0 [accessed 2024-06-25]
7. Guidelines for the structure and function of a state medical and osteopathic board. Federation of State Medical Boards. 2024. URL: https://www.fsmb.org/siteassets/advocacy/policies/guidelines-for-structure-function-of-state-medical-and-osteopathic-board-2024.pdf [accessed 2025-03-21]
8. Dahlawi S, Menezes RG, Khan MA, Waris A, Saifullah, Naseer MM. Medical negligence in healthcare organizations and its impact on patient safety and public health: a bibliometric study. F1000Res 2021 Mar 3;10:174 [FREE Full text] [doi: 10.12688/f1000research.37448.1] [Medline: 34113436]
9. Wardle J. Holding unregistered health practitioners to account: an analysis of current regulatory and legislative approaches. J Law Med 2014 Dec;22(2):350-375. [Medline: 25715537]
10. Chaudhry HJ, Rhyne J, Cain FE, Young A, Crane M, Bush F. Maintenance of licensure: protecting the public, promoting quality health care. J Med Regul 2010 Jun;96(2):13-20. [doi: 10.30770/2572-1852-96.2.13]
11. Irby DM, Cooke M, O Brien BC. The important role of medical licensure in the United States. Acad Med 2010 Nov;85(11):1657-1658. [doi: 10.1097/ACM.0b013e3181f5ce3c]
12. Xu B, Hu YN, Zhang Q, Wang H, Wang Z, Yu LY, et al. Prevalence of unlicensed medical practice in China: a retrospective analysis. Lancet 2016 Oct;388:S41. [doi: 10.1016/S0140-6736(16)31968-7]
13. Jeon HJ, Kim SH, Seo YS. 10 years of civil electronic litigation: achievements and prospects- focusing on civil lawsuits. Judicial Policy Research Institute, Republic of Korea. 2022. URL: https://jpri.scourt.go.kr/post/postView.do?boardSeq=32&lang=en&menuSeq=35&seq=1476 [accessed 2025-12-21]

XSL•FO
RenderX

14. Yoon K, Kim MS, Jang J, Jonghyun P, Seonmi J, Nahyun K, et al. Digitalization of South Korea's justice system: enhancing access and efficiency. World Bank Group. 2025. URL: https://documents.worldbank.org/en/publication/documents-reports/documentdetail/099012125231020407 [accessed 2026-01-03]
15. Hwang W, Lee D, Cho K, Lee H, Seo M. A multi-task benchmark for Korean legal language understanding and judgement prediction. arXiv Preprint posted online on June 10, 2022. [doi: 10.48550/arXiv.2206.05224]
16. Thomaidou MA, Berryessa CM. Mental illness as a sentencing determinant: a comparative case law analysis based on a machine learning approach. Crim Justice Behav 2023 May 05;50(7):976-995. [doi: 10.1177/00938548231170801]
17. Muñoz-Soro JF, del Hoyo Alonso R, Montañes R, Lacueva F. A neural network to identify requests, decisions, and arguments in court rulings on custody. Artif Intell Law 2024 Jan 09;33:101-135. [doi: 10.1007/s10506-023-09380-9]
18. Legal epidemiology. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/cardiovascular-resources/php/toolkit/legal-epidemiology.html [accessed 2026-02-28]
19. Barros JM, Duggan J, Rebholz-Schuhmann D. The application of internet-based sources for public health surveillance (infoveillance): systematic review. J Med Internet Res 2020 Mar 13;22(3):e13680 [FREE Full text] [doi: 10.2196/13680] [Medline: 32167477]
20. Li J, Xu Q, Cuomo R, Purushothaman V, Mackey T. Data mining and content analysis of the Chinese social media platform Weibo during the early COVID-19 outbreak: retrospective observational infoveillance study. JMIR Public Health Surveill 2020 Apr 21;6(2):e18700 [FREE Full text] [doi: 10.2196/18700] [Medline: 32293582]
21. LBox Corporation. URL: https://lbox.kr/v2 [accessed 2024-04-27]
22. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol 2008 Apr;61(4):344-349. [doi: 10.1016/j.jclinepi.2007.11.008] [Medline: 18313558]
23. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. arXiv Preprint posted online on June 12, 2017. [doi: 10.48550/arXiv.1706.03762]
24. Devlin J, Chang MW, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. arXiv Preprint posted online on October 11, 2018. [doi: 10.48550/arXiv.1810.04805]
25. Hwang W, Eom S, Lee H, Park HJ, Seo M. Data-efficient end-to-end information extraction for statistical legal analysis. arXiv Preprint posted online on November 3, 2022. [doi: 10.48550/arXiv.2211.01692]
26. Xue L, Constant N, Roberts A, Kale M, Al-Rfou R, Siddhant A, et al. mT5: a massively multilingual pre-trained text-to-text transformer. In: Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2021 Presented at: NAACL 2021; Jun 6-11, 2021; Virtual Event. [doi: 10.18653/v1/2021.naacl-main.41]
27. OpenAI. GPT-4 technical report. arXiv Preprint posted online on March 15, 2023. [doi: 10.48550/arXiv.2303.08774]
28. Parnami A, Lee M. Learning from few examples: a summary of approaches to few-shot learning. arXiv Preprint posted online on March 7, 2022. [doi: 10.48550/arXiv.2203.04291]
29. Sentencing guidelines. Sentencing Commission. URL: https://sc.scourt.go.kr/sc/krsc/criterion/down/standard_down.jsp [accessed 2026-03-08]
30. Tkachenko M, Malyuk M, Holmanyuk A, Liubimov N. Label Studio: data labeling software. GitHub. URL: https://github.com/HumanSignal/label-studio [accessed 2024-05-03]
31. Hooley C, Graaf G, Gopalan G. Scaling up evidence-based treatments in youth behavioral healthcare: social work licensing influences on task-shifting opportunities. Hum Serv Organ Manag Leadersh Gov 2021;45(5):375-388 [FREE Full text] [doi: 10.1080/23303131.2021.1970069] [Medline: 35284593]
32. De Maeseneer J, Bourek A, McKee M, Barry M, Brouwer W, Kringos D, et al. Task shifting and health system design: report of the expert panel on effective ways of investing in health (EXPH). Publications Office of the European Union. 2019. URL: https://op.europa.eu/en/publication-detail/-/publication/fd24874a-a9cd-11e9-9d01-01aa75ed71a1/language-en [accessed 2026-04-14]
33. Burris S, Cloud LK, Penn M. The growing field of legal epidemiology. J Public Health Manag Pract 2020;26 Suppl 2, Advancing Legal Epidemiology:S4-S9. [doi: 10.1097/PHH.0000000000001133] [Medline: 32004217]
34. Ribeiro AL, Araújo OR, Oliveira LB, Inácio M. The Executive Branch decisions in Brazil: a study of administrative decrees through machine learning and network analysis. PLoS One 2022 Jul 21;17(7):e0271741 [FREE Full text] [doi: 10.1371/journal.pone.0271741] [Medline: 35862360]
35. Buil-Gil D, Moretti A, Langton SH. The accuracy of crime statistics: assessing the impact of police data bias on geographic crime analysis. J Exp Criminol 2021 Mar 26;18:515-541. [doi: 10.1007/s11292-021-09457-y]
36. Hong T, Kim D, Ji M, Hwang W, Nam D, Park S. BROS: a pre-trained language model focusing on text and layout for better key information extraction from documents. Proc AAAI Conf Artif Intell 2022 Jun 28;36(10):10767-10775. [doi: 10.1609/aaai.v36i10.21322]

Abbreviations

aOR: adjusted odds ratio

XSL•FO
RenderX

aRRR: adjusted relative risk ratio
ML: machine learning
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
UMP: unlicensed medical practice

Edited by A Mavragani, T Sanchez; submitted 31.Jan.2026; peer-reviewed by AMW Lin, A Kodan; comments to author 12.Mar.2026; revised version received 21.May.2026; accepted 30.May.2026; published 15.Jun.2026.

Please cite as:
Kim MJ, Choi M
Legal Infoveillance of Unlicensed Medical Practices in South Korea Through Criminal Court Decisions Using Machine Learning: Retrospective Observational Study
JMIR Public Health Surveill 2026;12:e92591
URL: https://publichealth.jmir.org/2026/1/e92591
doi:10.2196/92591
PMID:

©Min Ji Kim, Min Choi. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 15.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Encouraging Adults at Risk for Type 2 Diabetes to Enroll in Diabetes Prevention Programs Through a Media Campaign in Hawai'i: Cross-Sectional Study

Kara Saiki[1], MPH; Eunjung Lim[2], MS, PhD; Blythe Nett[3], MPH; Gail Ogawa[3], BS; Lance K Ching[3], MPH, PhD; Riana M Kawasaki[1,3], MPH; Nianest Alers Barreto[3], MPH; Katherine Inoue[3], MPH, JD; Lola H Irvin[3], MEd; Kiana R Kaupiko[3], MPH; L Brooke Keliikoa[1], DrPH; Meghan D McGurk[1], MPH

[1]Healthy Hawai'i Evaluation Team, Department of Public Health Sciences, University of Hawai'i at Mānoa, 1960 East-West Rd, Biomed D-210, Honolulu, HI, United States

[2]Department of Quantitative Health Sciences, John A Burns School of Medicine, University of Hawai'i at Mānoa, Honolulu, HI, United States

[3]Chronic Disease Prevention and Health Promotion Division, Hawai'i State Department of Health, Honolulu, HI, United States

Corresponding Author:
Kara Saiki, MPH
Healthy Hawai'i Evaluation Team, Department of Public Health Sciences, University of Hawai'i at Mānoa, 1960 East-West Rd, Biomed D-210, Honolulu, HI, United States

Abstract

Background: The National Diabetes Prevention Program (DPP) is an evidence-based intervention proven to delay or prevent progression to type 2 diabetes, yet most at-risk people do not enroll. In Hawai'i, Native Hawaiian and Other Pacific Islander (NHOPI) and Filipino adults experience disproportionately high rates of prediabetes and diabetes but have low DPP enrollment. From July to October 2024, the Hawai'i State Department of Health launched Beat Diabetes, a statewide media campaign encouraging DPP enrollment among at-risk adults, with a focus on NHOPI and Filipino communities.

Objective: This evaluation assessed whether campaign exposure was associated with self-reported likelihood of joining a DPP among Hawai'i adults at risk for diabetes, particularly NHOPI or Filipino adults.

Methods: A postcampaign cross-sectional online survey was conducted from October to December 2024, with Hawai'i residents aged 35-64 years who reported at least 1 diabetes risk factor. NHOPI or Filipino adults were oversampled to determine campaign effectiveness among the target audience. The survey measured self-reported likelihood of joining a lifestyle change program (main outcome), campaign recall (main exposure), demographic characteristics, diabetes risk factors, and beliefs that could affect DPP enrollment likelihood, including intrinsic motivation, perceived inevitability of developing diabetes, and perceived health benefits of DPP participation. Three general linear regression models examined the association between campaign exposure and DPP enrollment likelihood ratings, adjusted for demographic characteristics, diabetes risk factors, and belief variables. A sensitivity analysis among just those diagnosed with prediabetes was conducted.

Results: A total of 860 adults completed the survey, with 34.7% (298/860) and 12.2% (105/860) self-identifying as NHOPI and Filipino, respectively. In total, 40% (346/860) reported campaign exposure. Exposed individuals had higher mean DPP enrollment likelihood ratings and higher inevitability belief scores than those not exposed. A large proportion of exposed respondents reported that enrolling in a DPP would "improve their health a lot." No significant differences in campaign exposure were observed across ethnicities. All 3 regression models showed a significant positive association between campaign exposure and DPP enrollment likelihood ratings. In the final adjusted model controlling for all covariates, significant predictors included campaign exposure (β=.52, P<.001), male gender (β=.34, P=.01), residence outside Honolulu County (β=.31, P=.02), motivation index scores (β=.38, P<.001), inevitability belief (β=.20, P<.001), and the belief that DPP improves health "a little" (β=.76, P<.001) or "a lot" (β=1.63, P<.001). The sensitivity analysis among those diagnosed showed exposure was not associated with likelihood ratings (β=.30, P=.27).

Conclusions: Campaign exposure was associated with higher ratings of likelihood to join a DPP among at-risk adults with no prediabetes diagnosis. Perceived positive health impact of DPP participation was the strongest contributor to likelihood ratings. Campaigns aiming to increase awareness of DPP and intentions to join should promote DPP effectiveness and the urgency of preventative actions.

(JMIR Public Health Surveill 2026;12:e90880) doi:10.2196/90880

XSL•FO
RenderX

KEYWORDS

prediabetes; diabetes prevention program; public health; intention; motivation; media campaign

Introduction

Background

In the United States, approximately 97.6 million adults have prediabetes [1], placing more than one-third of the adult population at risk for developing type 2 diabetes. Although the statewide prevalence of diagnosed prediabetes in Hawai'i is lower than the national estimate (14.9% vs 38%) [1,2], disaggregated data reveal notable differences among racial and ethnic groups. Native Hawaiian (17.2%), Other Pacific Islander (16.9%), and Filipino (17.3%) adults have higher prevalence of diagnosed prediabetes than non-Hispanic White (9.0%) adults [2]. These groups, along with non–Hispanic Black and Hispanic adults, are recognized nationally as having elevated risk for type 2 diabetes [3].

Established risk factors for prediabetes and diabetes include being older than 40 years, having a family history of diabetes, having high blood pressure, physical inactivity, a history of gestational diabetes, and being overweight [4,5]. In 2021, the US Preventive Services Task Force lowered the recommended prediabetic screening age for adults who are overweight or obese from 40 to 35 years, citing these risk factors as the strongest predictors for progressing to diabetes [6]. Identifying individuals at high risk for prediabetes and intervening early is critical to preventing progression to type 2 diabetes.

The National Diabetes Prevention Program (DPP) is an evidence-based lifestyle intervention that can reduce the risk of progressing from prediabetes to type 2 diabetes by up to 58% [7]. Despite the success of the program, less than 1% of US adults with prediabetes enroll in the program [8]. Enrollment data show that certain groups such as Native Hawaiian and Other Pacific Islander (NHOPI) and Filipino adults are disproportionately affected by prediabetes [9] and yet enroll in DPPs at lower rates than Hispanic, non–Hispanic Black, and non–Hispanic White adults [8].

Multiple factors contribute to low enrollment. Extrinsic barriers include program cost, limited time to participate, long wait periods before enrollment, lack of online options, and competing health conditions [10-17]. Intrinsic barriers also play a role. Low autonomous motivation, defined as low internal drive to adopt healthier behaviors [18], and belief that diabetes is inevitable due to family history may reduce individuals' willingness to engage in prevention efforts [19,20]. Together, these barriers can hinder adults with prediabetes from taking steps to prevent disease progression.

Campaign Development

In 2024, the Hawai'i State Department of Health (HDOH) conducted focus groups to better understand why adults diagnosed with prediabetes or at high risk for prediabetes do not engage in lifestyle changes and what messages might motivate enrollment in a DPP. Focus groups were segmented by the largest racial and ethnic groups in Hawai'i: Native Hawaiian, Other Pacific Islander, Filipino, and Other (White and Japanese) adults with focused recruitment of NHOPI or Filipino adults due to their high rates of prediabetes. Across groups, most participants reported that they did not have time to prioritize their health. Some participants felt little urgency to join a lifestyle change program, stating that they would be more likely to seek help only after receiving a diabetes diagnosis. Others were uncertain about how to start making a change but responded positively to messages that were direct and encouraging. Participants preferred campaign materials featuring people from diverse backgrounds, including NHOPI or Filipino adults.

Guided by these findings, the HDOH developed the Beat Diabetes media campaign to increase urgency around diabetes prevention and provide a clear call to action by joining DPP, a lifestyle change program. The campaign aimed to increase DPP enrollment overall, particularly for NHOPI or Filipino adults at risk for prediabetes and diabetes, by addressing intrinsic barriers such as low motivation and perceived inevitability of diabetes. The campaign consisted of 6 still advertisements and 1 video advertisement featuring NHOPI or Filipino adults delivering messages such as: "I'm not waiting. I'm preventing diabetes. Join the program." and "Tomorrow's a hard nope. I'm preventing diabetes today. Join the program." (Multimedia Appendix 1). Each advertisement contained a link to the landing page for the Beat Diabetes website, which provided more information about prediabetes, diabetes risk test, frequently asked questions, and information on how to enroll in a DPP [21]. The campaign ran from July to October 2024 across statewide television, radio, digital platforms, print outlets, and mall posters.

To evaluate the effect of the campaign on self-reported likelihood of joining a lifestyle change program, particularly among the campaign's target audience of NHOPI or Filipino adults at risk for diabetes, the HDOH collaborated with the Healthy Hawai'i Evaluation Team (HHET) to develop a postcampaign cross-sectional survey. A market research firm was contracted to administer the survey.

Methods

Study Design

This cross-sectional study followed the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) reporting guideline and checklist to draft this manuscript (Checklist 1) [22] and the CHERRIES (Checklist for Reporting Results of Internet E-Surveys) statement to report on web-based surveys (Checklist 2) [23].

Survey Development

The survey was developed by the HDOH and HHET teams, and programming of the screening and core survey questions was checked and revised by the market research firm. The online survey was pretested with 25 eligible respondents before fielding. It took approximately 10 - 12 minutes to complete and was composed of 7 screening questions and approximately 50

XSL•FO
RenderX

core questions in English. The survey could be viewed multiple ways: on a mobile phone showing 1 question per screen or on a desktop computer displaying multiple questions at once for a total of 24 screens.

Participant Recruitment

Data collection occurred between October and December 2024. A convenience sample of adults aged 35 - 64 years was recruited from 3 proprietary statewide market research consumer panels. The market research firm emailed eligible panelists an invitation to complete a closed online survey via a secured survey link, following the conclusion of the Beat Diabetes campaign. Consumer panels were selected over other sampling methods as they are a more cost-effective method to evaluate health department media campaigns to provide timely feedback for campaign improvement [24]. Inclusion criteria included residency in Hawai'i for at least 6 months and a diagnosis of either prediabetes or borderline diabetes, or at least 1 recognized prediabetes risk factor as determined by a diabetes risk test. Individuals with a prior diabetes diagnosis were excluded, as they are ineligible for DPP enrollment and may instead join a diabetes self-management education and support program. A minimum of 700 responses were required, with a goal of 900 responses determined by the budget set aside for recruitment. NHOPI and Filipino adults were oversampled to ensure that they made up 50% of the end sample, a large enough sample size to examine campaign recall and effectiveness within the campaign's target audience.

Ethical Considerations

This evaluation was reviewed by the HDOH Institutional Review Board, who confirmed it qualified as nonhuman subjects research under the revised Common Rule of 2018. Informed consent identified the purpose of the survey and who was conducting the investigation, identified the market research firm assisting with data collection, and gave details on the survey length and time, confidentiality and deidentification of data, and how participation was voluntary before eligible participants accessed the online survey. All survey responses were kept confidential on password-protected computers and deidentified prior to analysis. Incentives were provided by the marketing firm via each respondent's email address and ranged from points to gift cards valued at US $10, depending on the panel. The marketing firm retained email addresses in a separate working file used solely for incentive distribution, with no link to survey responses so that no personal identifier could be tracked.

Survey Measures

The full postsurvey instrument is provided in Multimedia Appendix 2. Survey measures included the following:

1. DPP enrollment likelihood rating: The main evaluation outcome was participants' self-reported likelihood of joining a lifestyle change program. Participants were given a short description of such a program and asked, "How likely are you to consider joining a lifestyle change program such as *Beat Diabetes*?" Responses were recorded on a 7-point Likert scale ranging from 1 (*not at all likely*) to 7 (*very likely*). DPP enrollment likelihood ratings were treated as continuous in all analyses.
2. Campaign exposure: Campaign exposure was defined by self-reported recall of any of the campaign advertisements. Participants were shown 4 different still advertisements and 1 video advertisement and asked whether they had seen or heard each. Advertisements were rotated to reduce response bias. If participants had not seen or heard any of these advertisements, they were shown 2 additional still advertisements in a single question. Recall of any of the 7 advertisements was coded as exposure to the campaign.
3. Demographic characteristics: Participants reported their age, gender (woman and man), education, household income, and county of residence (Honolulu County and other counties). Health insurance status and type (private, Medicare, Medicaid, and TRICARE/other) were also collected. Participants reported the group that best represented their race or ethnicity. Race or ethnicity was coded into 3 categories: Filipino, NHOPI, and Other. Participants indicating Filipino alone or along with another race or ethnicity were categorized as Filipino. Participants reporting at least one of the following ethnicities alone or with any other race or ethnicity (ie, multiethnic participants)—Native Hawaiian or Part Hawaiian and Other Pacific Islander (including Samoan, Tongan, Guamanian or Chamorro, Fijian, and peoples of Micronesia and Melanesia)—were categorized as NHOPI. All other racial and ethnic groups were categorized as "Other."
4. Diabetes risk factors: Screening questions were used to identify diabetes risk factors, including self-perceived overweight status, physical inactivity (less than 3 times per week), hypertension diagnosis, and prediabetes diagnosis. Family history risk was determined by responses to "My family members or relatives have diabetes" in the screening question and "Does your family have a history of diabetes?" in the main survey. Gestational diabetes risk combined 2 items: "I have given birth to a baby that weighed more than 9 pounds" and "I had gestational diabetes in a past pregnancy." Only those indicating woman as gender were asked about gestational diabetes risk.
5. Motivation index score: Four items adapted from the Treatment Self-Regulation Questionnaire [18] assessed participants' autonomous motivation to prevent diabetes. Participants rated each statement on a 7-point scale from 1 (*not at all true*) to 7 (*very true*). Each statement started with, "The reason I would take steps to help prevent diabetes is that..." (1) "It is very important for being as healthy as possible," (2) "I personally believe it is the best thing for my health," (3) "I feel that I want to take responsibility for my own health," and (4) "It is consistent with my life goals." All items were treated as continuous, with interitem correlations ranging from 0.54 to 0.72. The 4 items were averaged to create a motivation index score for each participant. The scale demonstrated good reliability with a Cronbach α of 0.86 (95% CI 0.8 - 0.9).
6. Inevitability belief: Participants' perceptions of the inevitability of developing diabetes were measured using the statement, "It is only a matter of time before I get diabetes," rated on the same 7-point scale for the motivation items. Scores were treated as continuous.

XSL•FO
RenderX

7. Perceived DPP effectiveness: Participants' beliefs about the effectiveness of DPP were measured with the question, "If you joined a lifestyle program such as Beat Diabetes, do you think your health would...*decline*, *stay the same*, *improve a little*, or *improve a lot*." For analyses, *decline* and *stay the same* responses were combined to indicate negative beliefs about the effectiveness of DPP.

Data Integrity

Data quality was continuously monitored throughout collection. Through the market research firm's data-cleaning protocols, surveys that took less than 5 minutes to complete were excluded from the completed surveys, as were those with duplicate IP addresses. Completeness of each survey was monitored by the market research firm and the number of completed surveys was reported in weekly field status updates alongside numbers of emailed contacts, those ineligible after screening, and surveys abandoned postscreening questions. Respondents were not able to review or change their answers once they submitted the survey. A final data cleaning was done to remove responses suspected to be bots due to open-ended responses that were nonsensical or inconsistent in their responses to multiple questions. The remaining responses were included in analysis.

Data Analysis

Descriptive statistics were reported as frequencies and percentages for categorical variables and as means and standard deviations for continuous variables. Bivariate analyses were conducted to examine potential associations between campaign exposure, DPP enrollment likelihood ratings, and covariates using 2-sample *t* tests or chi-square tests, as appropriate. Missing data for gender and education were excluded from analysis.

Three general linear regression models were used to assess the association between campaign exposure and DPP enrollment likelihood ratings, controlling for covariates. Model 1 included demographic variables only, model 2 added diabetes risk factors to model 1, and model 3 further adjusted for motivation index score, inevitability belief, and perceived DPP effectiveness. Graphical tests were conducted to assess model assumptions. Because slight deviations from normality were observed, sensitivity analyses using ordered logistic regression (analyses not shown) were conducted to identify potential deviations from the general linear regression results. No significant deviations were detected and thus general linear regression models are presented. Multicollinearity among independent variables was evaluated using the variance inflation factor. Subgroup analyses were conducted to examine race in a disaggregated manner using model 3 stratified by race or ethnicity. A sensitivity analysis was also conducted using model 3 among participants diagnosed with prediabetes to evaluate consistency with the overall findings. All analyses were conducted using R (version 4.5.1; R Core Team). A *P* value less than .05 was considered statistically significant.

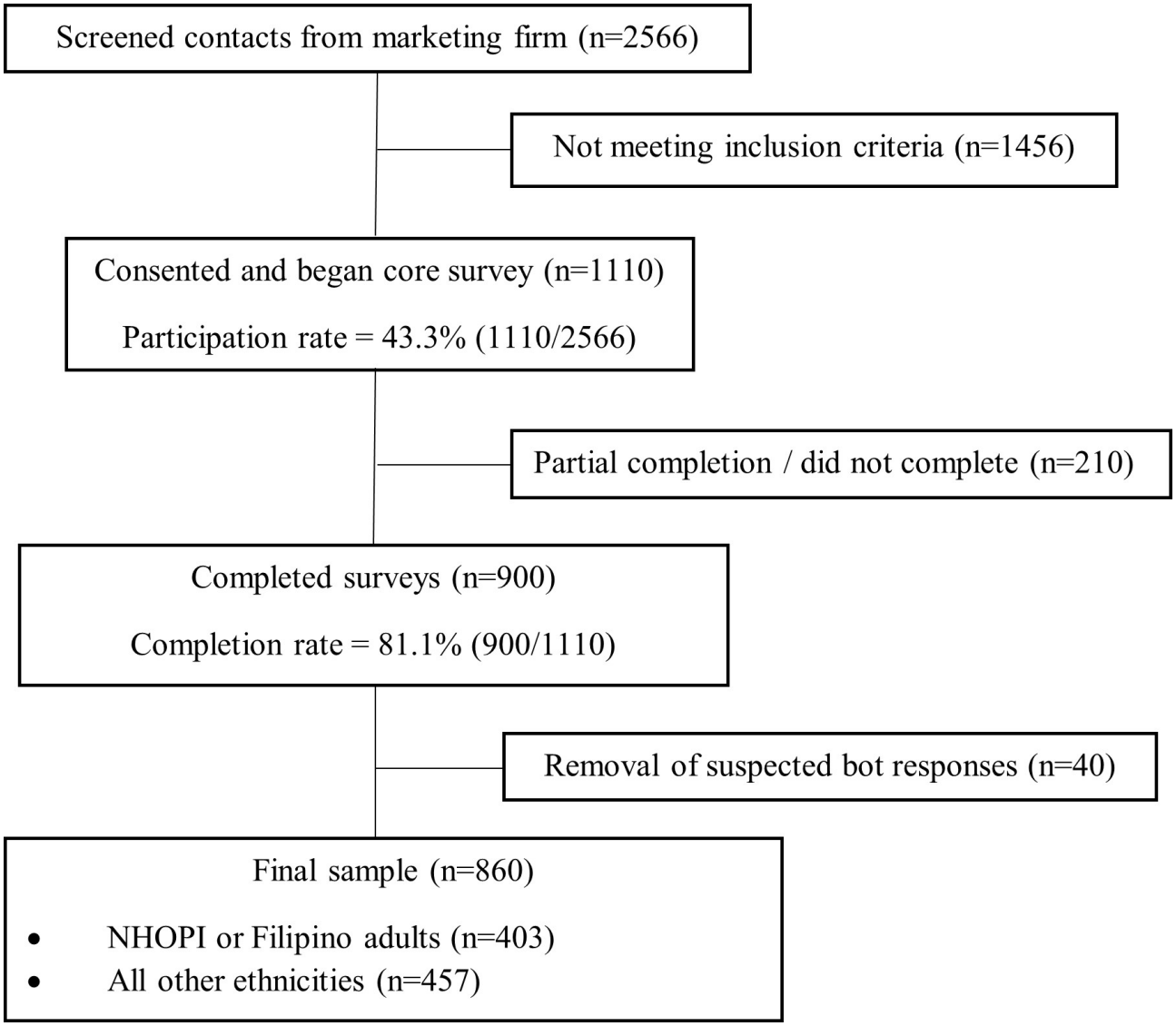
Results

Final Sample

A total of 2566 panel participants responded to the email invitation and were screened. Of these, 1456 were excluded for not meeting inclusion criteria and 210 were excluded for incomplete surveys, resulting in 900 completed surveys. An additional 40 responses were removed due to suspected bot activity, yielding a final analytic sample of 860 respondents (Figure 1).

XSL•FO
RenderX

Figure 1. Flow diagram of sample development for a cross-sectional study of associations between campaign exposure and ratings of likelihood to join a National Diabetes Prevention Program among adults at risk for diabetes. NHOPI: Native Hawaiian and Other Pacific Islander.



Sample Characteristics

The mean age was 49.2 (SD 8.5 years), with 64.7% (556/860) identifying as women, reflecting an overrepresentation of women compared with the state's demographic composition [25]. Approximately half the sample, or 46.9% (403/860), self-identified as NHOPI or Filipino adults. Most participants reported a family history of diabetes (570/860, 66.3%) or perceived themselves as overweight (455/860, 52.9%). Despite all participants being classified as at risk based on screening criteria, only 22.7% (195/860) reported a prior diagnosis of prediabetes (Table 1).

XSL•FO
RenderX

Table . Participant characteristics by overall sample and those exposed and not exposed to the media campaign[a].

Variable	Overall (N=860)	Exposed (n=346)	Not exposed (n=514)	*P* value[b]
Age (years), mean (SD)	49.2 (8.5)	49.4 (8.6)	49.0 (8.5)	.48
Gender, n (%)				.40
Woman	556 (64.7)	217 (39.0)	339 (61.0)	
Man	293 (34.1)	123 (42.0)	170 (58.0)	
Missing[c]	11 (1.3)	6	5	
Race or ethnicity, n (%)				.22
Filipino	105 (12.2)	34 (32.4)	71 (67.6)	
NHOPI[d]	298 (34.7)	123 (41.3)	175 (58.7)	
Other	457 (53.1)	189 (41.4)	268 (58.6)	
Education, n (%)				.16
HS[e] or lower	151 (17.7)	50 (33.1)	101 (66.9)	
Some college	279 (32.6)	113 (40.5)	166 (59.5)	
College graduate	263 (30.8)	117 (44.5)	146 (55.5)	
Post college graduate	162 (18.9)	65 (40.1)	97 (59.9)	
Missing[c]	5 (0.6)	1	4	
Household income (US dollars), n (%)				.052
<$25,000	105 (12.2)	40 (38.1)	65 (61.9)	
$25,000-$50,000	148 (17.2)	44 (29.7)	104 (70.3)	
$50,001-$100,000	260 (30.2)	118 (45.4)	142 (54.6)	
$100,001-$150,000	173 (20.1)	75 (43.4)	98 (56.6)	
>$150,000	133 (15.5)	51 (38.3)	82 (61.7)	
Don’t know/missing	41 (4.8)	18 (43.9)	23 (56.1)	
Insurance type, n (%)				.13
Private	514 (59.8)	222 (43.2)	292 (56.8)	
Medicare	62 (7.2)	23 (37.1)	39 (62.9)	
Medicaid	151 (17.6)	57 (37.7)	94 (62.3)	
TRICARE/Other	89 (10.3)	33 (37.1)	56 (62.9)	
None/Missing	44 (5.1)	11 (25.0)	33 (75.0)	
County, n (%)				.02
Honolulu County	631 (73.4)	239 (37.9)	392 (62.1)	
All other counties	229 (26.6)	107 (46.7)	122 (53.3)	
Diabetes risk factors, n (%)				
Family history				<.001
Yes	570 (66.3)	257 (45.1)	313 (54.9)	
No	290 (33.7)	89 (30.7)	201 (69.3)	
Overweight perception				.03
Yes	455 (52.9)	199 (43.7)	256 (56.3)	
No	405 (47.1)	147 (36.3)	258 (63.7)	
Not physically active (<3 times per week)				.27

XSL•FO
RenderX

Variable	Overall (N=860)	Exposed (n=346)	Not exposed (n=514)	*P* value[b]
Yes	319 (37.1)	136 (42.6)	183 (57.4)	
No	541 (62.9)	210 (38.8)	331 (61.2)	
Gestational diabetes				.56
Yes	78 (9.1)	29 (37.2)	49 (62.8)	
No	782 (90.9)	317 (40.5)	465 (59.5)	
Diagnosed with high blood pressure				.83
Yes	237 (27.6)	94 (39.7)	143 (60.3)	
No	623 (72.4)	252 (40.4)	371 (59.6)	
Diagnosed with prediabetes				.04
Yes	195 (22.7)	91 (46.7)	104 (53.3)	
No	665 (77.3)	255 (38.3)	410 (61.7)	
Motivation index, mean (SD)	6.3 (0.9)	6.4 (0.8)	6.3 (1.0)	.26
Inevitability belief, mean (SD)	3.4 (2.1)	3.7 (2.2)	3.2 (2.0)	<.001
Perceived DPP[f] effectiveness, n (%)				.04
Decline or stay the same	99 (11.5)	35 (35.4)	64 (64.6)	
Improve a little	302 (35.1)	110 (35.8)	192 (64.2)	
Improve a lot	364 (42.3)	167 (36.4)	197 (63.6)	
Don't know	95 (11.0)	34 (45.9)	61 (54.1)	
DPP enrollment likelihood ratings, mean (SD)	3.6 (1.9)	4.1 (2.0)	3.2 (1.8)	<.001

[a]Row percentages are used to compare Exposed and Not exposed.

[b]Two-sample *t* test or chi-square test was used.

[c]Missing on gender and education were excluded from analyses.

[d]NHOPI: Native Hawaiian and Other Pacific Islander.

[e]HS: High school.

[f]DPP: National Diabetes Prevention Program.

Sample Characteristics by Campaign Exposure

Table 1 also presents participant characteristics stratified by campaign exposure. Overall, 40.2% (346/860) of the sample reported exposure to the campaign. Among the 298 NHOPI or 105 Filipino participants, 123 (41.3%) and 34 (32.4%) reported being exposed to the campaign, respectively. Exposure to the campaign did not differ significantly across NHOPI, Filipino, and other racial or ethnic participant groups.

Significant differences between those exposed and not exposed to the campaign were observed by county of residence, specific diabetes risk factors, inevitability belief, and perceived DPP effectiveness. A smaller proportion of exposed participants resided in Honolulu County (239/631, 37.1%) compared with those not exposed (392/631, 62.1%; *P*=.02). Additionally, exposed participants were less likely to report a family history of diabetes (257/570, 45.1% vs 313/570, 54.9%; *P*<.001), perceive themselves as overweight (199/455, 45.1% vs 256/455, 56.3%; *P*=.03), or have been previously diagnosed with prediabetes (91/195, 46.7% vs 104/195, 53.3%; *P*=.04). Participants exposed to the campaign reported higher mean likelihood ratings for joining a DPP than those not exposed (mean 4.1, SD 2.0 vs mean 3.2, SD 1.8; *P*<.001) and had significantly higher mean scores for inevitability belief (mean 3.7, SD 2.2 vs mean 3.2, SD 2.0; *P*<.001). Fewer exposed participants believed that participation in a program such as Beat Diabetes would *improve* their health *a lot* (167/364, 42.3% vs 197/364, 63.6%; *P*=.04).

Linear Regression Models on DPP Enrollment Likelihood Rating

Table 2 presents results from the 3 linear regression models. All models showed a significant positive effect of campaign exposure. In model 1, controlling for demographic characteristics, campaign exposure (β=.76, *P*<.001), age (β=−.02, *P*=.01), and insurance type (*P*=.02) were significantly associated with likelihood of joining a DPP.

XSL•FO
RenderX

Table . Linear regression models predicting National Diabetes Prevention Program enrollment likelihood ratings by campaign exposure, controlling for demographic characteristics, diabetes risk factors, and intrinsic factors.

	Model 1: demographic characteristics				Model 2: diabetes risk factors				Model 3: intrinsic factors			
Variable	β	SE	95% CI	*P* value	β	SE	95% CI	*P* value	β	SE	95% CI	*P* value
Intercept	4.50	0.54	3.43 to 5.57	<.001	3.20	0.58	2.05 to 4.34	<.001	-.55	0.68	-1.88 to 0.78	.42
Campaign exposure (ref[a]: no)				<.001				<.001				<.001
Yes	.76	0.13	0.50 to 1.03		.67	0.13	0.41 to 0.93		.52	0.12	0.28 to 0.75	
Age (years)	-.02	0.01	-0.04 to 0.00	.01	-.01	0.01	-0.03 to 0.01	.19	-.01	0.01	-0.03 to 0.00	.15
Gender (ref: women)				.16				.049				.01
Man	.20	0.14	-0.08 to 0.47		.28	0.14	0.00 to 0.56		.34	0.13	0.09 to 0.59	
Race/ethnicity (ref: Filipino)				.15				0.43				.19
NHOPI[b]	-.04	0.22	-0.47 to 0.40		.02	0.22	-0.41 to 0.44		.15	0.2	-0.24 to 0.54	
Other	-.29	0.21	-0.70 to 0.12		-.17	0.21	-0.57 to 0.24		-.10	0.19	-0.47 to 0.27	
Education (ref: HS[c] or lower)				.27				0.29				.13
Some college	.25	0.2	-0.15 to 0.64		.14	0.2	-0.25 to 0.52		.14	0.18	-0.21 to 0.49	
College graduate	.23	0.22	-0.20 to 0.65		.17	0.21	-0.25 to 0.59		.15	0.19	-0.23 to 0.53	
Post college graduate	.47	0.24	0.00 to 0.95		.44	0.24	-0.03 to 0.91		.47	0.22	0.05 to 0.90	
Household income, US dollars (ref: <$25,000)				.53				.75				.85
$25,000-$50,000	-.24	0.26	-0.74 to 0.26		-.10	0.25	-0.59 to 0.39		-.11	0.23	-0.55 to 0.34	
$50,001-$100,000	-.37	0.25	-0.86 to 0.13		-.24	0.25	-0.72 to 0.25		-.27	0.22	-0.71 to 0.17	
$100,001-$150,000	-.24	0.28	-0.79 to 0.31		-.13	0.28	-0.67 to 0.41		-.27	0.25	-0.76 to 0.22	
>$150,001	-.37	0.3	-0.95 to 0.21		-.26	0.29	-0.83 to 0.32		-.29	0.27	-0.81 to 0.23	
Don't know/missing	-.68	0.37	-1.41 to 0.06		-.52	0.37	-1.24 to 0.21		-.29	0.33	-0.95 to 0.36	

XSL•FO
RenderX

	Model 1: demographic characteristics				Model 2: diabetes risk factors				Model 3: intrinsic factors			
Variable	β	SE	95% CI	*P* value	β	SE	95% CI	*P* value	β	SE	95% CI	*P* value
Insurance type (ref: private)				.02				.22				.27
Medicare	.00	0.27	-0.54 to 0.54		-.01	0.27	-0.54 to 0.52		.17	0.25	-0.31 to 0.65	
Medicaid	-.50	0.22	-0.93 to -0.08		-.32	0.21	-0.74 to 0.10		-.34	0.2	-0.72 to 0.05	
TRICARE/Other	-.18	0.22	-0.62 to 0.25		-.03	0.22	-0.46 to 0.40		.05	0.2	-0.35 to 0.44	
None/missing	-.90	0.32	-1.54 to -0.27		-.66	0.32	-1.29 to -0.03		-.24	0.29	-0.81 to 0.33	
County (ref: Honolulu County)				.09				.17				.02
All other counties	.26	0.15	-0.04 to 0.56		.21	0.15	-0.09 to 0.50		.31	0.14	0.05 to 0.58	
Diabetes risk factors												
Family history	N/A[d]	N/A	N/A	N/A	.33	0.14	0.05 to 0.61	.02	.11	0.13	-0.15 to 0.37	.41
Overweight perception	N/A	N/A	N/A	N/A	.47	0.14	0.19 to 0.75	.001	.19	0.13	-0.07 to 0.45	.15
Not physically active (<3 x per week)	N/A	N/A	N/A	N/A	.28	0.14	0.00 to 0.55	.047	.19	0.13	-0.06 to 0.44	.14
Gestational diabetes	N/A	N/A	N/A	N/A	.15	0.23	-0.31 to 0.60	.53	.03	0.21	-0.39 to 0.44	.90
Diagnosed with high blood pressure	N/A	N/A	N/A	N/A	.01	0.15	-0.28 to 0.31	.94	-.11	0.14	-0.37 to 0.16	.43
Diagnosed with prediabetes	N/A	N/A	N/A	N/A	.37	0.16	0.05 to 0.69	.02	.16	0.15	-0.14 to 0.45	.30
Motivation index	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	.38	0.07	0.25 to 0.51	<.001
Inevitability belief	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	.20	0.03	0.14 to 0.26	<.001

XSL•FO
RenderX

	Model 1: demographic characteristics				Model 2: diabetes risk factors				Model 3: intrinsic factors			
Variable	β	SE	95% CI	*P* value	β	SE	95% CI	*P* value	β	SE	95% CI	*P* value
Perceived DPP[e] effectiveness (ref: Decline/stay the same)				N/A				N/A				<.001
Don't know	N/A	N/A	N/A		N/A	N/A	N/A		.38	0.25	-0.11 to 0.87	
Improve a little	N/A	N/A	N/A		N/A	N/A	N/A		.76	0.2	0.36 to 1.16	
Improve a lot	N/A	N/A	N/A		N/A	N/A	N/A		1.63	0.2	1.23 to 2.03	

[a]Ref: reference category.

[b]NHOPI: Native Hawaiian and Other Pacific Islander.

[c]HS: high school.

[d]N/A: not applicable.

[e]DPP: National Diabetes Prevention Program.

Model 2 added diabetes risk factors. The association between campaign exposure and DPP enrollment likelihood ratings decreased slightly but remained significant (β=.67, *P*<.001). Significant predictors included male gender (β=.28, *P*=.049), family history of diabetes (β=.33, *P*=.02), self-perceived overweight status (β=.47, *P*=.001), physical inactivity (β=.28, *P*=.047), and prediabetes diagnosis (β=.37, *P*=.02). Age and insurance type were no longer significant. History of gestational diabetes and hypertension were not significantly associated with DPP enrollment likelihood ratings.

Model 3 further adjusted for intrinsic factors. Campaign exposure remained significant, albeit slightly reduced (β=.52, *P*<.001). Significant predictors included male gender (β=.33, *P*=.01), residence outside Honolulu County (β=.31, *P*=.02), motivation index (β=.38, *P*<.001), inevitability belief (β=.20, *P*<.001), and perceived effectiveness of DPP ("improve a little," β=.76; "improve a lot," β=1.63, *P*<.001). Diabetes risk factors were no longer significant.

No multicollinearity was detected (variance inflation factor <2). The variance explained by the models increased across model 1 to model 3 but remained low (R^2=0.08, 0.13, and 0.23, respectively). Graphical assessment of model assumptions indicated slight deviation from the normality assumption for model 1. Sensitivity analyses using ordered logistic regression showed similar results (not shown).

Table 3 presents the results of the subgroup analyses by race or ethnicity using model 3. Campaign exposure was positively associated with the outcome DPP enrollment likelihood ratings among NHOPI and other racial or ethnic groups, but the association did not reach statistical significance among Filipino adults. Notably, the estimated effect for the Filipino subgroup was larger in magnitude than those observed for NHOPI and other racial or ethnic groups and approached statistical significance. Across all subgroups, motivation index scores, inevitability belief, and perceived DPP effectiveness were consistently significant predictors.

XSL•FO
RenderX

Table . Subgroup analysis—linear regression models on National Diabetes Prevention Program enrollment likelihood ratings stratified by race or ethnicity.

Variable	Filipino				NHOPI[a]				All other races or ethnicities			
	β	SE	95% CI	*P* value	β	SE	95% CI	*P* value	β	SE	95% CI	*P* value
Intercept	−2.35	1.87	−6.02 to 1.32	.21	−.74	1.05	−2.80 to 1.31	.48	−1.00	0.99	−2.94 to 0.94	.31
Exposed (ref[b]: no)				.09				.04				.01
Yes	.65	0.38	−0.09 to 1.40		.45	0.21	0.03 to 0.87		.45	0.17	0.11 to 0.79	
Age (years)	.00	0.02	−0.04 to 0.04	.95	.00	0.01	−0.03 to 0.03	.94	−.01	0.01	−0.03 to 0.01	.27
Gender (ref: women)				.51				.99				.003
Man	−.26	0.39	−1.03 to 0.51		.00	0.26	−0.50 to 0.50		.51	0.17	0.17 to 0.85	
Education (ref: HS[c] or lower)				.53				.15				.58
Some college	−.53	0.55	−1.60 to 0.55		.19	0.27	−0.34 to 0.72		.29	0.29	−0.28 to 0.86	
College graduate	−.43	0.60	−1.60 to 0.73		.44	0.32	−0.19 to 1.06		.15	0.30	−0.45 to 0.74	
Post college graduate	.10	0.66	−1.21 to 1.40		.88	0.40	0.10 to 1.66		.36	0.32	−0.27 to 0.99	
Household income, US dollars (ref: <$25,000)				.28				.77				.85
$25,000-$50,000	1.08	0.76	−0.41 to 2.57		−.19	0.37	−0.91 to 0.53		.18	0.34	−0.49 to 0.84	
$50,001-$100,000	.93	0.78	−0.61 to 2.46		−.29	0.38	−1.04 to 0.46		−.06	0.32	−0.68 to 0.56	
$100,001-$150,000	.11	0.82	−1.50 to 1.72		.11	0.46	−0.79 to 1.01		−.25	0.35	−0.92 to 0.43	
>$150,001	1.03	0.93	−0.79 to 2.86		−.44	0.49	−1.40 to 0.51		−.18	0.36	−0.89 to 0.53	
Don't know/missing	1.15	0.94	−0.68 to 2.98		−.32	0.58	−1.46 to 0.82		−.08	0.50	−1.06 to 0.90	
Insurance type (ref: private)				.95				.001				.44
Medicare	−.52	1.68	−3.81 to 2.76		1.48	0.43	0.63 to 2.33		−.43	0.32	−1.07 to 0.20	
Medicaid	−.01	0.51	−1.01 to 0.98		.03	0.32	−0.59 to 0.65		−.50	0.30	−1.09 to 0.09	

XSL•FO
RenderX

Variable	Filipino				NHOPI[a]				All other races or ethnicities			
	β	SE	95% CI	*P* value	β	SE	95% CI	*P* value	β	SE	95% CI	*P* value
TRICARE/Other	−.42	0.58	−1.56 to 0.73		.72	0.51	−0.28 to 1.71		−.05	0.25	−0.54 to 0.44	
None/missing	.06	0.94	−1.78 to 1.90		−.13	0.56	−1.23 to 0.96		−.33	0.39	−1.10 to 0.44	
County (ref: Honolulu County)				.08				.18				.21
All other counties	.80	0.45	−0.08 to 1.68		.32	0.23	−0.14 to 0.77		.24	0.19	−0.14 to 0.62	
Diabetes risk factors												
Family history	−.68	0.38	−1.42 to 0.06	.07	.17	0.26	−0.32 to 0.67	.49	.17	0.18	−0.17 to 0.52	.33
Overweight perception	−.03	0.37	−0.76 to 0.69	.93	.41	0.24	−0.05 to 0.87	.08	.17	0.19	−0.20 to 0.54	.37
Not physically active (<3x per week)	.65	0.40	−0.12 to 1.43	.10	.05	0.22	−0.38 to 0.49	.81	.25	0.19	−0.11 to 0.61	.18
Gestational diabetes	−.77	0.65	−2.04 to 0.49	.23	−.18	0.34	−0.84 to 0.48	.59	.29	0.33	−0.35 to 0.94	.37
Diagnosed with high blood pressure	.33	0.35	−0.36 to 1.02	.35	−.08	0.23	−0.54 to 0.38	.72	−.26	0.20	−0.65 to 0.13	.19
Diagnosed with prediabetes	.20	0.40	−0.58 to 0.99	.61	−.05	0.26	−0.55 to 0.45	.84	.36	0.22	−0.07 to 0.78	.10
Motivation index	.64	0.19	0.27 to 1.01	.001	.35	0.11	0.13 to 0.57	.002	.39	0.10	0.19 to 0.59	<.001
Inevitability belief	.15	0.09	−0.03 to 0.33	.11	.20	0.05	0.10 to 0.30	<.001	.20	0.05	0.11 to 0.29	<.001
Perceived DPP[d] effectiveness (ref: decline/stay the same)				<.001				<.001				<.001
Don't know	−.55	0.82	−2.15 to 1.04		.32	0.43	−0.52 to 1.15		.61	0.36	−0.10 to 1.32	
Improve a little	.14	0.64	−1.11 to 1.38		.50	0.37	−0.23 to 1.22		1.00	0.28	0.46 to 1.54	

XSL•FO
RenderX

Variable	Filipino				NHOPI[a]				All other races or ethnicities			
	β	SE	95% CI	*P* value	β	SE	95% CI	*P* value	β	SE	95% CI	*P* value
Improve a lot	1.73	0.65	0.46 to 3.00		1.29	0.36	0.60 to 1.99		1.77	0.29	1.21 to 2.33	

[a]NHOPI: Native Hawaiian and Other Pacific Islander.

[b]Ref: reference category.

[c]HS: high school.

[d]DPP: National Diabetes Prevention Program.

Table 4 presents the results of the sensitivity analysis restricted to participants with prediabetes to better assess the robustness of the main findings. In this restricted sample, campaign exposure was not significantly associated with the outcome, and the magnitude of the estimated effect was attenuated (β=.30, *P*=.27). However, several key predictors remained statistically significant, including male gender (β=.72, *P*=.02), motivation index (β=.39, *P*=.02), inevitability belief (β=.16, *P*=.03), and perceived effectiveness of DPP ("improve a little," β=1.19; "improve a lot," β=2.28, *P*<.001).

XSL•FO
RenderX

Table . Sensitivity analysis—linear regression models on National Diabetes Prevention Program enrollment likelihood ratings among participants diagnosed with prediabetes.

Variable	β	SE	95% CI	*P* value
Intercept	−1.12	1.90	−4.84 to 2.60	.56
Exposed (ref[a]: no)				.27
Yes	.30	0.27	−0.23 to 0.84	
Age (years)	.00	0.02	−0.04 to 0.03	.84
Gender (ref: women)				.02
Man	.72	0.30	0.12 to 1.31	
Race or ethnicity (ref: Filipino)				.91
NHOPI[b]	.14	0.40	−0.65 to 0.93	
Other	.17	0.39	−0.60 to 0.95	
Education (ref: HS[c] or lower)				.51
Some college	−.20	0.44	−1.07 to 0.66	
College graduate	−.17	0.48	−1.10 to 0.77	
Post college graduate	.36	0.55	−0.71 to 1.42	
Household income, US dollars (ref: <$25,000)				.95
$25,000-$50,000	.31	0.56	−0.79 to 1.41	
$50,001-$100,000	.15	0.52	−0.86 to 1.16	
$100,001-$150,000	−.06	0.57	−1.18 to 1.07	
>$150,001	.29	0.61	−0.90 to 1.48	
Don’t know/missing	.07	0.82	−1.53 to 1.68	
Insurance type (ref: private)				.97
Medicare	.26	0.47	−0.66 to 1.18	
Medicaid	.32	0.54	−0.73 to 1.38	
TRICARE/Other	−.04	0.59	−1.18 to 1.11	
None/missing	.23	1.11	−1.95 to 2.42	
County (ref: Honolulu County)				.22
All other counties	.40	0.32	−0.24 to 1.03	
Diabetes risk factors				
Family history	−.12	0.36	−0.83 to 0.58	.73
Overweight perception	.23	0.34	−0.43 to 0.90	.49
Not physically active (<3x per week)	−.06	0.27	−0.60 to 0.48	.82
Gestational diabetes	−.39	0.39	−1.16 to 0.38	.32
Diagnosed with high blood pressure	−.32	0.28	−0.87 to 0.24	.26
Motivation index	.39	0.17	0.05 to 0.73	.02
Inevitability belief	.16	0.07	0.01 to 0.30	.03
Perceived DPP[d] effectiveness (ref: decline/stay the same)				<.001

XSL•FO
RenderX

Variable	β	SE	95% CI	*P* value
Don't know	1.23	0.63	0.00 to 2.45	
Improve a little	1.19	0.49	0.23 to 2.15	
Improve a lot	2.28	0.46	1.38 to 3.18	

[a]Ref: reference category.

[b]NHOPI: Native Hawaiian and Other Pacific Islander.

[c]HS: high school.

[d]DPP: National Diabetes Prevention Program.

Discussion

Principal Findings

This evaluation found that exposure to the Beat Diabetes campaign was associated with higher ratings of likelihood to enroll in a DPP among those at risk for prediabetes, after adjusting for demographic and other relevant factors. However, the campaign was not significantly associated with likelihood ratings among those medically diagnosed with prediabetes in the sensitivity analysis, which highlights the limits of including people with any risk of prediabetes in the cross-sectional study, not just those at highest risk. Despite this limitation, these findings show that a campaign with clear, action-oriented messaging may serve as a useful public health approach for targeting the approximately 80% of people at risk for prediabetes but unaware of their status [26].

The campaign builds on prior HDOH efforts, including the Prevent Diabetes, Hawai'i campaign, which focused on raising awareness of prediabetes among Native Hawaiian and Filipino adults [27]. Beat Diabetes extended this work by targeting a broader group disproportionately affected by diabetes and by emphasizing an immediate behavioral step—enrolling in a DPP. Although the campaign was designed to resonate among NHOPI or Filipino adults by featuring only NHOPI or Filipino individuals in all 7 advertisements, these groups were not more likely to report exposure, and their DPP enrollment likelihood ratings did not differ significantly from all other groups. In the subgroup analysis stratified by race or ethnicity, campaign exposure was statistically significant for NHOPI and all other races or ethnicities, while it was close to significant for Filipino adults. However, only 34 Filipino adults recalled the campaign, so it is possible that with a larger sample size, their exposure to the campaign could result with statistical significance. Additionally, grouping diverse races or ethnicities may have masked some campaign effects. Furthermore, the "other" group represents a combination of the remaining races or ethnicities outside of NHOPI and Filipino adults, including races or ethnicities with both high and low risk of prediabetes, such as African American and White adults, respectively. Aggregating them into a single group was important to explore the campaign exposure and likelihood ratings among the target populations of NHOPI and Filipino adults. Larger studies that enable disaggregation by race or ethnicity are needed to determine whether tailored messaging resonates more effectively within these target populations.

The regression models highlight the complex relationship between diabetes risk awareness and likelihood to enroll in a DPP. In model 2, several diabetes risk factors (family history, overweight perception, physical activity, and prediabetes diagnosis) were associated with higher likelihood ratings. However, in model 3, these were no longer significantly associated with likelihood to enroll in DPP. All 3 intrinsic factors—motivation, inevitability belief, and belief in DPP effectiveness—remained significantly associated with DPP enrollment likelihood ratings. This aligns with previous research that shows how additional factors, such as motivational or cognitive supports, may be necessary to translate contemplation to action [12,28]. Furthermore, belief in DPP effectiveness emerged as one of the strongest drivers, surpassing even the impact of campaign exposure. Participants who believed that DPP would improve their health "a little" had DPP enrollment likelihood ratings 0.75 points higher, and those who believed that it would improve their health "a lot" had ratings 1.63 points higher, compared with those who believed that DPP would have a neutral or negative effect. These differences were larger than the effect of campaign exposure (0.52-point increase), underscoring that confidence in DPP's benefits strongly influences willingness to enroll. The sensitivity analysis of model 3 among those with a prediabetes diagnosis further supports this finding. It showed that the perception that DPP is effective was the most significant factor associated with likelihood to enroll in a DPP. In fact, among those with a prediabetes diagnosis who believed that DPP would improve their health "a lot" had a 2.28-point higher likelihood rating than those who thought that DPP would have no effect on their health and those who thought their health would decline. These findings are consistent with those of Joiner and colleagues [29], who reported that promoting the benefits of lifestyle interventions increased DPP enrollment among adults with prediabetes. Future campaigns may benefit from spotlighting participant success stories and emphasizing tangible health improvements associated with DPP participation.

Interestingly, in the main analysis and sensitivity analysis, inevitability belief scores were significantly associated with DPP enrollment likelihood ratings. Participants who believed that their diabetes progression was inevitable had higher DPP enrollment likelihood ratings. This finding was unexpected and counterintuitive based on findings by Ross et al [19], who interviewed a diverse set of British adults who were diagnosed with nondiabetic hyperglycemia and found that those who viewed diabetes progression as inevitable were less likely to pursue behavior change and questioned DPP's value. It is unclear whether differences in study design or confounding are the source of the diverging findings. Furthermore, with a cross-sectional study design, it is unknown whether being aware

XSL•FO
RenderX

of one's diabetes risk influences feelings of inevitability of diabetes progression or whether feelings of inevitability influence one's diabetes risk awareness. Both should be explored outside of this study design to fully understand whether a campaign can influence intention to enroll in DPP.

Individuals who identify as men reported higher DPP enrollment likelihood ratings, even in sensitivity analyses among those with diagnosed prediabetes. It is possible that men enrolled in the study just had higher baseline likelihood to enroll in DPP. It is also important to note that only 34.5% (293/860) of the sample identified as men, and of those men, only 42% (123/293) were exposed to the campaign. Given that men historically have had lower DPP participation rates [8], the association between gender and intention warrants future evaluation studies using more rigorous study designs and enrolling a higher portion of men, in particular, men of NHOPI or Filipino race or ethnicity, to explore whether media campaigns can be effective at changing intentions to enroll among men.

Despite these significant associations, the models' fit statistics were relatively low, indicating that additional unmeasured factors likely play a major role in shaping enrollment intentions. Extrinsic structural and logistical barriers, such as the time commitment required for DPP or caregiving conflicts, have been well documented as barriers to DPP participation [12,17] but were not captured in this evaluation. This evaluation shows that for population-based efforts to reach all people with prediabetes risk, a campaign to encourage enrollment in DPP is only one small piece of the puzzle. A multifaceted strategy that pairs motivational messaging (eg, highlighting program effectiveness and success stories) with practical supports (eg, flexible schedules, virtual delivery, and culturally tailored programming) may be necessary to convert intention into enrollment for those at risk for prediabetes and those with a prediabetes diagnosis. Together, these findings help clarify which types of messaging may be most effective in increasing interest in DPP and point to the importance of pairing motivational communication with practical supports to translate intention into action.

Limitations

This evaluation has several limitations. First, the cross-sectional study design can show only associations between DPP enrollment likelihood ratings and advertisement recall and not definitive causal impacts of the campaign. Also, without baseline data, it is impossible to know whether the associations observed between campaign recall and likelihood ratings are true or due to confounding. Individuals who saw the campaign may simply have higher baseline intentions to join a lifestyle change program or may have been more likely to recall the advertisement due to their higher perceived personal risk of diabetes, or for reasons unrelated to the campaign. Despite the inherent limitations of a cross-sectional study design, they are the most cost-effective and feasible method of evaluating the real-world implementation of a health department media campaign [24]. Second, the sample was drawn from market research panels, which comprised both consumer data and participants with interest and time to participate in market research. Regardless of efforts by the market research firm to ensure a diverse database representative of the state's population and efforts to recruit participants from each county, the use of consumer panels limits the generalizability of the findings to the state's population. Third, although the survey incentive was of a nominal amount appropriate for a 10‑ to 12-minute survey, incentives can bias participation, encouraging more participation by those of lower rather than higher socioeconomic status or enticing participants to repeatedly complete a survey to receive multiple incentives. Although our sample included participants from a range of household incomes, over half of the sample (513/860, 59.6%) had household incomes lower than Hawai'i's median household income of US $100,745 [30], indicating the generalizability to the state's population is limited. Also, to avoid repeated participation, the market research firm used various data integrity checks, such as removing responses from duplicate IP addresses and removing responses that were completed too quickly. Fourth, because the purpose of the study was to evaluate a statewide media campaign aimed at reaching as many people at risk for developing diabetes as possible, the inclusion criterion was any individual with at least 1 recognized prediabetes risk factor. This resulted in a sample that included individuals with very low absolute risk of developing diabetes (eg, only being aged 40 years or older), which limits the generalizability of the data to only population-based strategies aiming to reach large segments of the population, not studies aiming to reach only those at highest risk of diabetes. Finally, the evaluation relied on self-reported likelihood to join a DPP rather than actual enrollment behavior. It is unclear whether higher DPP enrollment likelihood ratings will translate into real-world participation. During survey development, HHET looked for a psychometrically validated survey measure of campaign impact on likelihood to join a DPP, but none existed. The closest validated measure was an 11-item survey scale measuring autonomous and controlled motivation [18], which did not fully align with the campaign's goal of encouraging people to take urgent action. As a result, HHET developed a custom DPP enrollment likelihood question to measure intentions to change. However, because this measure is not psychometrically validated, its ability to accurately capture true likelihood to enroll is uncertain. Psychometric testing is warranted in the future to establish the validity of this item for measuring true likelihood to enroll in a DPP.

Conclusions

The Beat Diabetes media campaign was designed to encourage people at risk for diabetes to take immediate action by joining a DPP. The campaign focused on NHOPI or Filipino adults who are disproportionately affected by prediabetes and diabetes in Hawai'i. However, the cross-sectional design of the study used a posttest survey that could observe associations only between campaign exposure and the likelihood or intent to join DPP. Results showed that NHOPI or Filipino adults did not report higher DPP enrollment likelihood ratings. Overall exposure was associated with higher likelihood ratings after controlling for multiple factors such as demographic characteristics, diabetes risk factors, and intrinsic motivations and beliefs. This association indicates that the campaign may be effective at encouraging people to consider joining a DPP. However, increasing one's likelihood is the first step in the pathway to

XSL•FO
RenderX

actual enrollment, and further evaluation is needed to determine whether media campaigns boost DPP enrollment and participation. Future campaigns aiming to boost DPP participation could be strengthened by promoting DPP effectiveness to increase DPP enrollment likelihood and emphasizing the urgency of taking preventive action at the prediabetes stage. More importantly, using motivational messaging alone is insufficient without addressing structural and logistical barriers to accessing DPPs.

Acknowledgments

The authors would like to acknowledge Elia Titiimaea for his leadership as the diabetes project coordinator and his contributions to the development and launch of the Beat Diabetes campaign. The authors would like to thank all respondents who participated in the survey. They would also like to thank the National Diabetes Prevention Program providers who helped support campaign awareness. No artificial intelligence technologies were used to conduct this evaluation or the preparation of this manuscript. Finally, the authors would like to note that RMK was affiliated with the Healthy Hawai'i Evaluation Team at the University of Hawai'i at Mānoa at the time of data collection but is now currently affiliated with the Chronic Disease Prevention and Health Promotion Division at the Hawai'i State Department of Health.

Funding

This evaluation study was funded by a contract to the University of Hawai'i's Healthy Hawai'i Evaluation Team from the Hawai'i State Department of Health's Chronic Disease Prevention and Health Promotion Division. Data collection for the evaluation was paid for using Hawai'i State Department of Health's general funds. Support for statistical analyses was partially given by PIKO (U54GM138062) from the National Institute of General Medical Sciences and Ola HAWAII (U54MD007601) from the National Institute on Minority Health and Health Disparities. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funders.

Data Availability

The data collected and analyzed during this study may be made available upon a reasonable request to the authors.

Authors' Contributions

Conceptualization: KS (lead), BN (equal), GO (equal), LKC (equal), MDM (equal)
Data curation: KS (lead), MDM (equal)
Formal analysis: EL (lead), KS (supporting), MDM (supporting)
Funding acquisition: LBK (lead), LHI (equal), EL (equal), MDM (supporting), KI (supporting)
Investigation: KS (lead), BN (equal), GO (equal), LKC (equal), MDM (equal), RMK (supporting), NAB (supporting)
Methodology: KS (lead), BN (equal), GO (equal), LKC (equal), MDM (equal), EL (equal), NAB (supporting)
Project administration: KS (lead), BN (equal), GO (equal)
Resources: LBK (lead), LHI (equal), EL (equal), MDM (supporting), KI (supporting)
Supervision: MDM (lead), LKC (equal)
Validation: KS
Visualization: KS (lead), EL (equal), MDM (supporting)
Writing – original draft: KS (lead), MDM (supporting), EL (supporting), LKC (supporting)
Writing – review & editing: KS (lead), EL (supporting), BN (supporting), GO (supporting), LKC (supporting), RMK (supporting), NAB (supporting), KI (supporting), LHI (supporting), KRK (supporting), LBK (supporting), MDM (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
A sample of still advertisements featured in the Beat Diabetes media campaign that ran from July to October 2024.
[PNG File, 610 KB - publichealth_v12i1e90880_app1.png]

Multimedia Appendix 2
Postcampaign cross-sectional online survey conducted from October to December 2024.
[DOCX File, 7704 KB - publichealth_v12i1e90880_app2.docx]

Checklist 1
STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist.

XSL•FO
RenderX

[DOCX File, 33 KB - publichealth_v12i1e90880_app3.docx]

Checklist 2
CHERRIES (Checklist for Reporting Results of Internet E-Surveys) checklist.
[DOCX File, 23 KB - publichealth_v12i1e90880_app4.docx]

References

1. National diabetes statistics report. Diabetes. Centers for Disease Control and Prevention. 2024 Jan 8. URL: https://www.cdc.gov/diabetes/php/data-research/index.html [accessed 2025-05-21]
2. Query result for Hawaii's Behavioral Risk Factor Surveillance System (BRFSS) data—diabetes—prediabetes prevalence, age adjusted, 2021. Hawaii State Department of Health, Hawaii Health Data Warehouse. 2025 Mar 25. URL: https://hhdw.org/report/query/result/brfss/PreDiabetes/PreDiabetesAA11_.html [accessed 2025-05-21]
3. Formagini T, Brooks JV, Roberts A, et al. Prediabetes prevalence and awareness by race, ethnicity, and educational attainment among U.S. adults. Front Public Health 2023;11:1277657. [doi: 10.3389/fpubh.2023.1277657] [Medline: 38164446]
4. Bang H, Edwards AM, Bomback AS, et al. Development and validation of a patient self-assessment score for diabetes risk. Ann Intern Med 2009 Dec 1;151(11):775-783. [doi: 10.7326/0003-4819-151-11-200912010-00005] [Medline: 19949143]
5. Heikes KE, Eddy DM, Arondekar B, Schlessinger L. Diabetes risk calculator. Diabetes Care 2008 May 1;31(5):1040-1045. [doi: 10.2337/dc07-1150]
6. US Preventive Services Task Force, Davidson KW, Barry MJ, et al. Screening for prediabetes and type 2 diabetes: US Preventive Services Task Force Recommendation Statement. JAMA 2021 Aug 24;326(8):736-743. [doi: 10.1001/jama.2021.12531] [Medline: 34427594]
7. Diabetes prevention program. National Institute of Diabetes and Digestive and Kidney Diseases. 2021 Aug. URL: https://www.niddk.nih.gov/about-niddk/research-areas/diabetes/diabetes-prevention-program-dpp [accessed 2025-09-11]
8. Cannon MJ, Ng BP, Lloyd K, Reynolds J, Ely EK. Delivering the National Diabetes Prevention Program: assessment of enrollment in in-person and virtual organizations. J Diabetes Res 2022;2022:2942918. [doi: 10.1155/2022/2942918] [Medline: 35118160]
9. Uchima O, Wu YY, Browne C, Braun KL. Disparities in diabetes prevalence among Native Hawaiians/Other Pacific Islanders and Asians in Hawai'i. Prev Chronic Dis 2019 Feb 21;16:E22. [doi: 10.5888/pcd16.180187] [Medline: 30789820]
10. Whittemore R. A systematic review of the translational research on the Diabetes Prevention Program. Behav Med Pract Policy Res 2011 Sep;1(3):480-491. [doi: 10.1007/s13142-011-0062-y]
11. Mensa-Wilmot Y, Bowen SA, Rutledge S, et al. Early results of states' efforts to support, scale, and sustain the National Diabetes Prevention Program. Prev Chronic Dis 2017 Dec 7;14:E130. [doi: 10.5888/pcd14.170478] [Medline: 29215975]
12. Kent-Marvick J, Knysheva M, Gibson B, Simonsen SE. Why do people choose to enroll or not enroll in the National Diabetes Prevention Program lifestyle change program? A mixed-methods analysis from a sample of adults with a prediabetes diagnosis. J Prim Care Community Health 2024;15:21501319241282862. [doi: 10.1177/21501319241282862] [Medline: 39305089]
13. Subramanian K, Midha I, Chellapilla V. Overcoming the challenges in implementing type 2 diabetes mellitus prevention programs can decrease the burden on healthcare costs in the United States. J Diabetes Res 2017;2017:2615681. [doi: 10.1155/2017/2615681] [Medline: 28890899]
14. Baucom KJW, Pershing ML, Dwenger KM, Karasawa M, Cohan JN, Ozanne EM. Barriers and facilitators to enrollment and retention in the national diabetes prevention program: perspectives of women and clinicians within a health system. Womens Health Rep (New Rochelle) 2021;2(1):133-141. [doi: 10.1089/whr.2020.0102] [Medline: 34036296]
15. Carroll J, Winters P, Fiscella K, et al. Process evaluation of practice-based diabetes prevention programs: what are the implementation challenges? Diabetes Educ 2015 Jun;41(3):271-279. [doi: 10.1177/0145721715572444] [Medline: 25759431]
16. Chambers EC, Gonzalez JS, Marquez ME, Parsons A, Rehm CD. The reach of an urban hospital system-based diabetes prevention program: patient engagement and weight loss characteristics. Diabetes Educ 2019 Dec;45(6):616-628. [doi: 10.1177/0145721719880503] [Medline: 31608798]
17. Saiki K, Shalaby A, Nett B, et al. Recruitment and retention in the National Diabetes Prevention Program lifestyle change program in two federally qualified health centers in rural Hawai'i. Prev Chronic Dis 2024 Oct 31;21:E85. [doi: 10.5888/pcd21.240156] [Medline: 39481012]
18. Rodriguez LA, Thomas TW, Finertie H, Turner CD, Heisler M, Schmittdiel JA. Psychosocial and diabetes risk factors among racially/ethnically diverse adults with prediabetes. Prev Med Rep 2022 Jun;27:101821. [doi: 10.1016/j.pmedr.2022.101821] [Medline: 35656212]
19. Ross J, Cotterill S, Bower P, Murray E. Influences on patient uptake of and engagement with the national health service digital diabetes prevention programme: qualitative interview study. J Med Internet Res 2023 Feb 28;25:e40961. [doi: 10.2196/40961] [Medline: 36853751]

XSL•FO
RenderX

20. Vaja I, Umeh KF, Abayomi JC, Patel T, Newson L. A grounded theory of type 2 diabetes prevention and risk perception. British J Health Psychol 2021 Sep;26(3):789-806. [doi: 10.1111/bjhp.12503]
21. Beat Diabetes. Hawai i State Department of Health. 2025. URL: https://livinghealthy.hawaii.gov/beat-diabetes/ [accessed 2025-04-22]
22. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med 2007 Oct 16;147(8):573-577. [doi: 10.7326/0003-4819-147-8-200710160-00010]
23. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34. [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
24. McGurk MD, Ogawa G, Inoue K, et al. Sweet Lies! Lessons learned from Hawai'i's sweetened fruit drink countermarketing campaign. J Health Commun 2025 Mar 28;30(sup1):14-27. [doi: 10.1080/10810730.2025.2461588] [Medline: 40162925]
25. Query results for Hawaii's Behavioral Risk Factor Surveillance System (BRFSS) data—sex. Hawaii State Department of Health, Hawaii Health Data Warehouse. 2025 Jun 2. URL: https://hhdw.org/report/query/result/brfss/SexCat/SexCatCrude11_.html [accessed 2025-06-03]
26. A US report card. : Centers for Disease Control and Prevention; 2026 Mar 16 URL: https://www.cdc.gov/diabetes/communication-resources/diabetes-statistics.html [accessed 2026-04-23]
27. Sinclair-White B, Nett B, Ilagan L, et al. Insights in public health: development, implementation, and evaluation of the prevent diabetes, Hawai'i campaign. Hawaii J Health Soc Welf 2020 Mar 1;79(3):86-90. [Medline: 32190841]
28. Kullgren JT, Knaus M, Jenkins KR, Heisler M. Mixed methods study of engagement in behaviors to prevent type 2 diabetes among employees with pre-diabetes. BMJ Open Diabetes Res Care 2016;4(1):e000212. [doi: 10.1136/bmjdrc-2016-000212] [Medline: 27738513]
29. Joiner KL, McEwen LN, Hurst TE, Adams MP, Herman WH. Domains from the health belief model predict enrollment in the National Diabetes Prevention Program among insured adults with prediabetes. J Diabetes Complicat 2022 Jul;36(7):108220. [doi: 10.1016/j.jdiacomp.2022.108220]
30. Income and poverty. US Census Bureau. 2024. URL: https://data.census.gov/profile/Hawaii?g=040XX00US15#income-and-poverty [accessed 2026-04-23]

Abbreviations

CHERRIES: Checklist for Reporting Results of Internet E-Surveys
DPP: National Diabetes Prevention Program
HDOH: Hawai'i State Department of Health
HHET: Healthy Hawai'i Evaluation Team
NHOPI: Native Hawaiian and Other Pacific Islander
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Mavragani, T Sanchez; submitted 05.Jan.2026; peer-reviewed by J Zhang, S Isfandari; revised version received 30.Apr.2026; accepted 01.May.2026; published 18.Jun.2026.

Please cite as:
Saiki K, Lim E, Nett B, Ogawa G, Ching LK, Kawasaki RM, Barreto NA, Inoue K, Irvin LH, Kaupiko KR, Keliikoa LB, McGurk MD Encouraging Adults at Risk for Type 2 Diabetes to Enroll in Diabetes Prevention Programs Through a Media Campaign in Hawai'i: Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e90880
URL: https://publichealth.jmir.org/2026/1/e90880
doi:10.2196/90880

© Kara Saiki, Eunjung Lim, Blythe Nett, Gail Ogawa, Lance K Ching, Riana M Kawasaki, Nianest Alers Barreto, Katherine Inoue, Lola H Irvin, Kiana R Kaupiko, L Brooke Keliikoa, Meghan D McGurk. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 18.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Impact on Audience Engagement of Coordinating a Public Health Campaign on Antimicrobial Resistance Through a Network of Health Content Creators: Longitudinal Observational Study

Fangyue Chen, MB BChir, MA, MPH; Jack Cooper, BA, MScR; Amish Acharya, MBBS, PhD; Simon Dryden, MBChB, MPH, PhD; Ara Darzi, MBBS, MD; Kate Grailey, BA, BMBCh, PhD

Fleming Initiative, Institute of Global Health Innovation, Imperial College London, QEQM Building, St Mary's Hospital, South Wharf Road, London, United Kingdom

Corresponding Author:
Fangyue Chen, MB BChir, MA, MPH
Fleming Initiative, Institute of Global Health Innovation, Imperial College London, QEQM Building, St Mary's Hospital, South Wharf Road, London, United Kingdom

Abstract

Background: Antimicrobial resistance (AMR) is a significant global health threat. Several public health campaigns aimed to raise AMR awareness and inspire related behavioral changes have been delivered in a time-specific, coordinated manner, while others have placed less emphasis on campaign timing. Social media platforms can be leveraged as key vehicles for delivering public health campaigns, particularly by collaborating with health content creators who serve as influential messengers. Increasingly, organizations such as the World Health Organization and TikTok have created health content creator networks; however, the impact of such networks in public health campaigns, especially when delivered in a coordinated, time-specific manner, remains uncertain.

Objective: This study aimed to investigate whether mobilizing an established health content creator network to create social media content on the topic of AMR, released in a coordinated, time-specific manner, can have an impact on audience engagement.

Methods: We conducted a longitudinal observational study evaluating the effect of a coordinated social media campaign (“Pulse”) on YouTube, delivered by an established health content creator network during an international event on AMR. Members of the network prepared and coordinated the release of AMR-related videos. Engagement analytics were evaluated 6 months after release. The engagement with each campaign video was compared with that for a similar noncoordinated video and with the average engagement of the same creators.

Results: Around the day of the Pulse campaign on September 26, 2024, 18 campaign videos were released across 14 YouTube channels. Compared with paired videos, Pulse videos were not associated with higher view counts (incidence rate ratio [IRR] 0.98, 95% CI 0.44-2.13; P=.95) or like counts (IRR 1.10, 95% CI 0.48-2.41; P=.81) but were associated with significantly higher comment counts (IRR 2.99, 95% CI 1.02-8.52; P=.03). When compared with the creators’ 12-month channel averages, campaign videos had a significantly higher comment count (IRR 15.5, 95% CI 5.5-24.0; P<.001) but no difference in view counts (IRR −82.0, 95% CI −190.3 to 58.5; P=.26) or like counts (IRR −0.50, 95% CI −6.3 to 10.5; P=.93).

Conclusions: Coordinating health content creators to release AMR-related videos on YouTube coinciding with an international AMR event increased audience interactivity but did not enhance reach. This study shows the need to better understand which AMR-specific content factors contribute toward greater traction and to assess audience needs among the wider public to discern how best to harness social media interventions as a tool to improve AMR-related outcomes.

(JMIR Public Health Surveill 2026;12:e86587) doi:10.2196/86587

KEYWORDS

health campaign; audience engagement; social media; antimicrobial resistance; health content creator

Introduction

Antimicrobial resistance (AMR) occurs when disease-causing microorganisms (including bacteria, viruses, fungi, and parasites) are no longer affected by antimicrobial agents [1]. In combating the rising threat of AMR, which is expected to result in 8-22 million deaths per year globally by 2050, the World Health Organization (WHO) has emphasized the importance of changing behaviors that contribute to AMR among both the public and health care professionals (HCPs) [2-4].

Global and regional AMR awareness campaigns have been designed to improve public awareness of AMR and inspire related behavioral changes [5]. A common mechanism for delivering such campaigns is through time-specific, coordinated

XSL•FO
RenderX

activities that can provide a surge of AMR coverage [6-8]. Such examples include the annual World AMR Awareness Week held in November, and beyond AMR, campaigns such as Mental Health Awareness Week, May Measuring Month (encouraging individuals to measure their blood pressure), and World Cancer Day (advocating for people-centered cancer care) [9-11]. Associated with specific health campaigns, organizations often release content packs or toolkits incorporating key campaign messaging and infographics that can be adopted and shared by individuals and organizations to amplify campaign objectives [12,13].

On the other hand, campaigns such as the United Kingdom Health Security Agency's "Keep Antibiotics Working" campaign are not designed to be released at a specific time but have nevertheless gathered significant traction and high campaign recognition, including through a social media video with more than 1 million views globally [14]. Other examples of non–time-specific campaigns include the US Food and Drug Administration's "Little Lungs," a digital campaign within which a video series delivering antismoking messaging has gathered more than 17 million views on the YouTube platform alone [15].

The past decade has seen great enthusiasm for leveraging the potential of social media as a key vehicle in driving and amplifying health campaigns [12]. Certain HCPs have generated a large network of followers through their active social media presence and have been referred to as "health content creators," "clinician content creators," or "health influencers" [16-19]. On the largest video-based social media platform, YouTube, a number of such individuals have gathered more subscribers (individuals who follow their channels and regularly watch their videos) than international and national medical organizations such as the WHO (approximately 874,000 subscribers) and the United States Centers for Disease Control and Prevention (approximately 637,000 subscribers) [20-22]. Increasingly, health organizations have sought to endorse their dedicated networks of health content creators who can serve as credible sources for broadcasting organization-specific health information and tackling misinformation [23]. An example of such a network is the WHO's network of more than 800 health content creators—WHO Fides [24]—who are mobilized periodically to create content on specific health-related topics and to elevate evidence-based content. Similarly, social media platforms such as TikTok have launched health content creator networks, uniting accredited clinicians to share their insights into trending health topics [25].

Nevertheless, it is unknown how much audience engagement a network of health content creators can generate when they act as messengers for health organizations in broadcasting organization-specific health information. Moreover, it remains uncertain whether delivering the campaign messaging in a time-specific, coordinated manner can generate greater audience traction than the sporadic release of content. Coordinated campaigns have the potential to enhance audience engagement—defined by the number of views, likes, comments, and shares—through repeated exposure, algorithmic amplification, and a sense of social momentum [26,27]. While such a strategy has not previously been explored, it may be a novel avenue for the impactful dissemination of public health information that policymakers and health professionals can use.

In this study, we aimed to answer the question "Can mobilizing an established health content creator network to create social media content on the topic of AMR, released in a coordinated, time-specific manner, have an impact on audience engagement?" Our hypothesis was that a coordinated mobilization of a health content creator network in creating AMR-related social media content can result in greater audience engagement.

Methods

Study Design

A longitudinal observational study was conducted to evaluate the impact on audience engagement, when coordinating a health content creator network to disseminate social media videos on the topic of AMR, coinciding with an international event on AMR. We named this coordinated release of social media content "Pulse."

We first established a health content creator network. A media pack (Multimedia Appendix 1) containing recommended AMR campaign messaging was developed and distributed to the network. Members were informed in advance to prepare AMR-related social media content in accordance with the media pack, to be released on YouTube, the largest video-based social media platform. A specific date was chosen for Pulse to take place, coinciding with a United Nations General Assembly (UNGA) high-level meeting on AMR [28].

The study was reported in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) statement for cohort studies, given the prospective and observational nature of the study [29].

Study Setting

We conducted the study online on the YouTube (Alphabet Inc) platform, as it is the largest video-based social media platform, and its associated work stream, YouTube Health, has dedicated its mission to delivering high-quality health information online [22]. Registered users of the platform can upload and openly share videos on their YouTube channels. Users can find videos of interest by entering key related terms or phrases into the dedicated YouTube search bar or through the platform's recommendations. In addition, users can save selected videos into a playlist that can be accessed later. Users can also subscribe to specific channels to receive regular updates. Each video has a unique code (video ID). Each playlist has an ID that is sourced from the web address (also known as URL).

Currently, there are 2 formats of videos on YouTube: traditional long-form videos that individuals browse on a YouTube web page and Shorts, defined as vertical videos that are less than 180 seconds in duration and are presented to users passively and sequentially, from which users can choose either to watch or to swipe away [30].

XSL•FO
RenderX

Mobilization of a Health Content Creator Network to Release Time-Specific Social Media Video Content on AMR

An international network of health content creators, communications experts, and entertainment professionals was previously created through an in-person launch event on April 29, 2024, and social media platforms such as LinkedIn (Microsoft Corp) and X (X Corp). The network was called CHAIN (Content, Health, and AMR Innovation Network) [31]. The network aims to harness the power and global reach of video-based online content to drive the behavioral changes needed to tackle AMR [31].

A media pack was developed and approved by all researchers over a period of 2 months to coordinate the social media messaging during the Pulse. This media pack was released to CHAIN members immediately following its completion through email 1 month prior to the Pulse date. The media pack incorporated key messages from official sources such as the United Kingdom Health Security Agency and the WHO, explaining what AMR is and how members of the public could change their behaviors related to the development of AMR [32,33].

In addition to key messages, specific messages were created for key niches that the health content creators occupied, including women's health, children's health, pharmacists, ear, nose, and throat, dentistry, and medical myths. Creators were given the flexibility in the remaining aspects of video production, such as video format, narration, and imagery, to enable congruency of the content with their existing audience, while maintaining the diversity of content. The media pack is attached in Multimedia Appendix 1.

The Pulse took place on a date that coincided with the UNGA high-level meeting on AMR [28]. Given that this meeting was the principal official, health-focused event during the UNGA high-level meeting week, as well as the second UN high-level meeting on AMR since 2016, we anticipated a heightened global attention toward AMR [34].

Evaluation of the Impact of the Coordinated Social Media Content Release on Audience Engagement

Six months after the Pulse day, we evaluated the audience engagement (views, likes, and comments) of the content released onto YouTube on the Pulse date using the publicly available engagement data. To discern whether coordinating the social media content had an impact on audience engagement, we undertook comparisons in 2 ways.

First, the engagement analytics of the Pulse videos were compared with those of a similar video that was sporadically released by the same health content creator. A variety of factors can contribute toward the level of engagement of a YouTube video, such as the video content, length of time since video release, thumbnail design and titles, and the channel's number of subscribers [35]. Choosing a video similar to the Pulse video from the same health content creator allowed for the control of confounding factors, such that a video released in a coordinated manner could be compared against a video released sporadically.

In addition, the engagement of the Pulse videos was compared with the creator's baseline average video engagement over a 12-month window around the Pulse date. This allowed for direct comparison of the engagement level of the Pulse videos with the average engagement level of the health content creator's channel, while minimizing time-related confounding factors.

Data Collection

Following the date of the Pulse, 2 researchers (FC and JC) screened the YouTube channels hosted by members of the CHAIN for all videos released on the date of the Pulse, on the topic of AMR or relating to the appropriate use of antibiotics or other antimicrobial agents. Such videos were saved in a playlist on the YouTube platform. Videos released within 1 day of the Pulse day were also included.

For each Pulse video released by an individual health content creator, a paired video hosted on the same channel was selected by one researcher (FC) and verified by another researcher (KG). When selecting the videos, the following criteria were considered in terms of similarity, ordered by priority: video format (long-form vs Shorts), video topic, recency of release dates, video length, and video style. Given that it was unclear whether the video topic or recency of release can have a greater impact on the level of video engagement, a sensitivity analysis was conducted. An additional video was selected to pair with each Pulse video using slightly different prioritization criteria, in which recency was prioritized over video topic. The paired videos were collated into 2 other YouTube playlists, labeled "paired by topic" and "paired by recency."

We collected the engagement data of the Pulse videos at the 6-month point from the time of release to allow for the cumulative growth in video engagement to plateau. Therefore, for the paired videos, the researcher collected the engagement analytics at the time point closest to 6 months where possible, although there were variations in the dates of release (eg, some paired videos had been published for 3 years at the time of data collection). The exact dates of data collection for each video are shown in Multimedia Appendix 2.

The engagement analytics were extracted from publicly available YouTube data using YouTube Data application programming interface (API; version 3; Google LLC), which allowed the automated, scaled-up extraction of YouTube data. Specifically, an API key was created on the Google Cloud Platform (Alphabet Inc) and restricted to YouTube Data API [36]. For each playlist (Pulse video, paired by topic and by recency), the specific playlist ID was sourced from the URL of the YouTube playlist. The playlist ID and the API key were then entered into predetermined code using R software (version 4.4.2; R Foundation for Statistical Computing) on the RStudio platform (version 2026.01.0+392; Posit PBC) to extract the data [37,38]. The following data were extracted: YouTube video ID, video title, channel title, video publication date, video duration, video view count, like count, and comment count.

For each health content creator's videos and engagement data, the channel ID and the date range (a 12-month window around the Pulse date) were entered into the predetermined codes using Python (version 3.12; Python Software Foundation) on VSCode

XSL•FO
RenderX

platform (version 1.103; Microsoft Corp) [39,40]. The following data were collected on April 24, 2024: video ID, title, video publication date, views, likes, comments, and video duration.

Ethical Considerations

Ethics approval for this study was not required as all collected data were publicly available. This was confirmed with the Imperial College London Research Governance and Integrity Team in accordance with local policies [41]. All data have been anonymized and deidentified.

Informed consent was not obtained because the study analyzed publicly accessible online content. In accordance with relevant ethical guidance for internet-based research, the study involved minimal risk and did not include any attempt to identify or contact individuals whose content was analyzed. Participant compensation was not applicable.

Statistical Analysis

Summary statistics were calculated for the Pulse videos, including engagement data and video duration. Engagement outcomes (views, likes, and comments) were analyzed using a staged analytical approach. First, unadjusted comparisons were conducted between Pulse videos and noncoordinated videos paired by topic from the same health content creator. To assess for normality, differences in engagement were first computed within each pair, and the difference scores were assessed using the Shapiro-Wilk test [42]. For metrics for which the assumption of normality was met ($P>.05$), a paired 2-tailed t test was performed [43]. For metrics that violated the normality assumption ($P\leq.05$), the Wilcoxon signed-rank test was used as a nonparametric alternative [44]. Results were presented as the Hodges-Lehmann median differences with 95% CIs [45]. Second, to account for residual confounding, adjusted analyses were conducted using negative binomial regression [46].

Engagement data were modeled as count outcomes, with Pulse status specified as the primary exposure. Models were adjusted a priori for video-level characteristics influencing engagement, including video format (Shorts vs long-form), video duration, and time since publication at the point of data extraction (expressed in months). To account for correlation of outcomes within creators, robust SEs clustered at the creator level were used. Results from regression analyses were presented as incidence rate ratios (IRRs) with 95% CIs. Statistical significance was set at $\alpha=.05$.

To evaluate how the Pulse videos compared with each creator's average performance, we calculated the median and IQR of engagement metrics (views, likes, and comments) for each creator's noncoordinated videos published within this 12-month interval. To assess normality, differences were calculated for each creator's Pulse video and median engagement of the baseline videos and were assessed using the Shapiro-Wilk test [42]. Given the anticipated highly skewed distribution of engagement metrics and the paired within-creator design, differences in views, likes, and comments between the Pulse videos and the median engagement values of the same creator were analyzed using the Wilcoxon signed-rank tests. In addition, the Pulse videos were compared with the mean performance of videos of the same format (long-form or Shorts) within the 12-month window around the coordinated release date, using the same method described. The Hodges-Lehmann median differences and 95% CIs were reported. Statistical significance was set at $\alpha=.05$.

Data collection and analysis were conducted in R software on the RStudio platform [37,38]. The following packages were used: *httr*, *httr2*, *jsonlite*, *here*, *dplyr*, *vosonSML*, *readxl*, *ggplot2*, *effsize*, *tidyverse*, *reshape2*, *MASS*, *sandwich*, *lmtest*, and *broom* [38,40,47-59].

Results

Overview

Around the Pulse date on September 26, 2024, the health content creator network released 18 videos from 14 YouTube channels hosted by members of the network. Six months after the Pulse, the audience engagement data for all 18 coordinated videos (V1-V18) were successfully collected, with no missing data. Overall, the coordinated videos gathered a total of 6153 views, 262 likes, and 76 comments. The campaign media pack and the complete video engagement datasets are available in Multimedia Appendices 1 and 2, respectively.

Evaluating the Impact of a Coordinated Release on Video Engagement

Comparison of the Pulse Videos With Paired Videos

The chosen paired videos prioritizing topic choice (P1-P18) and recency (R1-R18) are shown in Multimedia Appendix 2. All engagement analytics were successfully collected with no missing data. The summary characteristics of the 3 video groups (Pulse, paired by topic, and paired by recency) are shown in Table 1 and Table S1 in Multimedia Appendix 3. Assessment of normality using the Shapiro-Wilk normality test is shown in Table S2 in Multimedia Appendix 3.

XSL•FO
RenderX

Table . Summary characteristics of the "Pulse" videos and the paired videos by the same health content creators[a].

Engagement metrics	Pulse videos, median (IQR)	Paired by topic, median (IQR)	Hodges-Lehmann estimator, median difference (95% CI)	*P* value
View count	123 (39-580)	72 (36-367)	58.5 (–31.5 to 237.0)	.15
Like count	10 (4-31)	7 (2-12)	8.0 (0.5-19.0)	.04
Comment count	2 (1-5)	1 (0-2)	2.0 (1.0-7.0)	.04
Duration (s)	95 (60-247)	134 (59-238)	__[b]	—
Time since publication (mo)	6 (6-6)	6.5 (6-15)	—	—

[a]The paired videos were chosen by prioritizing video topics. Paired differences in engagement (views, likes, and comments) were assessed using the Wilcoxon signed-rank test.

[b]Not applicable.

The differences in engagement between the video pairs were assessed using the Wilcoxon signed-rank test. When comparing the Pulse videos (V1-V18) with the videos paired by topic (P1-P18), no significant difference was found between the view counts (IRR 58.5, 95% CI –31.5 to 237.0; *P*=.15). However, the Pulse videos appeared to have significantly more like counts (IRR 8.0, 95% CI 0.5-19.0; *P*=.04) and comment counts (IRR 2.0, 95% CI 1.0-7.0; *P*=.04). Sensitivity analysis pairing videos by recency showed concordant findings (Table S1 in Multimedia Appendix 3).

After adjustment for confounding (Table 2), Pulse videos were not associated with higher view counts or like counts but were associated with significantly higher comment counts (IRR 2.99, 95% CI 1.02-8.52; *P*=.03). In sensitivity analyses pairing videos by recency rather than topic (Table S3 in Multimedia Appendix 3), effect estimates were directionally consistent, although associations were attenuated and no longer statistically significant.

Table . Adjusted association between "Pulse" videos and audience engagement[a].

Association	Views		Likes		Comments	
	IRR[b] (95% CI)	*P* value	IRR (95% CI)	*P* value	IRR (95% CI)	*P* value
Pulse (vs noncoordinated)	0.98 (0.44-2.13)	.95	1.10 (0.48-2.41)	.81	2.99 (1.02-8.52)	.03
Shorts (vs long-form content)	7.79 (2.98-20.49)	<.001	4.81 (1.97-11.93)	<.001	1.42 (0.45-4.53)	.51
Video duration (per 60 s)	1.06 (0.92-1.25)	.35	1.01 (0.88-1.20)	.82	1.00 (0.81-1.25)	.99
Time since publication (per mo)	1.03 (0.97-1.19)	.40	1.00 (0.93-1.07)	.93	1.00 (0.91-1.11)	.97

[a]Negative binomial regression models were used. Models were adjusted for video format, video duration, and time since publication. Robust SEs clustered at the creator level were used to account for within-creator correlation. Videos were paired based on topics.

[b]IRR: incidence rate ratio.

Comparison of the Pulse Video With the Creator's Average Level of Engagement

The summary characteristics of all creators' baseline videos are shown in Multimedia Appendix 2. When comparing the Pulse videos with each creator's baseline performance, there was no significant difference in the number of views (IRR −82.0, 95% CI −190.3 to 58.5; *P*=.26) or like counts (IRR −0.50, 95% CI −6.3 to 10.5; *P*=.93). However, the Pulse videos had a significantly greater number of comments (IRR 15.5, 95% CI 5.5-24.0; *P*<.001; Table 3). When comparing the Pulse videos against creators' baseline videos of the same format, no significant differences were observed (Table S4 and S5 in Multimedia Appendix 3).

XSL•FO
RenderX

Table . "Pulse" videos compared with creator's baseline videos released over a 12-month window around the Pulse date[a].

Engagement metrics	Hodges-Lehmann estimator, median difference (95% CI)	*P* value
View count	−82.0 (−190.3 to 58.5)	.26
Like count	0.50 (−6.3 to 10.5)	.93
Comment count	15.5 (5.5 to 24.0)	<.001

[a]Values are shown as Hodges-Lehmann median difference (95% CI). Paired differences in engagement (views, likes, and comments) were assessed using the Wilcoxon signed-rank test.

Discussion

Principal Findings

This study examined whether coordinating health content creators to release AMR-related videos simultaneously on YouTube, timed with an AMR-centered international event, could enhance audience engagement compared to sporadic, uncoordinated releases. The principal findings were that coordinating health content creators to release AMR-related videos on YouTube coinciding with an international event on AMR, defined as a Pulse, increased audience interactivity in terms of the number of comments but did not enhance overall reach compared with sporadic releases.

When comparing the Pulse videos with similar videos and baseline videos from the same health content creator, coordinated content gathered significantly greater comment counts. This suggests that coordinated campaigns may enhance interactive engagement among audiences, potentially through the perception of collective momentum or amplified algorithmic exposure. However, the lack of a significant difference in view counts and like counts after adjustment between coordinated and sporadic videos suggests that coordination alone may not be sufficient to drive greater overall reach.

It is noteworthy that Shorts, compared with long-form videos, have been found to be significantly associated with greater engagement in terms of views and likes ($P<.001$). While YouTube has traditionally been the main hub for long-form video content, the introduction of YouTube Shorts has transformed how individuals consume social media content. The audience tends to spend more time watching shorter videos, which provides them with bite-sized content, while long-form videos have seen a significant decrease in engagement [60].

Several possible explanations exist for the principal findings. First, although coordination can enhance visibility within a narrow window, it does not necessarily equate to broader audience interest or long-term viewership. Second, the relatively small size of the coordinated cohort (18 videos across 14 channels) may not have generated the critical mass required to trigger meaningful algorithmic amplification. Third, the media pack was developed based on accuracy of messaging rather than the enhancement of audience engagement. Currently, there remains a lack of research on how to design video content, particularly AMR-specific content, to improve audience engagement.

From a campaign design perspective, these findings suggest that while coordination can enhance engagement quality and audience interactivity (as seen in likes and comments), it may not inherently improve reach. Therefore, coordinated campaigns may be best seen as complementary tools within a broader health communication strategy rather than as standalone solutions.

Previous HCP-led social media campaigns have mostly aimed at HCPs, focusing on advocacy for particular specialties such as surgery, radiology, and nursing [61-64]. The remaining HCP-led social media campaigns took place during the COVID-19 pandemic, relating to vaccination and travel restrictions [64-66]. No studies have evaluated the impact of collaborating with HCPs in delivering time-coordinated social media campaigns relating to other global health challenges outside the COVID-19 pandemic period. To our knowledge, this study is the first evaluation of the impact of collaboration with health content creators during an international public health campaign. Our findings provide important insights into the dynamics of health communication on social media platforms and the potential role of health content creator networks in supporting public health campaigns [67,68].

Limitations

Several limitations should be acknowledged. The sample size of creators and videos was relatively small, limiting statistical power and generalizability. The observational design precludes causal inference. While we matched video pairs based on format, topic, and recency, not all engagement data were collected at the 6-month point from their time of release. Some creators posted relatively infrequently, such that their most recent video was posted 24 months ago. Although we welcomed creators internationally, most participating health content creators were based in the United Kingdom, meaning that the health challenge of AMR was competing with other country-specific challenges. Moreover, engagement metrics such as views, likes, and comments do not necessarily reflect a deeper impact on audience such as changes in knowledge, changes in attitudes, or AMR-related behavioral changes, which were not explored in this study.

Opportunities for Future Work

Future work should explore how to optimize coordinated health campaigns on social media platforms to support health policymakers in designing more effective public health strategies. This may include testing different messaging styles, content formats, and durations of such coordinated activity. Second, it will be valuable to assess the impact of such campaigns on AMR-related outcomes such as individual knowledge, attitudes, and behavioral intentions, which can be collected from YouTube through qualitative analysis of the

XSL•FO
RenderX

audience comments, as well as through embedded survey methodologies as part of the campaign content. Finally, the study calls for the need to better understand which AMR-specific content factors contribute toward greater audience traction and for assessment of audience needs among the wider public, to discern how best to harness social media interventions as a tool to improve AMR-related outcomes, including inspiring behavioral changes.

Conclusions

This study highlights both the potential and the limitations of using social media to engage the public on AMR. While campaigns led by health content creators increased audience interactivity, they did not substantially improve overall reach, reflecting the broader challenge of promoting AMR-related content in crowded social media environments. These findings highlight the need to understand the nuances behind audience behavior, preferences, and engagement dynamics within social media platforms, such that the content can objectively influence knowledge, attitudes, and behaviors. Importantly, our results underscore the need for innovative evaluation methods that capture the real-world impact of social media campaigns on AMR-related outcomes. By integrating content strategy with behavioral insights, future campaigns could more effectively leverage digital platforms to support antimicrobial stewardship and inform policy on health communication in the social media era.

Acknowledgments

The authors declare the use of generative artificial intelligence (GenAI) in the research and writing process. According to the GenAI Delegation Taxonomy (2025), the following tasks were delegated to GenAI tools under full human supervision: code generation and code optimization. The GenAI tool used was ChatGPT (version 4.5; OpenAI). Responsibility for the final manuscript lies entirely with the authors. The GenAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Funding

The authors declared that no financial support was received for this study.

Data Availability

The data used in the study, including codes, are available from Open Science Framework [69].

Authors' Contributions

Conceptualization: AD, KG, AA, FC
Methodology: AD, KG, FC, JC
Investigation: FC, KG, JC
Data curation: FC
Formal analysis: FC, KG, SD
Visualization: FC
Writing—original draft: FC
Writing—review and editing: KG, AA, SD
Supervision: AD, KG

Conflicts of Interest

None declared.

Multimedia Appendix 1
Antimicrobial resistance campaign media pack.
[PDF File, 1198 KB - publichealth_v12i1e86587_app1.pdf]

Multimedia Appendix 2
Pulse videos and paired videos selected by topic and by recency.
[XLSX File, 17 KB - publichealth_v12i1e86587_app2.xlsx]

Multimedia Appendix 3
Supplementary tables.
[DOCX File, 30 KB - publichealth_v12i1e86587_app3.docx]

References

XSL•FO
RenderX

1. Antimicrobial resistance (AMR). NHS England. URL: https://www.england.nhs.uk/ourwork/prevention/antimicrobial-resistance-amr/ [accessed 2026-03-20]
2. Antimicrobial resistance. World Health Organization. 2023. URL: https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance [accessed 2026-3-20]
3. GBD 2021 Antimicrobial Resistance Collaborators. Global burden of bacterial antimicrobial resistance 1990–2021: a systematic analysis with forecasts to 2050. Lancet 2024 Sep 28;404(10459):1199-1226. [doi: 10.1016/S0140-6736(24)01867-1] [Medline: 39299261]
4. People-centred approach to addressing antimicrobial resistance in human health: WHO core package of interventions to support national action plans. World Health Organization. 2023. URL: https://www.who.int/publications/i/item/9789240082496 [accessed 2026-03-20]
5. Gilham EL, Pearce-Smith N, Carter V, Ashiru-Oredope D. Assessment of global antimicrobial resistance campaigns conducted to improve public awareness and antimicrobial use behaviours: a rapid systematic review. BMC Public Health 2024 Feb 6;24(1):396. [doi: 10.1186/s12889-024-17766-w] [Medline: 38321479]
6. Defining health campaigns and health campaign effectiveness. Health Campaign Effectiveness Coalition. 2023. URL: https://campaigneffectiveness.org/defining-health-campaigns-and-health-campaign-effectiveness [accessed 2026-3-20]
7. Vernon E, Gottesman Z, Warren R. The value of health awareness days, weeks and months: a systematic review. Soc Sci Med 2021 Jan;268:113553. [doi: 10.1016/j.socscimed.2020.113553] [Medline: 33280924]
8. Rhode C. The effect of health awareness days, weeks, and months: can they change health behavior? J Consum Health Internet 2023;27(4):431-442. [doi: 10.1080/15398285.2023.2271308]
9. About mental health awareness week. Mental Health Foundation. URL: https://www.mentalhealth.org.uk/our-work/public-engagement/mental-health-awareness-week/about [accessed 2026-3-20]
10. Beaney T, Kerr GK, Kiru G, et al. May Measurement Month 2022: results from the global blood pressure screening campaign. BMJ Glob Health 2024 Dec 2;9(12):e016557. [doi: 10.1136/bmjgh-2024-016557] [Medline: 39622541]
11. Kioumarsi H, Liu B. Celebrating World Cancer Day: innovative biological approaches to cancer and their alignment with sustainable development goals (SDGs). Biol Proced Online 2025 Feb 4;27(1):3. [doi: 10.1186/s12575-025-00263-8] [Medline: 39901124]
12. Kite J, Chan L, MacKay K, et al. A model of social media effects in public health communication campaigns: systematic review. J Med Internet Res 2023 Jul 14;25:e46345. [doi: 10.2196/46345] [Medline: 37450325]
13. Campaign resource centre. Department of Health & Social Care, United Kingdom. URL: https://campaignresources.dhsc.gov.uk/ [accessed 2026-3-20]
14. Antibiotic resistance advert - keep antibiotics working and take your doctor's advice. YouTube. 2017. URL: https://www.youtube.com/watch?v=zTbLai2GaQM [accessed 2026-3-20]
15. Brubach A. “Little Lungs” - using digital platforms, innovative creative, and audience passion points to prevent smoking among U.S. youth. Tob Induc Dis 2018 Mar;16(1). [doi: 10.18332/tid/84281]
16. Liu J, Ye Q, Wu H, Ma R, Guo S, Long H. How do health content creators perform well? An integration research of short video and livestream behaviors. Front Public Health 2024 Oct 2;12:1446247. [doi: 10.3389/fpubh.2024.1446247] [Medline: 39416935]
17. Hendry NA, Hartung C, Welch R. Health education, social media, and tensions of authenticity in the ‘ influencer pedagogy’ of health influencer Ashy Bines. Learn Media Technol 2022 Oct 2;47(4):427-439. [doi: 10.1080/17439884.2021.2006691]
18. Baduashvili A, Fainstad B, Kudron EL. Considerations for clinician-educators developing online educational content: a narrative review. Cureus 2024 Feb 27;16(2):e55048. [doi: 10.7759/cureus.55048] [Medline: 38550457]
19. What is the place of the clinical content creator, and how can we increase the quality of health information online? BMJ Future Health. 2024. URL: https://futurehealth.bmj.com/webinars-1/place-clinical-content-creator-increase-quality-health-information-online- [accessed 2026-03-20]
20. World Health Organization (WHO). YouTube. URL: https://www.youtube.com/@who [accessed 2026-03-20]
21. Centers for Disease Control and Prevention (CDC). YouTube. URL: https://www.youtube.com/@CDC [accessed 2026-03-20]
22. YouTube. URL: https://www.youtube.com [accessed 2026-3-20]
23. Powell J, Pring T. The impact of social media influencers on health outcomes: systematic review. Soc Sci Med 2024 Jan;340:116472. [doi: 10.1016/j.socscimed.2023.116472] [Medline: 38070305]
24. Fides - a network of healthcare influencers. World Health Organization. URL: https://www.who.int/teams/digital-health-and-innovation/digital-channels/fides [accessed 2026-03-20]
25. Sopel A. WHO and TikTok to collaborate on more science-based information on health and well-being. World Health Organization. 2024. URL: https://www.who.int/news/item/26-09-2024-who-and-tiktok-to-collaborate-on-more-science-based-information-on-health-and-well-being [accessed 2026-3-20]
26. Bouchaud P. Algorithmic amplification of politics and engagement maximization on social media. In: Cherifi H, Rocha LM, Cherifi C, Donduran M, editors. Complex Networks & Their Applications XII: Springer; 2024. [doi: 10.1007/978-3-031-53503-1_11]

XSL•FO
RenderX

27. Rozelle A. Building momentum: creating change. In: Wallace A, Luttrell R, editors. Strategic Social Media as Activism: Repression, Resistance, Rebellion, Reform: Routledge; 2023:87-99. [doi: 10.4324/9781003291855]
28. High-level meeting on antimicrobial resistance (AMR). General Assembly of the United Nations. 2024. URL: https://www.un.org/pga/79/antimicrobial/ [accessed 2026-03-20]
29. Altman DG, von Elm E, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. EQUATOR Network. URL: https://www.equator-network.org/reporting-guidelines/strobe/ [accessed 2026-03-20]
30. Oladipo T. Ask Buffer: what is short-form video, and how can you use it? Buffer. 2023. URL: https://buffer.com/resources/short-form-video/ [accessed 2026-03-20]
31. CHAIN: The Content, Health and AMR Innovation Network. Fleming Initiative. URL: https://www.fleminginitiative.org/chain [accessed 2026-03-20]
32. Preventing and controlling infections. UK Health Security Agency. 2025. URL: https://www.gov.uk/government/publications/health-protection-in-schools-and-other-childcare-facilities/preventing-and-controlling-infections [accessed 2026-03-20]
33. WASH and antimicrobial resistance. World Health Organization. URL: https://www.who.int/teams/environment-climate-change-and-health/water-sanitation-and-health/burden-of-disease/wash-and-antimicrobial-resistance [accessed 2026-03-20]
34. UN General Assembly High-Level Meeting on antimicrobial resistance 2024. World Health Organization. 2024. URL: https://www.who.int/news-room/events/detail/2024/09/26/default-calendar/un-general-assembly-high-level-meeting-on-antimicrobial-resistance-2024 [accessed 2026-03-20]
35. Shen J. The research of the factors that influence the popularity of YouTube videos. Theor Nat Sci 2024;51(1):187-193. [doi: 10.54254/2753-8818/51/2024CH0200]
36. Cloud computing services. Google Cloud. URL: https://cloud.google.com/ [accessed 2026-03-20]
37. Posit. URL: http://www.posit.co/ [accessed 2026-3-20]
38. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2022. URL: https://www.R-project.org/ [accessed 2026-03-20]
39. van Rossum G. Python Reference Manual. : Centrum voor Wiskunde en Informatica; 1995 URL: https://ir.cwi.nl/pub/5008 [accessed 2026-03-20]
40. Ooms J. The jsonlite package: a practical and consistent mapping between JSON data and R objects. arXiv. Preprint posted online on Mar 12, 2014. [doi: 10.48550/arXiv.1403.2805]
41. Ethics approval overview. Imperial College London. URL: https://www.imperial.ac.uk/research-ethics-committee/ethics-approval-overview/ [accessed 2026-03-20]
42. González-Estrada E, Cosmes W. Shapiro–Wilk test for skew normal distributions based on data transformations. J Stat Comput Simul 2019 Nov 22;89(17):3258-3272. [doi: 10.1080/00949655.2019.1658763]
43. Hsu H, Lachenbruch PA. Paired t test. In: Balakrishnan N, Colton T, Everitt B, Piegorsch W, Ruggeri F, Teugels JL, editors. Wiley StatsRef: Statistics Reference Online: John Wiley & Sons; 2014.
44. Rey D, Neuhäuser M. Wilcoxon-signed-rank test. In: Lovric M, editor. International Encyclopedia of Statistical Science: Springer; 2011:1658-1659. [doi: 10.1007/978-3-642-04898-2_616]
45. Rosenkranz GK. A note on the Hodges-Lehmann estimator. Pharm Stat 2010;9(2):162-167. [doi: 10.1002/pst.387] [Medline: 19718773]
46. Allison PD, Waterman RP. Fixed–effects negative binomial regression models. Sociol Methodol 2002;32(1):247-265. [doi: 10.1111/1467-9531.00117]
47. Wickham H, François R, Henry L, Müller K, Vaughan D. dplyr: a grammar of data manipulation. The Comprehensive R Archive Network. URL: https://cran.r-project.org/web/packages/dplyr/ [accessed 2026-03-20]
48. Wickham H. httr: tools for working with URLs and HTTP. The Comprehensive R Archive Network. URL: https://cran.r-project.org/package=httr [accessed 2026-03-20]
49. Gertzel BB, Ackland R, Graham T, Borquez F. Package ‘vosonSML’: collecting social media data and generating networks for analysis. The Comprehensive R Archive Network. 2025. URL: https://cran.r-project.org/web/packages/vosonSML/vosonSML.pdf [accessed 2026-03-20]
50. Müller K, Bryan J. here: a simpler way to find your files. The Comprehensive R Archive Network. 2025. URL: https://cran.r-project.org/web/packages/here/index.html [accessed 2026-3-20]
51. Wickham H, Bryan J. readxl: read Excel files. The Comprehensive R Archive Network. 2025. URL: https://cran.r-project.org/web/packages/readxl/index.html [accessed 2026-03-20]
52. Wickham H. ggplot2: Elegant Graphics for Data Analysis: Springer; 2016.
53. Torchiano M. effsize: efficient effect size computation. The Comprehensive R Archive Network. URL: https://cran.r-project.org/web/packages/effsize/index.html [accessed 2026-03-20]
54. Wickham H, Averick M, Bryan J, et al. Welcome to the Tidyverse. J Open Source Softw 2019;4(43):1686. [doi: 10.21105/joss.01686]
55. Wickham H. Reshaping data with the reshape package. J Stat Softw 2007;21(12):1-20. [doi: 10.18637/jss.v021.i12]
56. Venables WN, Ripley BD. Modern Applied Statistics With S, 4th edition: Springer; 2002. [doi: 10.1007/978-0-387-21706-2]

XSL•FO
RenderX

57. Zeileis A, Köll S, Graham N. Various versatile variances: an object-oriented implementation of clustered covariances in R. J Stat Softw 2020;95(1):1-36. [doi: 10.18637/jss.v095.i01]
58. Zeileis A, Hothorn T. Diagnostic checking in regression relationships. R J 2002;2(3) [FREE Full text]
59. Robinson D, Hayes A, Couch S, et al. broom: convert statistical objects into tidy tibbles. The Comprehensive R Archive Network. URL: https://cran.r-project.org/web/packages/broom/index.html [accessed 2026-03-20]
60. Wojcik G, Afseth J, Fagan R, Thomson F, Ring N. Patient and public understanding of antimicrobial resistance: a systematic review and meta-ethnography. JAC Antimicrob Resist 2024 Aug 7;6(4):dlae117. [doi: 10.1093/jacamr/dlae117] [Medline: 39114563]
61. Tsao SF, Chen H, Tisseverasinghe T, Yang Y, Li L, Butt ZA. What social media told us in the time of COVID-19: a scoping review. Lancet Digit Health 2021 Mar;3(3):e175-e194. [doi: 10.1016/S2589-7500(20)30315-0] [Medline: 33518503]
62. Rajendran PT, Creusy K, Garnes V. Shorts on the rise: assessing the effects of YouTube shorts on long-form video content. arXiv. Preprint posted online on Feb 28, 2024. [doi: 10.48550/arXiv.2402.18208]
63. Albert AA, Knoll MA, Doke K, et al. #WomenWhoCurie: leveraging social media to promote women in radiation oncology. Adv Radiat Oncol 2019;4(2):218-225. [doi: 10.1016/j.adro.2019.01.003] [Medline: 31011664]
64. Corsini EM, Boeck M, Hughes KA, et al. Global impact of social media on women in surgery. Am Surg 2020 Feb 1;86(2):152-157. [Medline: 32106909]
65. Middleton J. We want to know why you are #proud2nurse. Nursing Times. 2014. URL: https://www.nursingtimes.net/opinion/we-want-to-know-why-you-are-proud2nurse-17-09-2014/ [accessed 2026-01-21]
66. Ritter AZ, Aronowitz S, Leininger L, et al. Dear pandemic: nurses as key partners in fighting the COVID-19 infodemic. Public Health Nurs 2021 Jul;38(4):603-609. [doi: 10.1111/phn.12903] [Medline: 33876450]
67. Bradshaw AS. #DoctorsSpeakUp: exploration of hashtag hijacking by anti-vaccine advocates and the influence of scientific counterpublics on Twitter. Health Commun 2023 Oct;38(10):2167-2177. [doi: 10.1080/10410236.2022.2058159] [Medline: 35437069]
68. Breza E, Stanford FC, Alsan M, et al. Effects of a large-scale social media advertising campaign on holiday travel and COVID-19 infections: a cluster randomized controlled trial. Nat Med 2021 Sep;27(9):1622-1628. [doi: 10.1038/s41591-021-01487-3] [Medline: 34413518]
69. Can the creation of a coordinated public health campaign on antimicrobial resistance via a network of health content creators generate engagement? A prospective evaluation. OSF. 2026. URL: https://osf.io/8hxr4/overview [accessed 2026-03-20]

Abbreviations

AMR: antimicrobial resistance
API: application programming interface
CHAIN: Content, Health, and Antimicrobial Resistance Innovation Network
HCP: health care professional
IRR: incidence rate ratio
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
UNGA: United Nations General Assembly
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 27.Oct.2025; peer-reviewed by D Iryawati, NR Geyer; revised version received 03.Feb.2026; accepted 25.Feb.2026; published 27.Mar.2026.

Please cite as:
Chen F, Cooper J, Acharya A, Dryden S, Darzi A, Grailey K
The Impact on Audience Engagement of Coordinating a Public Health Campaign on Antimicrobial Resistance Through a Network of Health Content Creators: Longitudinal Observational Study
JMIR Public Health Surveill 2026;12:e86587
URL: https://publichealth.jmir.org/2026/1/e86587
doi:10.2196/86587

© Fangyue Chen, Jack Cooper, Amish Acharya, Simon Dryden, Ara Darzi, Kate Grailey. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 27.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Social Media Campaign to Promote COVID-19 Vaccination: Cost-Effectiveness Analysis

Michael William Long[1], SD, MPH; Jeffrey B Bingenheimer[1], PhD; Khadidiatou Ndiaye[1], PhD; Dante Donati[2], PhD; Nandan Rao[3], MS; Selinam Akaba[1], MPH; Sohail Agha[4,5], PhD; William Douglas Evans[1], PhD

[1]Department of Prevention and Community Health, Milken Institute School of Public Health, The George Washington University, Washington, DC, United States

[2]School of Business, Columbia University, New York, NY, United States

[3]Virtual Lab LLC, Corvallis, OR, United States

[4]Behavioral Insights Lab, Seattle, WA, United States

[5]Global Health Visions, Seattle, WA, United States

Corresponding Author:
Michael William Long, SD, MPH
Department of Prevention and Community Health
Milken Institute School of Public Health
The George Washington University
950 New Hampshire Ave NW
Washington, DC, 20052
United States
Phone: 1 202 994 8729
Email: michael_long@gwu.edu

Abstract

Background: Vaccine hesitancy has increased in recent decades internationally, which sets up a critical barrier to the rapid deployment of novel vaccines against infection with SARS-CoV-2.

Objective: This study used a quasi-experimental design to evaluate the cost-effectiveness of a social media intervention to reduce COVID-19 vaccine hesitancy implemented in Nigeria in 2022.

Methods: The intervention targeted health care providers and adults from the general population who were users of a specific social media platform. We used published estimates from a quasi-experimental evaluation of the campaign's effectiveness compared to the status quo across 6 intervention states and 31 comparison states over a 10-month period. We estimated the cost-effectiveness of the campaign in terms of cost (2022 US dollars) per person vaccinated using a decision tree analysis and probabilistic sensitivity analysis.

Results: On the basis of the quasi-experimental trial, the campaign led to a crude 6.4–percentage point increase (219/692, 31.6% vs 117/463, 25.3%; P=.045) in vaccination rates and an adjusted 7.8–percentage point increase (95% CI 1.68-14.2; P=.02) controlling for age group, gender, educational level, religion, and occupation among the 20% (1933/9607) of the overall sample who were unvaccinated and in the persuadable middle. Scaled to the overall population, the campaign led to a 1.57–percentage point (95% CI 0.337-2.87; P=.02) increase in the proportion of those vaccinated against COVID-19 among those reached by the social media campaign. The social media campaign resulted in 58.3 million impressions and 1.87 million people reached for a total societal cost of US $1.15 million, or US $0.61 per person reached. This resulted in an incremental cost-effectiveness ratio of US $54.70 (95% uncertainty interval US $20.90-$163) per person vaccinated.

Conclusions: A social media–based campaign to address COVID-19 vaccine hesitancy in 6 states in Nigeria resulted in an increase in vaccination rates. The cost-effectiveness of the campaign compared to no campaign is comparable to that of other campaigns promoting COVID-19 vaccine uptake. The cost per person vaccinated due to the social media campaign was 1% to 8% of the estimated cost per life year saved by vaccination against COVID-19 in low- and middle-income countries. Investing in social media campaigns would likely be a cost-effective approach to increase vaccine uptake and save lives.

(JMIR Public Health Surveill 2026;12:e84540) doi:10.2196/84540

XSL•FO
RenderX

KEYWORDS

COVID-19; vaccination; social media; cost-effectiveness; health promotion

Introduction

The COVID-19 pandemic led to the death of 15 to 20 million people worldwide up to 2021 [1,2]. In response to this threat, governments and private companies demonstrated high capacity for innovation; the rapid development and testing of multiple effective vaccines stands out as a critical success [3]. The pandemic also highlighted ongoing systemic failures in global and national public health systems, including limited capacity for surveillance, communication, and distribution of preventive materials and services [4]. These failures exacerbated existing health inequities within and between countries.

The potential impact of the successful development, manufacture, and distribution of effective vaccines was not fully realized due to the public health system's inability to communicate the safety and benefit of the new vaccines in the context of widespread mis- and disinformation about the pandemic and the public health response. Building on well-established antivaccine movements, COVID-19 vaccine hesitancy emerged as a major barrier to the control of the pandemic [5]. By November 2023, a total of 80% of people living in high-income countries had received at least one dose of a COVID-19 vaccine compared to 33% of people living in low-income countries [6]. In the years before the COVID-19 pandemic, researchers were evaluating the potential use of social media communication campaigns to address vaccine misinformation and increase vaccine uptake. Previous vaccine promotion campaigns addressing vaccine hesitancy have mostly targeted a narrow set of vaccines (eg, influenza and human papillomavirus in high-income countries and diphtheria, tetanus, pertussis, and polio in middle- and low-income countries) [7]. Reviews of health promotion campaigns covering communicable and noncommunicable diseases on social media have found limited or mixed evidence of reported or observed behavior changes (ie, high engagement) and more reports of interaction with posts or changes in knowledge and attitudes (ie, low to medium engagement) [8,9].

With this promising but mixed and limited research base, and accompanied by calls for development of theoretically based and practice-based social marketing strategies [10], funders and public health organizations rapidly implemented social media campaigns to promote COVID-19 vaccine uptake. Initial evaluations of efforts to promote COVID-19 vaccination or other disease control behaviors through social media campaigns have been positive but with low to moderate effects, leading the public health community to consider whether and how to invest in a sustainable public health social media communication infrastructure [11-14]. Social media campaigns have the potential to reach targeted audiences with tailored messages in ways that may improve both impact and efficiency compared to mass media campaigns [15].

We evaluated the cost-effectiveness of a targeted social media campaign to promote vaccination against COVID-19 among health care providers and other adults in their social environment in Nigeria in 2022. By May 2022, after recording 250,000 COVID-19 cases, Nigeria had received enough COVID-19 vaccines to cover 25% of the population and had administered the first dose to 13% and the second dose to 8% of the population [16]. High levels of vaccine acceptance (76%) in late 2020 were being reported to be much lower as more data were published in 2021 (40%-60%) [16,17]. The World Bank, which classifies Nigeria as a lower-middle–income country, reported that 38% of the Nigerian population accessed the internet in 2022 [18]. A rapid rise in the use of social media in Nigeria and its complex role in the response to COVID-19 had been reported by the time the social media campaign in this paper had been implemented [19].

In this analysis, we aimed to evaluate the cost of implementing a social influencer–based social media campaign and estimate the value of the campaign in terms of cost per person vaccinated, which can be compared to other campaigns targeting vaccine uptake.

Methods

Overview

The prospective economic analysis plan was included in the overall analysis plan submitted to the funder and has not been published elsewhere. This project followed the guidelines of the Second Panel on Cost-Effectiveness in Health and Medicine and the reporting guidelines from the Consolidated Health Economic Evaluation Reporting Standards checklist [20,21]. The data used in the model synthesis were collected from 2021 to 2022. The analysis was completed in 2023.

Intervention Description

This cost-effectiveness analysis is based on the implementation and quasi-experimental evaluation of a 10-month social media campaign promoting vaccination against COVID-19 in Nigeria among health care workers and those in their social networks in 2022 [22]. The campaign was designed and implemented by a team of designers and local organizations and delivered through Facebook and Instagram. The campaign included provaccination social norms and vaccine hesitancy reduction messages delivered by social influencers (eg, local celebrities, health care providers, and religious and business leaders). The campaign theory of change was based on the theory of diffusion of innovations; social norms theory; and the motivation, opportunity, and ability framework [23-25].

Study Population and Setting

The intervention was implemented in 6 states in Nigeria (Anambra, Bauchi, Lagos, Niger, Rivers, and Sokoto), with participants in the control condition recruited from the Federal Capital Territory and all other states. Participants were eligible if they were aged ≥18 years, had a Facebook account registered in Nigeria and received recruitment advertising in their live feed promoting a study on COVID-19 vaccination, had not been previously vaccinated against COVID-19, and were defined as members of the "persuadable middle" [22]. Those who

XSL•FO
RenderX

responded "Definitely" or "Definitely not" to the question "Would you take a COVID-19 vaccine that is approved for use in Nigeria if offered to you?" were excluded based on not being in the persuadable middle. While people in low- and middle-income countries (LMICs) generally have higher vaccine acceptance than those in high-income countries. Nigeria faced vaccine availability and other challenges that may have impacted vaccine hesitancy differently than in higher-income settings, including perceptions that safety and efficacy had not been adequately evaluated in that setting [26-28].

Cost Evaluation

We used the standard microcosting approach, for which we evaluated all component costs of the intervention instead of using a global project budget. Microcosting includes 3 main steps: identification, measurement, and valuation. To identify the resources used, we prospectively developed a detailed description of the intervention activities and identified necessary resources for each activity. Resources were measured and valued using actual reported expenditures from implementing partners and reported or estimated opportunity costs for the nonbudgeted time from implementing partners, influencer organizations, and participants. Direct costs were all reported in US dollars by the implementing partners and were adjusted for inflation to 2022 US dollars. Opportunity costs accrued in Nigeria were estimated in 2022 Nigerian naira. Nigerian currency was converted to purchasing power parities, with total costs reported in 2022 purchasing power parities, which is equivalent to 2022 US dollars. Costs were converted in 2023. As we did not assess health or economic benefits of vaccination, we did not include opportunity costs of individuals or direct health care sector costs for receipt of the vaccine.

Intervention Reach

The intervention included 245 distinct advertising campaigns implemented on the Facebook social media platform, which means that the campaigns may have included distinct creative content or audience-targeting and promotion methods and their unique individual reach could not be combined with that of other campaigns. For each of these campaigns, the platform reported the total number of unique individuals receiving campaign messages (reach), the total impressions (ie, the number of times the campaign message was displayed on the target audience member's screen), and a range of engagement metrics for each of these campaigns. Because we did not have access to the total number of unique individuals reached across all campaigns, we estimated reach based on the largest reported reach across all campaigns. Due to a lack of data on the degree of overlap within a targeted campaign, we based our reach estimate on a conservative assumption that there was complete audience overlap across campaigns.

Cost-Effectiveness Analysis

We used a societal and payer perspective, which captured both the budgetary costs of implementing a similar campaign in the future and the opportunity costs of implementing partners and individuals engaging with campaign messages. The comparator was the status quo (ie, the current state of affairs in the absence of this intervention), which was chosen based on the intervention design and effect estimate. The time horizon for this study was 1 year to capture planning and implementation; we did not have the capacity to model longer-term health and cost effects following a change in vaccination rates. We did not discount costs or benefits over the 1-year time horizon.

Outcome Measurement

The primary outcome for this study was vaccination against COVID-19. The incremental effect of exposure to the advertising campaign was estimated from a survey of 10,965 participants who were users of the Facebook social media platform. Of the initial 10,965 participants screened for eligibility, 6198 (56.5%) were excluded as already vaccinated, 1476 (13.5%) were excluded for not being in the persuadable middle, 675 (6.2%) were excluded for missing baseline data, 648 (5.9%) were excluded for not meeting the age criteria, and 35 (0.3%) were excluded for having a duplicate ID. The remaining 17.6% (1933/10,965) of the participants were enrolled in the study. Surveys were fielded to the same cohort, with baseline data collection taking place during the period from December 1 to 31, 2021; first follow-up data collection taking place during the period from March 1, 2022, to April 30, 2022; and second follow-up data collection taking place during the period from October 1 to 4, 2022. Of the 1933 participants enrolled in the study, 1155 (59.8%) completed the first follow-up, and 462 (23.9%) completed the second follow-up. Exposure was based on state of residence, with the intervention implemented in 6 states (Anambra, Bauchi, Lagos, Niger, Rivers, and Sokoto) and control participants recruited from all other states in Nigeria.

Participants were recruited through a social media–based research platform called Virtual Lab. Recruitment was stratified by whether participants were health care providers, with the goal of recruiting 50% of the sample from the health care provider community. COVID-19 vaccination uptake was measured through a single question: "Have you received a COVID-19 vaccine?" Participants could respond as follows: "Yes, a single-dose vaccine"; "Yes, the first dose of a two-dose regimen"; "Yes, both doses of a two-dose regimen"; and "No." Due to changes in the types of vaccines available, as well as recommendations for boosters, we collapsed the outcome into a binary "vaccinated or not vaccinated" outcome.

The effect of the intervention was estimated using a linear regression model predicting vaccination status at the midpoint and final survey. The primary independent variable in each model was exposure to the intervention. Adjusted models included the following control variables: age group, gender, educational level, religion, and occupation. We used clustered SEs to account for nesting within state of residence. Additional details on the evaluation of the intervention on vaccine uptake are reported elsewhere [22].

For the purposes of this cost-effectiveness analysis, we estimated the reach of the campaign in the intervention states based on the impressions reported by the Facebook social media platform. Impressions are defined as an individual user's exposure to specific content on the platform that may or may not result in active engagement, such as liking, commenting, or following the account that disseminated or originated the content [29].

XSL•FO
RenderX

Impressions have been shown to account for most of the information exposure on social media, have low correlation with active engagement or "expression," and be independently correlated with user-reported influence of a given information source [29].

Uncertainty Analyses

We conducted a probabilistic sensitivity analysis by sampling from the distributions of all parameters with measured uncertainty (Table 1). We included the following scenario analysis: instead of using the effect estimate from the first follow-up from the original outcome study [22], we used the effect estimate from the second follow-up period from the same study. We did not evaluate the heterogeneity of the intervention effect or distributional effects of the intervention. Decision tree models and the probabilistic sensitivity analysis were conducted using TreeAge Pro (R2.0; TreeAge Software, LLC).

Table 1. Summary of inputs for the cost-effectiveness analysis of the COVID-19 vaccine promotion social media campaign in Nigeria in 2022.

Variable	Source	Point estimate (95% uncertainty interval)	Distribution (parameters)
Target population already vaccinated at the start of the campaign (%)	Quasi-experimental trial data [22]	64.5 (63.5 to 65.5)	Binomial (p^a=0.645, n^a=9607)
Persuadable middle population among those unvaccinated (%)	Quasi-experimental trial data [22]	56.7 (55.1 to 58.3)	Binomial (p=0.567, n=3409)
Percentage point increase in vaccination status due to treatment among the persuadable middle	Quasi-experimental trial data [22]	7.8 (1.68 to 14.2)	Normal (mean 0.078, SD 0.032)
Campaign reach	Meta advertiser platform	1,870,000	—[b]
Average engagement time per media impression (s)	Publisher analysis [30]	1.7	—
Total campaign impressions	Meta advertiser platform	58,300,000	—
Total cost (US $)	Campaign microcosting	1,150,000	—
Cost per person reached (US $)	Calculation	0.613	—
Sensitivity and scenario analyses			
Scenario 1: percentage point increase in vaccination status due to treatment among the persuadable middle using the second follow-up	Quasi-experimental trial data [22]	11.0 (−0.00337 to 0.225)	Normal (mean 0.110, SD 0.058)

[a]Parameters of each named distribution, where p denotes the probability and n denotes the number of trials.

[b]Not applicable.

Ethical Considerations

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional or national research committee and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards. This evaluation was approved by the George Washington University Institutional Review Board (NCR213708), as well as by the National Health Research Ethics Committee in Nigeria (NHREC/01/01/2007-04/10/2021). No identifiable data were used in this study. All participants provided informed consent to participate in the study following the institutional review board–approved protocol. Participants were compensated with 400 naira (approximately US $1) for completion of the 40-item survey implemented through the Facebook Messenger chat function.

Results

The intervention generated 58,255,000 total impressions across 245 distinct advertising campaigns, which, on the Meta platform (the company that owns Facebook), included one or more sets of individual advertisements. Distinct campaigns were run to allow the intervention to best measure and optimize performance against advertising objectives. The mean reach (unique individuals generating one or more impressions) per campaign was 100,000 (SD 176,000; range 1000-1,873,000). On the basis of an assumption that there was complete overlap across distinct advertising campaigns, the intervention reached 1,873,000 unique individuals.

We summarize intervention costs by activity category in Table 2. Due to the use of marketing labor in the United States and the United Kingdom as well as dollar-denominated contracts with partners in Nigeria, the payer costs accounted for 93% of the total societal costs even though the paid hours to implement the project constituted 14% of the total person-time included in the societal perspective.

XSL•FO
RenderX

Table 2. Cost of the COVID-19 vaccine promotion social media campaign by activity in Nigeria in 2022.[a]

	Payer perspective (US $)	Societal perspective (US $)
Government liaison	73,400	73,400
Monitoring and evaluation	98,300	98,300
Campaign development	360,000	360,000
Advertising expenditure	102,000	102,000
Advertising campaign implementation	134,000	134,000
Stakeholder management	293,000	293,000
Participant engagement with advertising	—[b]	77,700
Influencer campaign implementation	—	7520
Total	1,060,000	1,150,000

[a]Costs may not add up due to rounding.

[b]There are no participant opportunity costs included in the payer perspective.

Across both the control and intervention samples (excluding those who were ineligible based on age, duplicate ID, and missing baseline data), 64.5% (6198/9607) of the participants were already vaccinated at baseline. The vaccination rate among this sample of Facebook users was substantially higher than the 13% single-dose uptake reported at a similar point in the rollout (eg, May 2022) [16]. Of the 3409 participants screened in the study who were not vaccinated and were otherwise eligible, 1933 (56.7%) were considered to be in the persuadable middle and were enrolled in the study. In a previous study, we estimated that the intervention led to a 7.8–percentage point increase (95% CI 1.68-14.2) in vaccine uptake controlling for demographic variables among those in the persuadable middle.

In the primary analysis, we estimated that the incremental cost of the intervention per person reached was US $0.63 and the incremental percentage point increase in vaccination prevalence was 0.0157 (95% uncertainty interval [UI] 0.00337-0.0287). This resulted in an incremental cost-effectiveness ratio of US $54.70 (95% UI US $20.90-$163), which means that it cost US $54.70 more than the status quo (ie, the current state of affairs without the intervention) for every additional vaccination.

In scenario analysis 1, we used the effect estimate from the second follow-up of the same study as the primary analysis. In this scenario, the larger percentage point increase in vaccinations per person (0.0221 vs 0.0157) than in the no-intervention condition reduced the incremental cost-effectiveness ratio almost by half (US $29.60, 95% UI negative to US $180; Table 3). The UI includes 0 due to the smaller sample at the second follow-up and resulting marginally significant coefficient reported in the evaluation study. We found that using this estimate resulted in 3% of all model iterations having a negative effect.

Table 3. Cost-effectiveness results of the COVID-19 vaccine promotion social media campaign in Nigeria in 2022.

	Mean (95% uncertainty interval)
Incremental cost per person reached (US $)	0.613 (0.613 to 0.613)
Incremental increase in COVID-19 vaccinations per person exposed to the campaign	0.0157 (0.00337 to 0.0287)
Incremental cost-effectiveness ratio (US $ per vaccination)	54.70 (20.90 to 163)
Scenario 1: incremental COVID-19 vaccination per person	0.0221 (−0.000649 to 0.0452)[a]
Scenario 1: incremental cost-effectiveness ratio (US $ per vaccination)	29.60 (negative to 180)[b]

[a]For scenario 1, we used an alternative estimate of the effectiveness of the intervention from the second follow-up period of the same intervention used for the primary analysis.

[b]A total of 3% of the model iterations were negative.

Discussion

Principal Findings

In this cost-effectiveness analysis of a social media campaign promoting vaccination against COVID-19 among health care workers and adults in their social environment in Nigeria in 2022, we found that the intervention increased vaccination rates among the target audience at a cost in line with similar efforts in the field.

Incremental cost-effectiveness estimates of media campaigns promoting vaccine uptake vary substantially. On the basis of an analysis of attitude changes as a result of social media campaigns run by 174 public health organizations during the COVID-19 pandemic and another study linking attitudes to vaccination outcomes, Athey et al [31] estimated that the campaigns cost US $5.68 per person vaccinated. The study by Athey et al [31] only incorporated the cost of advertising, which accounted for only 12% of the total costs of running and participating in the campaign in our study. This suggests that

XSL•FO
RenderX

our estimate of US $54.70 is likely consistent with that of the analysis by Athey et al [31] (which estimated that it would cost US $48 per person vaccinated assuming a similar cost structure) and highlights the importance of incorporating as many relevant costs as feasible when presenting the cost-effectiveness of social media campaigns.

Because there is no willingness-to-pay threshold for the cost of an incremental person vaccinated, it may be useful to integrate the findings of this study with those of others that have measured the cost per year of life saved (YLS) or cost per disability-adjusted or quality-adjusted life year. A study estimating health benefits and donor costs of increase in COVID-19 vaccination rates in 91 LMICs found that spending on vaccination would cost between US $670 per YLS and US $7820 per YLS depending on the level of vaccination achieved [32]. The authors noted that the cost per YLS for COVID-19 vaccination was similar to the cost for antiretroviral therapy for HIV under the President's Emergency Plan for AIDS Relief, which they estimated at US $4310 per YLS using the total budget and life years saved from the President's Emergency Plan for AIDS Relief 2004 to 2013 [31]. The cost per person vaccinated in this study (US $54.70) was between 1% and 8% of the estimated cost per YLS by vaccination against COVID-19 in the 91 LMICs in the aforementioned study [32]. To further contextualize the value of the social media campaign evaluated in this study, vaccination against COVID-19 in LMICs was estimated to prevent 20.39 deaths per 10,000 people vaccinated; each death from COVID-19 was separately estimated to lead to 16 years of life lost [33,34]. This means that, for each person vaccinated, there was an average of 0.0326 (20.39 × 16/10,000) years of life lost prevented. On the basis of the estimates of the variable cost of vaccination delivery after rollout of a national campaign (US $10 for the vaccine and US $2.46 for delivery) and the cost of promotion obtained from this study (US $54.70), the marginal cost for each vaccination delivered would be US $67.16, leading to an estimate of US $2060 per year of life lost averted. The value of rapidly disseminating science-based vaccine promotion in terms of within-country health benefits likely underestimates the benefits of responding to shared global vulnerabilities with shared investments in mutually beneficial solutions such as vaccination. Baker et al [35] highlight this need for rapid collaboration as they paint an alarming picture of our new era of globally shared infectious disease risk caused by the confluence of climate change, urbanization, migration, travel, and intensifying trade of plants and animals.

Much of the work to prepare and launch this specific campaign to increase COVID-19 vaccine uptake could support other public health communications campaigns in Nigeria and potentially other countries. Moving the intervention to scale, such as all 37 states instead of the 6 in the intervention arm of this study, would spread fixed costs across a much larger population and reduce the cost per person vaccinated substantially. Goulbourne and Yanovitzky [36] argue that the COVID-19 pandemic clarified the role of health communication infrastructure as a social determinant of health and that public health organizations will need to invest in hyperlocal health communication capacity across populations to address health inequities. They suggest that training and providing ongoing technical support to trusted intermediaries is one approach to providing hyperlocal health communication at scale. The intervention evaluated in this study did implement the COVID-19 vaccine promotion social media campaign through 12 local health organizations and 10 other local influencers. The involvement of local influencers to shape and deliver health messages was considered an essential component of the campaign. This approach could limit the degree to which the intervention could be scaled at a lower marginal cost.

A primary limitation of this cost-effectiveness analysis is that we were not able to obtain a specific estimate of the total unique individuals reached by the intervention on the Meta platform. To be conservative, we estimated a total intervention reach of 1.87 million unique users based on the reach of the largest single campaign and not the 24.5 million reached if we summed the reported reach estimates for all campaigns. Our estimated US $0.61 per person reached by the campaign would instead be US $0.05, shifting the cost per vaccination from US $54.70 to US $2.98. This order of magnitude difference in the cost-effectiveness of the intervention emphasizes the importance of understanding how social media reach metrics are reported and how studies estimating the same cost-effectiveness outcomes (eg, cost per person vaccinated against COVID-19) are using these metrics. The lack of comparability across studies may be further compounded when studies only use active engagement or expression as a measure of campaign reach [29].

The extent of competing social media and other communication campaigns promoting vaccination against COVID-19, as well as the high levels of mis- and disinformation about the pandemic and the vaccines on the same social media platforms, created another limitation. The incremental effect of the intervention campaigns on the message environment was lower than it would have been in a nonpandemic context. We were not able to assess any competing or synergistic effects of the campaign due to variation in individual or community media environments, nor were we able to evaluate how the campaign interacted with other public health campaigns on the same platform or across channels. Extrapolation of findings from this study period to future pandemics may be limited by the rapidly changing nature of the social media landscape, including as it relates to platform responsibility to address public health misinformation. The recent divergence in the degree of regulatory control over content moderation between the European Union's Digital Services Act requirement that platforms address the systemic risks posed by misinformation and American jurisprudence's strengthening of free speech protections of content moderation means that mostly American corporations will potentially pursue jurisdictionally fragmented approaches to misinformation during the next pandemic [37].

We used a self-reported measure of vaccination, which could potentially overestimate the effect of the intervention. Stephenson et al [38] reported that, among a sample of approximately 2000 patients with both self-reported and recorded COVID-19 vaccination status in a hospital setting, the self-reported and recorded vaccination status matched for 95% of the participants. While we used existing studies on the cost-effectiveness of vaccination in similar settings [32], we did not directly estimate how the campaign affected health

XSL•FO
RenderX

outcomes, which may vary based on, among other factors, the vaccination level in the community, underlying demography and health status of the population, type of vaccine used, and health care system cost and effectiveness. Incorporating these factors within evaluation of new health communication and other strategies is likely infeasible for most interventions but could be accomplished by partnering with modeling groups that do address these factors or through sustained support of modeling consortia that could share modeling capacity more rapidly during future pandemics [39].

Conclusions

We found that a local influencer–based social media campaign implemented in 6 states in Nigeria during the COVID-19 pandemic increased COVID-19 vaccination rates among those exposed to the campaign. The campaign demonstrated comparable cost-effectiveness to that of other COVID-19 vaccination campaigns when accounting for differences in cost data included across studies. When combined with existing estimates of the effect of vaccination against COVID-19 on mortality and years of life lost per death due to COVID-19, this intervention achieved a lower cost per year of life lost averted (US $2060) than debated but recognized thresholds of 3 times the national gross domestic product per year of life lost averted [40]. Boosting the reach of vaccination efforts through influencer-based social media campaigns such as the one implemented in this study is likely to be a cost-effective approach to save lives.

Acknowledgments

The social media campaign evaluated in this study was led by M+C Saatchi, UpSwell, and a consortium of Nigerian organizations that published the social media content with approval from the Nigerian government. The authors thank these organizations for their collaboration.

Funding

This study was funded by the Bill & Melinda Gates Foundation (INV-033413).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: MWL, JBB, and WDE
Data curation: MWL
Formal analysis: MWL
Funding acquisition: WDE
Investigation: MWL, JBB, KN, DD, NR, S Akaba, S Agha, and WDE
Methodology: MWL, JBB, and WDE
Project administration: S Akaba
Supervision: WDE
Writing—original draft: MWL
Writing—review and editing: MWL, JBB, KN, DD, NR, S Akaba, S Agha, and WDE

Conflicts of Interest

NR has ownership interests in Virtual Lab, LLC, a company that uses open-source software described in this paper (Virtual Lab) to provide paid services. S Agha, the project officer, who was an employee of the funder at the time this project was initiated, participated in discussions about the study design and data collection. At the time of this manuscript preparation, S Agha was no longer with the funder. S Agha did participate in the development of the manuscript as noted in the Authors' Contributions section.

References

1. COVID-19 Excess Mortality Collaborators. Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020-21. Lancet 2022 Apr 16;399(10334):1513-1536 [FREE Full text] [doi: 10.1016/S0140-6736(21)02796-3] [Medline: 35279232]
2. Msemburi W, Karlinsky A, Knutson V, Aleshin-Guendel S, Chatterji S, Wakefield J. The WHO estimates of excess mortality associated with the COVID-19 pandemic. Nature 2023 Jan;613(7942):130-137 [FREE Full text] [doi: 10.1038/s41586-022-05522-2] [Medline: 36517599]
3. Bloom DE, Cadarette D, Ferranna M, Hyer RN, Tortorice DL. How new models of vaccine development For COVID-19 have helped address an epic public health crisis. Health Aff (Millwood) 2021 Mar;40(3):410-418. [doi: 10.1377/hlthaff.2020.02012] [Medline: 33539191]

XSL•FO
RenderX

4. Jha A. System failure: America needs a global health policy for the pandemic age. Foreign Affairs. URL: https://www.foreignaffairs.com/articles/united-states/2021-02-16/system-failure [accessed 2025-05-29]
5. Rozek LS, Jones P, Menon A, Hicken A, Apsley S, King EJ. Understanding vaccine hesitancy in the context of COVID-19: the role of trust and confidence in a seventeen-country survey. Int J Public Health 2021 May 14;66:636255 [FREE Full text] [doi: 10.3389/ijph.2021.636255] [Medline: 34744589]
6. Global dashboard for vaccine equity: UNDP data futures platform. United Nations Development Programme (UNDP). URL: https://data.undp.org/vaccine-equity/ [accessed 2025-05-29]
7. Jarrett C, Wilson R, O'Leary M, Eckersberger E, Larson HJ, SAGE Working Group on Vaccine Hesitancy. Strategies for addressing vaccine hesitancy - a systematic review. Vaccine 2015 Aug 14;33(34):4180-4190 [FREE Full text] [doi: 10.1016/j.vaccine.2015.04.040] [Medline: 25896377]
8. Giustini D, Ali SM, Fraser M, Kamel Boulos MN. Effective uses of social media in public health and medicine: a systematic review of systematic reviews. Online J Public Health Inform 2018;10(2):e215 [FREE Full text] [doi: 10.5210/ojphi.v10i2.8270] [Medline: 30349633]
9. Kite J, Chan L, MacKay K, Corbett L, Reyes-Marcelino G, Nguyen B, et al. A model of social media effects in public health communication campaigns: systematic review. J Med Internet Res 2023 Jul 14;25:e46345 [FREE Full text] [doi: 10.2196/46345] [Medline: 37450325]
10. Evans WD, French J. Demand creation for COVID-19 vaccination: overcoming vaccine hesitancy through social marketing. Vaccines (Basel) 2021 Apr 01;9(4):33 [FREE Full text] [doi: 10.3390/vaccines9040319] [Medline: 33915695]
11. Hunt IV, Dunn T, Mahoney M, Chen M, Nava V, Linos E. A social media based public health campaign encouraging COVID-19 vaccination across the United States. Am J Public Health 2022 Sep;112(9):1253-1256. [doi: 10.2105/ajph.2022.306934]
12. Breza E, Stanford FC, Alsan M, Alsan B, Banerjee A, Chandrasekhar AG, et al. Effects of a large-scale social media advertising campaign on holiday travel and COVID-19 infections: a cluster randomized controlled trial. Nat Med 2021 Sep 19;27(9):1622-1628 [FREE Full text] [doi: 10.1038/s41591-021-01487-3] [Medline: 34413518]
13. Di Mauro A, Di Mauro F, De Nitto S, Rizzo L, Greco C, Stefanizzi P, et al. Social media interventions strengthened COVID-19 immunization campaign. Front Pediatr 2022 Apr 5;10:869893 [FREE Full text] [doi: 10.3389/fped.2022.869893] [Medline: 35450111]
14. Elgersma IH, Fretheim A, Indseth T, Munch AT, Johannessen LB, Hansen CE. The evaluation of a social media campaign to increase COVID-19 testing in migrant groups: cluster randomized trial. J Med Internet Res 2022 Mar 24;24(3):e34544 [FREE Full text] [doi: 10.2196/34544] [Medline: 35285811]
15. Kisa S, Kisa A. A comprehensive analysis of COVID-19 misinformation, public health impacts, and communication strategies: scoping review. J Med Internet Res 2024 Aug 21;26:e56931 [FREE Full text] [doi: 10.2196/56931] [Medline: 39167790]
16. Eguavoen A, Larson HJ, Chinye-Nwoko F, Ojeniyi T. Reducing COVID-19 vaccine hesitancy and improving vaccine uptake in Nigeria. J Public Health Afr 2023 Apr 30;14(5):2290 [FREE Full text] [doi: 10.4081/jphia.2023.2290] [Medline: 37492424]
17. Babatope T, Ilyenkova V, Marais D. COVID-19 vaccine hesitancy: a systematic review of barriers to the uptake of COVID-19 vaccine among adults in Nigeria. Bull Natl Res Cent 2023 Mar 21;47(1):45 [FREE Full text] [doi: 10.1186/s42269-023-01017-w] [Medline: 36970323]
18. Individuals using the internet (% of population). World Bank Group. URL: https://data.worldbank.org/indicator/IT.NET.USER.ZS?type=shaded&view=map&year=2022 [accessed 2025-05-29]
19. Ahmed MO, Msughter AE. Assessment of the spread of fake news of COVID-19 amongst social media users in Kano State, Nigeria. Comput Hum Behav Rep 2022 May;6:100189 [FREE Full text] [doi: 10.1016/j.chbr.2022.100189] [Medline: 35372708]
20. Neumann PJ, Sanders GD, Russell LB, Siegel JE, Ganiats TG. Cost-Effectiveness in Health and Medicine. Oxford, UK: Oxford University Press; 2016.
21. Husereau D, Drummond M, Augustovski F, de Bekker-Grob E, Briggs AH, Carswell C, CHEERS 2022 ISPOR Good Research Practices Task Force. Consolidated health economic evaluation reporting standards 2022 (CHEERS 2022) statement: updated reporting guidance for health economic evaluations. Clin Ther 2022 Feb;44(2):158-168. [doi: 10.1016/j.clinthera.2022.01.011] [Medline: 35168801]
22. Evans WD, Bingenheimer JB, Long M, Ndiaye K, Donati D, Rao NM, et al. Outcomes of a social media campaign to promote COVID-19 vaccination in Nigeria. PLoS One 2023;18(9):e0290757 [FREE Full text] [doi: 10.1371/journal.pone.0290757] [Medline: 37713381]
23. Agha S. Use of a practitioner-friendly behavior model to identify factors associated with COVID-19 vaccination and other behaviors. Vaccines (Basel) 2022 Aug 05;10(8):1261 [FREE Full text] [doi: 10.3390/vaccines10081261] [Medline: 36016149]
24. Cislaghi B, Heise L. Using social norms theory for health promotion in low-income countries. Health Promot Int 2019 Oct 01;34(5):1069 [FREE Full text] [doi: 10.1093/heapro/day065] [Medline: 30188998]
25. Rogers EM. Diffusions of Innovations. 5th edition. New York, NY: Simon & Schuster; 2003.

XSL•FO
RenderX

26. Bhopal S, Nielsen M. Vaccine hesitancy in low- and middle-income countries: potential implications for the COVID-19 response. Arch Dis Child 2021 Feb;106(2):113-114. [doi: 10.1136/archdischild-2020-318988] [Medline: 32912868]
27. Moola S, Gudi N, Nambiar D, Dumka N, Ahmed T, Sonawane IR, et al. A rapid review of evidence on the determinants of and strategies for COVID-19 vaccine acceptance in low- and middle-income countries. J Glob Health 2021;11:05027 [FREE Full text] [doi: 10.7189/jogh.11.05027] [Medline: 34912550]
28. Patwary MM, Alam MA, Bardhan M, Disha AS, Haque MZ, Billah SM, et al. COVID-19 vaccine acceptance among low- and lower-middle-income countries: a rapid systematic review and meta-analysis. Vaccines (Basel) 2022 Mar 11;10(3):e62 [FREE Full text] [doi: 10.3390/vaccines10030427] [Medline: 35335059]
29. Liu R, Greene KT, Liu R, Mandic M, Valentino BA, Vosoughi S, et al. Using impression data to improve models of online social influence. Sci Rep 2021 Aug 16;11(1):16613 [FREE Full text] [doi: 10.1038/s41598-021-96021-3] [Medline: 34400698]
30. Capturing attention in feed: the science behind effective video creative. Meta. 2016. URL: https://www.facebook.com/business/news/insights/capturing-attention-feed-video-creative [accessed 2025-05-29]
31. Athey S, Grabarz K, Luca M, Wernerfelt N. Digital public health interventions at scale: the impact of social media advertising on beliefs and outcomes related to COVID vaccines. Proc Natl Acad Sci U S A 2023 Jan 31;120(5):e2208110120 [FREE Full text] [doi: 10.1073/pnas.2208110120] [Medline: 36701366]
32. Siedner MJ, Alba C, Fitzmaurice KP, Gilbert RF, Scott JA, Shebl FM, et al. Cost-effectiveness of coronavirus disease 2019 vaccination in low- and middle-income countries. J Infect Dis 2022 Nov 28;226(11):1887-1896 [FREE Full text] [doi: 10.1093/infdis/jiac243] [Medline: 35696544]
33. Watson OJ, Barnsley G, Toor J, Hogan AB, Winskill P, Ghani AC. Global impact of the first year of COVID-19 vaccination: a mathematical modelling study. Lancet Infect Dis 2022 Sep;22(9):1293-1302 [FREE Full text] [doi: 10.1016/S1473-3099(22)00320-6] [Medline: 35753318]
34. Pifarré I Arolas H, Acosta E, López-Casasnovas G, Lo A, Nicodemo C, Riffe T, et al. Years of life lost to COVID-19 in 81 countries. Sci Rep 2021 Feb 18;11(1):3504 [FREE Full text] [doi: 10.1038/s41598-021-83040-3] [Medline: 33603008]
35. Baker RE, Mahmud AS, Miller IF, Rajeev M, Rasambainarivo F, Rice BL, et al. Infectious disease in an era of global change. Nat Rev Microbiol 2022 Apr 13;20(4):193-205 [FREE Full text] [doi: 10.1038/s41579-021-00639-z] [Medline: 34646006]
36. Goulbourne T, Yanovitzky I. The communication infrastructure as a social determinant of health: implications for health policymaking and practice. Milbank Q 2021 Mar;99(1):24-40 [FREE Full text] [doi: 10.1111/1468-0009.12496] [Medline: 33528043]
37. Elkin-Koren N. Regulating content moderation for democracy: a transatlantic divide. annual review of law and social science. Annu Rev Law Soc Sci 2025;21:107-123 [FREE Full text] [doi: 10.1146/annurev-lawsocsci-062124-122736]
38. Stephenson M, Olson SM, Self WH, Ginde AA, Mohr NM, Gaglani M, et al. Ascertainment of vaccination status by self-report versus source documentation: impact on measuring COVID-19 vaccine effectiveness. Influenza Other Respir Viruses 2022 Nov;16(6):1101-1111 [FREE Full text] [doi: 10.1111/irv.13023] [Medline: 35818721]
39. Reddy KP, Shebl FM, Foote JH, Harling G, Scott JA, Panella C, et al. Cost-effectiveness of public health strategies for COVID-19 epidemic control in South Africa: a microsimulation modelling study. Lancet Glob Health 2021 Feb;9(2):e120-e129. [doi: 10.1016/s2214-109x(20)30452-6]
40. Bertram MY, Lauer JA, De Joncheere K, Edejer T, Hutubessy R, Kieny M, et al. Cost-effectiveness thresholds: pros and cons. Bull World Health Organ 2016 Dec 01;94(12):925-930 [FREE Full text] [doi: 10.2471/BLT.15.164418] [Medline: 27994285]

Abbreviations

LMIC: low- or middle-income country
UI: uncertainty interval
YLS: year of life saved

Edited by A Mavragani, T Sanchez; submitted 21.Sep.2025; peer-reviewed by M Bennett, T Nighbor; comments to author 07.Nov.2025; revised version received 28.Nov.2025; accepted 15.Dec.2025; published 07.Jan.2026.

Please cite as:
Long MW, Bingenheimer JB, Ndiaye K, Donati D, Rao N, Akaba S, Agha S, Evans WD
A Social Media Campaign to Promote COVID-19 Vaccination: Cost-Effectiveness Analysis
JMIR Public Health Surveill 2026;12:e84540
URL: https://publichealth.jmir.org/2026/1/e84540
doi:10.2196/84540
PMID:

XSL•FO
RenderX

©Michael William Long, Jeffrey B Bingenheimer, Khadidiatou Ndiaye, Dante Donati, Nandan Rao, Selinam Akaba, Sohail Agha, William Douglas Evans. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 07.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Physical Activity Indicators Among Children and Adolescents in Lebanon, Qatar, and the United Arab Emirates: Comparative Synthesis of Active Healthy Kids Report Card Data From 1998 to 2022

Ameneh Baghestani[1,2], MBBS; Lina Majed[3], PhD; Merilyn Lock[3], PhD; Ali Alrahma[4,5], PhD; Patrick Abi Nader[6], MS, MPH, PhD; Suzan Sayegh[7,8], MPH; Abdulla Saeed Al-Mohannadi[9], BSc, MPH; Javaid Nauman[5,10], PhD; Salomé Aubert[11], BSc, MSc, PhD; Mark S Tremblay[11,12], MSc, PhD; Tom Loney[1], PhD

[1]College of Medicine, Mohammed Bin Rashid University of Medicine and Health Sciences, Dubai Health, Dubai, United Arab Emirates

[10]Department of Circulation and Medical Imaging, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Trondheim, Trøndelag, Norway

[11]Active Healthy Kids Global Alliance, Ottawa, ON, Canada

[12]Children's Hospital of Eastern Ontario Research Institute, Ottawa, ON, Canada

[2]Department of Sports and Computer Science, Universidad Pablo de Olavide, Seville, Andalusia, Spain

[3]College of Health and Life Sciences, Hamad Bin Khalifa University, Doha, Qatar

[4]School of Health and Environmental Studies, Hamdan Bin Mohammed Smart University, Dubai, United Arab Emirates

[5]Institute of Public Health, College of Medicine and Health Sciences, United Arab Emirates University, Al Ain, United Arab Emirates

[6]Faculté des sciences de l'activité physique, Université de Sherbrooke, Sherbrooke, QC, Canada

[7]Modern University for Business and Science, Beirut, Lebanon

[8]Shift Wellness Institute, Vancouver, WA, United States

[9]World Innovation Summit for Health (WISH), Qatar Foundation, Doha, Qatar

Corresponding Author:
Tom Loney, PhD
College of Medicine, Mohammed Bin Rashid University of Medicine and Health Sciences, Dubai Health, Dubai, United Arab Emirates

Abstract

Background: Physical inactivity and sedentary behavior are important modifiable risk factors for noncommunicable diseases. High prevalences of physical inactivity among children and adolescents continue to represent a significant public health challenge globally, with approximately two-thirds of children worldwide not achieving the recommended daily amount of physical activity (PA). Countries in the Middle East exhibit some of the highest levels of physical inactivity and sedentary behavior, which contribute to the increasing rates of obesity among children and adolescents.

Objective: This study aims to provide a comparative synthesis of PA indicators among children and adolescents in Lebanon, Qatar, and the United Arab Emirates (UAE) based on the Active Healthy Kids Global Alliance (AHKGA) PA Report Cards and further compare the findings with regional and global trends.

Methods: Data were synthesized from previous AHKGA PA Report Cards published by researchers from Lebanon, Qatar, the UAE in the years 2016 (Global Matrix 2.0; 1998 - 2014 data), 2018 (Global Matrix 3.0; 2016 - 2017 data) and 2022 (Global Matrix 4.0; 2017 - 2022 data). We evaluated 10 key PA indicators across these countries to identify trends and gaps in PA levels among children and adolescents. These findings were further compared with regional and global data gathered and published in previous iterations of the AHKGA Global Matrix.

Results: Based on data collected between 1998 and 2022, less than one-third (15% - 33%) of children and adolescents in Lebanon, Qatar, and the UAE achieved the recommended daily average of 60 minutes of moderate- to vigorous-intensity PA. Additionally, more than one-half (45% - 74%) of children and adolescents exceeded the recommended limit of 2 hours of recreational screen time per day. Overall, boys were more physically active than girls; however, PA levels declined with increasing age. Other behavioral indicators such as participation in organized sports and active transportation revealed insufficient PA levels. The results were slightly better for sources of influence indicators especially with the opportunities provided by schools and governments. Compared with global estimates, PA levels in the Middle Eastern countries were similar to the averages observed

XSL•FO
RenderX

across Asian countries participating in the AHKGA; however, they were generally lower than PA levels in other regions of the world.

Conclusions: Data from a 25-year period show consistently low levels of PA and high levels of sedentary behavior among children and adolescents from these 3 Arab Middle Eastern countries. Despite governmental investments in implementation of PA initiatives, there seems to be a lag in eliciting increases in PA at the population level. Evidence points to a critical need for behavioral and lifestyle modifications among children and adolescents. These concerns are exacerbated by a lack of national surveillance systems and evidence-based policy interventions to improve PA levels.

(JMIR Public Health Surveill 2026;12:e85998) doi:10.2196/85998

KEYWORDS

physical activity surveillance; physical education; sedentary lifestyle; sports; children and adolescents; Middle East

Introduction

Background and Rationale

Physical activity (PA) is an important factor for preventing and managing noncommunicable diseases, as well as enhancing mental health, quality of life, and overall well-being [1]. The high prevalence of physical inactivity among children and adolescents continues to represent a global health challenge, with approximately two-thirds of children worldwide failing to achieve the minimum recommended level of daily PA (ie, at least an average of 60 minutes of moderate- to vigorous-intensity PA [MVPA]) [2]. Moreover, countries in the Middle East report some of the highest levels of physical inactivity and sedentary behavior [3,4]. This trend is further reflected in the increasing rates of obesity among children, adolescents, and adults in the region [5]. Given that physical inactivity is an independent risk factor for various chronic diseases, PA surveillance and intervention should be a top priority for public health agencies, policymakers, and governments [6,7].

In 2005, a team of Canadian researchers developed a Report Card that provided a comprehensive assessment of PA among Canadian children and adolescents, with the aim of identifying areas of concern and advocating for action to improve their PA levels [8,9]. The Report Card included indicators related to PA in different domains, and each indicator was assigned a letter grade (A to F) based on the best available evidence, with “A” representing the highest level of achievement and “F” representing the lowest. For 20 years, Active Healthy Kids Canada and ParticipACTION have worked to inspire the country to engage all children and adolescents in PA, and the Canadian Report Card was released annually between 2005 and 2016 and periodically thereafter.

Building on the success of the Canadian model, the Active Healthy Kids Report Card initiative expanded globally to an international project encompassing 15 countries producing PA Report Cards as part of the Global Matrix 1.0 in 2014 [10]. Since then, the project has continued to grow and include more countries from all regions (Global Matrix 2.0, 2016, 38 countries [11]; Global Matrix 3.0, 2018, 49 countries [12]; Global Matrix 4.0, 2022, 57 countries [13]). The Active Healthy Kids Global Alliance (AHKGA) leads the Global Matrix initiative and has made significant efforts to gather evidence-based data on PA indicators among children and adolescents worldwide [14,15]. As a result, the individual country Report Cards have formed a source of consolidated research knowledge on PA, with contributions to academia, research, education, health, policy, and related fields [16,17]. The primary goal of the AHKGA “Global Matrix” initiative is to raise awareness, develop capacity, and exchange knowledge and strategies, as well as expand PA surveillance of children and adolescents aged 5 years through 17 years across different jurisdictions, countries, and continents. The initiative follows a systematic assessment process; it involves reviewing key indicators and influencing factors using a predefined methodology and grading rubric. This process, steered by the AHKGA, is conducted periodically (every 2 years to 4 years). The AHKGA Global Matrix is currently the only worldwide initiative that comprehensively assesses PA among children and adolescents, covering a diverse range of indicators. This study presents the first comparative synthesis across participating Arab Middle Eastern AHKGA countries.

Data gathered from various participating countries over the past decade (ie, Global Matrix 2.0, 3.0, and 4.0) have consistently indicated low levels of PA across most domains [18]. Without significant changes in behaviors and lifestyles, current and future generations may face increased risks of noncommunicable diseases as they transition into adulthood [3]. In Asia, the number of countries taking part in the AHKGA initiative has increased over the years (9 in Global Matrix 2.0, 12 in Global Matrix 3.0, and 16 in Global Matrix 4.0), demonstrating the region’s enthusiasm for the promotion of PA. However, the overall PA levels have been relatively low throughout the years in most Asian countries [19,20]. As such, gaining additional insights on the region’s PA landscape is important for shaping targeted interventions and informing policy decisions to enhance PA levels and outcomes among children and adolescents. Cross-country analysis of data gathered by the AHKGA can help identify trends and influencing factors specific to the region, as well as reveal gaps in existing surveillance systems [21]. Within the Middle East region, participation in the AHKGA initiative has so far been limited to a few countries, with Lebanon, Qatar, and the United Arab Emirates (UAE) being the only Arab nations that have independently developed Report Cards.

Objectives

This study aimed to synthesize and compare evidence from the PA Report Cards developed by these Arab Middle Eastern countries. Our study specifically focused on Arab nations, with the objective of evaluating trends and identifying gaps in PA

XSL•FO
RenderX

indicators among children and adolescents in countries with similar cultural backgrounds but different socioeconomic and political climates. It also examined the relationships between socioeconomic and geographic factors and the levels of PA, thereby enabling a comparison of findings with those of other regions and global averages.

Methods

Study Design

We used a harmonized comparative approach to evaluate PA indicators among children and adolescents across 3 Arab countries in the Middle East region (Lebanon, Qatar, UAE) using previous PA Report Card outcomes published as part of the AHKGA Global Matrix initiative. We further compared these results with regional and global findings.

Variables

We evaluated 10 key indicators of PA related to children's and adolescents' behaviors, fitness, and sources of influence in this study. The indicators, as defined by the AHKGA, included (1) Overall Physical Activity, (2) Organized Sport and Physical Activity, (3) Active Play, (4) Active Transportation, (5) Physical Fitness (added in Global Matrix 3.0 onward), (6) Sedentary Behavior, (7) Family and Peers, (8) School, (9) Community and Environment, and (10) Government.

Data Sources

The AHKGA Report Card archives were checked for all Arab Middle Eastern countries that were involved in any Global Matrix initiative (1.0 to 4.0). Relevant publications outlining the results of each of the published Report Cards were identified by collaborating authors who were involved in the original projects. This study included data from the Report Cards published in Global Matrix 2.0 2016 (data from 1998 - 2014) [22,23], Global Matrix 3.0 2018 (data from 2016 - 2017) [24-26], and Global Matrix 4.0 2022 (data from 2017 - 2022) [27,28] by Lebanon, Qatar, and the UAE. Participation among these countries varied across the years; the UAE participated in 3 iterations of the Global Matrix (2.0 to 4.0), Qatar participated in Global Matrix 2.0 and 3.0, and Lebanon participated in Global Matrix 3.0 and 4.0. A total of 7 PA Report Cards were identified, and the results were evaluated in our study.

During the development of the Report Cards for each respective iteration (2016, 2018, 2022), teams from each country systematically gathered the most valid and reliable evidence and relevant data from the given time periods (ie, Global Matrix 2.0 data from pre-2016, Global Matrix 3.0 data from 2016 - 2017, and Global Matrix 4.0 data from 2017 - 2022). The countries' Report Card teams included researchers, clinicians, and policymakers from universities, educational authorities, public health departments, and government health authorities from each country. Data sources included local, national, and international studies; national surveys; and official reports. Data were then synthesized by Report Card teams, and a consensus was established regarding the grading of each indicator. The grading was based on specific benchmarks, with the assigned grades corresponding to the proportion of children who achieved the benchmark (Table 1). Grading in the 2016 Report Cards was as follows: A=81% to 100%; B=61% to 80%; C=41% to 60%; D=21% to 40%; F=0% to 20%; INC =incomplete data. However, from 2018 onward, each grade band was further subdivided (Table 2). Report Card grades for each country, including rationale and references, were audited and reviewed by the AHKGA prior to final approval and release.

XSL•FO
RenderX

Table . Indicators and benchmarks used to guide the Report Card grading for Global Matrix 2.0 in 2016 (data from 1998 - 2014), Global Matrix 3.0 in 2018 (data from 2016 - 2017), and Global Matrix 4.0 in 2022 (data from 2017 - 2022).

Indicator	Benchmark
Overall Physical Activity	• % of children and adolescents who meet the Global Recommendations on Physical Activity for Health, which recommend that children and adolescents accumulate at least 60 minutes of moderate- to vigorous-intensity PA[a] per day on average • % of children and adolescents meeting the guidelines on at least 4 days a week (when an average cannot be estimated)
Organized Sport and Physical Activity	• % of children and adolescents who participate in organized sport or PA programs
Active Play	• % of children and adolescents who engage in unstructured or unorganized active play at any intensity for more than 2 hours a day • % of children and adolescents who report being outdoors for more than 2 hours a day
Active Transportation	• % of children and adolescents who use active transportation to get to and from places (eg, school, park, mall, friend's house)
Sedentary Behavior	• % of children and adolescents who meet the Canadian Sedentary Behavior Guidelines for 5- to 17-year-olds, which is no more than 2 hours of recreational screen time per day
Physical Fitness	• Average percentile achieved on certain physical fitness indicators based on the normative values published by Tomkinson et al [29]
Family and Peers	• % of family members (eg, parents, guardians) who facilitate PA and sport opportunities for their children (eg, volunteering, coaching, driving, paying for membership fees and equipment) • % of parents who meet the Global Recommendations on PA for Health, which recommend that adults accumulate at least 150 minutes of moderate-intensity aerobic PA throughout the week or do at least 75 minutes of vigorous-intensity aerobic PA throughout the week or an equivalent combination of moderate- and vigorous-intensity PA • % of family members (eg, parents, guardians) who are physically active with their kids • % of children and adolescents with friends and peers who encourage and support them to be physically active • % of children and adolescents who encourage and support their friends and peers to be physically active
School	• % of schools with active school policies (eg, daily PE[b], daily PA, recess, "everyone plays" approach, bike racks at school, traffic calming on school property, outdoor time) • % of schools where the majority (≥80%) of students are taught by a PE specialist • % of schools where the majority (≥80%) of students are offered the mandated amount of PE (for the given state, territory, region, or country) • % of schools that offer PA opportunities (excluding PE) to the majority (>80%) of their students • % of parents who report their children and adolescents have access to PA opportunities at school in addition to PE classes • % of schools with students who have regular access to facilities and equipment that support PA (eg, gymnasium, outdoor playgrounds, sporting fields, multipurpose space for PA, equipment in good condition)

XSL•FO
RenderX

Indicator	Benchmark
Community and Environment	• % of children or parents who perceive their community or municipality is doing a good job at promoting PA (eg, variety, location, cost, quality). • % of communities or municipalities that report they have policies promoting PA • % of communities or municipalities that report they have infrastructure (eg, sidewalks, trails, paths, bike lanes) specifically geared toward promoting PA • % of children or parents who report having facilities, programs, parks, and playgrounds available to them in their community • % of children or parents who report living in a safe neighborhood where they can be physically active • % of children or parents who report having well-maintained facilities, parks, and playgrounds in their community that are safe to use
Government	• Evidence of leadership and commitment in providing PA opportunities for all children and adolescents • Allocated funds and resources for the implementation of PA promotion strategies and initiatives for all children and adolescents • Demonstrated progress through the key stages of public policy making (ie, policy agenda, policy formation, policy implementation, policy evaluation, and decisions about the future)

[a]PA: physical activity.

[b]PE: physical education.

Table . Report Card grading rubric used in the Global Matrix 2.0[a] in 2016 (data from 1998 - 2014), Global Matrix 3.0 in 2018 (data from 2016 - 2017), and Global Matrix 4.0 in 2022 (data from 2017 - 2022).

Grade	Interpretation
A+	94% - 100%
A	87% - 93% (We are succeeding with a large majority of children and adolescents.)
A–	80% - 86%
B+	74% - 79%
B	67% - 73% (We are succeeding with well over half of children and adolescents.)
B–	60% - 66%
C+	54% - 59%
C	47% - 53% (We are succeeding with about half of children and adolescents.)
C–	40% - 46%
D+	34% - 39%
D	27% - 33% (We are succeeding with less than half but some children and adolescents.)
D–	20% - 26%
F	<20% (We are succeeding with very few children and adolescents.)
INC[b]	Incomplete—insufficient or inadequate information to assign a grade

[a]Grading in the 2016 Report Cards was as follows: A=81%-100%; B=61%-80%; C=41%-60%; D=21%-40%; F=0%-20%.

[b]INC: incomplete data.

In this study, we incorporated references to the published results of the Report Cards from each country over the years [22-28]. We also relied on the interpretations made by the Report Card teams to generate a comprehensive table with the assigned grades and percentages based on the previous publications. In certain instances, we cited additional sources that did not play a direct role in determining the indicator's grade at the time that the Report Card was issued (eg, governmental policies, future plans, and suggested strategies for improvement); however, they remain relevant to the context of the grading. This approach

XSL•FO
RenderX

helped us provide thorough interpretations while ensuring that the grading process was based on information from the published Report Cards. Additionally, the most recent Report Cards (2022) included the top 3 priorities for improving 10 key PA indicators, as identified by Report Card teams in each country. Building on these priorities, we provided key applicable recommendations for the indicators to support the improvement of PA within the 3 Arab Middle Eastern countries.

Sociodemographic characteristics of the countries were obtained from the latest (2023 and 2024) Human Development Reports issued by the United Nations and the World Bank [30,31], including the Human Development Index (HDI), gross domestic product (GDP) per capita, Gender Inequality Index, Gini Index, expenditure on health, life expectancy at birth, mean years of schooling, total population, population density, urban population, and average mean temperature [32]. They were used to assess and interpret grade variations between these countries and in relation to other regions of the world.

Ethical Considerations

This study used secondary data from published papers and did not require institutional review board review as the data were publicly available. Specifically, data were collected from Global Matrix and Report Card papers published in the Journal of Physical Activity and Health [11-13] as well as the AHKGA website [14] for Global Matrix 2.0, 3.0, and 4.0. This article presents a comparative secondary data analysis of the aforementioned publications for the 3 participating countries. As such, no ethical approval was required for conducting the study. The referenced publications included data from local, national, and international studies; national surveys; and official reports. References to these data sources can be found in the relevant original Report Card publications as cited in the Methods section [22-28]. Further information about the process of Report Card development in each participating country can be found in the original Global Report Card publications [11-13].

Results

Overview

To contextualize the findings across these Arab Middle Eastern countries, sociodemographic and economic characteristics were first examined (Table 3). Qatar (0.886) and the UAE (0.940) had very high (≥0.800) HDIs [30,33]; Lebanon (0.752) had a high (0.700 - 0.799) HDI [30,34]. As of the year 2024, the UAE had the largest population among the 3 countries, with approximately 11 million residents, and the lowest population density (142 people/km^2 of land area) and urban population percentage (88%). Public health expenditure as a percentage of GDP was the highest in Lebanon (5.74%) and the lowest in Qatar (2.18%). Life expectancy at birth ranged from 77.8 years in Lebanon to 82.9 years in the UAE. The lowest Gini Index (least income inequality) was observed in the UAE (26.4), and the highest (greatest income inequality) was in Qatar (35.1) [31].

Table . Sociodemographic characteristics of the 3 participating Arab Middle Eastern countries from the most recent available data.

Sociodemographic variable	Country		
	Lebanon	Qatar	UAE[a]
Human Development Index	0.752	0.886	0.940
GDP[b] per capita (US $)	3477.70	127,711.10	49,377.60
Gender Inequality Index	0.036	0.195	0.040
Gini Index	31.8	35.1	26.4
Expenditure on health (% of GDP)	5.74	2.18	4.68
Life expectancy at birth (years), mean	77.8	82.4	82.9
Schooling duration (years), mean	10.4	10.8	13.0
Population (millions), n	5.8	2.9	10.9
Population density (people/km^2)	562	231	142
Urban population, %	90	99	88
Temperature (°C), mean	16.0	28.3	28.9

[a]UAE: United Arab Emirates.

[b]GDP: gross domestic product.

In total, 37 letter grades (37/66, 56%) and 29 incomplete (INC) grades (29/66, 44%) were assigned across all 7 Report Cards (Table 4). The indicator with the highest grade was Government (average of B+), and the indicator with the lowest grade was Overall Physical Activity (average of D/F). Active Play, Physical Fitness, and Community and Environment were marked INC throughout all Report Cards. The grades are explained in the following sections, except for indicators with INC grades, which are not included in the rest of the Results section.

XSL•FO
RenderX

Table . Grades assigned to the 10 common physical activity (PA) indicators for Global Matrix 2.0 in 2016 (data from 1998-2014), Global Matrix 3.0 in 2018 (data from 2016-2017), and Global Matrix 4.0 in 2022 (data from 2017-2022) for Lebanon, Qatar, and the United Arab Emirates (UAE).

Common PA indicators	GM[a] 2.0 (2016)[b],grade (%)		GM 3.0 (2018), grade (%)			GM 4.0 (2022), grade (%)	
	UAE (1998-2014)	Qatar (2004-2014)	UAE(2016)	Qatar(2016)	Lebanon (2016-2017)	UAE (2017-2021)	Lebanon (2018-2022)
Overall Physical Activity	F–/D– (17 - 28)	F (15)	F (16)	D (32)	D (33)	F (19)	D– (21)
Organized Sport and Physical Activity	INC[c]	D (30)	INC	D+ (48)[d]	F (5)	INC	INC
Active Play	INC	INC	INC	INC	INC	INC	INC
Active Transportation	F–/D– (19 - 21)	N/A	INC	N/A	D (26 - 37)[d]	F (5)	D+ (37)
Sedentary Behavior	C– (49 - 62)[d]	D (<30)	C– (40)	D+ (35 - 45)[d]	C– (46)	D– (26)	C (55)[d]
Physical Fitness	__[e]	—	INC	INC	INC	INC	INC
School	D (28)	INC	D– (26)	C (61 - 67)[d]	D (22 - 58)[d]	A– (100)[d]	D (32)
Family and Peers	C– (62)[d]	D	INC	INC	INC	D– (20 - 26)	INC
Community and Environment	INC	INC	INC	INC	INC	INC	INC
Government	B+	B	B+	B+	C+	B+	D

[a]GM: Global Matrix.

[b]Minus sign in the 2016 Report Card grading scheme indicates variations in subgroups based on age, gender, nationality, socioeconomic status, and geographic location.

[c]INC: incomplete.

[d]Percentages may not correspond to final grades, as final grades were determined by expert consensus.

[e]Not applicable: The Physical Fitness indicator was added in Global Matrix 3.0 onward.

Additionally, during each iteration of the Global Matrix (2016, 2018, 2022), participating countries had the choice to include supplementary indicators in their respective Report Cards, as deemed beneficial to enhance the value of their results. Due to this, some variations in the indicators were observed across the different countries and Report Cards. For example, Qatar's Report Card team decided not to evaluate Active Transportation and considered it "Not Applicable" due to road safety issues and the prevalence of high temperatures during much of the year. Instead, they included Obesity and Dietary Habits as additional indicators in their 2016 Report Card. Likewise, Lebanon included Weight Status in addition to the common indicators in their 2018 Report Card and Sleep and Body Mass Index in their 2022 Report Card.

Overall Physical Activity

Across all 3 countries, the proportions of children and adolescents meeting the recommended PA levels (ie, minimum of 60 minutes of MVPA per day on average) varied from 15% to 33%, with the average grades ranging from F to D. In the UAE, data from 1998 to 2021 showed that most children and adolescents (72% - 84%) were physically inactive. The first UAE Report Card (2016), which was based on the largest amount of data collected (1998 - 2014), revealed 17% to 28% of children and adolescents achieved the recommended daily PA, with a noticeable difference between boys (23% - 24%) and girls (13% - 16%). The World Health Organization (WHO) Global School-based Health Survey (GSHS) [35] and other studies with smaller sample sizes were used to inform this indicator, resulting in a split grade of D−/F−, with the minus sign indicating variations in subgroups. Overall Physical Activity remained low in subsequent UAE Report Cards, with only 16% of children and adolescents (21% of boys; 11% of girls) in 2018 and 19% of children and adolescents (27% of boys; 8% of girls) in 2022 accumulating sufficient PA. Likewise, in Qatar, only 15% of children and adolescents (30% of boys; 10% of girls) met the recommended daily PA as reported in their first and most comprehensive Report Card issued in 2016. There was a notable increase in the following Qatar Report Card (2018), with PA levels ranging between 25% and 39% among children and adolescents. The final average grade assigned to Overall Physical Activity in the 2018 Report Card by Qatar was D (32%). The Lebanese Report Cards for 2018 and 2022 reported comparable results, with average grades of D (33%) and D− (21%) assigned in those respective years. Like Qatar and the UAE, their evaluation primarily relied on the WHO GSHS data using self-report PA estimates. Boys had higher levels of PA (29%) than girls (14%), and students attending private schools (23%) appeared to be more active than students in public schools (17%) in Lebanon. In all 3 countries, the Overall Physical Activity estimates were mostly derived from

XSL•FO
RenderX

self-reported questionnaires, while device-based (accelerometer) assessments were limited.

Organized Sport and Physical Activity

In Lebanon, the proportion of children and adolescents participating in events or competitions administered by the Ministry of Youth and Sports was extremely low, leading to an F grade (5%) in 2018. This may be due to a lack of public sport infrastructure and investment in community-level sport programs. Their subsequent Report Card (2022) graded this indicator INC. In Qatar and the UAE, governmental support and investment in infrastructure were and continue to be higher than in Lebanon. In 2016, Qatar's GSHS national survey found that approximately 30% of children and adolescents regularly participated in at least one organized sport annually. Qatar also held a National Sports Day and had several initiatives aimed at promoting sport participation, which contributed to a grade D in this area. This grade improved to D+ in 2018, when nearly 48% of children and adolescents participated in organized sports. Additionally, 58% of boys and 42% of girls in Qatar were registered with sport clubs. The UAE marked this indicator as INC throughout all Report Cards.

Active Transportation

Of the 3 included countries, Lebanon reported the highest active transportation use (26% - 37%; grade D). The proportion remained similar (37%), but the grade increased from D (2018) to D+ in 2022. Additionally, sex-based differences were observed (42% of boys and 32% of girls) in the 2017 GSHS national survey results. The UAE's evaluation of Active Transportation in 2016 was derived from 2005 and 2010 GSHS national reports, when data showed that 19% to 21% of adolescents walked or cycled to school at least once during the week (grade F−/D−). The minus sign added to the grade in the 2016 Report Card was due to sex-based differences (29% of boys and 16% of girls in 2005; 29% of boys and 15% of girls in 2010). No data were available for the following (2018) UAE Report Card; however, the 2022 UAE Report Card used data from an online survey run by the Knowledge and Human Development Authority in Dubai. The survey of 33,784 parents indicated that only 5% of children and adolescents in Dubai used active transportation modes to travel to and from school (4.8% walked; 0.2% cycled) [27]. Based on these findings, the indicator received an F in 2022. Qatar did not include Active Transportation in their Report Cards, as they considered it to be not applicable for assessment in their country due to road safety issues and the predominantly hot climate.

Sedentary Behavior

All 3 countries have faced ongoing challenges in reducing sedentary behavior. The UAE Report Card showed 62% (2005) and 49% (2010) of children achieving the recommended Canadian Sedentary Behavior Guidelines of screen time exposure (ie, no more than 2 hours of recreational screen time per day), with variations according to age and sex. The indicator was assigned a C− in the UAE in the first two Report Cards (2016 and 2018), with the minus sign indicating variations in subgroups in 2016. The numbers declined over time to 40% in 2018 (C–) and only 26% in 2022 (D−). In Qatar, the 2016 Report Card showed that less than one-third of children and adolescents met the recommended daily recreational screen time, which corresponded to a D grade. These figures improved slightly to 35% for children and 45% for adolescents (D+) in 2018. Qatar did not participate in the 2022 Report Card; therefore, updated data were unavailable. The 2018 Lebanese Report Card showed that approximately 40% of children and 52% of adolescents met the screen time recommendations, resulting in an average grade of C− (46%). Similar findings were reported in 2022, with 55% of adolescents meeting the recommendations.

School

In the UAE, due to limited access to national data on school distribution, the benchmark used for this indicator varied between (1) the proportion of children and adolescents attending ≥3 physical education (PE) classes per week (approximately 150 minutes) and (2) the proportion of schools where the majority (≥80%) of students were taught by PE specialists. In 2016 and 2018, the UAE reported that slightly more than one-quarter (26% - 28%) of adolescents adhered to the mandated PE requirements and that participation decreased as age increased. The 2022 Report Card reflected updated figures, with an A− grade; however, this grade was based solely on data confirming that all children and adolescents in Dubai received PE classes taught by specialists. The minus sign in the 2022 Report Card was added to acknowledge the difference in the frequency of weekly PE classes across various schools. Although several components of the UAE schools (such as all PE teachers holding bachelor's degrees in PE and the adequate availability of PA facilities) would typically warrant a high grade for this indicator, only 79% of children reported receiving more than 100 minutes of PE classes per week. In Qatar, School was marked INC in 2016 due to insufficient data. In 2018, it was graded C, indicating that 67% of children participated in PE classes and 61% of schools offered support for PA programs or initiatives such as "Qatar Active Schools" and "Health Promoting Schools." It is worth noting here that, at the time, many PE teachers in Qatar were not classified as specialists. Therefore, schools in these circumstances would not have been counted, even if their PE classes were delivered adequately. In Lebanon, School was graded D in 2018 (22%) and 2022 (32%). In both instances, the results were based on self-reported participation of school students in the number of weekly PE classes. Additionally, there was a notable sex discrepancy observed, with 28% of boys participating in PE classes compared with only 18% of girls.

Family and Peers

This indicator relied on data showing the influence of family members and peers in encouraging or limiting PA opportunities and participation among children and adolescents. Based on parental PA levels in the UAE, the Family and Peers indicator was assigned C− (62%) in 2016 and D− (20% - 26%) in 2022. In the 2016 Qatar Report Card, it was graded D based on the overall prevalence of physical inactivity, suboptimal dietary habits, and a high rate of obesity among the population. Lebanon did not have any data to inform this indicator; hence, it was marked INC.

XSL•FO
RenderX

Government

This indicator was evaluated based on any evidence of governmental support in promoting and providing PA opportunities for children and adolescents through policy, legislation, or regulation. All 3 countries relied mainly on unpublished data to grade this indicator. Qatar assigned its highest grade to the Government indicator in both 2016 (B) and 2018 (B+). Similarly, in the UAE, the indicator was graded B+ across all 3 Report Cards. However, Lebanon's grades were C+ in 2018 and D in 2022.

Discussion

Summary of Main Findings

This study addressed an important research gap by offering the first regional comparative synthesis of PA indicators in 3 Arab Middle Eastern countries with shared cultural and climatic contexts. It provides insights into a critical public health concern based on 25 years of data (1998 - 2022). The findings revealed that less than one-third (15% - 33%) of children and adolescents participated in the recommended daily amount of PA (ie, an average of 60 minutes of moderate- to vigorous-intensity PA). Moreover, a large proportion (45% - 74%) exhibited high levels of sedentary behavior (ie, ≥2 hours of recreational screen time per day). Qatar and the UAE had comparable trends due to their similar demographics, geographic location, climate, and cultural characteristics. Compared with global estimates, PA levels in the 3 Arab Middle Eastern countries were close to those observed across other Asian countries participating in the AHKGA; however, they were generally lower than PA levels in other regions of the world.

Interpretations, Implications, and Comparisons With Existing Literature

Human Development Index

Countries worldwide are classified by their level of development using the HDI, which measures a country's success in 3 key dimensions: life expectancy, education levels, and the standard of living as measured by income per capita [30]. Global analysis of PA levels by HDI showed that countries with high and very high HDI generally had lower grades in behavioral indicators (ie, Overall Physical Activity, Organized Sports, Active Play, Active Transportation, and Sedentary Behavior) but higher grades in sources of influence indicators (ie, School, Family and Peers, Community and Environment, and Government) [36]. This pattern of lower PA and higher HDI was evident in most indicators (except for the Government indicator) across the 3 Middle Eastern countries included in our study.

On a global scale, the overall PA grades for countries with very high (≥0.800) HDI approximated a C, with the behavioral indicators near D and the source of influence indicators around B [33]. The UAE and Qatar grades were not far from the global averages of PA among other very high HDI countries participating in the Global Matrix initiative. Similarly, Lebanon's grades were consistent with other participating countries classified as having a high HDI (eg, Thailand, China, Bulgaria, and several nations in South America), with most results predominantly around a grade of D [34].

Behavioral Indicators

The latest global insights from the Global Matrix 4.0 (2022) revealed an Overall Physical Activity average of D (27% - 33%) based on data gathered from 57 countries [13]. A study evaluating further temporal trends of Asian countries from Global Matrices 2.0 - 4.0 showed consistently lower grades for Overall Physical Activity, with an average grade of D− (20% - 26%) across the years [20]. Countries in the Middle East have had similar and sometimes lower results (F to D) than regional (Asian) averages; however, the grades vary across the specific indicators.

The assigned grades for Sedentary Behavior across the 3 participating Arab Middle Eastern countries ranged between D− and C−, which is in line with the global findings observed over the past decade [37]. This was expected given the advancement in technologies and the widespread use of digital media. Moreover, the Organized Sport and Physical Activity indicator was marked INC in nearly all Asian countries (including Lebanon and the UAE) in Global Matrix 4.0 [13]. Considering its role in the overall health and PA of children and adolescents, it is important to incorporate organized activities into PA questionnaires with further specification of the type, intensity, frequency, and duration of each activity.

Active Play, known for its numerous health benefits in children and adolescents, remains one of the least researched and documented areas of PA [38]. This indicator is a critical component of children's and adolescents' physical, social, and cognitive well-being; however, it often goes unrecorded due to its unstructured nature [39]. In the most recent global reports (Global Matrix 4.0, 2022), Active Play was the indicator with the most INC grades (27/57, 47% INC; 30/57, 53% graded) [13]. Moreover, the average grades differed across countries participating in the Global Matrix in relation to the HDI; low-medium HDI countries scored around a C+, while high HDI countries scored around a D− [13]. This highlights the impact of the modern world regarding replacing unstructured PA with sedentary behavior. Country-specific barriers and enablers to assessing this indicator are worth exploring to better understand children's total PA and design interventions that promote healthy, playful, and active lifestyles.

Low levels of Active Transportation in the 3 Middle Eastern countries may be partially explained by the transport infrastructure and urban design. None of the 3 countries have sufficient walking or cycling lanes dedicated to traveling long distances between residential communities and schools. Most cycling and pedestrian tracks are primarily designed for recreational use, rather than serving as active transportation routes. The region's hot climate also adds to the barriers of outdoor PA in general, especially in Qatar and the UAE where temperatures can peak at 40.0 °C to 45.0 °C [32]. In Lebanon, the weather throughout the academic year is generally moderate (mean temperature 16.0 °C) [32], facilitating better conditions for active transportation. Additionally, the noticeable difference in Active Transportation between boys and girls may be due to cultural factors, such as the acceptability of female adolescents

XSL•FO
RenderX

walking and cycling to and from school or other places. A study conducted in the Middle East and North Africa region identified a number of barriers and facilitators to PA among different subgroups of the population [40]. The most common barriers among children and adolescents were being female, older age, lack of time, and obesity. Other less common barriers included high screen time, inadequate school PE programs, lack of sports facilities, peer support, sleep, and dietary habits [40].

These findings highlight the relationship between different behavioral indicators and their joint effect on achieving optimal daily PA. Therefore, in order to improve PA levels among children and adolescents (girls and boys), the first step would be to develop and implement a systematic tool for the assessment of PA across different domains [41-43].

Physical Fitness Indicator

The dearth of data on physical fitness levels presents a critical public health concern, as it is an indicator that offers important insights into the overall health of children and adolescents and reveals early risk factors for chronic diseases that can develop later in life [44,45]. Routine evaluation and monitoring of physical fitness can also serve as a baseline national indicator of the efficacy of PA policies, educational interventions, and environmental support for promoting active lifestyles among children and adolescents. Since its introduction in the Global Matrix 3.0 (2018), Physical Fitness has been reported as INC in over one-half of the participating countries. Most grades for this indicator were assigned to countries classified as having high or very high HDIs, with a global average of C− [12,13].

Within the Middle East, Qatar has exhibited preliminary progress, where the Ministry of Education and Higher Education initiated a “Physical Fitness Evaluation for Public School Students” and published their report for the 2019‑2020 academic year online [46]. The results included 104,191 students (50% Qatari nationals) who were evaluated for cardiorespiratory fitness, strength endurance, flexibility, speed, and consistency. Based on these findings, normative values were established for each fitness component. Qatar did not participate in the latest Global Matrix; therefore, these data were not included previous Report Cards and will likely be included in future ones.

Sources of Influence Indicators

Globally, countries with very high HDIs typically scored the highest in the School indicator [36]. The most common grades assigned to this indicator among the 3 Arab Middle Eastern countries were D− and D, which are lower than the Asian and global averages (C+). It is worth noting that the School grades were partially affected by the variations in benchmarks used for grading throughout the Global Matrix iterations. Hence, it is difficult to directly compare this indicator across countries, regionally and globally. Having said that, many countries in Asia are still not performing well with school-related PA [19,20]. Although PE is mandated in most school curricula, the duration or frequency of classes does not always provide adequate amounts of PA to meet international guidelines. Additionally, the heavy focus on academic grades and achievements often causes students to neglect PE and overlook the importance of sports [20]. Given the benefits of PA to the overall well-being and academic performance of children and adolescents, it is essential for schools to place greater importance on PE classes and other sports within their curricula.

The influence of Families and Peers is multidimensional and demands extensive research for better understanding and assessment across regions and cultures [47]. Globally, this indicator had great variations (from INC to F to A+), with Asian countries globally scoring an average of C− [13,20]. Despite the recognized importance of social influences, the available data on this indicator remain limited and mostly incomplete in the 3 Arab Middle Eastern countries, mainly due to a lack of valid assessment tools. Community and Environment is another indicator that strongly influences children’s and adolescents’ engagement in PA and is assessed based on various factors such as local policies, infrastructure, the perception of safety, and access to green and blue spaces. Globally, the grades for this indicator varied between very high HDI countries (B) versus low, medium, and high HDI countries (C−) [36]. Further variations were observed across different regions such as Africa (D+), Asia-Pacific (C), Europe (B−), and the Anglosphere (B−) [13]. In Lebanon, Qatar, and the UAE, this indicator was marked INC across all Report Cards. This is due to a lack of consistent monitoring in these countries, which highlights a critical gap in understanding how the built and social environments support or hinder active lifestyles. To standardize the evaluation of these indicators, it is necessary to conduct periodic public household surveys aimed at assessing social and environmental factors related to PA. Addressing this gap is essential for informing urban planning and public health interventions and policies [48].

Finally, the government plays an important role in facilitating PA opportunities. In the UAE, the data suggested strong governmental support and leadership in promoting PA through national strategies, policies, and local initiatives. The “National Sports Strategy 2031” [49] aims to increase sports participation to 71% of the UAE population by enhancing access to both community-level and competitive sports. Further evidence included the “National Policy for Promoting Healthy Lifestyles” [50], “National Strategy for Wellbeing 2031” [51], and “Masar Project” [52]. Likewise, Qatar’s highest grades were assigned to the Government indicator in both Report Cards, reflecting the country’s commitment to supporting PA in the population. Some of Qatar’s strategies that prioritized PA for children and adolescents included the “Qatar National Sports Day” [53], “Qatar National Physical Activity Guidelines” [54], and “National Health Strategy 2018-2022” [55]. Additionally, the WHO-Qatar partnership (2021‑2024) on the FIFA (Fédération Internationale de Football Association) World Cup and health legacy played a significant role in advancing PA through global awareness campaigns [56]. Lebanon’s results for the Government indicator reflected moderate efforts from both governmental and nongovernmental stakeholders, based on data from the 2018 Report Card [26]. Although existing policies focus on enhancing PA in schools, the implementation and coverage remained limited, and there was no indication of sustainable funding, evaluation, or national legislation mandating these efforts. As a result, Government grades declined in the latest Report Card (2022) for Lebanon. To maximize the benefits of these governmental initiatives and

XSL•FO
RenderX

programs, it is essential for all 3 countries to prioritize public education on the importance of PA across different domains.

Report Card Team Recommendations for Improvement of Physical Activity Levels in the Region

In Global Matrix 3.0 (2018), Report Card members provided their top 3 priorities to improve the 10 common PA indicators in their respective countries. Based on these, we developed recommendations that are relevant to all 3 countries and the broader Middle East region for improving PA levels among children and adolescents.

Overall Physical Activity

We recommend developing and implementing PA surveillance programs using nationally representative samples, with the aim of measuring subjective and objective estimates of PA across different domains.

Organized Sport and Physical Activity

Our recommendations include improving access to affordable, well-structured organized sports within schools and communities; establishing talent identification programs and athlete development pathways that increase mass participation at the grass-roots level; and promoting PA participation among girls and children and adolescents with disabilities or chronic conditions.

Active Play

For this PA domain, we recommend the integration of active play into urban planning by creating safe, accessible, and culturally appropriate play environments and the development of relevant surveillance tools.

Active Transportation

Accelerating the development of safe and accessible infrastructures (eg, walking, running, and cycling path networks) within communities and cities could promote active transportation.

Physical Fitness

We recommend developing and implementing nationwide standardized physical fitness tests and monitoring programs in school curricula, paired with targeted interventions to improve fitness and physical performance levels.

Sedentary Behavior

Integration of regular active breaks in school lessons can interrupt long periods of sedentary behavior, and implementation of awareness campaigns could promote reductions in nonessential screen time.

School

We recommend the development of physical literacy among students and maximizing the time spent in MVPA (minimum of 150 minutes per week during school time across all schools and grades) while offering fun, enjoyable, and skill-developing activities for both boys and girls.

Family and Peers

Establishment of family-centered community PA initiatives and peer-led activity clubs in collaboration with municipalities and sports clubs are recommendations in this area.

Community and Environment

Our recommendations include continued development and maintenance of accessible green and blue active spaces (ie, parks and beaches) for all children and adults to interact socially as well as the development and implementation of related surveillance systems.

Government

Establishment of national multisectoral PA monitoring and accountability frameworks could support the implementation of health policies, PA guidelines, and school-based PA programs.

Strengths and Limitations

To the best of our knowledge, this is the first study to synthesize and compare trends in PA indicators among children and adolescents in all Arab Middle Eastern countries participating in the AHKGA. The results were further compared with regional and global trends, offering valuable insights for policymakers. The study also highlighted major research and surveillance gaps and provided recommendations for the various PA indicators. Despite PA being an emerging area of research within the region, teams of PA researchers in each country retrieved and synthesized multiple data sources to grade several PA indicators over a 25-year period. When evaluating these grades and comparing between countries, we identified differences in data sources (ie, PA assessment was not harmonized across and within countries). This was carefully considered when comparing results and drawing conclusions.

The main limitation of this study was the lack of high-quality data for the majority of PA indicators (7/10, 70%) and the complete absence of data sources for the remaining 3 indicators (ie, Active Play, Physical Fitness, and Community and Environment). Specifically, there was a dearth of studies collecting accelerometer-derived objective estimates of PA using nationally representative samples. Certain indicators (eg, School, Family and Peers) were difficult to compare across the countries due to variations in the benchmarks used for grading. These challenges further point to the gaps in PA research within the region and the critical need for standardized and comprehensive PA surveillance studies.

Conclusions

Children and adolescents in Lebanon, Qatar, and the UAE had low PA levels overall coupled with high levels of sedentary behavior. These results were consistent with regional findings and global trends. Overall, there was a dearth of data from representative samples in all 3 countries. There were significant gaps in specific indicators such as Active Play, Physical Fitness, and Community and Environment. Addressing the current challenges requires a thoughtful approach that considers the unique context of each country and promotes equitable access to opportunities for active living across the lifespan. There is a clear need for a standardized and comprehensive PA surveillance

XSL•FO
RenderX

strategy that assesses both self-report and objective estimates of PA among children and their parents. School remains an important opportunistic arena for increasing PA levels. Enhancing the quantity and quality of PE programs and extracurricular sports would maximize the proportion of children and adolescents receiving the health benefits of regular PA. Stakeholder collaboration is essential for developing evidence-informed, locally relevant plans that address the current implementation gap between policy and practice. Specifically, this requires educational authorities, municipalities, environmental agencies, and sports councils to implement a harmonized approach to providing PA opportunities for all children and adolescents.

Acknowledgments

The authors would like to acknowledge the Active Healthy Kids Global Alliance (AHKGA) for organizing and leading the international initiative, developing the Physical Activity Report Card methodology, and standardizing the benchmarks and grading rubrics. We are also thankful to Report Card team members from Lebanon, Qatar and the United Arab Emirates (UAE) for all their dedicated hard work.

Responsibility for the final manuscript lies entirely with the authors. Generative artificial intelligence tools are not listed as authors and do not bear responsibility for the final outcomes. Declaration submitted by all authors of the manuscript (AB, LM, MJL, AMA, PAN, SS, ASA, JN, SA, MST, TL).

Funding

No external financial support nor grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this study. All authors are salaried staff of their respective institutions. The authors would like to thank Mohammed Bin Rashid University of Medicine and Health Sciences for payment of the article processing fee.

Data Availability

All data reported and analyzed in this study were obtained from published sources and publicly available datasets as cited in the reference list.

Authors' Contributions

Conceptualization: AB, LM, ML, AA, PAN, SS, ASA, JN, SA, MST, TL
Data curation: AB, LM, ML, AA, PAN, SS, ASA, JN, SA, MST, TL
Formal analysis: AB, LM, ML, AA, PAN, SS, ASA, JN, SA, MST, TL
Methodology: AB, LM, ML, AA, PAN, SS, ASA, JN, SA, MST, TL
Supervision: TL
Validation: AB, LM, ML, TL
Visualization: AB, TL
Writing – original draft: AB, TL
Writing – review & editing: AB, LM, ML, AA, PAN, SS, ASA, JN, SA, MST, TL

Conflicts of Interest

None declared.

References

1. Mahindru A, Patil P, Agrawal V. Role of physical activity on mental health and well-being: a review. Cureus 2023 Jan;15(1):e33475. [doi: 10.7759/cureus.33475] [Medline: 36756008]
2. Global status report on physical activity 2022. World Health Organization. URL: https://www.who.int/publications/i/item/9789240059153 [accessed 2025-08-12]
3. Rahim HFA, Sibai A, Khader Y, et al. Non-communicable diseases in the Arab world. Lancet 2014 Jan 25;383(9914):356-367. [doi: 10.1016/S0140-6736(13)62383-1] [Medline: 24452044]
4. Bull F, Dvorak J. Tackling chronic disease through increased physical activity in the Arab world and the Middle East: challenge and opportunity. Br J Sports Med 2013 Jul;47(10):600-602. [doi: 10.1136/bjsports-2012-092109] [Medline: 23624465]
5. World health statistics 2025: monitoring health for the SDGs, sustainable development goals. World Health Organization. URL: https://www.who.int/publications/i/item/9789240110496 [accessed 2025-08-15]
6. Lee IM, Shiroma EJ, Lobelo F, et al. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet 2012 Jul 21;380(9838):219-229. [doi: 10.1016/S0140-6736(12)61031-9] [Medline: 22818936]

XSL•FO
RenderX

7. Ding D, Lawson KD, Kolbe-Alexander TL, et al. The economic burden of physical inactivity: a global analysis of major non-communicable diseases. Lancet 2016 Sep 24;388(10051):1311-1324. [doi: 10.1016/S0140-6736(16)30383-X] [Medline: 27475266]
8. Barnes JD, Tremblay MS. Changes in indicators of child and youth physical activity in Canada, 2005-2016. Can J Public Health 2017 Mar 1;107(6):e586-e589. [doi: 10.17269/cjph.107.5645] [Medline: 28252380]
9. Tremblay MS. Major initiatives related to childhood obesity and physical inactivity in Canada: the year in review. Can J Public Health 2012;103(3):164-169. [doi: 10.1007/BF03403806] [Medline: 22905632]
10. Tremblay MS, Gray CE, Akinroye K, et al. Physical activity of children: a global matrix of grades comparing 15 countries. J Phys Act Health 2014 May;11 Suppl 1:S113-S125. [doi: 10.1123/jpah.2014-0177] [Medline: 25426906]
11. Tremblay MS, Barnes JD, González SA, et al. Global Matrix 2.0: Report Card grades on the physical activity of children and youth comparing 38 countries. J Phys Act Health 2016 Nov;13(11 Suppl 2):S343-S366. [doi: 10.1123/jpah.2016-0594] [Medline: 27848745]
12. Aubert S, Barnes JD, Abdeta C, et al. Global Matrix 3.0 Physical Activity Report Card grades for children and youth: results and analysis from 49 countries. J Phys Act Health 2018 Nov 1;15(S2):S251-S273. [doi: 10.1123/jpah.2018-0472] [Medline: 30475137]
13. Aubert S, Barnes JD, Demchenko I, et al. Global Matrix 4.0 Physical Activity Report Card grades for children and adolescents: results and analyses from 57 countries. J Phys Act Health 2022 Nov 1;19(11):700-728. [doi: 10.1123/jpah.2022-0456] [Medline: 36280233]
14. Active Healthy Kids Global Alliance. URL: https://www.activehealthykids.org/about/ [accessed 2025-09-12]
15. Tremblay MS, Barnes JD, Demchenko I, et al. Active Healthy Kids Global Alliance Global Matrix 4.0-A resource for physical activity researchers. J Phys Act Health 2022 Nov 1;19(11):693-699. [doi: 10.1123/jpah.2022-0257] [Medline: 36280231]
16. Tremblay MS, Barnes JD, Cowie Bonne J. Impact of the Active Healthy Kids Canada report card: a 10-year analysis. J Phys Act Health 2014 May;11 Suppl 1(s1):S3-S20. [doi: 10.1123/jpah.2014-0167] [Medline: 25426911]
17. Aubert S, Barnes JD, Forse ML, et al. The international impact of the Active Healthy Kids Global Alliance Physical Activity Report Cards for children and youth. J Phys Act Health 2019 Sep 1;16(9):679-697. [doi: 10.1123/jpah.2019-0244] [Medline: 31412317]
18. Reilly JJ, Barnes J, Gonzalez S, et al. Recent secular trends in child and adolescent physical activity and sedentary behavior internationally: analyses of Active Healthy Kids Global Alliance Global Matrices 1.0 to 4.0. J Phys Act Health 2022 Nov 1;19(11):729-736. [doi: 10.1123/jpah.2022-0312] [Medline: 36280229]
19. Huang WY, Aubert S, Tremblay MS, Wong SH. Global Matrix 4.0 Physical Activity Report Cards grades for children and adolescents: a comparison among 15 Asian countries and regions. J Exerc Sci Fit 2022 Oct;20(4):372-381. [doi: 10.1016/j.jesf.2022.10.002] [Medline: 36311170]
20. Lee EY, Shih AC, Collins M, et al. Report card grades on physical activity for children and adolescents from 18 Asian countries: patterns, trends, gaps, and future recommendations. J Exerc Sci Fit 2023 Jan;21(1):34-44. [doi: 10.1016/j.jesf.2022.10.008] [Medline: 36408204]
21. Aubert S, Brazo-Sayavera J, González SA, et al. Global prevalence of physical activity for children and adolescents; inconsistencies, research gaps, and recommendations: a narrative review. Int J Behav Nutr Phys Act 2021 Jun 29;18(1):81. [doi: 10.1186/s12966-021-01155-2] [Medline: 34187486]
22. Zaabi MA, Shah SM, Sheek-Hussein M, Abdulle A, Junaibi AA, Loney T. Results from the United Arab Emirates' 2016 Report Card on physical activity for children and youth. J Phys Act Health 2016 Nov;13(11 Suppl 2):S299-S306. [doi: 10.1123/jpah.2016-0312] [Medline: 27848750]
23. Al-Kuwari MG, Ibrahim IA, Hammadi EMA, Reilly JJ. Results from Qatar's 2016 Active Healthy Kids Report Card on physical activity for children and youth. J Phys Act Health 2016 Nov;13(11 Suppl 2):S246-S250. [doi: 10.1123/jpah.2016-0397] [Medline: 27848747]
24. Paulo MS, Nauman J, Abdulle A, et al. Results from the United Arab Emirates' 2018 Report Card on Physical Activity for children and youth. J Phys Act Health 2018 Nov 1;15(S2):S419-S421. [doi: 10.1123/jpah.2018-0543] [Medline: 30475123]
25. Ibrahim I, Al Hammadi E, Sayegh S, et al. Results from Qatar's 2018 Report Card on Physical Activity for children and youth. J Phys Act Health 2018 Nov 1;15(S2):S400-S401. [doi: 10.1123/jpah.2018-0502] [Medline: 30475130]
26. Abi Nader P, Majed L, Sayegh S, et al. First Physical Activity Report Card for children and youth in Lebanon. J Phys Act Health 2019 Jun 1;16(6):385-396. [doi: 10.1123/jpah.2018-0473] [Medline: 31023183]
27. Alrahma AM, Al Suwaidi H, AlGurg R, et al. Results from the United Arab Emirates 2022 Report Card on Physical Activity for children and adolescents. J Exerc Sci Fit 2023 Apr;21(2):218-225. [doi: 10.1016/j.jesf.2023.02.002] [Medline: 36923208]
28. Abi Nader P, Hadla R, Majed L, Mattar L, Sayegh S. Results from the Lebanese 2022 Report Card on Physical Activity for children and youth. J Exerc Sci Fit 2023 Jan;21(1):14-19. [doi: 10.1016/j.jesf.2022.10.007] [Medline: 36349302]
29. Tomkinson GR, Carver KD, Atkinson F, et al. European normative values for physical fitness in children and adolescents aged 9-17 years: results from 2 779 165 Eurofit performances representing 30 countries. Br J Sports Med 2018 Nov;52(22):1445-1456. [doi: 10.1136/bjsports-2017-098253] [Medline: 29191931]

XSL•FO
RenderX

30. Human Development Index. United Nations Development Programme. URL: https://hdr.undp.org/data-center/human-development-index#/indicies/HDI [accessed 2025-09-12]
31. World Bank Open Data. World Bank Group. URL: https://data.worldbank.org [accessed 2025-09-12]
32. Climate change knowledge portal. The World Bank. URL: https://climateknowledgeportal.worldbank.org [accessed 2026-03-08]
33. Aubert S, Barnes JD, Aguilar-Farias N, et al. Report Card grades on the physical activity of children and youth comparing 30 very high Human Development Index countries. J Phys Act Health 2018 Nov 1;15(S2):S298-S314. [doi: 10.1123/jpah.2018-0431] [Medline: 30475144]
34. González SA, Barnes JD, Abi Nader P, et al. Report Card grades on the physical activity of children and youth from 10 countries with high Human Development Index: Global Matrix 3.0. J Phys Act Health 2018 Nov 1;15(S2):S284-S297. [doi: 10.1123/jpah.2018-0391] [Medline: 30475136]
35. Global school-based student health survey. World Health Organization. URL: https://www.who.int/teams/noncommunicable-diseases/surveillance/systems-tools/global-school-based-student-health-survey [accessed 2025-08-20]
36. Silva DAS, Aubert S, Ng K, et al. Association Between Physical Activity Indicators and Human Development Index at a National Level: Information From Global Matrix 4.0 Physical Activity Report Cards for Children and Adolescents. J Phys Act Health 2022 Nov 1;19(11):737-744. [doi: 10.1123/jpah.2022-0321] [Medline: 36280230]
37. Carson V, Kuzik N, Hunter S, et al. Systematic review of sedentary behavior and cognitive development in early childhood. Prev Med 2015 Sep;78:115-122. [doi: 10.1016/j.ypmed.2015.07.016] [Medline: 26212631]
38. Lee EY, Shih AC, Tremblay MS. Exploring the world of active play: a comprehensive review of global surveillance and monitoring of active play based on the global matrix data. J Exerc Sci Fit 2024 Jul;22(3):254-265. [doi: 10.1016/j.jesf.2024.03.008] [Medline: 38577389]
39. Lee EY, de Lannoy L, Kim YB, et al. 2025 position statement on active outdoor play. Int J Behav Nutr Phys Act 2025 Sep 25;22(1):117. [doi: 10.1186/s12966-025-01813-9] [Medline: 40999404]
40. Chaabane S, Chaabna K, Doraiswamy S, Mamtani R, Cheema S. Barriers and facilitators associated with physical activity in the Middle East and North Africa region: a systematic overview. Int J Environ Res Public Health 2021 Feb 9;18(4):1647. [doi: 10.3390/ijerph18041647] [Medline: 33572229]
41. Brazo-Sayavera J, Silva DR, Lang JJ, et al. Physical fitness surveillance and monitoring systems inventory for children and adolescents: a scoping review with a global perspective. Sports Med 2024 Jul;54(7):1755-1769. [doi: 10.1007/s40279-024-02038-9] [Medline: 38710913]
42. Loney T, Standage M, Thompson D, Sebire SJ, Cumming S. Self-report vs. objectively assessed physical activity: which is right for public health? J Phys Act Health 2011 Jan;8(1):62-70. [doi: 10.1123/jpah.8.1.62] [Medline: 21297186]
43. Larouche R, Neves de Oliveira S, Rostami Haji Abadi M, et al. Content validation of the Global Adolescent and Child Physical Activity Questionnaire (GAC-PAQ) in low-, middle-, and high-income countries across 6 continents. J Sport Health Sci 2025 Jul 5;15:101074. [doi: 10.1016/j.jshs.2025.101074] [Medline: 40623649]
44. Lang JJ, Zhang K, Agostinis-Sobrinho C, et al. Top 10 international priorities for physical fitness research and surveillance among children and adolescents: a twin-panel Delphi study. Sports Med 2023 Feb;53(2):549-564. [doi: 10.1007/s40279-022-01752-6] [Medline: 36001291]
45. Fraser BJ, Blizzard L, Tomkinson GR, Dwyer T, Venn AJ, Magnussen CG. Added predictive value of childhood physical fitness to traditional risk factors for adult cardiovascular disease. Eur J Prev Cardiol 2025 Feb 24:zwaf102. [doi: 10.1093/eurjpc/zwaf102] [Medline: 39993170]
46. National Health Strategy 2024-2030. State of Qatar Ministry of Public Health. URL: https://www.moph.gov.qa/English/NHS/Pages/default.aspx [accessed 2025-08-30]
47. Rhodes RE, Guerrero MD, Vanderloo LM, et al. Development of a consensus statement on the role of the family in the physical activity, sedentary, and sleep behaviours of children and youth. Int J Behav Nutr Phys Act 2020 Jun 16;17(1):74. [doi: 10.1186/s12966-020-00973-0] [Medline: 32539730]
48. Alnaqbi SE, Sohail R, Radwan HM, et al. Physical activity, screen time, dietary habits, and health outcomes among children and adolescents in the Middle East and North Africa region: a narrative review. Front Public Health 2025;13:1628904. [doi: 10.3389/fpubh.2025.1628904] [Medline: 41069803]
49. National Sports Strategy - 2031. United Arab Emirates. URL: https://u.ae/en/about-the-uae/strategies-initiatives-and-awards/strategies-plans-and-visions/sports/national-sports-strategy-2031 [accessed 2025-08-30]
50. Ministry of Health and Prevention. National Policy for Promoting Healthy Lifestyles. URL: https://www.uaelegislation.gov.ae/en/policy/details/lsy-s-lotny-ltaazyz-anm-t-lhy-lshy-fy-ldol [accessed 2025-09-12]
51. National Strategy for Wellbeing 2031. United Arab Emirates. URL: https://u.ae/en/about-the-uae/strategies-initiatives-and-awards/strategies-plans-and-visions/social-affairs/national-strategy-for-wellbeing-2031 [accessed 2025-09-12]
52. MoHAP unveils latest developments and achievements for “Masar” project at Arab Health 2025. United Arab Emirates Ministry of Health and Prevention. URL: https://mohap.gov.ae/en/w/mohap-unveils-latest-developments-and-achievements-for-masar-project-at-arab-health-2025 [accessed 2025-09-12]

XSL•FO
RenderX

53. National Sports Day: a pioneering initiative to promote physical health and sports awareness. State of Qatar Ministry of Sports and Youth. URL: https://www.msy.gov.qa/en/news/national-sports-day-a-pioneering-initiative-to-promote-physical-health-and-sports-awareness [accessed 2025-09-15]
54. Qatar National Physical Activity Guidelines. Aspetar. URL: https://www.aspetar.com/en/professionals/aspetar-clinical-guidelines/qatar-national-physical-activity-guidelines [accessed 2025-09-15]
55. National Health Strategy 2018-2022. State of Qatar Ministry of Public Health. URL: https://www.moph.gov.qa/english/strategies/National-Health-Strategy-2018-2022/Pages/default.aspx [accessed 2025-09-15]
56. Healthy FIFA World Cup Qatar 2022: Creating Legacy for Sport and Health. World Health Organization. URL: https://www.who.int/initiatives/sports-and-health/healthy-2022-world-cup [accessed 2025-09-15]

Abbreviations

AHKGA: Active Healthy Kids Global Alliance
FIFA: Fédération Internationale de Football Association
GDP: gross domestic product
GM: Global Matrix
GSHS: Global School-based Health Survey
HDI: Human Development Index
INC: incomplete
MVPA: moderate- to vigorous-intensity physical activity
PA: physical activity
PE: physical education
UAE: United Arab Emirates
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 16.Oct.2025; peer-reviewed by CK Chang, PA Okooboh; revised version received 30.Dec.2025; accepted 02.Jan.2026; published 24.Mar.2026.

Please cite as:
Baghestani A, Majed L, Lock M, Alrahma A, Abi Nader P, Sayegh S, Al-Mohannadi AS, Nauman J, Aubert S, Tremblay MS, Loney T Physical Activity Indicators Among Children and Adolescents in Lebanon, Qatar, and the United Arab Emirates: Comparative Synthesis of Active Healthy Kids Report Card Data From 1998 to 2022
JMIR Public Health Surveill 2026;12:e85998
URL: https://publichealth.jmir.org/2026/1/e85998
doi:10.2196/85998

© Ameneh Baghestani, Lina Majed, Merilyn Lock, Ali Alrahma, Patrick Abi Nader, Suzan Sayegh, Abdulla Saeed Al-Mohannadi, Javaid Nauman, Salomé Aubert, Mark S Tremblay, Tom Loney. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 24.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Depressive and Anxiety Symptoms and Their Comorbidity Across Industries: Multi-Industry Cross-Sectional Study in China

Lei Cao[1], MSc; Ruihong Ran[1], BSc; Huadong Zhang[1*], MSc; Heling Bao[2*], PhD

[1]Institute of Public Health and Safety Surveillance, Chongqing Center for Disease Control and Prevention, Chongqing, China

[2]Institute of Medical Information, Chinese Academy of Medical Sciences and Peking Union Medical College, No.3, Yabao Road, Chaoyang District, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Heling Bao, PhD
Institute of Medical Information, Chinese Academy of Medical Sciences and Peking Union Medical College, No.3, Yabao Road, Chaoyang District, Beijing, China

Abstract

Background: Depressive and anxiety symptoms are primary drivers of China's mental health burden, with workplace psychological issues growing increasingly pronounced. However, evidence on industry-specific patterns of depressive and anxiety symptoms and their occupational determinants remains limited. Understanding these sector-specific differences is critical for pinpointing high-risk occupations and designing effective, tailored intervention strategies.

Objective: Our objective was to assess the prevalence, determinants, and industry-specific patterns of depressive, anxiety, and comorbid symptoms across multiple industries.

Methods: We conducted a cross-sectional survey among 10,112 workers from 11 industries in all districts or counties of Chongqing, China. Depressive and anxiety symptoms were measured using the 9-item Patient Health Questionnaire (PHQ-9) and the 7-item Generalized Anxiety Disorder scale (GAD-7), with comorbidity defined as meeting criteria for both. Weighted prevalence and 95% CIs were calculated using Taylor series linearization accounting for cluster effects. Mixed-effects logistic regression models were applied to examine associations between demographic, health, and occupational factors and depressive and anxiety symptoms, as well as to compare these outcomes across industries. Multinomial mixed-effects models were used to examine patterns of comorbidity by modeling the 4-category outcome as follows: no symptoms, depression only, anxiety only, and comorbid depressive-anxiety symptoms.

Results: Overall, 19.0% (95% CI 16.7% - 21.3%) of workers reported depressive symptoms, 7.8% (95% CI 6.4% - 9.1%) reported anxiety symptoms, and 5.9% (95% CI 4.8% - 7.1%) experienced depressive-anxiety comorbidity. Younger age, being single, recent sick leave, multisite bodily pain, working hours, occupational stress, and health literacy were strongly associated with all outcomes, with the largest effects observed for the comorbid symptoms. Significant disparities were observed across industries, with the highest prevalence in education, health care, express delivery, and construction, and the lowest in traditional manufacturing and mining. These disparities persisted significantly after comprehensive adjustment for demographic, health-related, and occupational factors.

Conclusions: Depressive, anxiety, and comorbid symptoms are prevalent among Chinese workers, exhibiting significant variation across industries and distinct links to modifiable occupational and health factors. Targeted, sector-specific interventions and integration of mental health promotion into occupational health systems are vital to alleviate work-related psychological burdens and propel national mental health initiatives.

(JMIR Public Health Surveill 2026;12:e91782) doi:10.2196/91782

KEYWORDS

depression; anxiety; comorbidity; occupational stress; epidemiology

Introduction

Depression and anxiety disorders have emerged as leading contributors to the global burden of disease, with the prevalence substantially increasing since 2010 and exacting a heavy toll on individuals and society through suffering, reduced productivity, and economic costs [1]. Recent global estimates suggest a continuing rise in disability-adjusted life years due to these disorders [2]. In China, meta-analytic data indicate that up to 31% of the population experience depressive symptoms and 29% report anxiety symptoms [3]. Critically, comorbidity between depression and anxiety is common [4]. Individuals suffering from both tend to have greater functional impairment

XSL•FO
RenderX

and poorer quality of life than those with only one condition [5,6]. This convergence of high prevalence and co-occurrence has intensified the public health urgency.

A substantial body of evidence demonstrates that work-related factors—including long working hours, high job demands, low job control, inadequate social support, and emotional labor—are consistently associated with elevated risks of depression and anxiety across diverse occupational settings [7-9]. As these psychosocial risks have become increasingly recognized, many countries have incorporated work-related mental health into policy frameworks. For example, the European Union has mandated psychosocial risk management within occupational safety regulations to safeguard workers' psychological well-being [10], while the United States has implemented initiatives targeting work stress, fatigue, and organizational climate [11].

China has also elevated mental health to a national priority. The Healthy China Action (2019 - 2030) plan includes goals to reduce the prevalence of depression and anxiety and strengthen access to mental health services [12]. However, the mental health impact of occupational exposures has received far less scrutiny [7,13]. Evidence on industry-specific risks, patterns of depressive and anxiety symptoms, and work-related determinants remains limited. Industry-level data are urgently needed to inform targeted interventions and guide the development of occupational mental health prevention strategies.

This study aimed to assess depressive symptoms, anxiety symptoms, and their comorbidity across multiple industries using a large, representative sample of workers in Chongqing, China. Our objective was to identify industry-specific significant differences in the prevalence of mental health outcomes and to examine the demographic, health-related, and occupational factors associated with them. In addition, we sought to characterize the patterns of comorbidity between depression and anxiety across industries, providing evidence to inform targeted workplace mental health interventions.

Methods

Study Design

This cross-sectional study was a part of the 2024 Occupational Health Literacy Surveillance Program in Chongqing, China. We followed the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) reporting guideline and checklist to draft this manuscript (Checklist 1).

Survey Development

From January 1 to December 31, 2024, we conducted a survey to assess occupational health literacy, behavior, and mental health among urban workers across all 38 districts/counties in the Chongqing province (Multimedia Appendix 1). According to the industrial characteristics and Industrial Classification for National Economic Activities (GB/T 4754 - 2017), we selected 11 industries from the secondary and tertiary sectors, including the leather, fur, feather and their products, and footwear manufacturing industry (C-1900), automotive manufacturing industry (C-3600), electrical equipment manufacturing industry (C-3800), construction industry (E-0000), non-metallic mineral mining industry (B-1000), computer and communication equipment manufacturing industry (C-3900), health care and social work industry (Q-8400), education industry (P-8300), environmental sanitation industry (N-7820), transportation industry (G-5400), and express delivery and food delivery industry (G-6000).

Participant Recruitment

Eligible participants were front-line workers aged 16 - 59 years with at least 6 months of work experience in the selected industries, whereas administrative staff were excluded. We applied a multistage stratified cluster sampling design and at least 800 workers were required in each industry. For the secondary industries, enterprises served as primary sampling units and were stratified by size (large, medium, small/micro). Within each stratum, enterprises were randomly selected, and target sample sizes were allocated proportionally, with approximately 160, 240, and 400 workers recruited from large, medium, and small/micro enterprises, respectively. Large enterprises were limited to no more than 10% of the sampled worksites. For the tertiary industries, 5 districts (Yuzhong, Banan, Kaizhou, Jiulongpo, and Yongchuan) were selected, and industry-specific clusters were sampled within each district. Health care workers were recruited from 1 - 2 tertiary/secondary hospitals and 2 - 3 primary hospitals per district; teachers were sampled from 1 - 3 middle schools and 1 - 3 primary schools; sanitation workers from 1 - 3 environmental service agencies; transportation workers from 2 - 3 taxi, bus, freight, or ride-hailing companies; and delivery workers from 1 - 3 courier or food-delivery companies. Each district aimed to recruit at least 200 participants per industry. Finally, 10,112 workers were invited to completed an anonymous, self-administered electronic questionnaire.

Survey Measures

The primary outcomes were the prevalence of depressive symptoms, anxiety symptoms, and their comorbidity. Depression and anxiety were assessed using the validated Chinese versions of the 9-item Patient Health Questionnaire (PHQ-9) and the 7-item Generalized Anxiety Disorder scale (GAD-7) [14,15]. The PHQ-9 evaluates depressive symptoms over the past 6 months, and the GAD-7 assesses anxiety symptoms over the past 2 weeks. Both use a 4-point Likert scale and have demonstrated strong reliability and validity in Chinese populations [16,17]. Higher scores indicate greater symptom severity. Clinically relevant depressive and anxiety symptoms were defined using a cutoff value of ≥10 on each scale, and comorbidity was defined as meeting both thresholds. In this study, the scale demonstrated a Cronbach α of 0.89 for PHQ-9 and 0.91 for GAD-7.

Occupational stress was measured using the Core Occupational Stress Scale [18], developed for Chinese occupational settings and assessing social support, remuneration, dedication, and autonomy. A total score ≥50 indicated occupational stress (Cronbach α=0.94). Workload indicators included weekly working hours and night or rotating shift work. Occupational health literacy—assessing legal knowledge, protective knowledge, skills, and healthy workplace behaviors—was measured according to the national surveillance protocol.

XSL•FO
RenderX

Additional covariates included age, sex, marital status, education, sick leave days, and multisite bodily pain.

Statistical Analysis

Categorical variables were summarized as frequencies and percentages, and continuous variables as median with interquartile ranges. Differences between groups were assessed using chi-square or Mann-Whitney *U* tests. Standardized prevalence estimates were derived using sampling weights to represent the occupational population aged 16‑59 years in Chongqing, with poststratification adjustments for age and sex based on the Seventh National Census. Crude prevalence was reported for sensitivity analysis. Variances and 95% CIs were calculated using Taylor series linearization accounting for cluster effects.

Generalized linear regression models were used to examine associations between industry types, occupational stress, workload, occupational health literacy, health status, and mental health outcomes, adjusting for demographic characteristics. Random intercepts for district/county were included to account for the hierarchical sampling structure. For the comorbid outcome, the joint distribution of depression and anxiety was modeled as a 4-category variable (none, anxiety only, depression only, comorbidity). A Bayesian multinomial mixed-effects logistic regression was fitted using the *brms* package in R with a categorical logit link, with parameter variances and 95% credible intervals derived from Hamiltonian Monte Carlo posterior sampling. Results were presented as odds ratios (ORs) with 95% CIs. To further explore factors that may account for occupational differences in mental health, we fitted a series of sequentially adjusted models by progressively introducing demographic, health-related, and occupational variables. Changes in industry-specific ORs across models were examined to evaluate the extent to which these factors accounted for the observed heterogeneity across industries. All analyses were conducted in R version 4.4.1 (R Foundation for Statistical Computing). Two-sided *P* values <.05 were considered statistically significant.

Ethical Considerations

This study was approved by the Ethics Committee of the Chongqing Academy of Preventive Medicine (approval no. KY-2025-026-1) before participant recruitment. Written informed consent was obtained from all participants prior to questionnaire completion. Participation was voluntary, and all data were collected anonymously, stored securely, and analyzed in deidentified form to protect participant confidentiality. The study was conducted in accordance with the ethical principles of the Declaration of Helsinki.

Results

Participant Characteristics

A summary of participant characteristics is presented in Table 1. A total of 10,112 workers from 11 industries were included, comprising 5366 (53.0%) men and 4746 (47.0%) women. Women were more often employed in education and health care, while men predominated in transportation, construction, and mining sectors. The mean age was 40.5 (SD 10.3) years. Most participants were married, and two-thirds had degrees from senior middle school or a higher institution. About half of the participants were determined to have low occupational health literacy. Approximately a quarter of participants worked more than 55 hours per week, and 45.4% (n=4595) reported night or rotating shifts. One-fourth experienced occupational stress. Mean scores were 6.7 (SD 4.6) for PHQ-9 and 3.2 (SD 4.1) for GAD-7.

XSL•FO
RenderX

Table . Sociodemographic and occupational characteristics of participants from in a cross-sectional survey in Chongqing, China (cross-sectional survey, January 1-December 31, 2024; N=10,112), stratified by sex.

	Overall	Men	Women
Overall, n (%)	10,112 (100)	5366 (53.0)	4746 (47.0)
Age (years)			
Mean (SD)	40.5 (10.3)	40.9 (10.7)	40.0 (9.8)
Age group, n (%)			
16 - 39	4957 (49.0)	2558 (47.7)	2399 (50.6)
40 - 59	5155 (51.0)	2808 (52.3)	2347 (49.5)
Relationship status, n (%)			
Single	1692 (16.7)	1067 (19.9)	625 (13.2)
Married	7887 (78.0)	4003 (74.6)	3884 (81.8)
Other	533 (5.3)	296 (5.5)	237 (5.0)
Education, n (%)			
Junior middle school and lower	3772 (37.3)	2014 (37.5)	1758 (37.0)
Senior middle school	3928 (38.8)	2503 (46.7)	1425 (30.0)
Undergraduate	2412 (23.9)	849 (15.8)	1563 (32.9)
Sick leave days in the past year, n (%)			
None	7903 (78.2)	4266 (79.5)	3637 (76.6)
Less than 3 days	1461 (14.5)	719 (13.4)	742 (15.6)
More than 3 days	748 (7.4)	381 (7.1)	367 (7.7)
Weekly working hours, n (%)			
35 - 40 hours	2400 (23.7)	1193 (22.2)	1207 (25.4)
40 - 55 hours	5258 (52.0)	2613 (48.7)	2645 (55.7)
Over 55 hours	2454 (24.3)	1560 (29.1)	894 (18.8)
Multisite bodily pain, n (%)			
None	4494 (44.4)	2621 (48.8)	1873 (39.5)
Less than 3	3355 (33.2)	1669 (31.1)	1686 (35.5)
More than 3	2263 (22.4)	1076 (20.1)	1187 (25.0)
Night shift, n (%)			
Yes	4595 (45.4)	2791 (52.0)	1804 (38.0)
No	5517 (54.6)	2575 (48.0)	2942 (62.0)
Occupational stress, n (%)			
Yes	2484 (24.6)	1449 (27.0)	1035 (21.8)
No	7628 (75.4)	3917 (73.0)	3711 (78.2)
Occupational health literacy, n (%)			
Low	5165 (51.1)	2716 (50.6)	2449 (51.6)
High	4947 (48.9)	2650 (49.4)	2297 (48.4)
Industry categories, n (%)			
P-8300: education industry	946 (9.4)	221 (4.1)	725 (15.3)
Q-8400: health care and social work industry	941 (9.3)	179 (3.3)	762 (16.1)
G-6000: express delivery and food delivery industry	917 (9.1)	668 (12.5)	249 (5.3)

XSL•FO
RenderX

	Overall	Men	Women
G-5400: transportation industry	866 (8.6)	712 (13.3)	154 (3.2)
N-7820: environmental sanitation industry	888 (8.8)	257 (4.8)	631 (13.3)
C-3600: automotive manufacturing industry	919 (9.1)	631 (11.8)	288 (6.1)
C-3800: electrical equipment manufacturing industry	933 (9.2)	372 (6.9)	561 (11.8)
C-3900: computer and communication equipment manufacturing industry	921 (9.1)	417 (7.8)	504 (10.6)
B-1000: non-metallic mineral mining industry	890 (8.8)	757 (14.1)	133 (2.8)
C-1900: leather, fur, feather and their products, and footwear manufacturing industry	953 (9.4)	405 (7.6)	548 (11.6)
E-0000: construction industry	938 (9.3)	747 (13.9)	191 (4.0)
Mental health measures, mean (SD)			
PHQ-9[a] score	6.7 (4.6)	6.8 (4.7)	6.6 (4.5)
GAD-7[b] score	3.2 (4.1)	3.0 (4.1)	3.3 (4.1)

[a]PHQ-9: 9-item Patient Health Questionnaire.

[b]GAD-7: 7-item Generalized Anxiety Disorder scale.

Prevalence of Depressive and Anxiety Symptoms

The standardized and crude prevalence of depressive symptoms, anxiety symptoms, and their comorbidity is detailed in Table 2 and Table S1 in Multimedia Appendix 2, respectively. Overall, 19.0% (95% CI 16.7% - 21.3%) of workers reported depressive symptoms, 7.8% (95% CI 6.4% - 9.1%) reported anxiety symptoms, and 5.9% (95% CI 4.8% - 7.1%) experienced comorbid depressive-anxiety symptoms. Younger workers (16 - 39 years) and single individuals consistently exhibited higher comorbidity prevalence than their comparison groups. Participants who had more sick leave days had a markedly greater burden of comorbidity compared with those reporting less days or none (16.0% vs 10.1% and 4.2%, respectively). The prevalence of depressive and anxiety symptoms increased in a graded manner with longer working hours, more sites of body pain, night or rotating shifts, and higher scores of occupational stress. In addition, workers with lower occupational health literacy showed substantially elevated comorbidity prevalence compared with their counterparts (7.7% vs 4.1%).

XSL•FO
RenderX

Table . Weighted prevalence of depressive symptom, anxiety symptom, and comorbid symptoms among participants in Chongqing, China (cross-sectional survey, January 1-December 31, 2024; N=10,112), overall and by subgroups.

	Depressive symptoms, % (95% CI)	Anxiety symptoms, % (95% CI)	Comorbid symptoms, % (95% CI)		
			Depression only	Anxiety only	Comorbidity
Overall	19.0 (16.7 - 21.3)	7.8 (6.4 - 9.1)	13.1 (11.7 - 14.5)	1.8 (1.5 - 2.1)	5.9 (4.8 - 7.1)
Sex					
Men	19.3 (16.7 - 21.8)	7.7 (6.3 - 9.1)	13.4 (11.7 - 15.2)	1.9 (1.4 - 2.3)	5.8 (4.7 - 6.9)
Women	18.7 (16.1 - 21.2)	7.9 (6.3 - 9.4)	12.5 (11.2 - 13.9)	1.7 (1.4 - 2.0)	6.1 (4.6 - 7.6)
Age group (years)					
16 - 39	23.8 (21.6 - 26.0)	9.9 (8.1 - 11.6)	16.0 (14.6 - 17.4)	2.0 (1.6 - 2.4)	7.8 (6.3 - 9.4)
40 - 59	14.7 (12.3 - 17.2)	5.9 (4.7 - 7.1)	10.5 (8.9 - 12.0)	1.6 (1.3 - 2.0)	4.3 (3.2 - 5.3)
Relationship status					
Single	26.5 (23.9 - 29.1)	9.8 (8.3 - 11.3)	18.3 (16.3 - 20.4)	1.7 (1.2 - 2.2)	8.1 (6.8 - 9.5)
Married	16.7 (14.5 - 18.9)	6.9 (5.6 - 8.2)	11.6 (10.3 - 12.9)	1.8 (1.5 - 2.1)	5.1 (4.0 - 6.2)
Other	23.8 (19.3 - 28.2)	11.8 (7.5 - 16.1)	14.1 (10.2 - 17.9)	2.1 (0.8 - 3.3)	9.7 (5.5 - 13.9)
Education					
Junior middle school and lower	13.9 (11.9 - 15.9)	6.3 (5.0 - 7.5)	9.3 (8.2 - 10.4)	1.7 (1.3 - 2.0)	4.6 (3.6 - 5.7)
Senior middle school	20.5 (17.9 - 23.1)	7.8 (6.7 - 8.9)	14.7 (12.6 - 16.8)	2.0 (1.5 - 2.5)	5.8 (4.9 - 6.7)
Undergraduate	25.0 (23.3 - 26.6)	10.1 (8.1 - 12.2)	16.6 (15.5 - 17.6)	1.7 (1.0 - 2.5)	8.4 (6.9 - 10.0)
Sick leave days in the past year					
None	15.4 (13.2 - 17.7)	5.7 (4.5 - 6.9)	11.2 (9.8 - 12.6)	1.5 (1.2 - 1.8)	4.2 (3.2 - 5.3)
Less than 3 days	28.3 (25.4 - 31.1)	12.8 (10.0 - 15.5)	18.2 (15.7 - 20.7)	2.7 (1.5 - 3.8)	10.1 (7.9 - 12.3)
More than 3 days	38.9 (35.7 - 42.2)	19.7 (16.3 - 23.1)	22.9 (20.0 - 25.8)	3.7 (2.1 - 5.2)	16.0 (13.0 - 19.0)
Weekly working hours					
35 - 40 hours	15.8 (13.3 - 18.4)	6.6 (5.2 - 8.0)	10.9 (8.9 - 12.9)	1.7 (1.1 - 2.3)	5.0 (3.8 - 6.1)
40 - 55 hours	17.8 (15.0 - 20.6)	6.8 (5.3 - 8.4)	12.5 (10.9 - 14.2)	1.5 (1.2 - 1.9)	5.3 (3.9 - 6.6)
Over 55 hours	24.5 (21.2 - 27.9)	10.8 (7.9 - 13.7)	16.3 (13.9 - 18.6)	2.5 (1.6 - 3.4)	8.3 (6.0 - 10.6)
Multisite bodily pain					
No	8.8 (7.5 - 10.2)	3.0 (2.3 - 3.8)	6.7 (5.6 - 7.8)	1.0 (0.6 - 1.3)	2.1 (1.5 - 2.7)
Less than 3	21.9 (19.4 - 24.5)	8.6 (7.3 - 9.9)	15.5 (13.5 - 17.5)	2.1 (1.5 - 2.7)	6.5 (5.5 - 7.4)
More than 3	35.5 (33.0 - 37.9)	16.2 (14.3 - 18.1)	22.4 (20.7 - 24.2)	3.2 (2.6 - 3.7)	13.0 (11.2 - 14.9)
Night shift					
Yes	22.7 (19.9 - 25.4)	9.5 (7.8 - 11.2)	15.2 (13.3 - 17.0)	2.1 (1.3 - 2.8)	7.5 (6.2 - 8.8)
No	17.2 (14.8 - 19.5)	6.9 (5.6 - 8.2)	12.0 (10.6 - 13.4)	1.7 (1.5 - 1.9)	5.2 (4.0 - 6.3)
Occupational stress					
Yes	37.4 (34.2 - 40.5)	17.6 (15.1 - 20.0)	23.4 (21.1 - 25.6)	3.5 (3.0 - 4.1)	14.0 (11.6 - 16.4)
No	13.0 (11.3 - 14.6)	4.5 (3.7 - 5.4)	9.7 (8.5 - 10.9)	1.2 (0.9 - 1.6)	3.3 (2.7 - 3.9)
Occupational health literacy					
Low	21.9 (19.5 - 24.2)	10.2 (8.9 - 11.4)	14.2 (12.5 - 15.9)	2.5 (2.1 - 2.9)	7.7 (6.6 - 8.7)
High	16.0 (13.3 - 18.8)	5.2 (3.7 - 6.8)	11.9 (10.1 - 13.7)	1.1 (0.8 - 1.4)	4.1 (2.9 - 5.4)

XSL•FO
RenderX

Table 3 presents the associations of demographic, health, and occupational factors with depressive, anxiety, and comorbid symptoms in the fully adjusted mixed-effects models. Younger workers had consistently higher odds of depression only, anxiety only, and comorbid symptoms, and being single was associated with higher odds of depression only symptoms and comorbidity. Having more than 3 days of sick leave and experiencing pain at more than 3 body sites were strongly associated with all mental health outcomes, with the strongest associations observed for comorbid symptoms (OR 3.35, 95% CI 2.61 - 4.29 and OR 5.66, 95% CI 4.36 - 7.37, respectively). Occupational stress exhibited significant associations across all outcomes and was associated with approximately threefold higher odds of depression only (OR 2.99, 95% CI 2.62 - 3.39) and anxiety only (OR 3.51, 95% CI 2.59 - 4.76) symptoms, and nearly fivefold higher odds of comorbid symptoms (OR 4.78, 95% CI 3.98 - 5.74). Long working hours (>55 hours/week) were also associated with higher odds of depression only (OR 1.42, 95% CI 1.17 - 1.70), anxiety only (OR 1.63, 95% CI 1.09 - 2.51), and comorbid symptoms (OR 1.64, 95% CI 1.27 - 2.13). In addition, low occupational health literacy was independently associated with higher odds of all outcomes, whereas night or rotating shift work was not significantly associated with any outcome.

XSL•FO
RenderX

Table . Associations between explanatory factors and depressive, anxiety, and comorbid symptoms among participants in Chongqing, China (cross-sectional survey, January 1-December 31, 2024; N=10,112).

Variables	Depressive symptom		Anxiety symptom		Comorbid symptoms					
					Depression only		Anxiety only		Comorbidity	
	OR[a] (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Sex										
Men	1.15 (1.02 - 1.29)	.02	1.05 (0.89 - 1.24)	.56	1.18 (1.03 - 1.34)	.02	1.14 (0.83 - 1.56)	.42	1.09 (0.90 - 1.32)	.40
Women	Ref[b]		Ref		Ref		Ref		Ref	
Age group (years)										
16 - 39	1.32 (1.16 - 1.49)	<.001	1.43 (1.20 - 1.71)	<.001	1.26 (1.09 - 1.46)	.002	1.42 (1.03 - 2.00)	.03	1.54 (1.25 - 1.87)	<.001
More than 40	Ref		Ref		Ref		Ref		Ref	
Relationship status										
Single	1.39 (1.19 - 1.61)	<.001	1.10 (0.89 - 1.36)	.37	1.40 (1.18 - 1.65)	<.001	0.81 (0.51 - 1.27)	.39	1.35 (1.06 - 1.70)	.01
Married	Ref		Ref		Ref		Ref		Ref	
Other	1.44 (1.14 - 1.81)	.002	1.57 (1.15 - 2.12)	.004	1.29 (0.98 - 1.69)	.07	1.17 (0.59 - 2.12)	.64	1.85 (1.30 - 2.54)	<.001
Education										
Junior school and lower	Ref		Ref		Ref		Ref		Ref	
Senior middle school	1.36 (1.18 - 1.56)	<.001	1.09 (0.89 - 1.33)	.41	1.45 (1.22 - 1.69)	<.001	1.21 (0.84 - 1.76)	.31	1.20 (0.95 - 1.50)	.13
Under-graduate	1.66 (1.40 - 1.95)	<.001	1.44 (1.14 - 1.81)	.002	1.63 (1.35 - 1.98)	<.001	1.19 (0.77 - 1.86)	.43	1.78 (1.37 - 2.32)	<.001
Sick leave days in the past year										
None	Ref		Ref		Ref		Ref		Ref	
Less than 3 days	1.41 (1.22 - 1.63)	<.001	1.48 (1.22 - 1.81)	<.001	1.33 (1.13 - 1.56)	<.001	1.27 (0.84 - 1.90)	.25	1.70 (1.34 - 2.12)	<.001
More than 3 days	2.33 (1.96 - 2.78)	<.001	2.48 (2.00 - 3.09)	<.001	2.11 (1.73 - 2.58)	<.001	2.38 (1.50 - 3.61)	<.001	3.35 (2.61 - 4.29)	<.001
Weekly working hours										
35 - 40 hours	Ref		Ref		Ref		Ref		Ref	
40 - 55 hours	1.15 (1.00 - 1.32)	.06	1.09 (0.88 - 1.34)	.44	1.17 (0.99 - 1.38)	.06	1.11 (0.74 - 1.64)	.63	1.14 (0.90 - 1.46)	.27
Over 55 hours	1.45 (1.23 - 1.70)	<.001	1.50 (1.20 - 1.89)	<.001	1.42 (1.17 - 1.70)	<.001	1.63 (1.09 - 2.51)	.02	1.64 (1.27 - 2.13)	<.001
Multisite bodily pain										

XSL•FO
RenderX

Variables	Depressive symptom		Anxiety symptom		Comorbid symptoms					
					Depression only		Anxiety only		Comorbidity	
	OR[a] (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value
No	Ref		Ref		Ref		Ref		Ref	
Less than 3 sites	2.53 (2.19 - 2.91)	<.001	2.48 (1.99 - 3.10)	<.001	2.48 (2.11 - 2.92)	<.001	2.65 (1.76 - 3.98)	<.001	3.00 (2.34 - 3.91)	<.001
More than 3 sites	4.02 (3.46 - 4.67)	<.001	3.91 (3.13 - 4.88)	<.001	3.76 (3.18 - 4.48)	<.001	3.95 (2.57 - 5.95)	<.001	5.66 (4.36 - 7.37)	<.001
Night shift										
Yes	1.07 (0.95 - 1.21)	.27	1.00 (0.85 - 1.19)	.96	1.07 (0.93 - 1.22)	.36	0.91 (0.65 - 1.24)	.52	1.06 (0.88 - 1.28)	.53
No	Ref		Ref		Ref		Ref		Ref	
Occupational stress										
Yes	3.25 (2.89 - 3.64)	<.001	3.36 (2.87 - 3.94)	<.001	2.99 (2.62 - 3.39)	<.001	3.51 (2.59 - 4.76)	<.001	4.78 (3.98 - 5.74)	<.001
No	Ref		Ref		Ref		Ref		Ref	
Occupational health literacy										
Low	1.32 (1.18 - 1.49)	<.001	1.69 (1.43 - 2.00)	<.001	1.25 (1.09 - 1.42)	<.001	2.14 (1.55 - 2.99)	<.001	1.69 (1.40 - 2.05)	<.001
High	Ref		Ref		Ref		Ref		Ref	

[a]OR:odds ratio.

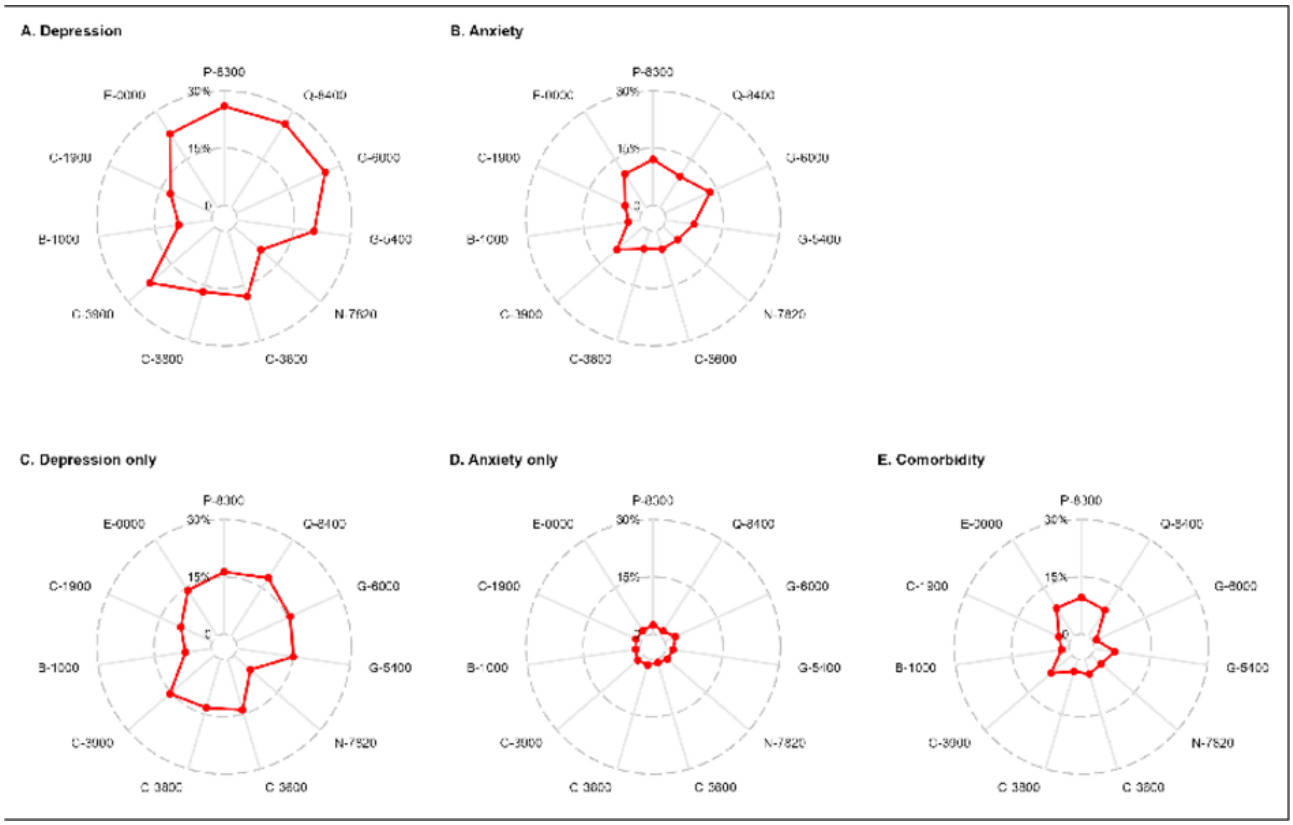
[b]Ref: reference.

Variations in the Prevalence of Depressive and Anxiety Symptoms Across Industries

Striking variations in the prevalence of depressive symptom, anxiety symptom, and comorbid symptoms across the 11 occupational groups are shown in Figure 1 and Table S2 in Multimedia Appendix 2. Depressive symptoms were most common in the health care and social work sector (26.1%, 95% CI 24.3% - 28.0%) and least common in the non-metallic mineral mining industry (8.7%, 95% CI 6.1% - 11.4%), whereas anxiety symptoms were highest among express and food delivery workers (13.1%, 95% CI 9.8% - 16.3%). Comorbid depressive-anxiety symptoms were particularly elevated in education (9.6%, 95% CI 7.5% - 11.8%), express delivery (10.0%, 95% CI 7.3% - 12.6%), and construction (8.7%, 95% CI 5.5% - 11.9%) workers, and lowest in mining (1.9%, 95% CI 1.0% - 2.8%) and leather, fur, and feather manufacturing (3.1%, 95% CI 2.7% - 3.4%). Depression only symptoms were more frequent in health care (18.1%, 95% CI 15.1% - 21.0%) and education (16.3%, 95% CI 14.3% - 18.3%), while anxiety only symptoms were relatively more common in express delivery (3.1%, 95% CI 1.8% - 4.4%).

XSL•FO
RenderX

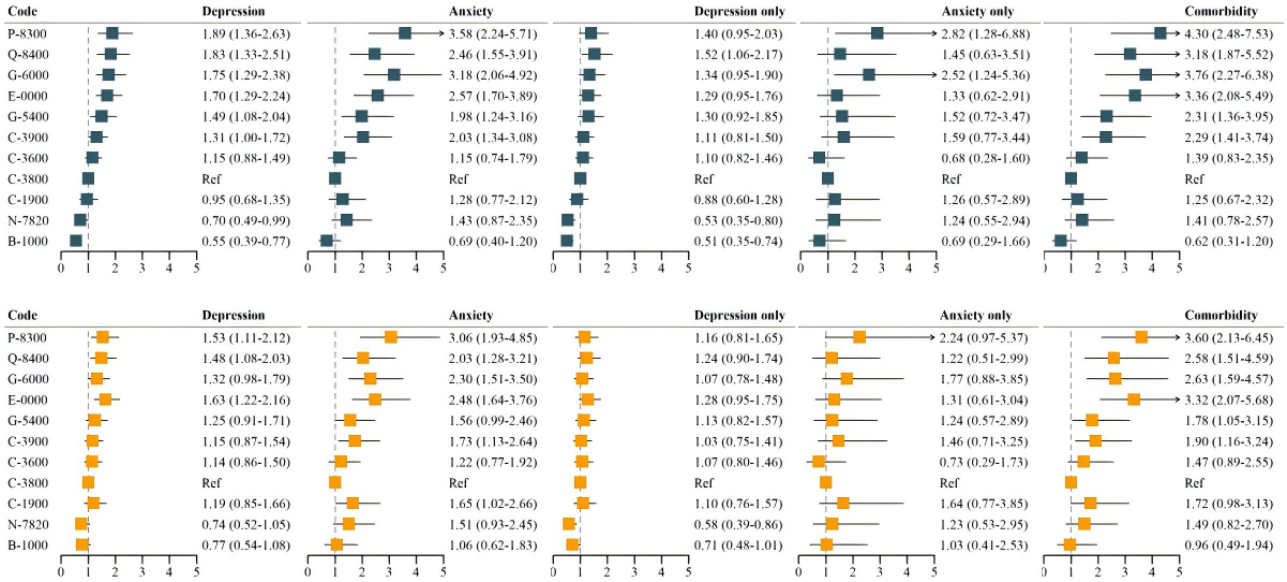
Figure 1. Prevalence of (A) depressive, (B) anxiety, (C) depression only, (D) anxiety only, and (E) comorbid symptoms across 11 industries in Chongqing, China (cross-sectional survey, January 1-December 31, 2024; N=10,112). C-1900: leather, fur, feather and their products, and footwear manufacturing industry; C-3600: automotive manufacturing industry; C-3800: electrical equipment manufacturing industry; E-0000: construction industry; B-1000: non-metallic mineral mining industry; C-3900: computer and communication equipment manufacturing industry; Q-8400: health care and social work industry; P-8300: education industry; N-7820: environmental sanitation industry; G-5400: transportation industry; G-6000: express delivery and food delivery industry.



Multivariable analyses reinforced differences in mental health outcomes across industries (Figure 2 and Table S3 in Multimedia Appendix 2). In models controlling for demographic factors, workers in the education, health care, express delivery, and construction industries had significantly higher odds of depression only, anxiety only, and comorbid symptoms than those in the reference industry. These associations remained after further adjustment for demographic, health, and occupational characteristics, although effect estimates were attenuated. Workers in the education, health care, express delivery, and construction industries had the highest odds of comorbid depression-anxiety symptoms among all industries. In contrast, workers in traditional production sectors, particularly non-metallic mineral mining (OR 0.68, 95% CI 0.47 - 0.97), showed lower odds of such mental health symptoms, whereas the association for the leather and footwear manufacturing sector was not statistically significant (OR 1.15, 95% CI 0.81 - 1.62).

XSL•FO
RenderX

Figure 2. Associations between industries and depressive, anxiety, and comorbid symptoms, adjusting for covariates among participants in Chongqing, China (cross-sectional survey, January 1-December 31, 2024; N=10,112). Blue represents results adjusting for only demographic factors and yellow represents results adjusting for all covariates. C-1900: leather, fur, feather and their products, and footwear manufacturing industry; C-3600: automotive manufacturing industry; C-3800: electrical equipment manufacturing industry; E-0000: construction industry; B-1000: non-metallic mineral mining industry; C-3900: computer and communication equipment manufacturing industry; Q-8400: health care and social work industry; P-8300: education industry; N-7820: environmental sanitation industry; G-5400: transportation industry; G-6000: express delivery and food delivery industry.



Sequential adjustment analyses (Table S4 in Multimedia Appendix 2) showed different explanatory patterns across mental health outcomes. Physical health indicators markedly attenuated differences in comorbid and depressive symptoms, whereas occupational stress accounted for the largest reductions in anxiety. Demographic characteristics and night shift work explained little of the observed occupational heterogeneity.

Discussion

Principal Findings

In this large, multi-industry survey of Chinese workers, we identified a substantial burden of depressive symptoms, anxiety symptoms, and their comorbidity. Nearly 1 in 5 participants reported depressive symptoms, and more than 1 in 8 experienced comorbidity, with younger workers exhibiting the highest prevalence across outcomes. Occupational stress and multisite bodily pain showed the strongest associations with all mental health outcomes; long working hours, low occupational health literacy, and poorer physical health were also independently associated with higher odds of depressive symptoms, anxiety symptoms, and comorbidity. Marked heterogeneity was observed across industries: workers in education, health care, express delivery, and construction sectors had the highest prevalence and odds of adverse mental health outcomes, whereas those in traditional manufacturing and mining industries exhibited substantially lower prevalence. Collectively, these findings suggest that work-related mental health problems represent a substantial and heterogeneous public health challenge and underscore the importance of industry-specific occupational mental health strategies and supportive policy development.

The prevalence estimates in this study fall within the upper range of recent findings in China. A meta-analysis by Bin et al [3] reported pooled rates around 31% for depressive symptoms and 29% for anxiety symptoms. Globally, the prevalence of depressive symptoms in working populations from high-income countries typically ranges from 10% to 20%, and anxiety from 5% to 15%, with comorbidity often representing a more severe clinical subgroup [19]. The higher comorbidity observed among younger workers in our study echoes global evidence that early-career adults face elevated psychological distress under job insecurity, high performance demands, and limited autonomy [20]. Moreover, the industry-specific pattern we observed—particularly high burdens in education, health care, service delivery and other high-demand sectors—is consistent with prior research identifying teachers, health professionals, and service workers as high-risk groups due to emotional labor, time pressure, and intensive interpersonal demands [13]. These parallels reinforce the plausibility and generalizability of our findings and suggest shared structural determinants of work-related mental health risk across national labor contexts.

Several individual and occupational factors were independently associated with higher odds of depressive symptoms, anxiety symptoms, and their comorbidity among workers. Occupational stress showed the strongest and most consistent associations across all outcomes, consistent with the job demand-control and effort-reward imbalance models [21,22], which propose that sustained job demands combined with low control or inadequate reward may contribute to emotional exhaustion and poor mental health. Longer sick leave was also associated with higher odds of all symptom profiles, supporting previous evidence that physical and mental health are closely interconnected and may reinforce one another over time. Although working more than 55 hours per week showed relatively modest associations, this finding is consistent with accumulating evidence linking long working hours to burnout, depression symptoms, and impaired cognitive recovery. Lower occupational health literacy was independently associated with

XSL•FO
RenderX

higher odds of adverse mental health outcomes, possibly because limited knowledge of occupational health and workplace protections may reduce workers' capacity to recognize psychological symptoms, adopt effective coping strategies, and seek timely support. By contrast, night or rotating shift work was not independently associated with mental health outcomes after adjustment, suggesting that the observed associations between shift work and mental health reported in previous studies may be partly explained by related factors, such as sleep disruption, occupational stress, and other co-occurring workplace exposures, rather than shift schedules alone [23].

The pronounced industry differences observed in our study likely reflect divergent psychosocial exposures, work organization, and labor conditions. Education and health care settings are classic high emotional labor sectors where workers routinely manage intense interpersonal demands, high expectations, and mounting administrative workloads, factors linked to elevated stress and burnout in multiple reviews. These pressures, compounded by chronic resource constraints in public service settings, may contribute to sustained psychological strain [24]. Platform-based delivery workers face sustained time pressure, algorithmic management, income instability, and limited institutional protections—conditions associated with higher anxiety and depressive symptoms in recent Chinese studies of couriers and delivery riders [25]. Construction workers' high comorbidity risk may reflect a confluence of hazardous environments, heavy physical load, injury-related pain, precarious/migrant employment, and constrained access to social and health supports [26]. In contrast, lower prevalence in some manufacturing and mining settings may be related to the more routinized tasks, stronger team cohesion, or selection/work-survival effects, although measurement and reporting differences could also be contributors. Importantly, comorbid depressive-anxiety presentations were the most disparate outcome across sectors, implying that psychosocially demanding occupations not only raise single-disorder risk but also foster more complex, clinically significant co-occurring illnesses [27,28].

Sequential adjustment analyses provided further insight into the mechanisms underlying occupational differences. For depressive-anxiety comorbidity, both physical health (eg, sick leave and multisite bodily pain) and occupational stress may contribute to severe mental health outcomes. For depression only, physical health indicators explained a substantial proportion of the between-industry variation, whereas occupational stress contributed relatively little. In contrast, for anxiety only, occupational stress consistently produced the largest attenuation across industries. Demographic characteristics and night shift work explained little of the observed occupational heterogeneity across all outcomes, highlighting workplace-related factors rather than workforce composition as the principal contributors to industry differences. These findings suggest that different mental health outcomes may arise through distinct occupational pathways, with physical health playing a greater role in depression and comorbidity, whereas psychosocial stress appears more closely linked to anxiety.

These findings carry important implications for public health practice and occupational policy. Occupational stress and physical health indicators should be prioritized as key, modifiable targets within workplace mental health strategies. Evidence-based measures—including workload redesign, stress-management interventions, and organizational policies that ensure adequate rest—may collectively mitigate psychological risk [29]. Industry-specific approaches are also essential. In high emotional labor sectors such as education and health care, interventions might focus on reducing administrative burden, improving staffing levels, and strengthening resilience and support systems [30,31]. For platform-based delivery workers, regulatory oversight of task allocation, guaranteed rest periods, and safeguards against algorithm-driven overwork are critical [32]. Among construction workers, integrating routine mental health screening, health-literacy training, and onsite counseling into occupational health services may be particularly beneficial [33]. Strengthening occupational health literacy across the workforce could further enhance early symptom recognition, help-seeking, and adoption of protective behaviors. These findings underscore the need to embed mental health more formally within China's occupational health system and Healthy China initiatives, with emphasis on systematic monitoring and targeted prevention.

Strengths and Limitations

This study has several strengths, including its large, multi-industry sample, the use of validated mental health instruments, and the application of mixed-effects and multinomial models that allowed examination of both individual-level determinants and industry-level heterogeneity. However, several limitations should be noted. The cross-sectional design precludes causal inference, and reliance on self-reported symptoms introduces the possibility of reporting bias. This study was conducted in a single province-level municipality, Chongqing, one of China's largest municipalities that has undergone rapid industrialization, urbanization, and labor-force diversification. These characteristics make it an informative setting for investigating occupational mental health in rapidly developing urban environments. Nevertheless, caution is warranted when generalizing the findings to other regions, and future studies including multiple provinces and cities are needed. Although marital status was adjusted for, other family-related factors, including family support, work-family conflict, and caregiving responsibilities, were not measured. Future studies should examine how occupational and family environments jointly influence workers' mental health. Finally, important industry-specific psychosocial constructs—such as job control, organizational justice, and supervisor support—were not measured and should be incorporated into future research to improve explanatory depth and intervention relevance.

Conclusion

In conclusion, this study indicates substantial and heterogeneous burdens of depressive, anxiety, and comorbid symptoms among workers across multiple industries in China. The findings identify key modifiable risk factors and high-risk sectors that require policy attention. Integrating mental health promotion into occupational health systems and tailoring interventions to

XSL•FO
RenderX

sector-specific needs will be critical for reducing work-related psychological burdens and advancing national mental health goals.

Acknowledgments

We thank all research staff from the Centers for Disease Control and Prevention in the provinces, cities, districts, and counties for their hard work in the implementation of the program and collection of data. The authors declare the use of generative AI (GAI) in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to GAI tools under full human supervision: proofreading and editing and translation. The GAI tool used was ChatGPT-4.5 (OpenAI). Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Funding

The study was supported by the Chongqing Medical Scientific Research Project–Joint Project of Chongqing Health Commission and Science and Technology Bureau (NO.2025GDRC013) and IMICAMS Youth Talent Development Fund (NO.2024YT02). The funders had no role in the conduct of the study or decision to submit the manuscript for publication.

Data Availability

The anonymized dataset analyzed during this study is available from the corresponding author upon reasonable request.

Authors' Contributions

LC led the conceptualization, data curation, investigation, formal analysis, and writing (original draft of the paper). HB contributed to the scientific conceptualization, formal analysis, methodology, and preparation of data. HZ contributed to the project leadership, serving as supervisor, and preparation of data. RR conducted to the investigation and data preparation. All authors contributed to writing (review and editing of the paper), with HB leading revisions. HB and HZ are co-corresponding authors for this paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study design details.
[DOCX File, 74 KB - publichealth_v12i1e91782_app1.docx]

Multimedia Appendix 2
Additional tables on the prevalence of depressive symptoms, anxiety symptoms, and comorbid symptoms, and their associations with industries.
[DOCX File, 94 KB - publichealth_v12i1e91782_app2.docx]

Checklist 1
STROBE checklist.
[DOCX File, 33 KB - publichealth_v12i1e91782_app3.docx]

References

1. GBD 2023 Disease and Injury and Risk Factor Collaborators. Burden of 375 diseases and injuries, risk-attributable burden of 88 risk factors, and healthy life expectancy in 204 countries and territories, including 660 subnational locations, 1990-2023: a systematic analysis for the Global Burden of Disease Study 2023. Lancet 2025 Oct 18;406(10513):1873-1922. [doi: 10.1016/S0140-6736(25)01637-X] [Medline: 41092926]
2. Jang W, Cho H, Kim S, et al. Global burden of anxiety, depression, and self-harm (1990-2021) and the impact of the COVID-19 pandemic. J Affect Disord 2026 Jan 1;392:120201. [doi: 10.1016/j.jad.2025.120201] [Medline: 40915503]
3. Bin X, Qu KY, Wang YH, et al. Prevalence of depression, anxiety in China during the COVID-19 pandemic: an updated systematic review and meta-analysis. Front Public Health 2023;11:1267764. [doi: 10.3389/fpubh.2023.1267764] [Medline: 38249392]
4. Kaiser T, Herzog P, Voderholzer U, Brakemeier EL. Unraveling the comorbidity of depression and anxiety in a large inpatient sample: network analysis to examine bridge symptoms. Depress Anxiety 2021 Mar;38(3):307-317. [doi: 10.1002/da.23136] [Medline: 33465284]

XSL•FO
RenderX

5. Pappa S, Shah M, Young S, Anwar T, Ming T. Health-care utilisation and associated economic burden in major depressive disorder and treatment-resistant depression in a longitudinal population-based cohort study of 110.000 patients. J Psychiatr Res 2024 Dec;180:511-518. [doi: 10.1016/j.jpsychires.2024.11.022] [Medline: 39550927]
6. Alwhaibi M, Alhawassi TM. Humanistic and economic burden of depression and anxiety among adults with migraine: a systematic review. Depress Anxiety 2020 Nov;37(11):1146-1159. [doi: 10.1002/da.23063] [Medline: 32579794]
7. van der Molen HF, Nieuwenhuijsen K, Frings-Dresen MHW, de Groene G. Work-related psychosocial risk factors for stress-related mental disorders: an updated systematic review and meta-analysis. BMJ Open 2020 Jul 5;10(7):e034849. [doi: 10.1136/bmjopen-2019-034849] [Medline: 32624469]
8. Theorell T, Hammarström A, Aronsson G, et al. A systematic review including meta-analysis of work environment and depressive symptoms. BMC Public Health 2015 Aug 1;15(738):738. [doi: 10.1186/s12889-015-1954-4] [Medline: 26232123]
9. Wahrendorf M, Sembajwe G, Zins M, Berkman L, Goldberg M, Siegrist J. Long-term effects of psychosocial work stress in midlife on health functioning after labor market exit--results from the GAZEL study. J Gerontol B Psychol Sci Soc Sci 2012 Jul;67(4):471-480. [doi: 10.1093/geronb/gbs045] [Medline: 22546992]
10. Fioritti A, Jònasson H, de Winter L, Van Audenhove C, van Weeghel J. Mental health and work: a European perspective. Epidemiol Psychiatr Sci 2024 Apr 5;33:e20. [doi: 10.1017/S2045796024000246] [Medline: 38576243]
11. Leso V, Carugno M, Carrer P, et al. The Total Worker Health® (TWH) approach: a systematic review of its application in different occupational settings. BMC Public Health 2024 Jul 30;24(1):2037. [doi: 10.1186/s12889-024-19500-y] [Medline: 39080673]
12. The “Healthy China 2030” planning outline. : China’s Central Party Committee and the State Council; 2016 URL: https://www.gov.cn/zhengce/2016-10/25/content_5124174.htm [accessed 2026-08-15]
13. Sussell AL, Yeoman K, Nixon CT, Scott KA, Robinson TS, Poplin GS. US workers’ self-reported mental health outcomes by industry and occupation. JAMA Netw Open 2025 Jun 2;8(6):e2514212. [doi: 10.1001/jamanetworkopen.2025.14212] [Medline: 40478574]
14. Costantini L, Pasquarella C, Odone A, et al. Screening for depression in primary care with Patient Health Questionnaire-9 (PHQ-9): a systematic review. J Affect Disord 2021 Jan 15;279:473-483. [doi: 10.1016/j.jad.2020.09.131] [Medline: 33126078]
15. Toussaint A, Hüsing P, Gumz A, et al. Sensitivity to change and minimal clinically important difference of the 7-item Generalized Anxiety Disorder Questionnaire (GAD-7). J Affect Disord 2020 Mar 15;265:395-401. [doi: 10.1016/j.jad.2020.01.032] [Medline: 32090765]
16. Wu Y, Levis B, Riehm KE, et al. Equivalency of the diagnostic accuracy of the PHQ-8 and PHQ-9: a systematic review and individual participant data meta-analysis. Psychol Med 2020 Jun;50(8):1368-1380. [doi: 10.1017/S0033291719001314] [Medline: 31298180]
17. Gong Y, Zhou H, Zhang Y, et al. Validation of the 7-item Generalized Anxiety Disorder scale (GAD-7) as a screening tool for anxiety among pregnant Chinese women. J Affect Disord 2021 Mar 1;282:98-103. [doi: 10.1016/j.jad.2020.12.129] [Medline: 33401129]
18. Wang J, Liu XM, Sun YY, Wang C, Li S, Yu SF. A preliminary study on establishment of the Core Occupational Stress Scale (COSS) norms for manufacturing and medical workers in China [Article in Chinese]. Zhonghua Yu Fang Yi Xue Za Zhi 2020 Nov 6;54(11):1190-1196. [doi: 10.3760/cma.j.cn112150-20200430-00674] [Medline: 33147915]
19. Hasan SI, Yee A, Rinaldi A, Azham AA, Mohd Hairi F, Amer Nordin AS. Prevalence of common mental health issues among migrant workers: a systematic review and meta-analysis. PLoS ONE 2021;16(12):e0260221. [doi: 10.1371/journal.pone.0260221] [Medline: 34855800]
20. Darvishmotevali M, Kim S(, Ning H. The impact of quantitative and qualitative job insecurity on employees’ mental health and critical work-related performance: exploring the role of employability and gender differences. Int J Hosp Manag 2024 Jan;116:103629. [doi: 10.1016/j.ijhm.2023.103629]
21. Wang H, Han J, Wang J, Wu Y, Cai H. The relationship between effort-reward imbalance and job satisfaction among primary care general practitioners in Gansu of China: the mediating role of public service motivation. Hum Resour Health 2025 Nov 27;24(1):2. [doi: 10.1186/s12960-025-01036-w] [Medline: 41310712]
22. Calnan M, Wadsworth E, May M, Smith A, Wainwright D. Job strain, effort--reward imbalance, and stress at work: competing or complementary models? Scand J Public Health 2004;32(2):84-93. [doi: 10.1080/14034940310001668] [Medline: 15255497]
23. Cyr M, Artenie DZ, Al Bikaii A, Borsook D, Olson JA. The effect of evening light on circadian-related outcomes: a systematic review. Sleep Med Rev 2022 Aug;64:101660. [doi: 10.1016/j.smrv.2022.101660] [Medline: 35753149]
24. Yu F, Raphael D, Mackay L, Smith M, Fernandez R. Personal and work-related factors associated with nurse resilience: an updated systematic review using meta-analysis and narrative synthesis. Int J Nurs Stud 2025 Jun;166:105054. [doi: 10.1016/j.ijnurstu.2025.105054] [Medline: 40147222]
25. Peng Y, Shao Y, Li Z, et al. Status and determinants of symptoms of anxiety and depression among food delivery drivers in Shanghai, China. Int J Environ Res Public Health 2022 Oct 13;19(20):13189. [doi: 10.3390/ijerph192013189] [Medline: 36293768]

XSL•FO
RenderX

26. Gómez-Salgado C, Camacho-Vega JC, Gómez-Salgado J, et al. Stress, fear, and anxiety among construction workers: a systematic review. Front Public Health 2023;11:1226914. [doi: 10.3389/fpubh.2023.1226914] [Medline: 37521990]
27. Lamers F, van Oppen P, Comijs HC, et al. Comorbidity patterns of anxiety and depressive disorders in a large cohort study: the Netherlands Study of Depression and Anxiety (NESDA). J Clin Psychiatry 2011 Mar;72(3):341-348. [doi: 10.4088/JCP.10m06176blu] [Medline: 21294994]
28. Luchman JN, González-Morales MG. Demands, control, and support: a meta-analytic review of work characteristics interrelationships. J Occup Health Psychol 2013 Jan;18(1):37-52. [doi: 10.1037/a0030541] [Medline: 23339747]
29. Joyce S, Modini M, Christensen H, et al. Workplace interventions for common mental disorders: a systematic meta-review. Psychol Med 2016 Mar;46(4):683-697. [doi: 10.1017/S0033291715002408] [Medline: 26620157]
30. West CP, Dyrbye LN, Shanafelt TD. Physician burnout: contributors, consequences and solutions. J Intern Med 2018 Jun;283(6):516-529. [doi: 10.1111/joim.12752] [Medline: 29505159]
31. Robbins R, Underwood P, Jackson CL, et al. A systematic review of workplace-based employee health interventions and their impact on sleep duration among shift workers. Workplace Health Saf 2021 Nov;69(11):525-539. [doi: 10.1177/21650799211020961] [Medline: 34284682]
32. Margerison CE, Joachim GE, Michling TR, Johnson CY. Platform-mediated gig work and health: a scoping review. Am J Ind Med 2025 Sep;68(9):729-739. [doi: 10.1002/ajim.70007] [Medline: 40662947]
33. Greiner BA, Leduc C, O'Brien C, et al. The effectiveness of organisational-level workplace mental health interventions on mental health and wellbeing in construction workers: a systematic review and recommended research agenda. PLoS ONE 2022;17(11):e0277114. [doi: 10.1371/journal.pone.0277114] [Medline: 36383613]

Abbreviations

GAD-7: 7-item Generalized Anxiety Disorder scale
OR: odds ratio
PHQ-9: 9-item Patient Health Questionnaire
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Mavragani, T Sanchez; submitted 20.Jan.2026; peer-reviewed by M Pritchard, Y Wu; revised version received 06.Aug.2026; accepted 08.Aug.2026; published 21.Aug.2026.

Please cite as:
Cao L, Ran R, Zhang H, Bao H
Depressive and Anxiety Symptoms and Their Comorbidity Across Industries: Multi-Industry Cross-Sectional Study in China
JMIR Public Health Surveill 2026;12:e91782
URL: https://publichealth.jmir.org/2026/1/e91782
doi:10.2196/91782

©Lei Cao, Ruihong Ran, Huadong Zhang, Heling Bao. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 21.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Associations of Health Care System Deficiency–Related Patient Deterioration With Burnout, Depressive Symptoms, and Sleep Problems Among Emergency Medical Service Providers: Cross-Sectional Study

Min Seung Choi[1*], MPH; Soo Jin Kim[1,2*], PhD; Yonghwan Moon[3,4], MPH; Hyekyung Woo[3], PhD

[1]Department of Emergency Medical Services, Graduate School of Public Health & Welfare, Eulji University, Seongnam, Republic of Korea

[2]Laboratory of Emergency Medical Services, Seoul National University Hospital Biomedical Research Institute, Seoul, Republic of Korea

[3]Department of Health Administration, Kongju National University, 56 Gongjudaehak-ro, Kongju-si, Republic of Korea

[4]Division of Research Planning, National Center for Mental Health, Mental Health Research Institute, Seoul, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Hyekyung Woo, PhD
Department of Health Administration, Kongju National University, 56 Gongjudaehak-ro, Kongju-si, Republic of Korea

Abstract

Background: On February 6, 2024, the Korean government announced an increase in medical school enrollment quotas, triggering a nationwide strike by medical residents that precipitated a severe health care crisis. Prehospital emergency medical service (EMS) providers were forced to transport patients to distant hospitals, resulting in some patients experiencing symptom deterioration or death due to delayed critical care. Recent studies have examined moral injury—a mental health response to events that undermine moral values—among EMS providers, with exposure to potentially morally injurious events increasing the likelihood of depression, anxiety, and posttraumatic stress disorder. However, studies specifically examining the association between health care–related patient deterioration experiences caused by health care service disruptions and mental health outcomes among EMS providers remain limited.

Objective: This study evaluated the associations of patient deterioration due to health care service disruptions (hereafter referred to as health care–related patient deterioration experiences [HPDEs]) with burnout, depressive symptoms, and sleep problems among prehospital EMS providers during government–medical community disputes.

Methods: We conducted a retrospective cross-sectional survey from November 1 to December 31, 2024, among EMS providers across South Korea using a web-based questionnaire. The independent variable was the presence or absence of HPDE, and the dependent variables were 3 mental health outcomes (personal burnout [PB], work-related burnout [WRB], and citizen-related burnout [CRB]), depressive symptoms, and sleep problems. χ^2 tests and multivariable logistic regression were used to assess associations and identify influencing factors after adjustment for potential confounders.

Results: Of 845 participants, 75% (n=626) reported HPDEs, which were associated with higher odds of PB, WRB, CRB, depressive symptoms, and sleep problems. Additionally, sex, age, educational attainment, and work region significantly influenced burnout, depressive symptoms, and sleep problems.

Conclusions: During government–medical community disputes, EMS providers exposed to HPDEs—a form of potentially morally injurious event resulting from emergency medical resource shortages—had significantly elevated odds of burnout, depressive symptoms, and sleep problems. These findings underscore the urgent need for systematic psychological interventions and regular mental health screenings, such as for moral injury, to mitigate adverse mental health outcomes during future health care service disruptions.

(JMIR Public Health Surveill 2026;12:e90130) doi:10.2196/90130

KEYWORDS

patient deterioration; health care service disruptions; burnout; depression; sleep disorders; emergency medical technicians

XSL•FO
RenderX

Introduction

On February 6, 2024, the Korean government announced an increase in medical school enrollment quotas to secure essential health care personnel. In response, beginning February 20, approximately 12,000 medical residents—more than 90% of all residents—initiated a strike by submitting their resignations [1]. This led to a sharp reduction in essential medical personnel across hospitals nationwide, including emergency medical centers, disrupting both the broader health care system and the emergency medical system severely.

Prehospital emergency medical service (EMS) providers were particularly affected. Although these providers are required to transport emergency patients to the nearest available hospital after administering initial treatment [2], many were frequently forced to seek hospitals farther away when nearby facilities could not accept patients. Emergency medical centers nationwide report bed availability and staffing levels in real time to the National Emergency Medical Center through the National Emergency Department Information System, and EMS providers use this data to select transport destinations. When a facility is unable to accept emergency patients due to staff shortages or other reasons, a "care restriction message" is displayed on the dashboard. In 2024, a total of 110,033 care restriction messages were recorded on the National Emergency Medical Center's integrated dashboard—an 88% increase compared to 58,520 cases in 2023, prior to the crisis. Notably, care restrictions attributed to staff shortages surged to 43,658 cases, representing a 132.8% increase over the previous year [3]. During this process, some patients experienced symptom deterioration or death due to delays in receiving critical emergency care that should have been promptly administered at hospitals. EMS providers witnessed these outcomes firsthand [4,5].

Recently, moral injury (MI) has been studied as a mental health issue among EMS providers and other health care professionals [6-8]. MI is a response to events that undermine an individual's beliefs and moral values [6,9], triggered by exposure to repetitive or severe potentially morally injurious events (PMIEs) [10]. Individuals exposed to PMIEs are 2.6 times more likely to develop mental health disorders such as depression, anxiety, and posttraumatic stress disorder (PTSD) compared with those who are not exposed [11]. EMS providers must make instantaneous decisions in life-and-death situations and often find themselves taking actions that conflict with their moral beliefs [12]. Due to the nature of their work, they are repeatedly exposed to circumstances that can trigger MI [8]. Frequent exposure to PMIEs increases the risk of developing MI [10,13,14]. Its core features—guilt, shame, and feelings of betrayal toward institutions or leadership—not only contribute to mental health issues such as depression and anxiety [15] but also carry significant occupational and social consequences. EMS providers who experience MI report stronger intentions to leave the profession [16], along with increased rates of medical incidents and declines in care quality [17]. Understanding EMS providers' exposure to PMIEs and their association with mental health outcomes is therefore essential for prevention and intervention.

During health care service disruptions, such as the COVID-19 pandemic and government–medical community disputes, emergency medical resources may become scarce and hospitals may be unable to accommodate incoming patients. In these contexts, EMS providers often face prolonged prehospital treatment times and witness patient deterioration helplessly. This study defines such specific and clinically critical incidents—in which EMS providers fail to transport patients in a timely manner due to health care service disruptions and are forced to witness patient deterioration—as health care–related patient deterioration experiences (HPDEs). While PMIEs represent a broad psychological concept encompassing all events that violate an individual's moral beliefs, HPDEs constitute a specific, occupationally distinct type of event that EMS providers may experience as PMIEs [18].

Although recent research has begun to explore MI among EMS providers and health care professionals, studies specifically examining the association between HPDE exposure and mental health outcomes remain limited. Accordingly, we evaluated associations of HPDEs with personal burnout (PB; the degree of physical and psychological fatigue and exhaustion experienced by an individual), work-related burnout (WRB; the degree of physical and psychological fatigue and exhaustion an individual feels in relation to their work), citizen-related burnout (CRB; the degree of physical and psychological fatigue and exhaustion an individual feels in relation to work involving interactions with citizens, ie, clients), depressive symptoms, and sleep problems among EMS providers in South Korea during health care service disruptions.

Methods

Study Design and Population

This was a retrospective cross-sectional study. The survey was designed as an open web-based survey accessible to any EMS provider through a distributed link. Among the 14,310 EMS providers working in South Korea, 855 voluntarily participated in the survey, representing 6% of the total population. After excluding 1 participant who did not consent to participate and 9 participants with missing data regarding age and years of service, 845 participants were included in the analysis.

Data Source and Collection

The survey was conducted over a 2-month period from November 1 to December 31, 2024, targeting EMS providers nationwide through an online questionnaire. To account for internet security access issues at some provincial fire departments, both Naver and Google platforms were used in parallel. To reflect the working environment and health status of EMS providers during health care service disruptions caused by government–medical community disputes, as well as practical issues in emergency medical care, the validity of the questionnaire was reviewed through literature searches and interviews with field experts in emergency medical services, including EMS providers, emergency dispatching and operating centers, and a prehospital EMS policy officer. The final survey items were refined after the content had been modified through a pilot study.

XSL•FO
RenderX

The questionnaire, titled Korean EMS Provider's Working Environment in Situations of Medical Disruption, included items regarding the following five domains:

1. Experience with prehospital emergency medical activities related to health care service disruptions (transport difficulties, delayed transfer times, and exposure to HPDEs)
2. Work life (workplace social support, organizational culture, and job satisfaction)
3. Mental health (psychological well-being, psychological stress, work-related exhaustion, burnout, depressive symptoms, suicidal ideation, and sleep problems)
4. General health status (self-assessed health and dietary habits)
5. Demographic and sociological characteristics of the study population (sex, age, marital status, and work region)

To increase response rates, participation was encouraged through official letters, standardized study information sheets, and participation links sent to institutions employing EMS providers and firefighter unions. Institutional coordinators distributed these links to their affiliated EMS providers via emails and SMS messages. EMS providers who received the link reviewed the study information and voluntarily decided whether to participate. To prevent duplicate responses, participation was restricted to once per person. Furthermore, the survey was designed so that data were saved to the server only after participants responded to all items and clicked the Submit button, meaning incomplete responses were automatically excluded from the analysis.

Variables and Measures

The independent variable was the presence or absence of HPDEs, whereas the 3 dependent variables were burnout, depressive symptoms, and sleep problems. The measurement methods for each variable were as follows.

HPDE

In this study, an HPDE was operationally defined as the specific clinical experience of EMS providers who are forced to helplessly witness patient deterioration due to structural deficiencies in the health care system and delays in care. HPDEs refer to EMS providers who witness patient deterioration due to health care–related delays in care, corresponding to "omission" in the PMIE classification [18]. In the survey, participants were asked whether they had ever witnessed patient deterioration due to a health care system failure resulting from health care service disruption caused by government–medical community disputes that occurred after March 2024. Responses were categorized as "yes" for having such an experience at least once and "no" for not having that experience.

Burnout

Burnout was assessed using the Korean version of the Copenhagen Burnout Inventory [19], which consists of 3 domains: PB, WRB, and CRB [20]. For each item, participants could choose from 5 options: "always/almost always" (100 points), "often/very much" (75 points), "sometimes/somewhat" (50 points), "seldom/a little bit" (25 points), and "never/very little" (0 points). Average scores were calculated for each domain; higher scores indicated higher levels of burnout. An average score of 50 points or higher in each domain was considered indicative of burnout [21]. The Cronbach α for burnout was 0.96.

Depressive Symptoms

Depressive symptoms were assessed using the Center for Epidemiologic Studies Depression Scale, which measures experiences over the past week and consists of 20 items. For each item, respondents could choose from "less than 1 day per week," "1 - 2 days per week," "3 - 4 days per week," or "5 or more days per week.'" Responses were converted into scores of 0 to 3 points according to the items. When the total score for the 20 items was 16 points or higher, the respondent was considered to have depressive symptoms [22]. The Cronbach α for depressive symptoms was 0.93.

Sleep Problems

Sleep problems were assessed by evaluating how frequently sleep-related issues occurred over the prior 12 months. The survey included 3 questions, with response options of "daily," "several times a week," "several times a month," "rarely," and "never." Each response was converted into 0 to 4 points, and respondents with a total score of 6 points or higher were considered to have experienced sleep problems [23]. The Cronbach α for sleep problems was 0.86.

Statistical Analysis

Continuous variables were reclassified into categorical boundaries for analysis; specifically, age was grouped into 5- to 10-year intervals (20 - 29, 30 - 34, 35 - 39, 40 - 44, and ≥45 years), and years of service were grouped into specific intervals (<4, 5 - 9, 10 - 14, and ≥15 years). All categorical variables are presented as frequencies and percentages (n, %). χ^2 tests were performed to examine differences between participant characteristics and dependent variables. Logistic regression was conducted to determine odds ratios and 95% CIs for burnout, depressive symptoms, and sleep problems according to HPDE exposure status. In building the multivariable model, key demographic and occupational variables frequently used in previous studies of EMS providers and health care workers were comprehensively included as potential confounders. To prevent multicollinearity issues beforehand, correlations among covariates were examined. As a result, the correlation coefficients among age, years of service, and position were 0.7 or higher; therefore, only age—the variable that best reflects the representativeness of the model—was included. Additionally, after constructing the final multivariable model, multicollinearity was diagnosed by calculating the variance inflation factor (VIF), and it was confirmed that there was no multicollinearity issue as all VIF values were less than 5. Furthermore, since this study aimed to estimate the independent main effects of HPDE exposure, potential interactions between the primary independent variable and covariates were exploratorily evaluated. Although significant interactions were observed in some models, a main effects model excluding interaction terms was adopted as the primary analysis in the final multivariable model to focus on the primary hypothesis of main effects and to maintain the intuitiveness and consistency of result interpretation across the 5 dependent variable models.

XSL•FO
RenderX

However, the identified interaction results are described as exploratory findings in the Results section. The final regression model included sex, age, marital status, household size, educational level, household income, job position, current assignment, work type, and work region as confounders. The goodness-of-fit of the multivariable logistic regression model was evaluated using the Hosmer-Lemeshow test and χ^2 tests; *c* statistics were measured to assess the discriminative ability of the final model. Analyses were performed using SAS (version 9.4; SAS Institute), and the statistical significance level was set at $P<.05$.

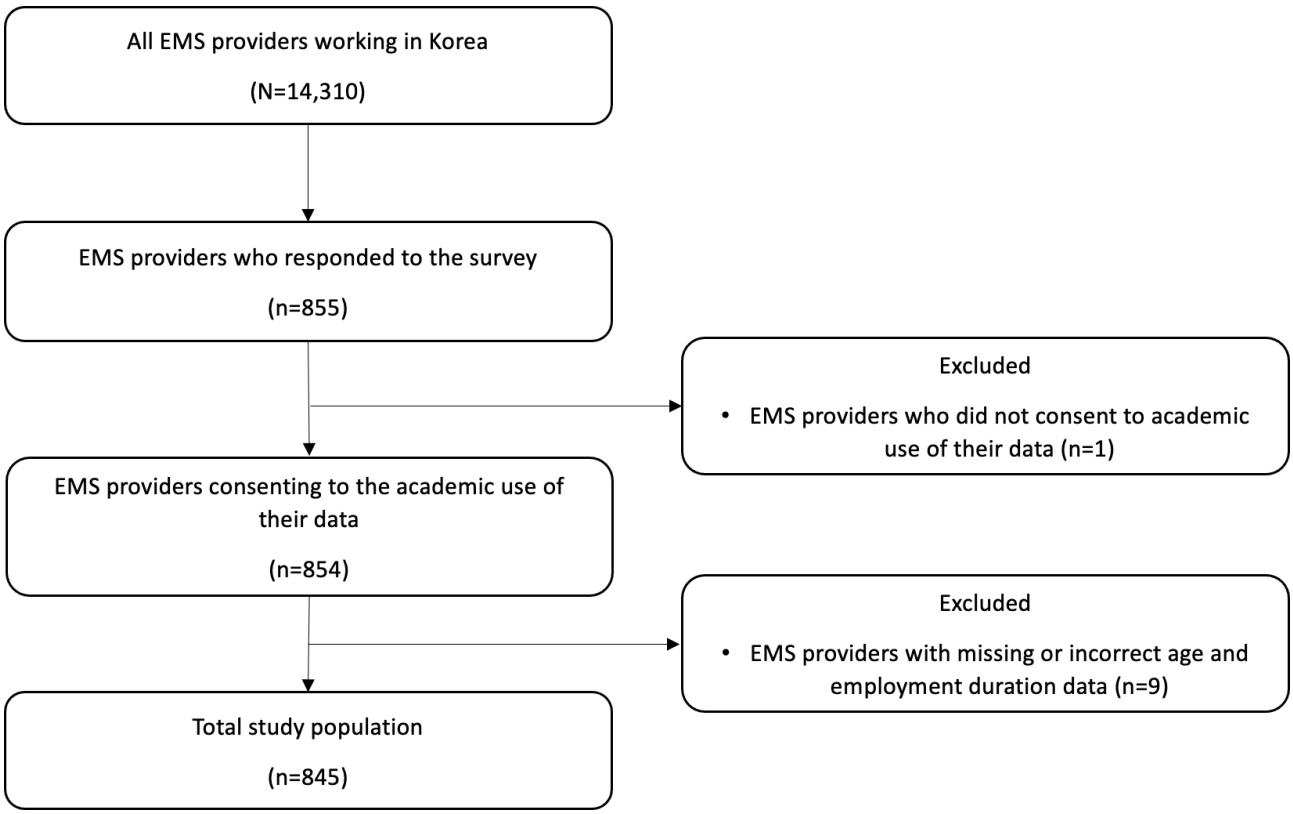
Ethical Considerations

This study was approved by the institutional review board of Eulji University (EUIRB2024–102). Before beginning the survey, all participants were provided with detailed information regarding the study purpose, procedures, data collection, and academic use. Electronic written informed consent was obtained from those who voluntarily agreed to participate and consented to the academic use of their data. The collected data were processed to be untraceable and unidentifiable to ensure the anonymity and confidentiality of the participants. All participants were recorded solely by a unique identification number assigned in the order of their participation. Research data were stored in encrypted files on a secure server with restricted access, and strict security measures were implemented to prevent anyone other than the research team from accessing the data. As a token of appreciation for survey participation, a small mobile beverage coupon was provided. The provision of contact information was optional and collected only from participants who wished to receive compensation. Ten participants were randomly selected from those who voluntarily provided their contact information to receive coupons. The collected contact information was managed separately from the anonymous research data and safely discarded once the purpose of compensation was fulfilled.

Results

Among the 845 study participants (Figure 1), 75% (n=626) reported HPDEs, which differed significantly by household size, current duties, and work region. The prevalences of mental health–related symptoms among all participants were as follows: PB 50.1% (n=423), WRB 46.9% (n=396), CRB 52.8% (n=446), depressive symptoms 19.8% (n=167), and sleep problems 46.8% (n=395). All of these outcomes were more prevalent in the HPDE group, with statistically significant differences (Table 1).

Figure 1. Flowchart of the study population selection among emergency medical service (EMS) providers in South Korea (November-December 2024).



XSL•FO
RenderX

Table . Demographic characteristics, occupational factors, and mental health status of emergency medical service (EMS) providers by health care–related patient deterioration experience (HPDE) in South Korea (November-December 2024).

Potential risk factors	Total (N=845), n (%)	HPDE (n=626), n (%)	No HPDE (n=219), n (%)	*P* value
Sex				.93
Male	635 (75.2)	470 (75.1)	165 (75.5)	
Female	210 (24.8)	156 (24.9)	54 (24.5)	
Age				.12
20 - 29	109 (12.9)	86 (13.7)	23 (10.5)	
30 - 34	335 (39.6)	234 (37.4)	101 (45.9)	
34 - 39	184 (21.8)	144 (23.0)	40 (18.2)	
40 - 44	154 (18.2)	112 (17.9)	42 (19.1)	
≥45	63 (7.5)	50 (8.0)	13 (6.3)	
Marital status				.52
Single	311 (36.8)	239 (38.2)	72 (32.9)	
Married	511 (60.5)	371 (59.2)	140 (63.9)	
Divorced, separated, or widowed	15 (1.8)	10 (1.6)	5 (2.3)	
Declined to answer	8 (0.9)	6 (1.0)	2 (0.9)	
Household size				.02
1 person	177 (20.9)	145 (23.2)	32 (14.6)	
2 persons	206 (24.3)	142 (22.7)	64 (29.2)	
3 persons	234 (27.9)	164 (26.1)	70 (32.0)	
4 persons	188 (22.2)	145 (23.2)	43 (19.6)	
5 or more persons	40 (4.7)	30 (4.8)	10 (4.6)	
Education level				.08
High school or less	24 (2.8)	21 (3.4)	3 (1.4)	
College	800 (94.8)	586 (93.6)	214 (98.2)	
Graduate school	18 (2.1)	17 (2.7)	1 (0.2)	
Declined to answer	3 (0.3)	2 (0.3)	1 (0.2)	
Household income (million ₩[d])				.72
<40	90 (10.7)	72 (11.5)	18 (8.2)	
40 - 49.99	156 (18.5)	119 (19.0)	37 (16.9)	
50 - 59.99	181 (21.4)	133 (21.2)	48 (21.9)	
60 - 69.99	110 (13.0)	79 (12.6)	31 (14.2)	
≥70	255 (30.2)	185 (29.6)	70 (32.0)	
Declined to answer	53 (6.2)	38 (6.1)	15 (6.8)	
Years of service				.23
<4	304 (35.9)	218 (34.8)	86 (39.1)	
5 - 9	292 (34.5)	214 (34.2)	78 (35.6)	
10 - 14	140 (16.6)	105 (16.8)	35 (15.9)	
≥15	109 (13.0)	89 (14.2)	20 (9.4)	
Job rank				.47
Sobang-sa[a] (firefighter)	255 (30.1)	181 (28.9)	74 (33.6)	

XSL•FO
RenderX

Potential risk factors	Total (N=845), n (%)	HPDE (n=626), n (%)	No HPDE (n=219), n (%)	*P* value
Sobang-gyo (fire engineer)	254 (30.0)	191 (30.5)	63 (28.6)	
Sobang-jang (fire lieutenant)	255 (30.1)	190 (30.4)	65 (29.5)	
Over sobang-wi (fire captain)	81 (9.8)	64 (10.2)	17 (8.3)	
Job position				.99
Ambulance crew	651 (77.0)	482 (77.0)	169 (76.8)	
Ambulance driver	164 (19.5)	122 (19.5)	42 (19.5)	
Other[b]	30 (3.5)	22 (3.5)	8 (3.7)	
Current assignment				.03
General EMS unit	619 (73.3)	445 (71.1)	174 (79.5)	
Special EMS unit	213 (25.2)	169 (27.0)	44 (20.0)	
EMS command and control center	13 (1.5)	12 (1.9)	1 (0.5)	
Work type				.31
Day shift	15 (1.8)	9 (1.5)	6 (2.7)	
Shift work (3 teams, 2 shifts)	450 (53.2)	340 (54.3)	110 (50.2)	
Other shift work[c]	380 (45.0)	277 (44.2)	103 (47.1)	
Work region				.003
Urban	137 (16.2)	113 (18.1)	24 (11.0)	
Suburban	343 (40.6)	262 (41.9)	81 (37.0)	
Rural	365 (43.2)	251 (40.0)	114 (52.0)	
Personal burnout				<.001
Yes	423 (50.1)	345 (55.1)	78 (35.9)	
No	422 (49.9)	281 (44.9)	141 (64.1)	
Work-related burnout				<.001
Yes	396 (46.9)	331 (52.9)	65 (30.0)	
No	449 (53.1)	295 (47.1)	154 (70.0)	
Citizen-related burnout				<.001
Yes	446 (52.8)	368 (58.8)	78 (35.9)	
No	399 (47.2)	258 (41.2)	141 (64.1)	
Depressive symptoms				<.001
Yes	167 (19.8)	142 (22.7)	25 (11.4)	
No	678 (80.2)	484 (77.3)	194 (88.6)	
Sleep problems				.001
Yes	395 (46.8)	313 (50.0)	82 (37.7)	
No	450 (53.2)	313 (50.0)	137 (62.3)	

[a]Lowest.

[b]Other includes tasks such as report reception, counseling, and administrative duties.

[c]Other shift work includes 4 teams and 2 shifts.

[d]1 ₩=US $0.00067 as of July 15, 2026

Health care service disruptions were categorized into 3 domains: EMS providers' workload burden, fire service support, and experiences of verbal violence (Table 2). In the workload burden domain, 98.8% (n=835) encountered difficulty arranging

XSL•FO
RenderX

hospital transport, whereas 95.5% (n=807) faced refusal from nearby hospitals and were consequently required to transfer patients to facilities located more than 10 km away. Among these transfers, 89.6% (n=758) occurred at a rate of 1 to 2 long-distance transports per day on average. Notably, EMS providers performing 5 or more long-distance transports daily exhibited more than a 10-fold increase in the odds of HPDE exposure compared with those performing fewer transports. Furthermore, 17.7% (n=150) experienced waiting times for hospital selection that exceeded predisruption conditions by more than 1 hour, and 95.3% (n=805) reported communication difficulties with hospital medical staff. Together, these findings indicate that increasing long-distance transfer frequencies and the resulting workload burden are significantly associated with elevated HPDE exposure likelihood.

In a multivariable logistic regression adjusted for sociodemographic variables, HPDE exposure was associated with significantly higher odds of mental health outcomes (Table 3). Among sociodemographic characteristics, sex and work region were significantly correlated with burnout, depressive symptoms, and sleep problems. Female providers showed elevated odds of PB, WRB, and depressive symptoms. Compared with rural areas, suburban providers demonstrated higher odds of PB, WRB, CRB, and sleep problems, whereas urban providers exhibited the greatest odds. Additionally, participants aged 30 to 34 years had higher odds of depressive symptoms. Detailed analysis results for all confounders included in the model are presented in Multimedia Appendix 1. Furthermore, exploratory analysis of interactions between the primary independent variable and covariates revealed no significant interactions in the 3 burnout models; however, significant interactions with HPDE exposure were observed for sex in the depressive symptom model (P=.007) and for age group in the sleep problem model (P=.046). Nevertheless, these interaction effects represented minor findings limited to only a few models. To clearly convey the independent main effects of HPDE exposure on the entire paramedic population—which is the primary objective of this study—and to maintain consistency in interpreting all models, the final results were presented based on the main effects model.

XSL•FO
RenderX

Table . Association between health care service disruption characteristics and health care–related patient deterioration experiences (HPDEs) among emergency medical service (EMS) providers in South Korea (November-December 2024).

Variable	Total (N=845), n (%)	HPDE (n=626), n (%)	No HPDE (n=219), n (%)	*P* value
Workload				
Difficulty in hospital selection				<.001
Yes	835 (98.8)	624 (99.7)	211 (96.4)	
No	10 (1.2)	2 (0.3)	8 (3.6)	
Long-distance transfer due to nearby hospital refusal				<.001
Yes	807 (95.5)	613 (97.9)	194 (75.0)	
No	38 (4.5)	13 (2.1)	25 (25.0)	
Average daily long-distance transfers				<.001
1 - 2 times	506 (59.8)	362 (57.8)	144 (65.8)	
3 - 4 times	161 (19.0)	140 (22.4)	21 (9.6)	
≥5 times	91 (10.8)	83 (13.3)	8 (3.7)	
Not applicable	87 (10.4)	41 (6.5)	46 (20.9)	
Delay over 1 hour for hospital selection				<.001
No	72 (8.5)	30 (4.8)	42 (19.1)	
<30%	365 (43.3)	244 (39.0)	121 (55.5)	
30% - 50%	258 (30.5)	217 (34.7)	41 (18.6)	
≥50%	150 (17.7)	135 (21.5)	15 (6.8)	
Prolonged return-to-base time				<.001
<1 hour	465 (55.0)	304 (48.6)	161 (73.5)	
≥1 hour	380 (45.0)	322 (51.4)	58 (26.5)	
Private ambulance request at scene				<.001
Yes	249 (29.4)	216 (34.5)	33 (15.0)	
No	596 (70.6)	410 (65.5)	186 (85.0)	
Reason for increased direct medical control				<.001
No increase	200 (23.6)	103 (16.5)	97 (44.3)	
Increased for critical patient emergency care	144 (17.0)	114 (18.2)	30 (13.7)	
Increased for hospital selection	312 (36.9)	250 (39.9)	62 (28.3)	
Increased for other reasons	189 (22.5)	159 (25.4)	30 (13.7)	
Urgent administrative task experience				<.001
Yes	661 (78.2)	534 (85.3)	127 (58.0)	
No	184 (21.8)	92 (14.7)	92 (42.0)	
Communication difficulty with hospital staff				<.001
Yes	805 (95.3)	620 (99.0)	185 (84.5)	
No	40 (4.7)	6 (1.0)	34 (15.5)	
Fire department institutional support				
Sufficiency of fire department institutional support				<.001
Insufficient	639 (75.6)	515 (82.3)	124 (56.6)	
Neutral	178 (21.1)	97 (15.5)	81 (37.0)	
Sufficient	28 (3.3)	14 (2.2)	14 (6.4)	

XSL•FO
RenderX

Variable	Total (N=845), n (%)	HPDE (n=626), n (%)	No HPDE (n=219), n (%)	*P* value
Most needed stage for fire department support				.97
From call to on-scene arrival	109 (12.9)	80 (12.8)	29 (13.2)	
From patient contact to hospital selection	691 (81.8)	513 (81.9)	178 (81.3)	
In other areas	45 (5.3)	33 (5.3)	12 (5.5)	
Verbal abuse from patients or caregivers				<.001
Yes	490 (58.0)	416 (66.5)	74 (34.1)	
No	355 (42.0)	210 (33.5)	145 (65.9)	

Table . Multivariable logistic regression analysis of factors associated with mental health problems among emergency medical service (EMS) providers in South Korea (November-December 2024). Adjusted for covariates including sex, age, marital status, household size, education level, household income, job position, current assignment, work type, and work region. The full regression model including all covariates is available in Table S1 in Multimedia Appendix 1. AOR: adjusted odds ratio; HPDE: health care–related patient deterioration experience.

HPDE[b]	Personal burnout, AOR[a] (95% CI)	Work-related burnout, AOR (95% CI)	Citizen-related burnout, AOR (95% CI)	Depressive symptoms, AOR (95% CI)	Sleep problems, AOR (95% CI)
No	Reference	Reference	Reference	Reference	Reference
Yes	2.05 (1.46-2.87)	2.66 (1.86-3.78)	2.74 (1.95-3.86)	2.22 (1.36-3.60)	1.63 (1.17-2.27)

[a]AOR: adjusted odds ratio.

[b]HPDE: health care system–related patient deterioration experience.

Discussion

This study showed that HPDEs among EMS providers during health care service disruptions have significant associations with burnout, depressive symptoms, and sleep problems. To understand the mechanism through which these clinical experiences inflict severe psychological impacts on EMS providers, it is necessary to interpret HPDEs within the theoretical framework of PMIEs.

PMIEs were first classified by Litz et al [10], who identified 3 common categories: commission, omission, and betrayal. Commission refers to actions that directly violate one's moral beliefs or values; omission denotes the failure to perform actions one believes to be morally right, often due to situational constraints; and betrayal occurs when one feels let down by others, such as supervisors or colleagues, who act against one's moral convictions or compel unethical behavior [18]. In this context, HPDEs—the core variable of this study—can be interpreted as complex PMIEs that encompass both the emotion of omission, which stems from failing to fulfill the duty of patient protection due to situational constraints, and the emotion of betrayal, rooted in the feeling that the state and health care system have neglected both EMS providers and patients [8,24,25]. During the strike, most EMS providers surveyed were unable to secure nearby receiving hospitals and were forced to transport patients to distant facilities, resulting in approximately 75% witnessing patient deterioration helplessly. These repeated experiences of long-distance transport and patient deterioration do not stem from a lack of individual competence among EMS providers, but rather originate from uncontrollable structural limitations. The repeated experience of having the primary mission of saving patient lives through rapid and effective treatment blocked by external factors ultimately undermines their moral beliefs and professional values.

EMS providers working in urban areas exhibited a higher prevalence of mental health issues, including burnout, depressive symptoms, and sleep problems, likely due to shift work and increased call volumes. Shift work, a common scheduling pattern among EMS personnel, contributes to both physical and psychological strain; populations engaged in shift work demonstrate elevated rates of depressive symptoms, sleep problems, anxiety, and fatigue compared with the general population [26,27]. Urban regions experience frequent late-night emergency incidents due to high population density and diverse lifestyle patterns, disproportionately exposing EMS providers to violent encounters and psychiatric emergencies. Frequent nighttime deployments impair sleep quality, whereas repeated exposure to high-risk scenes sustains heightened psychological arousal, thereby threatening overall mental health. The findings of this study support the conclusions of previous research that these unique urban occupational characteristics and heavy workloads act as core risk factors for PTSD, depressive symptoms, anxiety, burnout, and sleep problems [28].

Due to the nature of their work, EMS providers exhibit higher prevalence rates of mental health disorders compared with the general population, even in the absence of special situations such as the COVID-19 pandemic [29-32], and burnout and depressive symptoms are particularly prevalent [30,33,34]. Sleep problems due to shift work are associated with higher likelihoods of PTSD and depressive symptoms [27,35], whereas fatigue caused by sleep deprivation is linked to elevated rates of injury and medical errors for EMS providers [36]. In addition to this, exposure to MI and PMIEs further exacerbates these

XSL•FO
RenderX

underlying vulnerabilities [37]. Health care workers facing PMIEs experience more severe depressive symptoms, burnout, sleep problems, and anxiety [38,39], along with greater turnover rates [40].

Such exposure to PMIEs can be particularly exacerbated during health care service disruptions. During the COVID-19 pandemic, health care workers, including EMS providers, were exposed to severe PMIEs due to delays in patient acceptance caused by critical shortages of hospital beds and medical resources [20,37], resulting in a significant increase in the prevalence of burnout and depressive symptoms [32,41,42]. In this study, EMS providers exposed to HPDE reported significantly higher rates of burnout, depressive symptoms, and sleep problems relative to those without such exposure.

Although the COVID-19 pandemic and the 2024 government–medical community dispute arose from different causes, both situations are comparable in that they led to the identical consequence of medical institutions being unable to accept patients due to health care service disruptions. This suggests that PMIEs, which inflict critical impacts on EMS providers, are not unique phenomena occurring only during infectious disease outbreaks. Regardless of the underlying cause, if a health care service disruption occurs, EMS providers can be exposed to severe PMIEs at any time. Therefore, to minimize psychological damage to EMS providers during unavoidable health care service disruptions, it is required to establish a systematic framework capable of early detection and response to PMIEs. While EMS providers inherently work in unpredictable and uncontrollable environments, exposure to HPDEs and PMIEs driven by these structural factors can cause irreversible damage to their mental health. Therefore, to protect the mental health of EMS personnel and prevent turnover, proactive and systematic organizational psychological intervention strategies—rather than reactive measures—are essential. Specifically, an organizational safety net must be established to proactively incorporate moral injury screening tools—such as the Moral Injury Events Scale and Moral Injury and Distress Scale for tracking event exposure [43,44], alongside the Moral Injury Outcome Scale and Moral Injury Symptom Scale-Healthcare Professionals for clinical symptom assessment [45,46]—into the routine health examinations of EMS providers to identify high-risk individuals early [47] and immediately provide specialized psychological counseling effective in addressing guilt and shame to those identified. Such organizational interventions will contribute to reducing the prevalence of mental disorders and enhancing the professional sustainability of EMS providers.

Although this study holds significant academic and clinical value as an investigation into the mental health outcomes of EMS providers exposed to HPDEs during government–medical community disputes, and provides policy-relevant evidence to improve emergency medical system delivery and the working conditions of EMS providers in scenarios involving health care service disruptions, several limitations must be acknowledged. First, despite extensive efforts to increase participation, including collaboration with firefighter unions, the response rate remained low at approximately 6% of all EMS providers. This voluntary participation–based sampling method can introduce selection bias; it is possible that providers experiencing higher levels of occupational stress or psychological distress participated more actively in the survey, potentially leading to an overestimation of the prevalence of mental health problems. Additionally, the geographic concentration in metropolitan areas (Seoul and Gyeonggi Province) limits the generalizability of the results to the entire domestic EMS provider population. Second, reliance on self-reported data introduces potential recall bias. In particular, measuring HPDE exposure with a single self-reported item restricts reliability and validity, and the possibility of misclassification due to respondents' subjective memory and personal interpretation cannot be ruled out. Although this study used an intuitive item to rapidly capture urgent health care service disruption scenarios, it is limited by its inability to reflect more precise, multidimensional measurements such as the frequency, duration, and intensity of HPDE exposure. Third, the retrospective cross-sectional design prevents definitive conclusions about causal relationships between HPDE exposure and subsequent mental health outcomes. In future studies, the introduction of objective and multidimensional measurement tools to supplement the single item is needed, alongside large-scale, nationally representative studies to generate robust evidence for emergency care and working environment policy improvements.

In conclusion, regardless of the cause, unavoidable health care service disruptions can expose EMS providers to severe psychological crises, including MI. To prevent and minimize such occupational damage, it is urgently required to officially include moral injury screening tests, such as the Moral Injury Outcome Scale [45,47], in the routine health examinations of EMS providers and establish customized psychological treatment and management programs that support healing and recovery from MI.

Acknowledgments

The authors sincerely thank all the prehospital emergency medical service (EMS) providers who participated in the survey during the 2024 health care disruption.

Funding

This research was supported by 3 funding sources: the Eulji University Researcher Academic Research Fund in 2026 (EJRG-26-21), the National Research Foundation of Korea (NRF) funded by the Korean government (Ministry of Science and ICT)

XSL•FO
RenderX

(RS-2024-00350688), and the Research Grant of Kongju National University. The funders had no role in the study design, data collection, analysis, interpretation, or writing of the manuscript.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: MSC, SJK, HW
Data curation: MSC, SJK
Formal analysis: MSC, SJK
Funding acquisition: HW, SJK
Investigation: MSC, SJK
Methodology: MSC, HW, SJK, YM
Project administration: HW, SJK
Resources: SJK, HW
Supervision: HW, SJK
Validation: SJK
Visualization: MSC, SJK
Writing – original draft: MSC, SJK
Writing – review & editing: HW, SJK, YM.
MSC and SJK contributed equally as co-first authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Full multivariable logistic regression models for mental health outcomes, including health care–related patient deterioration experience and covariates, among emergency medical service providers in South Korea, November-December 2024.
[DOCX File, 44 KB - publichealth_v12i1e90130_app1.docx]

Checklist 1
STROBE checklist.
[DOCX File, 25 KB - publichealth_v12i1e90130_app2.docx]

References

1. Yonhap News Agency. Govt to cut trainee doctor quota unless hospitals accept resignations by next Monday. The Korea Herald. 2024 Jul 9. URL: https://www.koreaherald.com/article/3431116 [accessed 2025-06-15]
2. 119 emergency medical technicians' field care guidelines [Article in Korean]. National Fire Agency 119. 2024 Jan 23. URL: https://tinyurl.com/3ee2hvtw [accessed 2025-06-15]
3. Hong WK. Post-residency resignation aftermath: emergency room treatment restriction messages exceeded 110,000 last year [Article in Korean]. Doctors News. 2025 Oct 1. URL: https://www.doctorsnews.co.kr/news/articleView.html?idxno=161459 [accessed 2026-05-01]
4. Yoon MS. Elderly woman dies of heart attack after ER refusal. The Korea Herald. 2024 Mar 27. URL: https://www.koreaherald.com/article/3355161 [accessed 2025-06-15]
5. Lee SJ. Exhausted emergency rooms on verge of collapse amid protracted doctor shortage. Korea JoongAng Daily. 2024 Aug 28. URL: https://koreajoongangdaily.joins.com/news/2024-08-28/national/socialAffairs/Exhausted-emergency-rooms-on-verge-of-collapse-amid-protracted-doctor-shortage/2122766 [accessed 2025-06-17]
6. Dean W, Morris D, Manzur MK, Talbot S. Moral injury in health care: a unified definition and its relationship to burnout. Fed Pract 2024 Apr;41(4):104-107. [doi: 10.12788/fp.0467] [Medline: 39399281]
7. Weiner S. Moral injury harms providers and patients. Association of American Medical Colleges. 2023 Nov 5. URL: https://www.aamc.org/news/moral-injury-harms-providers-and-patients [accessed 2025-06-17]
8. Rodrigues S, Mercier JM, McCall A, Nannarone M, Hosseiny F. "Against everything that got you into the job": experiences of potentially morally injurious events among Canadian public safety personnel. Eur J Psychotraumatol 2023;14(2):2205332. [doi: 10.1080/20008066.2023.2205332] [Medline: 37170904]
9. Drescher KD, Foy DW, Kelly C, Leshner A, Schutz K, Litz B. An exploration of the viability and usefulness of the construct of moral injury in war veterans. Traumatology 2011 Mar 1;17(1):8-13. [doi: 10.1177/1534765610395615]
10. Litz BT, Stein N, Delaney E, et al. Moral injury and moral repair in war veterans: a preliminary model and intervention strategy. Clin Psychol Rev 2009 Dec;29(8):695-706. [doi: 10.1016/j.cpr.2009.07.003] [Medline: 19683376]

XSL•FO
RenderX

11. Lamb D, Gnanapragasam S, Greenberg N, et al. Psychosocial impact of the COVID-19 pandemic on 4378 UK healthcare workers and ancillary staff: initial baseline data from a cohort study collected during the first wave of the pandemic. Occup Environ Med 2021 Nov;78(11):801-808. [doi: 10.1136/oemed-2020-107276] [Medline: 34183447]
12. Knobloch LK, Owens JL. Moral injury among first responders: experience, effects, and advice in their own words. Psychol Serv 2024 Aug;21(3):500-508. [doi: 10.1037/ser0000838] [Medline: 38271025]
13. Bartlett B. Moral injury: what is it and how does it impact first responders? EMS1. 2023 Mar 15. URL: https://tinyurl.com/4263vmsz [accessed 2025-06-17]
14. Benatov J, Zerach G, Levi-Belz Y. Moral injury, depression, and anxiety symptoms among health and social care workers during the COVID-19 pandemic: the moderating role of belongingness. Int J Soc Psychiatry 2022 Aug;68(5):1026-1035. [doi: 10.1177/00207640221099421] [Medline: 35634658]
15. Jinkerson JD. Defining and assessing moral injury: a syndrome perspective. Traumatology 2016;22(2):122-130. [doi: 10.1037/trm0000069]
16. Powell JR, Gage CB, Crowe RP, et al. National evaluation of emergency medical services clinician burnout and workforce-reducing factors. J Am Coll Emerg Physicians Open 2025 Jan 10;6(1):100024. [doi: 10.1016/j.acepjo.2024.100024] [Medline: 40012660]
17. Spates CR, Callaway KL. Moral injury in healthcare: causes, consequences, and solutions. Moral Injury Spectrum. 2024. URL: https://moralinjuryspectrum.com/moral-injury-in-healthcare/ [accessed 2025-06-17]
18. Understanding moral injury. Government of Canada. 2025 Feb 19. URL: https://www.canada.ca/en/department-national-defence/corporate/reports-publications/health/understanding-moral-injury.html [accessed 2025-06-17]
19. Kristensen TS, Borritz M, Villadsen E, Christensen KB. The Copenhagen Burnout Inventory: a new tool for the assessment of burnout. Work & Stress 2005 Jul;19(3):192-207. [doi: 10.1080/02678370500297720]
20. Kim JH, Yoon J, Kim SJ, Kim JY, Bahk J, Kim SS. Lack of compensation for COVID-19-related overtime work and its association with burnout among EMS providers in Korea. Epidemiol Health 2023;45:e2023058. [doi: 10.4178/epih.e2023058] [Medline: 37340845]
21. Kim JH, Yoon J, Kim SJ, Kim JY, Kim SS. People-oriented culture and its association with burnout, depressive symptoms, and sleep problems during COVID-19 pandemic among EMS providers in Korea. J Occup Environ Med 2024 Aug 1;66(8):e359-e364. [doi: 10.1097/JOM.0000000000003154] [Medline: 38769075]
22. Lewinsohn PM, Seeley JR, Roberts RE, Allen NB. Center for Epidemiologic Studies Depression Scale (CES-D) as a screening instrument for depression among community-residing older adults. Psychol Aging 1997 Jun;12(2):277-287. [doi: 10.1037/0882-7974.12.2.277] [Medline: 9189988]
23. Cho SS, Kang MY. Association between occupational exposure to chemical or physical factors and sleep disturbance: an analysis of the fifth Korean Working Conditions Survey. Sleep Health 2022 Oct;8(5):521-527. [doi: 10.1016/j.sleh.2022.06.004] [Medline: 35927181]
24. Amsalem D, Lazarov A, Markowitz JC, et al. Psychiatric symptoms and moral injury among US healthcare workers in the COVID-19 era. BMC Psychiatry 2021 Nov 5;21(1):546. [doi: 10.1186/s12888-021-03565-9] [Medline: 34740357]
25. Senger AR, Torres D, Ratcliff CG. Potentially morally injurious events as a mediator of the association of gratitude and mindfulness with distress. Psychol Trauma 2024 Feb;16(2):262-269. [doi: 10.1037/tra0001233] [Medline: 35357878]
26. Courtney JA, Francis AJP, Paxton SJ. Caring for the country: fatigue, sleep and mental health in Australian rural paramedic shiftworkers. J Community Health 2013 Feb;38(1):178-186. [doi: 10.1007/s10900-012-9599-z] [Medline: 22843315]
27. Khan WAA, Conduit R, Kennedy GA, Jackson ML. The relationship between shift-work, sleep, and mental health among paramedics in Australia. Sleep Health 2020 Jun;6(3):330-337. [doi: 10.1016/j.sleh.2019.12.002] [Medline: 32223969]
28. Almasry H. Comparative analysis of mental health challenges and shift work impact on urban vs. rural emergency responders. AJHMN 2025 Apr 9;11(1):44-74. [doi: 10.47672/ajhmn.2669]
29. Eiche C, Birkholz T, Jobst E, Gall C, Prottengeier J. Well-being and PTSD in German emergency medical services - a nationwide cross-sectional survey. PLoS ONE 2019;14(7):e0220154. [doi: 10.1371/journal.pone.0220154] [Medline: 31335903]
30. Petrie K, Milligan-Saville J, Gayed A, et al. Prevalence of PTSD and common mental disorders amongst ambulance personnel: a systematic review and meta-analysis. Soc Psychiatry Psychiatr Epidemiol 2018 Sep;53(9):897-909. [doi: 10.1007/s00127-018-1539-5] [Medline: 29869691]
31. Popa F, Raed A, Purcarea VL, Lală A, Bobirnac G. Occupational burnout levels in emergency medicine–a nationwide study and analysis. J Med Life 2010;3(3):207-215. [Medline: 20945809]
32. Huang G, Chu H, Chen R, et al. Prevalence of depression, anxiety, and stress among first responders for medical emergencies during COVID-19 pandemic: a meta-analysis. J Glob Health 2022 Jul 25;12:05028. [doi: 10.7189/jogh.12.05028] [Medline: 35871411]
33. Bentley MA, Crawford JM, Wilkins JR, Fernandez AR, Studnek JR. An assessment of depression, anxiety, and stress among nationally certified EMS professionals. Prehosp Emerg Care 2013 Jul;17(3):330-338. [doi: 10.3109/10903127.2012.761307] [Medline: 23414106]
34. Reardon M, Abrahams R, Thyer L, Simpson P. Review article: prevalence of burnout in paramedics: a systematic review of prevalence studies. Emerg Med Australas 2020 Apr;32(2):182-189. [doi: 10.1111/1742-6723.13478] [Medline: 32067408]

XSL•FO
RenderX

35. Nguyen E, Meadley B, Harris R, et al. Sleep and mental health in recruit paramedics: a 6-month longitudinal study. Sleep 2023 Aug 14;46(8):zsad050. [doi: 10.1093/sleep/zsad050] [Medline: 36861384]
36. Patterson PD, Weaver MD, Frank RC, et al. Association between poor sleep, fatigue, and safety outcomes in emergency medical services providers. Prehosp Emerg Care 2012 Jan;16(1):86-97. [doi: 10.3109/10903127.2011.616261] [Medline: 22023164]
37. Coimbra BM, Zylberstajn C, van Zuiden M, et al. Moral injury and mental health among health-care workers during the COVID-19 pandemic: meta-analysis. Eur J Psychotraumatol 2023;15(1):2299659. [doi: 10.1080/20008066.2023.2299659] [Medline: 38189775]
38. Ritchie K, D'Alessandro-Lowe AM, Brown A, et al. The hidden crisis: understanding potentially morally injurious events experienced by healthcare providers during COVID-19 in Canada. Int J Environ Res Public Health 2023 Mar 9;20(6):4813. [doi: 10.3390/ijerph20064813] [Medline: 36981725]
39. Hegarty S, Lamb D, Stevelink SAM, et al. "It hurts your heart": frontline healthcare worker experiences of moral injury during the COVID-19 pandemic. Eur J Psychotraumatol 2022 Oct 18;13(2):2128028. [doi: 10.1080/20008066.2022.2128028] [Medline: 36276556]
40. Usset TJ, Baker LD, Griffin BJ, et al. Burnout and turnover risks for healthcare workers in the United States: downstream effects from moral injury exposure. Sci Rep 2024 Oct 22;14(1):24915. [doi: 10.1038/s41598-024-74086-0] [Medline: 39438496]
41. Lamas-Mendoza MDM, Fernandez-Alonso J, Ballesteros-Peña S, Gravina L. Factors related to sleep quality in emergency nedical technicians and their sleep habits during COVID-19 pandemic [Article in Spanish]. Rev Esp Salud Publica 2023 May 31;97:e202305043. [Medline: 37254874]
42. Kwon JD. Public health workers overworked and underpaid in struggle to fight COVID-19. Hankyoreh. 2020 Jul 4. URL: https://english.hani.co.kr/arti/english_edition/e_national/952235.html [accessed 2025-06-21]
43. Nash WP, Marino Carper TL, Mills MA, Au T, Goldsmith A, Litz BT. Psychometric evaluation of the Moral Injury Events Scale. Mil Med 2013 Jun;178(6):646-652. [doi: 10.7205/MILMED-D-13-00017] [Medline: 23756071]
44. Norman SB, Griffin BJ, Pietrzak RH, McLean C, Hamblen JL, Maguen S. The Moral Injury and Distress Scale: psychometric evaluation and initial validation in three high-risk populations. Psychol Trauma 2024 Feb;16(2):280-291. [doi: 10.1037/tra0001533] [Medline: 37347882]
45. Litz BT, Plouffe RA, Nazarov A, et al. Defining and assessing the syndrome of moral injury: initial findings of the Moral Injury Outcome Scale Consortium. Front Psychiatry 2022 Jul 5;13:923928. [doi: 10.3389/fpsyt.2022.923928] [Medline: 35873252]
46. Mantri S, Lawson JM, Wang Z, Koenig HG. Identifying moral injury in healthcare professionals: the Moral Injury Symptom Scale-HP. J Relig Health 2020 Oct;59(5):2323-2340. [doi: 10.1007/s10943-020-01065-w] [Medline: 32681398]
47. Jeong W, Kim YA, Song SY, et al. Cohort profile: the Registry-based Epidemiological Study of Cancer in Fire Unit and Emergency Officers (RESCUE) cohort. Int J Epidemiol 2024 Oct 13;53(6):dyae143. [doi: 10.1093/ije/dyae143] [Medline: 39460562]

Abbreviations

CRB: citizen-related burnout
EMS: emergency medical service
HPDE: health care–related patient deterioration experience
MI: moral injury
PB: personal burnout
PMIE: potentially morally injurious event
PTSD: posttraumatic stress disorder
VIF: variance inflation factor
WRB: work-related burnout

Edited by A Mavragani; submitted 22.Dec.2025; peer-reviewed by M Tusconi, S Ham, S Kim; revised version received 23.May.2026; accepted 30.Jun.2026; published 27.Jul.2026.

Please cite as:
Choi MS, Kim SJ, Moon Y, Woo H
Associations of Health Care System Deficiency–Related Patient Deterioration With Burnout, Depressive Symptoms, and Sleep Problems Among Emergency Medical Service Providers: Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e90130
URL: https://publichealth.jmir.org/2026/1/e90130
doi:10.2196/90130

XSL•FO
RenderX

© Min Seung Choi, Soo Jin Kim, Yonghwan Moon, Hyekyung Woo. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 27.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Impact of Different COVID-19 Pandemic Phases on Emergency Department Length of Stay and Intensive Care Unit Admission Risk: Single-Center Retrospective Cohort Study

Zehua Wang, MS; Xiangmei Zhao, PhD

Emergency Department, Henan Provincial People's Hospital, People's Hospital of Zhengzhou University, 7 Weiwu Road, Zhengzhou, Henan, China

Corresponding Author:
Xiangmei Zhao, PhD
Emergency Department, Henan Provincial People's Hospital, People's Hospital of Zhengzhou University, 7 Weiwu Road, Zhengzhou, Henan, China

Abstract

Background: COVID-19 profoundly affected emergency department (ED) services and patient outcomes. However, changes in ED length of stay (LOS) and intensive care unit (ICU) admission risk before, during, and after the pandemic remain poorly understood.

Objective: We investigated and compared ED LOS and ICU admission risk among patients in the ED resuscitation area in 2019, 2021, and 2023, respectively before, during, and after the pandemic.

Methods: This single-center retrospective cohort study enrolled patients in the ED resuscitation area of a regional emergency center in central China. ED LOS was compared across pandemic phases using the Kruskal-Wallis *H* test with Bonferroni correction. Cox proportional hazards models were applied to evaluate the association between phases and ICU admission risk. Sensitivity analyses included 2021 stratification, exclusion of extreme ED LOS values, and E-value analysis to validate result robustness.

Results: A total of 44,161 patients were included. Stratified analyses demonstrated significant temporal variations in ED LOS across disposition groups. Among discharged patients, median ED LOS increased from a prepandemic 248 (IQR 149-404.8) minutes to a midpandemic 262.0 (IQR 153-423.8) minutes (P=.047) and remained elevated in the postpandemic phase (median 258, IQR 152 - 451 minutes; P<.001). Among patients admitted to general wards, median ED LOS nearly doubled during the pandemic (from 186, IQR 96-431 to 350, IQR 188-610 minutes) and partially decreased to a postpandemic 293 (IQR 113-840) minutes (all P<.001). Among ICU-admitted patients, median ED LOS increased from a prepandemic 95 (IQR 57-169) minutes to a midpandemic 171 (IQR 77-306) minutes and a postpandemic 105 (IQR 50-215) minutes (all P<.003). The overall ICU admission rate was 15.8% (6982/44,161). Multivariate Cox regression with 2019 as the reference showed that the postpandemic phase was independently associated with an increased ICU admission risk (hazard ratio [HR] 1.079, 95% CI 1.012-1.150; P=.02), whereas the pandemic phase was not (HR 1.032, 95% CI 0.967-1.100; P=.34). In female patients, the postpandemic phase was associated with a higher ICU admission risk (HR 1.143, 95% CI 1.027-1.273; P=.02). All sensitivity analyses verified the robustness of the main findings (all P<.05).

Conclusions: The COVID-19 pandemic exerted differential effects on ED LOS across disposition groups. The postpandemic period was associated with an increased risk of ICU admission, especially among female patients, whereas no significant change was observed during the pandemic. As a novel gender-stratified analysis covering the full pandemic timeline, this study distinguishes itself from prior research by revealing the sustained impacts of pandemic control policies on emergency care. Our findings deliver key evidence for patient risk stratification and offer practical guidance to optimize health care resource allocation in the postpandemic era. Health care systems should strengthen early identification of high-risk ED patients and rationalize critical care resource deployment accordingly.

(JMIR Public Health Surveill 2026;12:e87597) doi:10.2196/87597

KEYWORDS

COVID-19 pandemic; emergency department; ED; length of stay; LOS; intensive care unit admission risk; ICU admission risk; retrospective cohort study

Introduction

Since 2020, COVID-19 has spread rapidly worldwide, triggering a global pandemic and posing a substantial threat to public health [1]. In response to the outbreak, governmental authorities implemented a series of public health interventions, including large-scale polymerase chain reaction screening, the establishment of fever clinics and isolation wards, the promotion

XSL•FO
RenderX

of telemedicine services, and the enforcement of traffic restrictions, social distancing, and home isolation policies [2]. Public health emergencies frequently alter individual health care–seeking behaviors and disease management patterns [3]. As a frontline component of the health care system, emergency medical services (EMS) and emergency department (ED) operations have undergone substantial restructuring during the COVID-19 pandemic, with notable shifts in patient demographics, risk of intensive care unit (ICU) admission, and clinical outcomes. Previous studies have demonstrated that the incidence of stroke presentations to EDs decreased significantly during the pandemic, accompanied by elevated mortality rates [4]. The pandemic has also been associated with a relative increase in EMS activations related to acute cardiac events [5]. In South Korea, the pandemic resulted in a marked reduction in the number of patients transferred to ICUs via EDs, along with increased in-hospital mortality [6]. Similarly, the rate of ED visits for ankle trauma decreased notably during the pandemic, whereas the admission rate of patients with severe trauma increased significantly [7]. Another study reported that the COVID-19 pandemic led to a significant increase in the incidence of out-of-hospital cardiac arrest (OHCA) and a corresponding decline in survival rates [8].

In China, the COVID-19 pandemic evolved through 3 distinct phases. Following confirmation of human-to-human transmission on January 20, 2020, the government implemented strict Class A infectious disease control measures, including citywide lockdowns and travel restrictions [9]. These interventions rapidly contained the initial outbreak by May 2020 but caused severe disruptions to health care services and a sharp decline in ED visits [10]. Calendar year 2021 represented a relatively stable period under China's "dynamic zero-COVID" strategy, characterized by standardized nonpharmaceutical interventions (NPIs) without the chaotic disruptions of the early outbreak or the extreme pressure of the Omicron surge [11]. In January 2023, the government adjusted the COVID-19 pandemic management measures to "Class B management," marking a new phase in epidemic prevention and control [12]. The evolving prevention and control measures across different phases of the COVID-19 pandemic exerted differential effects on ED management models, which may have led to substantial changes in ED length of stay (LOS), clinical disposition, risk of ICU admission, and long-term prognosis. In Xi'an, China, ED visits for facial injuries decreased by 17% during the postpandemic period compared with the prepandemic period, accompanied by a significant delay in the time from injury to hospital presentation [13]. A study from northern China focusing on patients with infective endocarditis (IE) reported a striking increase in in-hospital mortality from 7.7% in 2019 to 26.4% in 2023 [14]. Similarly, patients undergoing hemodialysis infected with the Omicron variant of SARS-CoV-2 exhibited an elevated mortality risk and prolonged hospital stays [15].

ICU resources play a critical role in the treatment of critical illness and in responding to public health emergencies [16]. ICU admission serves as a primary outcome measure for evaluating health care system responsiveness and resource allocation efficiency [17]. Multiple studies have confirmed gender disparities in patient outcomes during ICU care [18,19]. Yang et al [20] demonstrated that the biological and social determinants driving disease deterioration during the Omicron-predominant phase differ between the sexes. Nevertheless, current research has predominantly focused on changes in ED use patterns before and during the pandemic, with limited data available for the postpandemic phase after the lifting of public health restrictions. Moreover, most existing studies are limited by relatively small sample sizes and population-specific sociodemographic and cultural contexts, leading to restricted generalizability and interpretability.

Therefore, the objective of this study was to describe and compare the risk of ICU admission among patients with ED resuscitation area across 3 distinct periods: the prepandemic (2019), pandemic (2021), and postpandemic (2023) periods. Using a retrospective cohort design, we also assessed temporal changes in ED LOS and disposition patterns. Prespecified sex-stratified analyses were performed. No causal hypotheses were tested; the study was descriptive and hypothesis generating in nature.

Methods

Study Design

This was a retrospective observational cohort study that enrolled patients treated in the emergency resuscitation area across 3 predefined periods: the prepandemic, COVID-19 pandemic, and postpandemic periods. The study design was observational with no experimental intervention, and all data were derived from routine clinical electronic medical records (EMRs).

Study Setting and Population

The study was performed at a comprehensive emergency center serving a densely populated region in central China, covering approximately 90 million people within a 300-km radius. The annual ED volume is approximately 100,000 cases, with about 15,000 annual visits to the emergency resuscitation area. Patients were triaged by experienced nurses using a validated 5-level triage system. Only patients triaged as level I or II and admitted to the emergency resuscitation area were included [21]. During the pandemic period, all confirmed COVID-19 cases were transferred to designated hospitals and were not treated at the study ED.

We therefore selected 2021 as the representative pandemic period to reflect the typical impact of the pandemic under a consistent and sustained containment policy. This design provides a clear baseline for comparison with the prepandemic (2019) and postpandemic (2023) periods, when the country officially shifted to "Class B management" in January 2023 [12].

Inclusion and Exclusion Criteria

Patients were included if they had received treatment in the emergency resuscitation area; had triage level 1 or 2; admission between January 1 and December 31 in 2019, 2021, or 2023; and available data for the primary outcome and key covariates.

Patients were excluded if there was a missing primary outcome (ICU admission), a complete absence of key demographic or

XSL•FO
RenderX

laboratory covariates, or records with logical inconsistencies or biologically implausible values.

Participant Characteristics

A total of 44,161 patients were included in the final analysis after the exclusion of 767 invalid records. Participant characteristics included age (continuous, years), sex (male or female), ED LOS, vital signs, laboratory results, and clinical disposition. Baseline characteristics were compared across the 3 study periods.

Sampling Procedures and Data Collection

Participants were consecutively enrolled during 3 calendar years representing distinct public health phases: prepandemic phase (January 1, 2019, to December 31, 2019), COVID-19 pandemic phase (January 1, 2021, to December 31, 2021; stable dynamic zero-COVID phase), and postpandemic phase (January 1, 2023, to December 31, 2023; after implementation of Class B management). Data collection was retrospective, based on routine clinical care, with no participant self-selection bias. No additional interventions, examinations, or payments were provided to participants for research purposes.

Sample Size and Data Diagnostics

The final analytic sample included 44,161 patients (2019: n=13,351, 30.23%; 2021: n=14,528, 32.9%; and 2023: n=16,282, 36.87%). No formal sample size calculation was performed because this was a retrospective real-world study using all eligible patients. Patients with missing primary outcomes or completely missing key covariates were excluded. Only 38 (0.09%) cases had missing ED disposition; complete case analysis (CCA) was used due to the negligible proportion of missing data. As a sensitivity analysis, multiple imputation by chained equations (MICE) was performed, generating 20 imputed datasets under the assumption of missing at random (MAR). The imputation model included all analysis variables (outcome, exposure, covariates, and censoring indicator). Pooled results were derived using the Rubin rules. The results of the CCA and the imputed analysis were consistent, as shown in Table S1 in Multimedia Appendix 1. Statistical outliers and biologically implausible values were identified and excluded before analysis.

Measures and Covariates

The primary outcome was ICU admission, defined as transfer to an ICU for critical care management after initial treatment in the ED resuscitation area. ICU admission was treated as a binary time-to-event outcome in the Cox proportional hazards analysis.

ED LOS was defined as the interval from arrival at the resuscitation area to a predefined end point. For the primary time-to-event analysis (ICU admission), the end point was the first documentation of ICU admission; censoring events included ED discharge, transfer to a general ward, non-ICU death, or completion of the study period. For the secondary descriptive analysis of ED LOS, the end point was physical departure from the ED, irrespective of final disposition (discharge, general ward admission, ICU admission, or other outcomes).

Covariates were selected based on existing literature and clinical relevance. The following prespecified variables were forced into all multivariable models: demographic characteristics (age [continuous, years] and sex [male or female]) and laboratory markers (hemoglobin [g/L], lymphocyte percentage [%], albumin [g/L], creatinine [mmol/L], D-dimer [μg/mL], and bicarbonate [HCO_3^-, mmol/L]). All laboratory values were standardized to international units where applicable.

Data Collection and Quality Assurance

Data were retrospectively extracted from the hospital EMR system. Diagnoses were coded according to the *International Statistical Classification of Diseases, 10th revision* (*ICD-10*). Data quality was ensured by double independent entry, logical verification, and routine random review of 10% of all records. All data were fully anonymized and deidentified before analysis.

Statistical Analysis

Baseline characteristics across the 3 study periods (2019, 2021, and 2023) were compared using 1-way ANOVA for normally distributed continuous variables, the Kruskal-Wallis *H* test for nonnormally distributed variables, and the chi-square test for categorical variables. Cox proportional hazards models were constructed to evaluate the association between pandemic phases and ICU admission risk, with the time scale defined as the interval from ED arrival to ICU admission. The proportional hazards assumption was validated using Schoenfeld residuals (global test $P>.05$). Results are reported as hazard ratios (HRs) with corresponding 95% CIs. Sensitivity analyses were conducted to verify the robustness of the primary findings: (1) temporal stratification of 2021 (first half vs second half) to assess within-year heterogeneity, (2) CCA using only participants with complete covariate data, and (3) E-value analysis to quantify potential unmeasured confounding for the primary comparison. All statistical tests were 2-sided, with the significance level set at $\alpha=.05$. Statistical analyses were performed using R software (version 4.2; R Foundation for Statistical Computing) with the *survival*, *MICE*, and *EValue* packages.

Ethical Considerations

This study was approved by the Ethics Committee of Henan Provincial People's Hospital (2023-42). Informed consent was waived due to the retrospective observational design. All patient data were deidentified prior to analysis in compliance with the Health Insurance Portability and Accountability Act (HIPAA) Safe Harbor standard. Direct personal identifiers (full names, medical record numbers, admission IDs, and contact details) were permanently deleted and replaced with unique sequential study codes. The code-identifier linkage was stored on a password-protected institutional server, with access limited solely to the assigned data manager. Exact birth dates, free-text clinical notes, and identifiable images were excluded; age was retained as a continuous variable. The full deidentification process was audited by the hospital data security officer to ensure that no individual could be identified from the final dataset. No compensation was provided to participants. The manuscript contains no identifiable patient images or personal information.

XSL•FO
RenderX

Reporting Guidelines

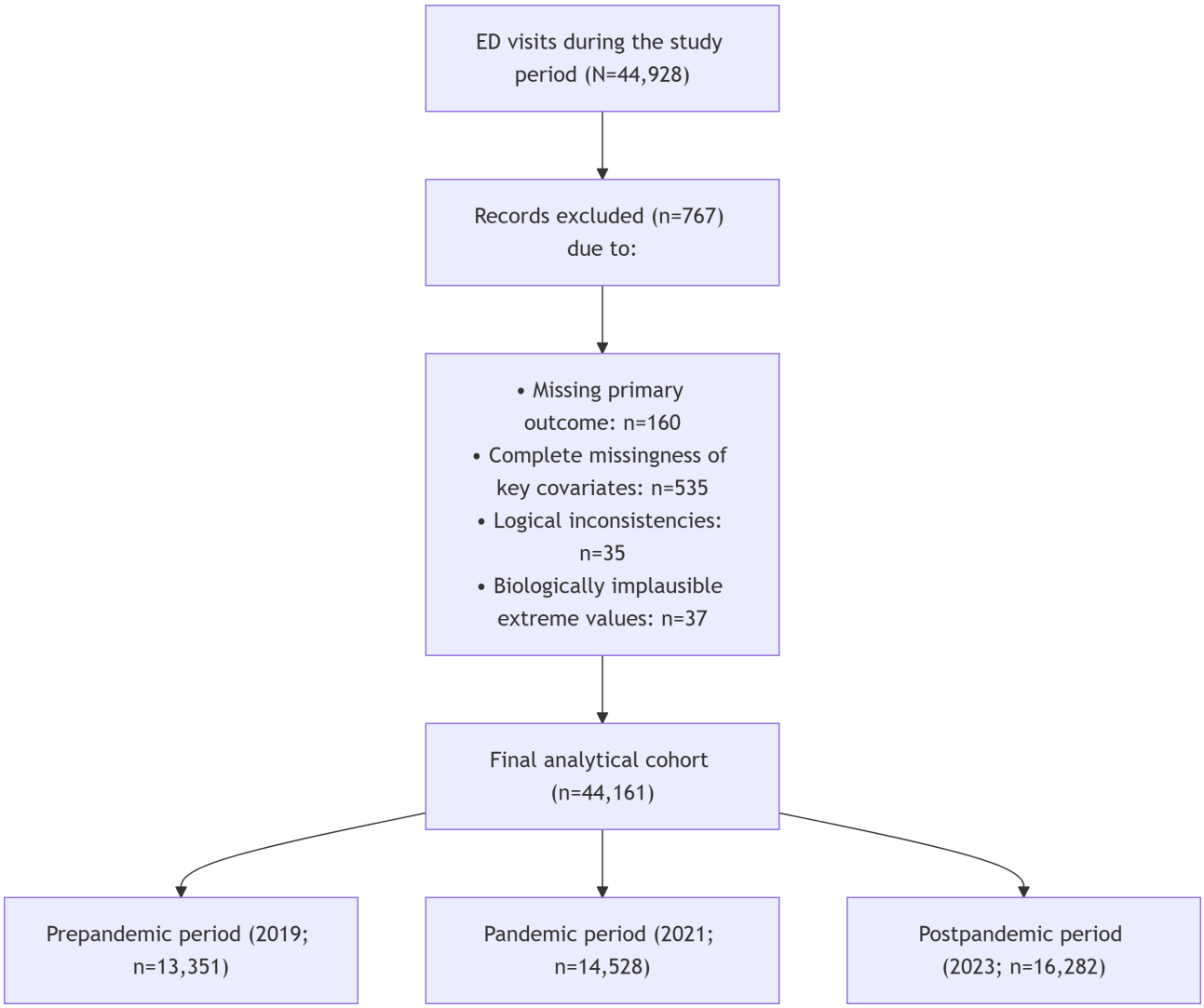
This study was reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement for observational cohort studies and the American Psychological Association (APA) Journal Article Reporting Standards (JARS)–Quant for quantitative research. Completed checklists are provided in the supplementary materials.

Results

Participant Characteristics

A total of 44,928 ED visits were recorded during the study period; 767 invalid records were excluded, yielding a final analytical cohort of 44,161 patients (prepandemic phase: n=13,351, 30.23%; pandemic phase: n=14,528, 32.9%; and postpandemic phase: n=16,282, 36.87%; Figure 1). Baseline characteristics, laboratory parameters, and clinical outcomes across the 3 periods are summarized in Table 1.

Figure 1. Study flow diagram. A total of 44,928 patients were initially admitted to the emergency department (ED); 767 were excluded because of a missing primary outcome, missing key covariates, logical inconsistencies, or biologically implausible extreme values. The final cohort of 44,161 patients was stratified into 3 periods: prepandemic (2019), pandemic (2021), and postpandemic (2023) periods.



XSL•FO
RenderX

Table . Baseline clinical and outcome characteristics of emergency department (ED) patients across COVID-19 prepandemic, pandemic, and postpandemic periods (N=44,161)[a].

Characteristics	All	2019 (prepandemic period; n=13,351)	2021 (pandemic period; n=14,528)	2023 (postpandemic period; n=16,282)	*P* value
Age (years), mean (SD)	57.12 (20.08)	56.75 (19.89)	56.98 (19.94)	57.55 (20.35)[b,c]	.002[b,c]
Male, n (%)	24,570 (55.64)	7344 (55.01)	8082 (55.63)	9144 (56.16)[b]	<.001[b]
Hemoglobin (g/L), mean (SD)	120.1 (24)	121.7 (22.8)	120.3 (24.3)[b]	118.7 (24.6)[b,c]	<.001[b,c]
Lymphocyte leukocytes percentage (%), mean (SD)	17.98 (6.67)	19.04 (7.2)	18.14 (6.56)[b]	16.97 (6.15)[b,c]	<.001[b,c]
Albumin (g/L), mean (SD)	36.6 (6.30)	36.68 (6.24)	36.65 (5.7)	36.48 (6.84)[b,c]	.009[b,c]
Creatinine (μmol/L), mean (SD)	91.94 (95.47)	85.75 (83.98)	92.82 (95.47)[b]	97.24 (105.2)[b,c]	<.001[b,c]
D-dimer (μg/mL), mean (SD)	0.89 (1.11)	0.86 (1)	0.89 (1.18)	0.9 (1.14)[b,c]	.002[b,c]
Bicarbonate (mmol/L), mean (SD)	20.64 (3.79)	21.16 (4.06)	21 (3.22)[b]	19.9 (3.91)[b,c]	<.001[b,c]
ED length of stay (minutes), mean (SD)	387.32 (504.63)	325.56 (445.7)	390.79 (418.35)[b]	434.88 (605.18)[b,c]	<.001[b,c]
In-hospital cardiac arrest, n (%)	355 (0.8)	94 (0.7)	104 (0.72)	157 (0.96)[b,c]	.02[b,c]
Admission, n (%)	29,958 (67.83)	8719 (65.3)	10,295 (70.86)[b]	10,944 (67.22)[b,c]	<.001[b,c]
General ward admission, n (%)	22,976 (52.03)	6870 (51.46)	7778 (53.34)[b]	8328 (51.15)[c]	<.001[b,c]
Intensive care unit admission, n (%)	6982 (15.81)	1849 (13.85)	2517 (17.33)[b]	2616 (16.07)[b]	<.001[b]
Discharge, n (%)	13,800 (31.25)	4494 (33.66)	4130 (28.43)[b]	5176 (31.79)[b,c]	<.001[b,c]
Death, n (%)	167 (0.38)	42 (0.31)	53 (0.36)	72 (0.44)	.20
Transfer, n (%)	198 (0.45)	96 (0.72)	47 (0.32)[b]	55 (0.34)[b]	<.001[b]

[a]Continuous variables are presented as mean (SD), and categorical variables are presented as n (%). Comparisons among the 3 groups were performed using 1-way ANOVA for continuous variables and the chi-square test for categorical variables.

[b]*P*<.05 vs the 2019 (prepandemic phase) group.

[c]*P*<.05 vs the 2021 (pandemic phase) group.

Patients in the postpandemic period were significantly older than those in the prepandemic and pandemic periods (mean 57.55, SD 20.35 vs mean 56.75, SD 19.89 and mean 56.98, SD 19.94 years; *P*=.002). The proportion of male patients was also higher in the postpandemic group (9144/16,282, 56.16%) than in the prepandemic group (7344/13,351, 55.01%; *P*<.001). All laboratory indices differed significantly across groups (all *P*<.05). Hemoglobin, lymphocyte percentage, albumin, and bicarbonate (HCO_3^-) decreased sequentially, whereas creatinine and D-dimer increased continuously from 2019 to 2023. ED LOS increased progressively over time (mean 325.56, SD 445.70 minutes; mean 390.79, SD 418.35 minutes; and mean 434.88, SD 605.18 minutes; *P*<.001). Hospital and ICU admission rates were significantly higher in the pandemic and postpandemic periods than in the prepandemic period (all *P*<.001). ED discharge rates were lowest during the pandemic and partially recovered during the postpandemic period (*P*<.001). Interhospital transfer rates declined markedly after 2019 (*P*<.001). The incidence of IHCA was significantly higher in the postpandemic period (157/16,282, 0.96%) than in the prepandemic (94/13,351, 0.7%) and pandemic (104/14,528, 0.72%) periods (*P*=.02). In-hospital mortality showed a nonsignificant upward trend across the 3 periods (42/13,351, 0.31%; 53/14,528, 0.36%; and 72/16,282, 0.44%; *P*=.20).

Analysis of ED Disposition Outcomes and LOS

A total of 44,123 eligible ED patients were stratified by pandemic phase and disposition: 13,351 (30.3%) in the prepandemic period, 14,525 (32.9%) during the pandemic, and 16,247 (36.8%) in the postpandemic period. Disposition groups included discharge (13,800/44,123, 31.3%), general admission (22,976/44,123, 52.1%), ICU admission (6982/44,123, 15.8%), and other dispositions (365/44,123, 0.8%).

XSL•FO
RenderX

The ED discharge rate decreased from 33.66% (4494/13,351; prepandemic phase) to 28.43% (4130/14,525; pandemic phase) and recovered to 31.86% (5176/16,247; postpandemic phase). The ICU admission rate rose from 13.85% to 17.33% during the pandemic period and then slightly declined to 16.1% during the postpandemic period. The proportion of general ward admissions remained stable across all periods.

Median ED LOS was compared using the Kruskal-Wallis *H* test with Bonferroni correction (Table 2).

Table . Emergency department (ED) disposition and median ED length of stay (LOS) across prepandemic, pandemic, and postpandemic periods (N=44,161)[a].

ED dispositions and variables	2019 (prepandemic period; n=13,351)	2021 (pandemic period; n=14,525)	2023 (postpandemic period; n=16,247)	*P* value[b]	Pairwise comparisons[c], *P* value		
					2019 vs 2021	2019 vs 2023	2021 vs 2023
Discharge				<.001	.047	<.001	.79
Patients, n (%)	4494 (33.66)	4130 (28.43)	5176 (31.86)				
ED LOS (minutes), median (IQR)	248 (149-404.75)	262 (153-423.75)	258 (152-451)				
General ward admission				<.001	<.001	<.001	<.001
Patients, n (%)	6870 (51.46)	7778 (53.55)	8328 (51.26)				
ED LOS (minutes), median (IQR)	186 (96-431)	350 (188-610)	293 (113-840)				
Intensive care unit admission				<.001	<.001	.003	<.001
Patients, n (%)	1849 (13.85)	2517 (17.33)	2616 (16.1)				
ED LOS (minutes), median (IQR)	95 (57-169)	171 (77-306)	105 (50-215)				
Other				.30	.38	>.99	.78
Patients, n (%)	138 (1.03)	100 (0.69)	127 (0.78)				
ED LOS (minutes), median (IQR)	220 (104-548)	182.5 (74.75-400.75)	226 (98.5-478)				

[a]Categorical data are presented as n (%), and continuous data are presented as median (IQR).

[b]Comparisons across the 3 study periods were performed using the Kruskal-Wallis *H* test.

[c]Post hoc pairwise comparisons were conducted using the Bonferroni correction.

In the discharge group, median ED LOS increased from 248 (IQR 149-404.8) minutes (prepandemic period) to 262 (IQR 153-423.8) minutes (pandemic period; *P*=.047) and remained elevated at 258 (IQR 152-451) minutes (postpandemic period; *P*<.001 vs the prepandemic period). In the general admission group, median ED LOS nearly doubled from 186 (IQR 96-431) minutes (prepandemic period) to 350 (IQR 188-610) minutes (pandemic period) and decreased to 293 (IQR 113-840) minutes (postpandemic period); all pairwise comparisons were significant (all *P*<.001). In the ICU admission group, median ED LOS increased from 95 (IQR 57-169) minutes (prepandemic period) to 171 (IQR 77-306) minutes (pandemic period) and returned to 105 (IQR 50-215) minutes (postpandemic period); all pairwise comparisons were significant (all *P*<.003). In the other groups, no significant interperiod difference in median ED LOS was observed (*P*=.30).

Cox Proportional Hazards Analysis for ICU Admission

Cox proportional hazards regression was performed to evaluate ICU admission risk across pandemic phases using 3 sequentially adjusted models, with the 2019 prepandemic period as the reference (Table 3). After excluding 38 cases with missing data, 44,123 patients were included in the overall analysis (female: n=19,572, 44.36%; male: n=24,542, 55.62%).

XSL•FO
RenderX

Table . Cox proportional hazards models for the risk of intensive care unit admission across the COVID-19 pandemic phases (N=44,123).

Models and variables	All patients		Female patients (n=19,572)		Male patients (n=24,542)	
	HR[a] (95% CI)	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value
Model 1 (unadjusted)						
Prepandemic phase (2019; reference)	1 (reference)	—[b]	1 (reference)	—	1 (reference)	—
Pandemic phase (2021)	1.045 (0.98 - 1.115)	.18	1.128 (1.013 - 1.256)	.03	0.993 (0.917 - 1.076)	.86
Postpandemic phase (2023)	1.098 (1.031 - 1.169)	.003	1.165 (1.048 - 1.294)	.005	1.055 (0.976 - 1.14)	.18
Model 2 (adjusted for age and sex)						
Prepandemic phase (2019; reference)	1 (reference)	—	1 (reference)	—	1 (reference)	—
Pandemic phase (2021)	1.035 (0.970 - 1.103)	.30	1.117 (1.003 - 1.244)	.04	0.991 (0.915 - 1.073)	.82
Postpandemic phase (2023)	1.083 (1.017 - 1.153)	.01	1.146 (1.032 - 1.274)	.01	1.05 (0.971 - 1.135)	.22
Male sex (vs female)	1.413 (1.341 - 1.487)	<.001	—	—	—	—
Age (per 1-year increase)	1.004 (1.003 - 1.005)	<.001	1.008 (1.006 - 1.01)	<.001	1.002 (1 - 1.003)	.03
Model 3 (fully adjusted)[c]						
Prepandemic phase (2019; reference)	1 (reference)	—	1 (reference)	—	1 (reference)	—
Pandemic phase (2021)	1.032 (0.967 - 1.1)	.34	1.112 (0.999 - 1.239)	.053	0.978 (0.902 - 1.059)	.59
Postpandemic phase (2023)	1.079 (1.012 - 1.15)	.02	1.143 (1.027 - 1.273)	.02	1.028 (0.95 - 1.113)	.53
Male sex (vs female)	1.413 (1.342 - 1.488)	<.001	—	—	—	—
Age (per 1-year increase)	1.004 (1.003 - 1.005)	<.001	1.005 (1.003 - 1.007)	<.001	1.003 (1.001 - 1.005)	<.001

[a]HR: hazard ratio.

[b]Not applicable.

[c]Fully adjusted for age, sex, hemoglobin (g/L), lymphocyte percentage (%), albumin (g/L), creatinine (mmol/L), D-dimer (μg/mL), and bicarbonate (mmol/L).

In the overall cohort, the postpandemic period (2023) was associated with an increased ICU admission risk in the unadjusted model (model 1: HR 1.098, 95% CI 1.031-1.169; *P*=.003), after adjustment for age and sex (model 2: HR 1.083, 95% CI 1.017-1.153; *P*=.01), and after full adjustment for laboratory covariates (model 3: HR 1.079, 95% CI 1.012-1.150; *P*=.02). No significant association was found for the pandemic period (2021) in any model (all *P*>.05).

In female patients, both the pandemic (2021) and postpandemic (2023) periods were linked to elevated ICU admission risk in most models. In model 1, HR values were 1.128 (95% CI 1.013 - 1.256; *P*=.03) for 2021 and 1.165 (95% CI 1.048 - 1.294; *P*=.005) for 2023. After full adjustment (model 3), the postpandemic period remained an independent risk factor (HR 1.143, 95% CI 1.027 - 1.273; *P*=.02).

In male patients, no significant association between pandemic phase and ICU admission risk was observed in any model (all *P*>.05). Male sex (HR 1.413, 95% CI 1.342 - 1.488; *P*<.001) and older age (HR 1.004, 95% CI 1.003 - 1.005; *P*<.001) were independent predictors of higher ICU admission risk. None of the laboratory parameters in the fully adjusted model showed a significant association with ICU admission in male patients.

Sensitivity Analyses

All sensitivity analyses (stratification of 2021 into 2 halves and exclusion of extreme ED LOS values) produced results

XSL•FO
RenderX

consistent with the primary analysis, with the postpandemic period (2023) remaining significantly associated with ICU admission (all $P<.05$). Detailed results are presented in Table S2 in Multimedia Appendix 2.

Discussion

Principal Findings

The present study aimed to investigate temporal changes in ED LOS, patient disposition, and ICU admission risk among ED resuscitation patients across the prepandemic, pandemic, and postpandemic periods in central China. Three core findings were identified: ED LOS was progressively prolonged and exhibited distinct patterns across different disposition groups; the postpandemic period was independently associated with an elevated risk of ICU admission, while the pandemic period showed no significant association; and the increased ICU admission risk was predominantly driven by female patients, with no period-related changes observed in male patients. These findings highlight the sustained and heterogeneous impacts of the COVID-19 pandemic on emergency care efficiency and critical care use.

The progressive prolongation of overall ED LOS throughout the study period is consistent with global observations of pandemic-related disruptions to emergency care workflows [22-25]. Notably, the stratified analysis revealed divergent LOS trajectories across disposition groups, which have rarely been reported in previous studies. Patients admitted to general wards experienced the most severe prolongation during the pandemic, with only partial recovery in the postpandemic period. This pattern is likely attributable to downstream bed shortages and inpatient admission bottlenecks, which are known to exacerbate ED crowding and boarding [26]. In contrast, ED LOS for ICU-admitted patients increased sharply during the pandemic but nearly returned to prepandemic levels during the postpandemic period, suggesting that critical care pathways were prioritized and optimized to mitigate delays for the most severely ill patients. These divergent trends indicate that pandemic-related strain affected noncritical and critical care pathways differently, with long-lasting impacts on general ward admissions.

Prolonged ED boarding for noncritically ill patients has well-documented adverse clinical consequences, including increased mortality and worse functional outcomes [27-29]. Studies have reported that sequential capacity-building interventions—such as staff training, space reorganization, and workflow optimization—successfully reduced ED crowding and waiting times during the pandemic [30]. The marked prolongation of ED LOS for patients admitted to general wards in our study reinforces the need for targeted strategies to alleviate inpatient bed bottlenecks and streamline ED-to-ward transitions, even after the acute pandemic phase. Such interventions are critical for preventing secondary harm from delayed inpatient care in the postpandemic era.

Our analysis confirmed that the postpandemic period was associated with a significantly higher risk of ICU admission, a finding that aligns with growing evidence of persistent pandemic-related deterioration in acute care outcomes worldwide [23]. Adverse consequences of the pandemic on acute care outcomes persisted well after official containment measures were lifted [31]. This sustained elevation in critical care risk suggests that the health care system had not fully recovered to its prepandemic baseline function by 2023, likely due to lingering increases in patient acuity and delayed health care–seeking behaviors consistent with reports of persistently excess mortality in many countries long after the acute pandemic phase subsided [32,33]. Notably, no significant association was observed between the pandemic period and ICU admission risk, which may reflect the stabilizing effect of China's dynamic zero-COVID policy in 2021, characterized by standardized NPIs and systematic patient triage, which may have attenuated the impact of pandemic-related disruptions on ICU admission risk [34]. During this period, strict containment measures may have reduced community transmission of other infectious diseases and maintained relatively stable ED care pathways for critically ill patients [2].

The most novel and clinically meaningful finding was the sex-specific disparity in ICU admission risk. Female patients exhibited persistently elevated ICU admission risk during both the pandemic and postpandemic periods, whereas male patients showed no period-related changes despite male sex being a strong independent predictor of ICU admission. This pattern suggests that women were disproportionately vulnerable to pandemic-related care disruptions, with effects persisting into the postpandemic period. Several mechanisms may explain this sex disparity. First, gendered caregiving roles may have constrained women's health care–seeking behavior during the pandemic [35,36]. Second, implicit triage bias during periods of health care strain may have contributed. Evidence suggests that women presenting with acute cardiovascular symptoms are often evaluated less aggressively than men [37]; such biases might have been amplified during pandemic surges, leading to later recognition of critical illness in female patients. These findings highlight the need for sex-specific risk stratification and targeted interventions to identify high-risk female patients in the ED. To assess the robustness of these sex-specific findings to unmeasured confounding, we calculated E-values. For female patients, the E-values were 1.51 for 2021 vs 2019 and 1.60 for 2023 vs 2019. These values fall into the moderate range, suggesting that residual confounding cannot be entirely ruled out. Future studies with richer covariate data, including detailed psychosocial measures and sex-disaggregated triage scores, are needed to confirm whether this sex-specific vulnerability is a reproducible phenomenon.

Limitations

Several limitations should be acknowledged. First, this single-center retrospective design limits the generalizability of the findings to other regions or health care systems with different pandemic responses. Second, residual confounding cannot be excluded, as detailed illness severity scores, ED crowding metrics, and socioeconomic factors were not available. Third, the borderline significance of pandemic period risk in female patients requires cautious interpretation and validation in future studies. Fourth, specific indications for ICU admission were not analyzed, which may limit mechanistic interpretation. Fifth,

XSL•FO
RenderX

because death from non-ICU causes prior to ICU admission acts as a competing risk, the traditional Cox model might overestimate the incidence of ICU admission. Future research should use the Fine-Gray model to validate our findings, as this issue was not addressed in our analysis. Finally, the postpandemic observation period may still reflect residual pandemic effects, and longer-term follow-up is needed to assess full recovery trends.

Conclusions

The COVID-19 pandemic exerted heterogeneous and sustained effects on ED LOS across disposition groups and increased ICU admission risk, particularly among female patients. ED workflow disruptions persisted most notably for patients who are not critically ill, while critical care pathways showed partial recovery. As a novel sex-stratified analysis covering the full pandemic timeline, this study differs from prior work by demonstrating the long-term impacts of pandemic control policies on emergency care. Our findings not only provide valuable evidence for patient risk stratification but also offer practical implications for optimizing health care resource allocation in the postpandemic era. These results support the implementation of sex-sensitive ED triage strategies, streamlined inpatient admission pathways, and targeted resource allocation for high-risk patient groups. Future multicenter prospective studies are warranted to validate these findings and clarify the mechanisms underlying sex-specific vulnerability to pandemic-related emergency care disruptions.

Acknowledgments

The authors thank the developers of SPSS software and its packages for their contributions. The authors declare the use of generative AI (GAI) in the research and writing process. According to the Generative Artificial Intelligence Delegation Taxonomy (2025), the task that was delegated to GAI tools under full human supervision was translation. The GAI tool used was Doubao (2026; ByteDance Ltd). Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes. Additionally, AI was used for language polishing and manuscript formatting. All content, data analysis, and scientific conclusions were fully reviewed, revised, and finalized by the authors.

Funding

This work was supported by the Joint Medical Science and Technology Project of Henan Province (grant LHGJ20230010).

Data Availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

XZ designed the study, acquired the data, and performed the analyses. ZW interpreted the findings and drafted the manuscript. Both authors contributed to the critical revision of the manuscript and approved the final version for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Cox proportional hazards regression analysis of risk factors for intensive care unit admission after multiple imputation for missing outcome data.
[XLSX File, 9 KB - publichealth_v12i1e87597_app1.xlsx]

Multimedia Appendix 2
Sensitivity analyses for the association between pandemic phases and risk of intensive care unit admission.
[XLSX File, 11 KB - publichealth_v12i1e87597_app2.xlsx]

References

1. Cheng ZJ, Shan J. 2019 novel coronavirus: where we are and what we know. Infection 2020 Apr;48(2):155-163. [doi: 10.1007/s15010-020-01401-y] [Medline: 32072569]
2. Cheng ZJ, Zhan Z, Xue M, et al. Public health measures and the control of COVID-19 in China. Clin Rev Allergy Immunol 2023 Feb;64(1):1-16. [doi: 10.1007/s12016-021-08900-2] [Medline: 34536214]
3. Elbay RY, Kurtulmuş A, Arpacıoğlu S, Karadere E. Depression, anxiety, stress levels of physicians and associated factors in Covid-19 pandemics. Psychiatry Res 2020 Aug;290:113130. [doi: 10.1016/j.psychres.2020.113130] [Medline: 32497969]
4. Ayubcha C, Smulowitz P, O'Malley J, et al. National trends in stroke presentation, treatments, and outcomes during the first 2 years of the COVID-19 pandemic. Neurol Clin Pract 2025 Apr;15(2):e200436. [doi: 10.1212/CPJ.0000000000200436] [Medline: 39830675]

XSL•FO
RenderX

5. Handberry M, Bull-Otterson L, Dai M, et al. Changes in emergency medical services before and during the COVID-19 pandemic in the United States, January 2018-December 2020. Clin Infect Dis 2021 Jul 15;73(Suppl 1):S84-S91. [doi: 10.1093/cid/ciab373] [Medline: 33956123]
6. Lee KS, Han C, Min HS, et al. Impact of the early phase of the COVID-19 pandemic on emergency department-to-intensive care unit admissions in Korea: an interrupted time-series analysis. BMC Emerg Med 2024 Apr 1;24(1):51. [doi: 10.1186/s12873-024-00968-1] [Medline: 38561666]
7. Mascio A, Greco T, Comisi C, et al. Foot and ankle trauma: epidemiology before, during, and post COVID-19 pandemic in a level I trauma center: a 5-year experience and data analysis. J Clin Med 2024 Dec 13;13(24):7585. [doi: 10.3390/jcm13247585] [Medline: 39768506]
8. Xu RH, Sun R, Fu SN. Out-of-hospital cardiac arrest before and during the COVID-19 pandemic in Hong Kong: registry-based study from 2017 to 2023. JMIR Public Health Surveill 2024 May 21;10:e56054. [doi: 10.2196/56054] [Medline: 38771620]
9. Xiong Y, Wang Y, Chen F, Zhu M. Spatial statistics and influencing factors of the COVID-19 epidemic at both prefecture and county levels in Hubei province, China. Int J Environ Res Public Health 2020 May 31;17(11):3903. [doi: 10.3390/ijerph17113903] [Medline: 32486403]
10. Shen J, Duan H, Zhang B, et al. Prevention and control of COVID-19 in public transportation: experience from China. Environ Pollut 2020 Nov;266(Pt 2):115291. [doi: 10.1016/j.envpol.2020.115291] [Medline: 32829124]
11. The State Council's Joint Prevention and Control Mechanism issued the "Guiding opinions on scientific prevention, precise policies, and differentiated and hierarchical implementation of COVID-19 prevention and control work." [Article in Chinese]. China Government Network. 2020 Feb 18. URL: http://www.gov.cn/xinwen/2020-02/18/content_5480514.htm [accessed 2025-01-10]
12. Notice on issuing the overall plan for implementing "Class B, Class B Management" for Novel Coronavirus infection [Article in Chinese]. China Government Network. 2022. URL: https://www.gov.cn/xinwen/2022-12/27/content_5733739.htm [accessed 2025-01-10]
13. Yin AA, Mu SQ, Guo WH, et al. Epidemiological and clinical characteristics of craniofacial soft tissue injuries in a plastic emergency department in Xi'an, China: a retrospective study of 22887 patients from 2019 to 2023. BMC Public Health 2025 Mar 4;25(1):854. [doi: 10.1186/s12889-025-22068-w] [Medline: 40033249]
14. Liu J, Liu J, Guo T, et al. Temporal trends, clinical characteristics and prognostic factors analysis of infective endocarditis: a multicenter ambispective cohort study in China. BMC Cardiovasc Disord 2025;25:182. [doi: 10.1186/s12872-025-04633-8]
15. Pan Y, Li D, Zhang Z, et al. Clinical characteristics and outcomes of maintenance hemodialysis patients with COVID-19 during the Omicron wave of the pandemic in Beijing: a single center retrospective study. BMC Nephrol 2024 Apr 22;25(1):143. [doi: 10.1186/s12882-024-03575-1] [Medline: 38649863]
16. Zhang W, Li X. A data-driven combined prediction method for the demand for intensive care unit healthcare resources in public health emergencies. BMC Health Serv Res 2024 Apr 17;24(1):477. [doi: 10.1186/s12913-024-10955-8] [Medline: 38632553]
17. Anesi GL, Admon AJ, Halpern SD, Kerlin MP. Understanding irresponsible use of intensive care unit resources in the USA. Lancet Respir Med 2019 Jul;7(7):605-612. [doi: 10.1016/S2213-2600(19)30088-8] [Medline: 31122898]
18. Todorov A, Kaufmann F, Arslani K, et al. Gender differences in the provision of intensive care: a Bayesian approach. Intensive Care Med 2021 May;47(5):577-587. [doi: 10.1007/s00134-021-06393-3] [Medline: 33884452]
19. Modra LJ, Higgins AM, Pilcher DV, Bailey M, Bellomo R. Sex differences in vital organ support provided to ICU patients. Crit Care Med 2024 Jan 1;52(1):1-10. [doi: 10.1097/CCM.0000000000006058] [Medline: 37846932]
20. Yang R, Guan X, Niu Z, et al. Establishment of sex-specific predictive models for critical illness in Chinese people with the Omicron variant. Front Microbiol 2023;14:1224132. [doi: 10.3389/fmicb.2023.1224132] [Medline: 38322760]
21. Shi DL, Liu X, Zhou Y. Expert consensus on emergency department triage [Article in Chinese]. Chin J Emerg Med 2018;27(6):599-604. [doi: 10.3760/cma.j.issn.1671-0282.2018.06.006]
22. McLaren JT, Taher AK, Chartier LB. Flattening the other curve: reducing emergency department STEMI delays during the COVID-19 pandemic. Am J Emerg Med 2021 Nov;49:367-372. [doi: 10.1016/j.ajem.2021.06.057] [Medline: 34246966]
23. Oskvarek JJ, Zocchi MS, Black BS, et al. Emergency department volume, severity, and crowding since the onset of the Coronavirus disease 2019 pandemic. Ann Emerg Med 2023 Dec;82(6):650-660. [doi: 10.1016/j.annemergmed.2023.07.024] [Medline: 37656108]
24. De la Rosa MJ, Duca A, Querci L, et al. Does calm always follow the storm? A comprehensive temporal analysis of emergency department visits in northern Italy before and after the COVID-19 pandemic. Epidemiologia (Basel) 2025 Mar 1;6(1):10. [doi: 10.3390/epidemiologia6010010] [Medline: 40136998]
25. Janke AT, Melnick ER, Venkatesh AK. Hospital occupancy and emergency department boarding during the COVID-19 pandemic. JAMA Netw Open 2022 Sep 1;5(9):e2233964. [doi: 10.1001/jamanetworkopen.2022.33964] [Medline: 36178691]
26. Hammer C, DePrez B, White J, Lewis L, Straughen S, Buchheit R. Enhancing hospital-wide patient flow to reduce emergency department crowding and boarding. J Emerg Nurs 2022 Sep;48(5):603-609. [doi: 10.1016/j.jen.2022.06.002] [Medline: 36084984]

XSL•FO
RenderX

27. Balen F, Routoulp S, Charpentier S, et al. Impact of emergency department length of stay on in-hospital mortality: a retrospective cohort study. Eur J Emerg Med 2024 Feb 1;31(1):39-45. [doi: 10.1097/MEJ.0000000000001079] [Medline: 37788143]
28. Wu L, Chen X, Khalemsky A, et al. The association between emergency department length of stay and in-hospital mortality in older patients using machine learning: an observational cohort study. J Clin Med 2023 Jul 18;12(14):4750. [doi: 10.3390/jcm12144750] [Medline: 37510865]
29. Wessman T, Ärnlöv J, Carlsson AC, et al. The association between length of stay in the emergency department and short-term mortality. Intern Emerg Med 2022 Jan;17(1):233-240. [doi: 10.1007/s11739-021-02783-z] [Medline: 34110561]
30. Saleem SG, Ali S, Akhtar A, et al. Impact of sequential capacity building on emergency department organisational flow during COVID-19 pandemic: a quasi-experimental study in a low-resource, tertiary care centre. BMJ Open 2023 Jul 20;13(7):e063413. [doi: 10.1136/bmjopen-2022-063413] [Medline: 37474172]
31. Eccleston C, Marsden E, Craswell A, et al. Impact of COVID-19 state-wide lockdown on older adult presentations to Queensland emergency departments. Emerg Med Australas 2023 Oct;35(5):799-808. [doi: 10.1111/1742-6723.14236] [Medline: 37160322]
32. Nørgaard SK, Nielsen J, Nordholm AC, et al. Excess mortality in Europe coincides with peaks of COVID-19, influenza and respiratory syncytial virus (RSV), November 2023 to February 2024. Euro Surveill 2024 Apr;29(15):2400178. [doi: 10.2807/1560-7917.ES.2024.29.15.2400178] [Medline: 38606570]
33. Minhas AM, Fudim M, Michos ED, Abramov D. Has mortality in the United States returned to pre-pandemic levels? An analysis of provisional 2023 data. J Intern Med 2024 Aug;296(2):168-176. [doi: 10.1111/joim.13811] [Medline: 38829151]
34. Liu J, Liu M, Liang W. The dynamic COVID-zero strategy in China. China CDC Wkly 2022 Jan 28;4(4):74-75. [doi: 10.46234/ccdcw2022.015] [Medline: 35186372]
35. Dey AK, Bhan N, Rao N, Ghule M, Chatterji S, Raj A. Factors affecting delayed and non-receipt of healthcare during the COVID-19 pandemic for women in rural Maharashtra, India: evidence from a cross-sectional study. EClinicalMedicine 2022 Nov;53:101741. [doi: 10.1016/j.eclinm.2022.101741] [Medline: 36411817]
36. White J, Cavenagh D, Byles J, Mishra G, Tooth L, Loxton D. The experience of delayed health care access during the COVID-19 pandemic in Australian women: a mixed methods exploration. Health Soc Care Community 2022 Jul;30(4):e1384-e1395. [doi: 10.1111/hsc.13546] [Medline: 34423499]
37. Cunha GS, Cerci RJ, Silvestre OM, Cavalcanti AM, Nadruz W, Fernandes-Silva MM. Sex- and age-related impact of the COVID-19 pandemic on emergency department visits for chest pain in Curitiba, Brazil. J Emerg Med 2022 Nov;63(5):656-660. [doi: 10.1016/j.jemermed.2022.08.003] [Medline: 36243615]

Abbreviations

APA: American Psychological Association
CCA: complete case analysis
ED: emergency department
EMR: electronic medical record
EMS: emergency medical services
HIPAA: Health Insurance Portability and Accountability Act
HR: hazard ratio
***ICD-10*:** *International Statistical Classification of Diseases, 10th revision*
ICU: intensive care unit
IE: infective endocarditis
JARS: Journal Article Reporting Standards
LOS: length of stay
MAR: missing at random
MICE: multiple imputation by chained equations
NPI: nonpharmaceutical intervention
OHCA: out-of-hospital cardiac arrest
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

XSL•FO
RenderX

Edited by S Brini; submitted 11.Nov.2025; peer-reviewed by J Zheng, P Kovacevic; revised version received 12.Jun.2026; accepted 17.Jun.2026; published 11.Aug.2026.

Please cite as:
Wang Z, Zhao X
Impact of Different COVID-19 Pandemic Phases on Emergency Department Length of Stay and Intensive Care Unit Admission Risk: Single-Center Retrospective Cohort Study
JMIR Public Health Surveill 2026;12:e87597
URL: https://publichealth.jmir.org/2026/1/e87597
doi:10.2196/87597

© Zehua Wang, Xiangmei Zhao. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 11.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Genetic Subtypes, Pretreatment Drug Resistance, and Associated Factors Among Individuals Newly Diagnosed With HIV-1 in Zhejiang Province From 2022 to 2024: Cross-Sectional Study

Wei Cheng, MPH; Qin Fan, MPH; Jiafeng Zhang, MPH; Jun Jiang, MPH; Jiezhe Yang, MPH; Yun Xu, MPH; Lin Chen, MPH; Chengliang Chai, MPH; Xiaohong Pan, MPH

Zhejiang Provincial Center for Disease Control and Prevention, 3399 Binsheng Road, Binjiang District, Hangzhou, Zhejiang, Hangzhou, China

Corresponding Author:
Chengliang Chai, MPH
Zhejiang Provincial Center for Disease Control and Prevention, 3399 Binsheng Road, Binjiang District, Hangzhou, Zhejiang, Hangzhou, China

Abstract

Background: Pretreatment drug resistance (PDR) poses an increasing threat to the effectiveness of antiretroviral therapy and treatment-as-prevention strategies. Zhejiang Province, an economically developed region in eastern China with high population mobility, experiences complex HIV-1 subtype dynamics and evolving resistance patterns. However, recent population-based evidence integrating HIV genetic subtypes, PDR prevalence, and associated factors remains limited.

Objective: This study aimed to characterize HIV-1 subtype distribution, estimate the prevalence of PDR, and identify associated demographic, virological, and prophylaxis-related factors among newly diagnosed individuals in Zhejiang Province from 2022 to 2024.

Methods: We conducted a cross-sectional study among adults newly diagnosed with HIV-1 in Zhejiang Province between 2022 and 2024. Demographic, epidemiological, laboratory, and exposure prophylaxis information was collected through the provincial HIV surveillance system. Partial Pol gene sequences were amplified and sequenced to determine HIV subtypes and drug resistance mutations. PDR was interpreted using the Stanford University HIV Drug Resistance Database. Logistic regression models were applied to identify factors associated with PDR.

Results: Among 8798 newly diagnosed individuals, the predominant subtypes were CRF07_BC (n=3913, 44.5%), CRF01_AE (n=2524, 28.7%), and CRF08_BC (n=954, 10.8%), with significant heterogeneity across age groups, sexes, transmission routes, and diagnosis years (P<.001). Overall PDR prevalence was 5.7% (n=504), with resistance mainly to nonnucleoside reverse-transcriptase inhibitors (n=352, 4.0%). The most frequently detected mutations were K103N, M184V, and M46L. Multivariable analysis showed that higher baseline CD4 cell counts, diagnosis in 2023 and 2024, and infection with the CRF01_AE subtype were independently associated with increased odds of PDR. Pre-exposure prophylaxis was not associated with overall PDR; however, drug-specific analyses demonstrated significant associations between exposure prophylaxis and pretreatment resistance to emtricitabine (P<.001) and lamivudine (P<.001).

Conclusions: HIV-1 genetic diversity in Zhejiang Province remains complex and dynamic, accompanied by an increasing prevalence of PDR. Although exposure prophylaxis was not linked to overall PDR, its association with emtricitabine-specific resistance highlighted the need for strengthened HIV-1 testing prior to prophylaxis initiation and enhanced resistance surveillance, but they require confirmation in future studies.

(JMIR Public Health Surveill 2026;12:e96139) doi:10.2196/96139

KEYWORDS

genetic subtypes; pretreatment drug resistance; exposure prophylaxis; associated factors; HIV-1

Introduction

Despite substantial progress in antiretroviral (ARV) therapy (ART) coverage, HIV transmission and AIDS-related mortality remain unacceptably high. According to the Joint United Nations Programme on HIV/AIDS, an estimated 1.3 million new HIV infections and 630,000 AIDS-related deaths occurred worldwide in 2024, far exceeding the global targets set for 2025 [1]. The epidemic continues to disproportionately affect low- and middle-income countries and key populations [1,2]. Consequently, HIV infection remains a major public health concern requiring sustained and adaptive interventions [1,2].

The concept of “treatment as prevention” has become a cornerstone of global HIV control strategies. Early initiation of ART effectively suppresses viral replication, reduces onward

XSL•FO
RenderX

transmission, and significantly improves survival and quality of life among people living with HIV [2]. Nevertheless, the effectiveness of treatment as prevention is increasingly threatened by the growing prevalence of pretreatment drug resistance (PDR) [3,4]. PDR refers to any drug-resistant virus detected in ARV drug–naive individuals initiating ART or individuals with previous ARV drug exposure initiating or reinitiating first-line ART [5]. PDR can compromise first-line ART regimens, increase the risk of virological failure, and impose substantial economic burdens on health care systems by necessitating the use of more expensive second-line therapies [3,4,6]. Therefore, continuous monitoring of PDR and identification of its associated factors are essential for optimizing treatment strategies and preventing the spread of resistant HIV strains.

In recent years, integrase strand transfer inhibitor (INSTI)–based regimens have become the preferred first-line treatment strategy worldwide [5]. Although the prevalence of INSTI resistance remains relatively low worldwide, the increasing use of INSTIs has been accompanied by the emergence of resistance-associated mutations in both treatment-experienced and, occasionally, treatment-naive populations [2,5,7]. In China, dolutegravir-based regimens have been progressively incorporated into national treatment guidelines; however, population-level surveillance data on INSTI resistance remain limited [8]. Long-acting ARV agents, including capsid inhibitors such as lenacapavir, are expanding the therapeutic landscape. Although resistance to capsid inhibitors remains uncommon, treatment-emergent capsid resistance mutations have been documented [9], highlighting the importance of proactive resistance surveillance as new drug classes are introduced into clinical practice.

Zhejiang Province, located in southeastern China, is an economically developed region characterized by high population mobility. Rapid socioeconomic development and population mobility may facilitate the frequent movement of HIV-positive individuals and ongoing viral recombination, potentially leading to a more complex HIV genotyping distribution and posing challenges to local prevention and control. Several studies have investigated HIV drug resistance in the Zhejiang Province. In Hangzhou, Zhang et al [10] reported a transmitted drug resistance (TDR) prevalence of 4.0% among newly diagnosed individuals in 2013, with CRF01_AE as the dominant subtype. In Ningbo, Hong et al [11] observed a TDR prevalence of 6.1% from 2018 to 2021, with K103N as the most frequent mutation and CRF07_BC as the major genotype. In Wenzhou, Zhang et al [12] reported a TDR prevalence of 5.7% between 2020 and 2023 and identified CRF08_BC as a subtype associated with a higher resistance risk. Recently, in Lishui, Li et al [13] observed a TDR prevalence of 12.1% from 2020 to 2023, with resistance predominantly to nonnucleoside reverse-transcriptase inhibitors (NNRTIs). These studies indicate that HIV-1 genetic diversity and drug resistance are present across Zhejiang but vary by city, period, and study population [10-13]. Therefore, a standardized, population-based study covering the entire province over a recent period is urgently needed to characterize HIV subtype distribution, estimate PDR prevalence using a consistent definition, and identify associated demographic, virological, and prophylaxis-related factors. Moreover, the expanding use of pre-exposure prophylaxis (PrEP) and postexposure prophylaxis (PEP) has raised concerns about the potential selection of drug resistance when these interventions are initiated during undiagnosed HIV infection or when adherence is suboptimal [5]. Nonadherence could allow for breakthrough viremia and the emergence of resistant strains, which might then be transmitted or persist as pretreatment resistance. However, empirical evidence from real-world settings in China remains scarce. Against this background, a systematic analysis integrating viral genetic subtypes, PDR, and associated demographic and behavioral factors is urgently needed.

Therefore, this study aimed to characterize HIV subtype distribution, estimate the prevalence of PDR, and identify associated demographic, virological, and prophylaxis-related factors among newly diagnosed individuals in Zhejiang Province from 2022 to 2024.

Methods

Study Participants and Data Collection

This cross - sectional study was conducted in Zhejiang Province. This province is home to over 65 million local residents spread across 11 metropolitan areas. Inclusion criteria were as follows: (1) new diagnosis with HIV-1 between January 1, 2022, and December 31, 2024; (2) age of 18 years or above; and (3) provision of informed consent for PDR testing and successful provision of blood samples. According to the law, all individuals newly diagnosed with HIV must be reported to the China Notifiable Infectious Disease Reporting Information System. After confirming the HIV diagnosis, staff at local Center for Disease Control and Prevention (CDC) conducted face-to-face interviews with the patients to collect baseline demographic characteristics (sex, age, educational level, and marital status) and epidemiological data (route of infection). Blood samples were obtained from patients who provided informed consent to record their baseline CD4 cell counts before the initiation of ART. All collected information was reported to the Zhejiang Province database of the national data information system for comprehensive HIV/AIDS control. During the study period, in addition to collecting common information, data on exposure prophylaxis history (ie, whether the patient had such a history and, if yes, further classification as PrEP or PEP) were collected through patient self-report. To ensure strict confidentiality of personal data, all blood samples were assigned numeric codes linked to corresponding sociodemographic information. During the study period, the total number of newly diagnosed cases in the province was 10,485, of whom 9668 (92.2%) had blood samples collected. The success rate for amplification and sequencing was 91.0% (8798/9668).

HIV-1 RNA Extraction, Amplification, and Sequencing

Total viral RNA was extracted from plasma specimens in accordance with the manufacturer's protocol using a viral RNA extraction kit (Tianlong Science and Technology). A Pol gene fragment (1316 bp, HXB2 coordinates 2147-3462) encoding entire protease and the first 299 residues of reverse transcriptase was amplified via reverse-transcriptase polymerase chain reaction (PCR; RT-PCR) followed by nested PCR. Commercially available amplification kits were used, including

XSL•FO
RenderX

the PrimeScript One Step RT-PCR kit version 2 (Takara Bio) and the Ex Taq kit (Takara Bio). Primer sequences and thermal cycling conditions were as previously described [10]. Amplification was performed in the laboratories of the 11 city-level CDC. PCR products were analyzed using 1% agarose gel electrophoresis. Target amplicons were subsequently purified and sequenced by Tsingke Biotech using a 3730XL DNA sequencer (Applied Biosystems) with 5 overlapping primers.

HIV Genotyping and Drug Resistance Testing

Sequence trimming, contig assembly, and interpretation of ambiguous bases were conducted using Sequencher (version 5.4.6; Gene Codes Corporation). In accordance with a previous study [14], ambiguous bases were considered if the signal intensity of the secondary peak exceeded 20% of the primary peak, whereas sequences containing 5% ambiguous bases or more were excluded from the study. All sequences were submitted to the HIV-1 sequence quality control tool for quality verification, which included the screening and exclusion of hypermutation. Subsequently, assembled sequences were aligned using ClustalW within BioEdit (version 7.2.0). HIV-1 subtypes were determined using 2 independent methods: the online automated subtyping tool COMET (Context-Based Modeling for Expeditious Typing) and phylogenetic analysis [15]. Phylogenetic analysis was used as the definitive determination when the 2 tools yielded inconsistent results. Phylogenetic trees were constructed in MEGA (version 6.0) using the neighbor-joining method based on the Kimura 2-parameter model, with bootstrap validation of 1000 replicates. Reference sequences encompassing major HIV-1 subtypes and circulating recombinant forms were obtained from the Los Alamos National Laboratory HIV sequence database. Potential intersubtype recombination events were analyzed using the Recombination Identification Program (version 3.0) [16]. Unclassified sequences were considered Unique Recombinant Form, and the recombination pattern was confirmed using a jumping profile hidden Markov model and SimPlot (version 3.5.1) [14].

Resistance interpretation for each ARV drug was performed using the Stanford University HIV Drug Resistance Database (version 9.8), with resistance defined as low level or higher (score of ≥15) [17,18]. The presence of one or more drugs with resistance was deemed to constitute PDR.

Statistical Analyses

Categorical variables are presented as frequencies and percentages. Between-group comparisons were performed using the Pearson chi-square test. Factors associated with PDR were identified using logistic regression analysis. Variables with a *P* value below .10 in univariate analysis, along with factors previously reported to be associated with PDR, were included in the multivariable model. We applied Firth logistic regression when the events per variable (EPV) were less than 10. All analyses were conducted using the R statistical software (version 4.1.2; R Foundation for Statistical Computing). A 2-sided *P* value below .05 was considered statistically significant in the final model.

Ethical Considerations

This study was a retrospective analysis of deidentified data from the Zhejiang provincial HIV surveillance system. The ethics committee of the Zhejiang provincial CDC branch approved the use of these data for research purposes (approval number 2025-077-01) and waived the requirement for individual informed consent because the data were anonymized and the study posed minimal risk. This study was conducted in accordance with the principles of the Declaration of Helsinki.

Results

Characteristics of Study Participants

A total of 8798 individuals newly diagnosed with HIV in Zhejiang Province from 2022 to 2024 were included. Most (n=7341, 83.4%) were male and aged 50 years or older (n=3274, 37.2%). More than half (n=5173, 58.8%) had educational levels of middle school or lower. Heterosexual transmission accounted for 54.8% (n=4820), whereas homosexual transmission represented 44.3% (n=3900). Over 50% of cases (n=4479, 50.9%) were identified through medical institutions, and 70.5% (n=6202) had CD4 counts below 350 cells/μL at diagnosis. In total, 1.0% (n=91) of individuals had a history of PEP or PrEP use. The predominant HIV subtypes were CRF07_BC (n=3913, 44.5%) and CRF01_AE (n=2524, 28.7%; Table 1).

XSL•FO
RenderX

Table . Characteristics of individuals with newly diagnosed HIV infection in Zhejiang Province (2022-2024; N=8798).

Variables	Individuals, n (%)
Age group (y)	
18-29	2277 (25.9)
30-39	1713 (19.5)
40-49	1534 (17.4)
≥50	3274 (37.2)
Sex	
Male	7341 (83.4)
Female	1457 (16.6)
Educational level	
Illiterate	517 (5.9)
Primary school	1912 (21.7)
Middle school	2744 (31.2)
High school or technical secondary school	1668 (19.0)
College degree or higher	1957 (22.2)
Marital status	
Unmarried, divorced, or widowed	5100 (58.0)
Married or cohabiting	3633 (41.3)
Other	65 (0.7)
Transmission route	
Heterosexual	4820 (54.8)
Homosexual	3900 (44.3)
Other	78 (0.9)
HIV diagnosis approach	
Medical institution	4479 (50.9)
VCT[a]	1871 (21.3)
STD[b] clinic	1015 (11.5)
Other	1433 (16.3)
CD4 cell count at diagnosis (cells/μL)	
0-199	3087 (35.1)
200-349	3115 (35.4)
350-499	1590 (18.1)
≥500	861 (9.8)
Missing	145 (1.6)
Confirmation year	
2022	2903 (33.0)
2023	3066 (34.8)
2024	2829 (32.2)
Exposure prophylaxis history	
PEP[c]	46 (0.5)
PrEP[d]	45 (0.5)
None	8707 (99.0)

XSL•FO
RenderX

Variables	Individuals, n (%)
HIV subtype	
B	168 (1.9)
C	119 (1.4)
CRF01_AE	2524 (28.7)
CRF07_BC	3913 (44.5)
CRF08_BC	954 (10.8)
CRF55_01	304 (3.5)
CRF85_BC	148 (1.7)
URF[e] (CRF01_AE/CRF07_BC)	334 (3.8)
Other	334 (3.8)

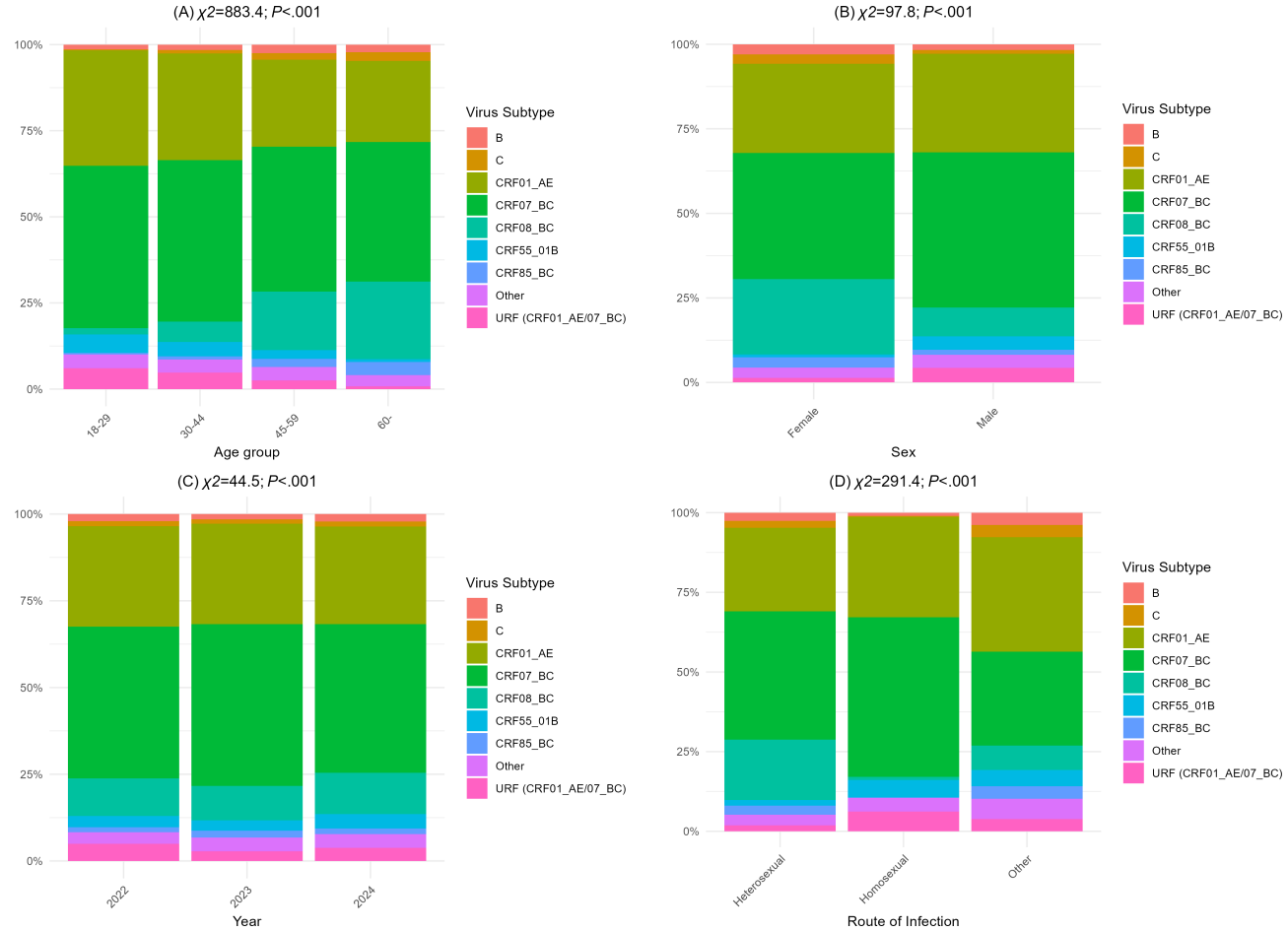
[a]VCT: voluntary counseling and testing.
[b]STD: sexually transmitted disease.
[c]PEP: postexposure prophylaxis.
[d]PrEP: pre-exposure prophylaxis.
[e]URF: Unique Recombinant Form.

HIV Subtype Distribution

Although CRF07_BC was the dominant circulating subtype, significant differences in subtype distribution were observed across age groups, sexes, routes of transmission, and years of diagnosis ($P<.001$ in all cases). The proportion of CRF07_BC and CRF01_AE decreased with increasing age, whereas the proportion of CRF08_BC increased. Compared to male individuals, female individuals showed a relatively higher proportion of CRF08_BC. Temporal analysis also revealed changes in subtype composition over the 3-year study period. In terms of transmission route, CRF07_BC and CRF01_AE were highly prevalent among homosexual transmissions, whereas CRF08_BC constituted a larger share among heterosexual transmissions (Figure 1).

XSL•FO
RenderX

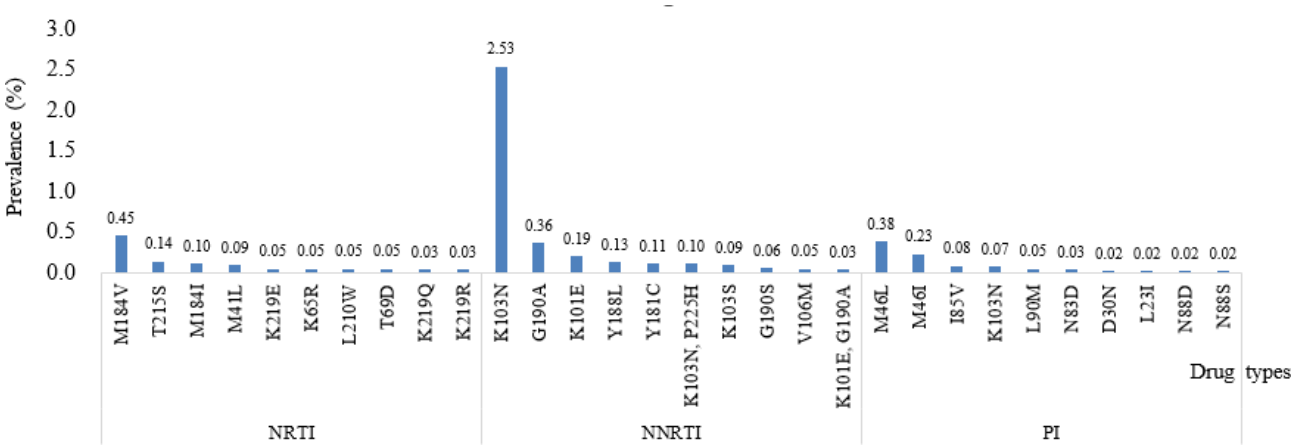
Figure 1. HIV genetic subtype distribution across (A) age groups, (B) sexes, (C) diagnosis years, and (D) transmission routes among newly diagnosed HIV infections in Zhejiang Province (2022-2024). URF: Unique Recombinant Form.



Drug Resistance Mutation Patterns

Of the 8798 patients involved in our study, 529 (6.0%) had drug resistance mutations. The proportion of participants carrying at least one nucleoside reverse-transcriptase inhibitor (NRTI), NNRTI, and protease inhibitor (PI)–associated mutation was 1.2% (n=113), 3.9% (n=346), and 1.1% (n=92), respectively. The 10 most prevalent drug resistance mutations across NRTIs, NNRTIs, and PIs are shown in Figure 2. Among detected resistance mutations, the most common NRTI mutation was M184V (n=40, 0.5%). The K103N mutation (n=223, 2.5%) predominated among NNRTI-resistant strains, whereas M46L (n=33, 0.4%) was the most frequently observed PI resistance mutation.

Figure 2. Prevalence of the top 10 detected HIV pretreatment drug resistance mutations in Zhejiang Province (2022-2024). NNRTI: nonnucleoside reverse-transcriptase inhibitor; NRTI: nucleoside reverse-transcriptase inhibitor; PI: protease inhibitor.



Prevalence of PDR

Overall, 5.7% (504/8798) of the individuals had PDR. The prevalence of PDR to NRTIs, NNRTIs, and PIs was 1.2% (101/8798), 4.0% (352/8798), and 0.9% (82/8798), respectively. PDR prevalence remained relatively low in most regions. However, nearly half of the municipalities exhibited an

XSL•FO
RenderX

increasing trend. Meanwhile, areas such as Shaoxing (19/190, 10%), Jinhua (40/377, 10.6%), and Taizhou (25/248, 10.1%) exhibited relatively higher PDR, with a prevalence of more than 10% in 2024 (Figure 3). The degree of resistance to each ARV drug is shown in Figure 4. Abacavir exhibited the highest resistance rate among NRTIs (65/8798, 0.7%), with emtricitabine (FTC) and lamivudine (3TC) both showing rates of 0.6% (60/8798). For NNRTIs, nevirapine (NVP) and efavirenz (EFV) resistance was the most common (351/8798, 4.0%), followed by dapivirine (108/8798, 1.2%). The percentages of high - level resistance to NVP and EFV were notably high at 3.8% (331/8798) and 3.2% (277/8798), respectively. Among PIs, nelfinavir resistance was the most common (69/8798, 0.8%), followed by tipranavir (18/8798, 0.2%). The PDR prevalence was significantly across different confirmation years and HIV subtypes (Table 2).

Figure 3. Prevalence of HIV pretreatment drug resistance across 11 metropolitan areas among newly diagnosed HIV infections in Zhejiang Province (2022-2024).

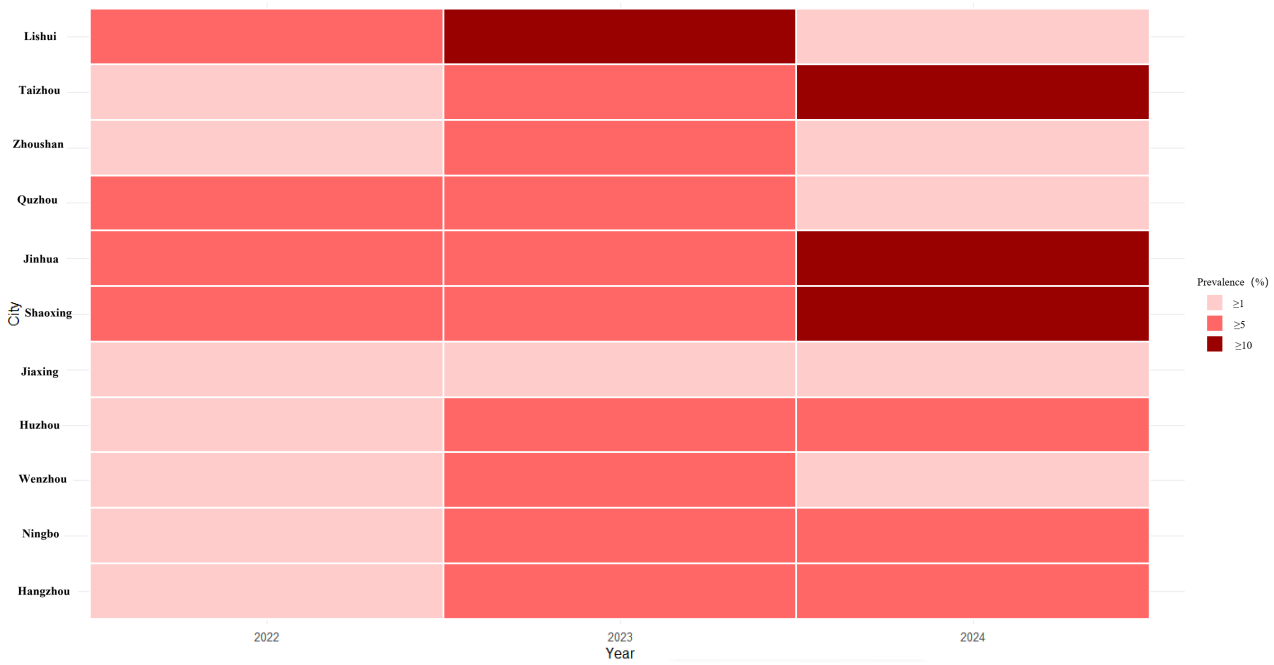
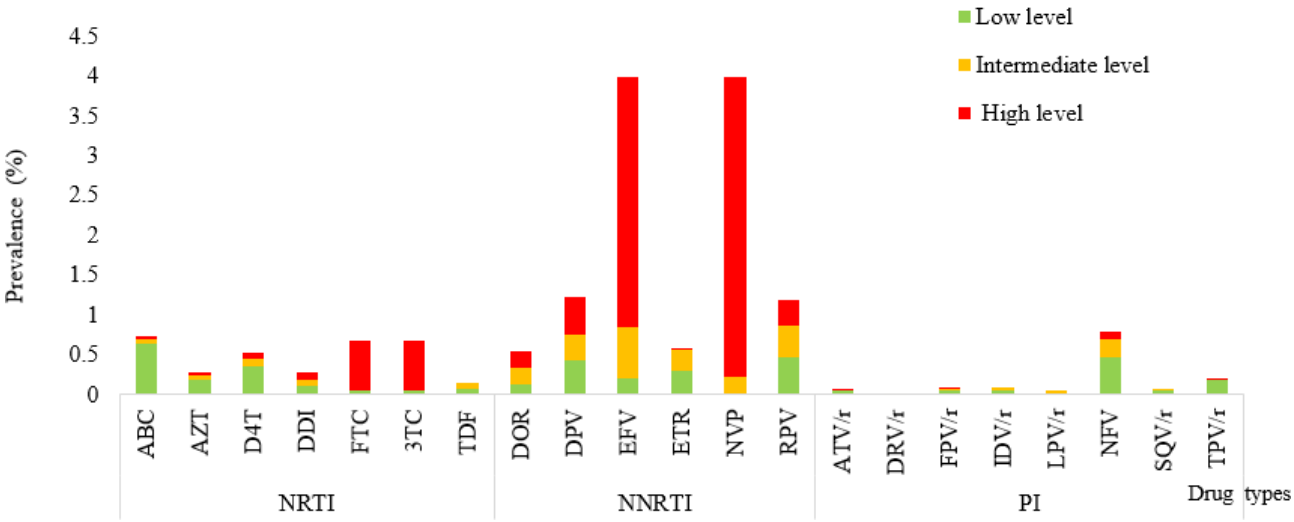


Figure 4. Prevalence and levels of HIV pretreatment drug resistance to different antiretroviral drugs in Zhejiang Province (2022-2024). 3TC: lamivudine; ABC: abacavir; ATV/r: atazanavir/r; AZT: zidovudine; D4T: stavudine; DDI: didanosine; DOR: doravirine; DPV: dapivirine; DRV/r: darunavir/r; EFV: efavirenz; ETR: etravirine; FPV/r: fosamprenavir/r; FTC: emtricitabine; IDV/r: indinavir/r; LPV: lopinavir/r; NFV: nelfinavir; NNRTI: nonnucleoside reverse-transcriptase inhibitor; NRTI: nucleoside reverse-transcriptase inhibitor; NVP: nevirapine; PI: protease inhibitor; RPV: rilpivirine; SQV/r: saquinavir/r; TDF: tenofovir; TPV/r: tipranavir/r.



XSL•FO
RenderX

Table . Prevalence of pretreatment drug resistance and associated factor analysis among newly diagnosed HIV infections in Zhejiang Province (2022-2024).

Variables	Individuals, (n=504) (%)	*P* value (univariate analysis)	aOR[a] (95% CI)	*P* value (multivariate analysis)
Age group (y)		.28		
18-29	124 (5.5)		0.89 (0.65-1.23)	.48
30-39	115 (6.7)		1.18 (0.88-1.57)	.26
40-49	85 (5.5)		0.95 (0.72-1.26)	.74
≥50	180 (5.5)		Reference	—[b]
Sex		.58		
Male	416 (5.7)		Reference	—
Female	88 (6.0)		1.10 (0.84-1.43)	.50
Educational level		.54		
Illiterate	24 (4.6)		Reference	—
Primary school	104 (5.4)		1.19 (0.75-1.88)	.46
Middle school	171 (6.2)		1.37 (0.86-2.17)	.18
High school or technical secondary school	90 (5.4)		1.14 (0.69-1.88)	.61
College degree or higher	115 (5.9)		1.20 (0.72-1.99)	.50
Marital status		.47		
Unmarried, divorced, or widowed	293 (5.8)		Reference	—
Married or cohabiting	205 (5.6)		0.99 (0.80-1.23)	.93
Other	6 (9.2)		1.71 (0.72-4.06)	.23
Transmission route		.70		
Heterosexual	272 (5.6)		Reference	—
Homosexual	229 (5.9)		1.03 (0.82-1.30)	.80
Other	3 (3.9)		0.6 (0.18-1.96)	.40
HIV diagnosis approach		.27		
Medical institution	118 (6.3)		Reference	—
VCT[c]	262 (5.9)		1.04 (0.81-1.32)	.77
STD[d] clinic	56 (5.5)		0.93 (0.69-1.27)	.65
Other	68 (4.8)		0.77 (0.59-1.02)	.07
CD4 cell count at diagnosis (cells/μL)		.05		
0-199	154 (5.0)		Reference	—
200-349	175 (5.6)		1.20 (0.96-1.51)	.11
350-499	111 (7.0)		1.58 (1.22-2.05)	.001
≥500	57 (6.6)		1.45 (1.05-2.01)	.02
Missing	7 (4.8)		0.98 (0.45-2.16)	.97
Confirmation year		.001		
2022	105 (3.6)		Reference	—
2023	206 (6.7)		1.96 (1.54-2.49)	.001
2024	193 (6.8)		1.97 (1.54-2.52)	.001
Exposure prophylaxis history		.10		
None	496 (5.7)		Reference	—

XSL•FO
RenderX

Variables	Individuals, (n=504) (%)	*P* value (univariate analysis)	aOR[a] (95% CI)	*P* value (multivariate analysis)
PEP[e]	2 (4.4)		0.62 (0.15-2.58)	.51
PrEP[f]	6 (13.3)		2.04 (0.84-4.96)	.11
HIV subtype		.001		
CRF07_BC	199 (5.1)		Reference	—
CRF01_AE	192 (7.6)		1.60 (1.30-1.96)	.001
CRF08_BC	53 (5.6)		1.18 (0.85-1.63)	.34
CRF55_01	14 (4.6)		0.92 (0.52-1.60)	.76
Other	46 (4.2)		0.84 (0.60-1.17)	.29

[a]aOR: adjusted odds ratio.

[b]Not applicable

[c]VCT: voluntary counseling and testing.

[d]STD: sexually transmitted disease.

[e]PEP: postexposure prophylaxis.

[f]PrEP: pre-exposure prophylaxis.

Factors Associated With PDR

Multivariable logistic regression analysis identified several factors independently associated with PDR. Individuals with a baseline CD4 cell count of 350 cells/µL or higher were more likely to have PDR than those with a baseline CD4 cell count of less than 50 cells/µL (350-499 cells/µL: adjusted odds ratio [aOR] 1.58, 95% CI 1.22-2.05; ≥500 cells/µL: aOR 1.45, 95% CI 1.05-2.01). Patients diagnosed in 2023 and 2024 had approximately 2-fold higher odds of having PDR than those diagnosed in 2022 (eg, 2024: aOR 1.97, 95% CI 1.54-2.52). In addition, HIV-1 subtype was significantly associated with PDR. Compared with CRF07_BC, infection with the CRF01_AE subtype was associated with a significantly increased risk of PDR (aOR 1.60, 95% CI 1.30-1.96). No significant associations were observed for age, sex, marital status, transmission route, or HIV diagnosis approach. No statistically significant differences in PDR prevalence were observed between individuals with a history of PrEP or PEP use and those without prior exposure prophylaxis.

Association Between Exposure Prophylaxis and Drug-Specific Pretreatment Resistance

We examined the association between prior exposure prophylaxis (PrEP and PEP) and resistance to tenofovir (TDF), FTC, and 3TC in the 8798 newly diagnosed individuals. Only 13 events of TDF resistance occurred in the entire study sample. Owing to an EPV value of less than 1, multivariate analysis was not performed, and only unadjusted estimates were presented (odds ratio [OR] 17.76, 95% CI 3.88 - 81.31). For FTC and 3TC, the prevalence of pretreatment resistance was 0.6% (55/8707) among those without prophylactic exposure and 5.5% (5/91) among those with prophylactic exposure. Univariate analysis showed a significant association (OR 9.15, 95% CI 3.57 - 23.41). After adjusting for age, sex, educational level, transmission route, marital status, HIV diagnosis approach, baseline CD4 cell count, diagnosis year, and HIV subtype using Firth logistic regression, the association remained significant (aOR 7.61, 95% CI 3.01 - 19.25; Table 3).

XSL•FO
RenderX

Table . Association between exposure prophylaxis and pretreatment tenofovir (TDF), emtricitabine (FTC), and lamivudine (3TC) resistance among newly diagnosed HIV infections in Zhejiang Province (2022-2024).

Exposure prophylaxis	PDR[a] prevalence, n/N (%)	Univariate analysis		Multivariate analysis	
		OR[b] (95% CI)	*P* value	aOR[c] (95% CI)	*P* value
TDF					
No	11/8707 (0.1)	Reference	__[d]	Reference	__[d]
Yes	2/91 (2.2)	17.76 (3.88-81.31)	.001	__[e]	__[e]
FTC					
No	55/8707 (0.6)	Reference	__[d]	Reference	__[d]
Yes	5/91 (5.5)	9.15 (3.57-23.41)	.001	7.61 (3.01-19.25)[f]	.001
3TC					
No	55/8707 (0.6)	Reference	__[d]	Reference	__[d]
Yes	5/91 (5.5)	9.15 (3.57-23.41)	.001	7.61 (3.01-19.25)[f]	.001

[a]PDR: pretreatment drug resistance.

[b]OR: odds ratio.

[c]aOR: adjusted odds ratio.

[d]Not applicable

[e]Not available because of events per variable <1.

[f]Adjusting for age, sex, educational level, transmission route, marital status, HIV diagnosis approach, CD4 cell count at diagnosis, confirmation year, and HIV subtype using Firth logistic regression.

Discussion

This study provided the first representative description of the PDR trend along with HIV-1 variant dynamics over 3 years in Zhejiang Province. This study showed that the most prevalent HIV genotype was CRF07_BC (3913/8798, 44.5%), followed by CRF01_AE (2524/8798, 28.7%) and CRF08_BC (954/8798, 10.8%). Moreover, HIV-1 subtypes showed heterogeneity in distribution across age groups, sexes, transmission routes, and diagnosis years. PDR was prevalent in 5.7% (504/8798) of the study population but presenting an upward trend, with the highest prevalence among NNRTIs (351/8798, 4.0%). In addition, baseline CD4 cell count, diagnosis year, and HIV subtype were independently associated with PDR. Although exposure prophylaxis was not associated with overall PDR, it was significantly associated with pretreatment resistance to FTC and 3TC.

We observed that CRF07_BC and CRF01_AE remained the dominant circulating HIV subtypes, consistent with previous reports from Tianjin, Chongqing, Beijing, Shenzhen, and Xi'an [19-23]. Age-specific differences in subtype composition suggest that distinct transmission dynamics operate across generations. Sex- and transmission route–specific variations in subtype distribution further emphasize the role of social and behavioral structures in shaping HIV epidemics. The statistically significant variation in subtype distribution by diagnosis year suggests that the local HIV epidemic is dynamic rather than static, with ongoing introductions, expansions, or contractions of specific viral lineages. Simultaneously, the complexity and diversity of HIV genotypes may be attributed to population flow and the spread of resistant strains [6]. The observed demographic and temporal heterogeneity in subtype distribution underscores the need to integrate molecular epidemiology into routine surveillance frameworks.

We observed notable geographical heterogeneity in PDR prevalence across Zhejiang Province, with some municipalities (eg, Jinhua and Taizhou) showing PDR prevalence exceeding 10% in 2024 (40/377, 10.6% and 25/248, 10.1%, respectively). This uneven distribution may be explained by the higher rates of CRF01_AE (which is associated with PDR) and clustered transmission of resistant strains [24], which requires further research. PDR has been reported in different regions of China, with a prevalence of 4.8% in Shanghai [25], 23.1% in Shenzhen [19], 7.4% in Guangzhou [17], 10.5% in Chongqing [26], and 18.3% in Xi'an [20]. The varying prevalence of PDR reported in different studies is likely attributed to multifactorial causes, including distinct definition, local differences in HIV subtype distribution, transmission dynamics, and accessibility and quality of ART programs. Therefore, continuous surveillance is essential to detect early warning signals and guide timely policy responses [19].

Notably, higher CD4 cell counts at diagnosis were associated with increased PDR risk. This finding may reflect the predominance of TDR among individuals diagnosed at earlier stages of infection, who are more likely to harbor resistant strains acquired from treated partners and probably no adherence to treatment. In addition, the significantly higher PDR prevalence among individuals diagnosed in 2023 and 2024 suggests an accelerating spread of drug-resistant HIV strains in recent years. These findings underscore the need for

XSL•FO
RenderX

comprehensive adherence support interventions, including peer educators, psychological counseling, and ensuring a stable drug supply, to prevent treatment failure and the subsequent spread of resistant strains. We also found that infection with the CRF01_AE subtype was independently associated with a higher risk of PDR compared with CRF07_BC. This observation aligns with previous studies reporting a higher burden of resistance mutations among CRF01_AE strains in other areas [20,27], which supports enhanced resistance surveillance in populations where CRF01_AE is prevalent as it may facilitate a better understanding of local resistance dynamics.

While prior PrEP or PEP use was not associated with overall PDR, drug-specific analyses revealed associations with pretreatment resistance to TDF, FTC, and 3TC—the backbone agents of most prophylaxis regimens. This selective pattern is biologically plausible and warrants careful interpretation. Initiation of PrEP or PEP during undiagnosed acute infection may expose replicating viruses to dual-NRTI pressure, facilitating selection of M184V-, M184I-, or K65R-related resistance pathways [28,29]. In addition, partial adherence to PrEP but not complete nonadherence may favor resistant variant emergence [28]. Alternatively, prophylaxis users may belong to higher-incidence sexual networks where resistant strains are more likely to circulate, reflecting confounding by transmission intensity rather than direct selection pressure. The magnitude of the association observed for TDF resistance was substantial, but the wide CI and inability to adjust for confounders due to the low event count (EPV<1) need to be confirmed in future studies.

Crucially, the potential emergence of drug-specific resistance must be weighed against the substantial epidemiological benefits of PrEP and PEP scale-up. Mathematical modeling and real-world implementation data consistently indicate that the potential risk of increased ARV drug resistance with high levels of PrEP scale-up is offset by decreased incident infections, which alleviates this key concern [30,31]. Consequently, the appropriate response to emerging resistance signals is not restriction of prophylaxis programs but refinement of delivery models, including enhanced diagnostic algorithms to exclude acute infection, adherence reinforcement, and integration of resistance surveillance into routine monitoring frameworks. Our findings, based on a limited number of exposed individuals, underscore the need for larger studies to precisely quantify the risk-benefit balance, but they do not challenge the continued rollout of PrEP and PEP as effective HIV prevention strategies.

Analysis of resistance mutation patterns revealed that M184V, K103N, and M46L were the most common mutations associated with NRTI, NNRTI, and PI resistance, respectively. These mutations have been widely reported in both national and international studies and are known to compromise the efficacy of traditional first-line regimens, particularly NNRTI-based therapies [6]. The M184V and M184I mutations confer high-level resistance to both 3TC and FTC [26]. The K103N mutation is one of the most prevalent NNRTI mutations observed in China. It is a persistent intrahost mutation that confers reduced susceptibility to EFV and NVP [11,19,20]. Moreover, 0.4% (33/8798) of individuals carried PI resistance related to the M46L mutation, which conferred TDR to nelfinavir [11]. Currently, 3TC and EFV are widely used free drugs in first-line ART regimens in China [20]. Long-term use of a limited number of drugs facilitated the development and spread of drug resistance. Therefore, persistent surveillance of drug-resistant strains is required to accurately track and update the prevalence rates of major resistance genes in our region.

Despite the transition toward INSTI-based treatment and the emergence of long-acting ARV agents, continuous surveillance of traditional NRTI and NNRTI resistance remains highly relevant. First, NRTIs continue to serve as the backbone of most first-line regimens, and mutations such as M184V, M184I, and K65R may reduce the effectiveness of companion drugs even when combined with high-barrier INSTIs. Second, historical NRTI and NNRTI resistance profiles remain critical for constructing effective second-line and salvage regimens as archived resistance mutations can influence future treatment strategies. Third, because TDF and FTC constitute the pharmacological foundation of most PrEP programs worldwide, surveillance of resistance to these agents is essential for maintaining the effectiveness of biomedical HIV prevention strategies. Therefore, resistance monitoring should evolve from a single drug class approach to an integrated surveillance framework encompassing INSTI, capsid inhibitor, and traditional reverse-transcriptase inhibitor resistance.

This study has some limitations. First, the cross-sectional design precludes causal inference. Second, behavioral and prophylaxis-related information was self-reported and may be subject to recall and social desirability biases. Third, the small number of individuals with PrEP or PEP use and the low number of drug-specific resistance events limited the precision of our estimates. Although we used Firth logistic regression to reduce bias, the results for FTC and 3TC should be considered exploratory, and the TDF findings could not be adjusted for confounding. Fourth, we used conventional Sanger sequencing, which detects only major drug-resistant variants. However, as a recent study demonstrated, deep sequencing uncovers many low-abundance resistance mutations missed by conventional methods [7]; thus, our PDR prevalence was likely underestimated. Therefore, future studies in Zhejiang Province should adopt next-generation sequencing to capture minority variants and enhance early warning systems. Furthermore, resistance testing for INSTIs was not conducted as part of this investigation. Given the rising use of INSTIs in China, it becomes essential to investigate the impact of HIV-1 drug resistance on INSTIs.

In conclusion, Zhejiang Province currently exhibits an increasing burden of PDR within a genetically diverse HIV-1 epidemic. The association between CRF01_AE infection and elevated PDR risk underscores subtype-specific vulnerability. Although exposure prophylaxis was not linked to overall resistance, its association with FTC- and 3TC-specific resistance highlights the need for enhanced testing strategies and targeted resistance surveillance. Sustained integration of molecular epidemiology, biomedical prevention, and optimized ART selection will be essential to maintain the long-term effectiveness of HIV control efforts.

XSL•FO
RenderX

Acknowledgments

The authors would like to thank the staff at the city- and county-level Center for Disease Control and Prevention across Zhejiang Province for their dedication to completing the countless data forms that made this work possible. The authors used DeepSeek solely for English-language grammar checking under full human supervision. No scientific content was generated by AI.

Funding

This work was supported by the National Science and Technology Major Project on Prevention and Control of Emerging and Major Infectious Diseases (2025ZD01904701).

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: CC, XP, JY, YX, LC
Data curation: WC, JY, YX, LC
Formal analysis: WC
Funding acquisition: CC, XP
Investigation: QF, JZ, JJ
Methodology: QF, JZ, JJ, JY, YX, LC
Project administration: CC, XP
Software: WC
Supervision: CC, XP
Validation: WC
Visualization: WC
Writing—original draft: WC
Writing—review and editing: QF, JZ, JJ, JY, YX, LC
All authors have read and agreed to the final version of the manuscript.

Conflicts of Interest

None declared.

References

1. UNAIDS global AIDS update 2025: AIDS, crisis and the power to transform. : UNAIDS; 2025 URL: https://www.unaids.org/en/UNAIDS-global-AIDS-update-2025 [accessed 2026-09-02]
2. Mody A, Sohn AH, Iwuji C, Tan RK, Venter F, Geng EH. HIV epidemiology, prevention, treatment, and implementation strategies for public health. Lancet 2024 Feb 3;403(10425):471-492. [doi: 10.1016/S0140-6736(23)01381-8] [Medline: 38043552]
3. Phillips AN, Stover J, Cambiano V, et al. Impact of HIV drug resistance on HIV/AIDS-associated mortality, new infections, and antiretroviral therapy program costs in sub-Saharan Africa. J Infect Dis 2017 May 1;215(9):1362-1365. [doi: 10.1093/infdis/jix089] [Medline: 28329236]
4. Zuo L, Liu K, Liu H, et al. Trend of HIV-1 drug resistance in China: a systematic review and meta-analysis of data accumulated over 17 years (2001-2017). EClinicalMedicine 2020 Jan;18:100238. [doi: 10.1016/j.eclinm.2019.100238] [Medline: 31922125]
5. HIV drug resistance – brief report 2024. : World Health Organization; 2024 URL: https://www.who.int/publications/i/item/9789240086319 [accessed 2026-09-02]
6. Pang X, Liang S, Tang K, et al. Disparity of HIV-1 pretreatment drug resistance in men who have sex with men and the heterosexual population in Guangxi, China. Open Forum Infect Dis 2023;10(2):ofad016. [doi: 10.1093/ofid/ofad016] [Medline: 36751650]
7. Yuan D, Zhao F, Liu Y, et al. Deep sequencing reveals underestimated burden of integrase inhibitor resistance mutations in people living with HIV: a global individual patient data meta-analysis. Emerg Microbes Infect 2025 Dec;14(1):2595793. [doi: 10.1080/22221751.2025.2595793] [Medline: 41293884]
8. National Free Antiretroviral Therapy Manual: National Center for AIDS/STD Control and Prevention, Chinese Center for Disease Control and Prevention; 2023.

XSL•FO
RenderX

9. Margot NA, Naik V, VanderVeen L, et al. Resistance analyses in highly treatment-experienced people with human immunodeficiency virus (HIV) treated with the novel capsid HIV inhibitor lenacapavir. J Infect Dis 2022 Nov 28;226(11):1985-1991. [doi: 10.1093/infdis/jiac364] [Medline: 36082606]
10. Zhang J, Guo Z, Yang J, et al. Genetic diversity of HIV-1 and transmitted drug resistance among newly diagnosed individuals with HIV infection in Hangzhou, China. J Med Virol 2015 Oct;87(10):1668-1676. [doi: 10.1002/jmv.24223] [Medline: 25899877]
11. Hong H, Tang C, Liu Y, Jiang H, Fang T, Xu G. HIV-1 drug resistance and genetic transmission network among newly diagnosed people living with HIV/AIDS in Ningbo, China between 2018 and 2021. Virol J 2023 Oct 13;20(1):233. [doi: 10.1186/s12985-023-02193-x] [Medline: 37833806]
12. Zhang T, Dou H, Ye H, et al. Transmitted drug resistance and molecular transmission network among treatment-naive HIV-1 patients in Wenzhou, China, 2020-2023. Virol J 2024 Oct 17;21(1):257. [doi: 10.1186/s12985-024-02528-2] [Medline: 39420358]
13. Li J, Mei J, Yu J, et al. Characteristics of molecular epidemiology and transmitted drug resistance among newly diagnosed HIV-1 infections in Lishui, China from 2020 to 2023. Virol J 2025 Apr 19;22(1):111. [doi: 10.1186/s12985-025-02734-6] [Medline: 40253369]
14. Fan Q, Zhang J, Pan X, et al. Insights into the molecular network characteristics of major HIV-1 subtypes in developed Eastern China: a study based on comprehensive molecular surveillance data. Infection 2025 Jun;53(3):905-920. [doi: 10.1007/s15010-024-02389-5] [Medline: 39325352]
15. Struck D, Lawyer G, Ternes AM, Schmit JC, Bercoff DP. COMET: adaptive context-based modeling for ultrafast HIV-1 subtype identification. Nucleic Acids Res 2014 Oct;42(18):e144. [doi: 10.1093/nar/gku739] [Medline: 25120265]
16. Siepel AC, Halpern AL, Macken C, Korber BT. A computer program designed to screen rapidly for HIV type 1 intersubtype recombinant sequences. AIDS Res Hum Retroviruses 1995 Nov;11(11):1413-1416. [doi: 10.1089/aid.1995.11.1413] [Medline: 8573400]
17. Lv S, Lan Y, He Y, et al. Pretreatment drug resistance among people living with HIV from 2018 to 2022 in Guangzhou, China. J Med Virol 2024 Oct;96(10):e29937. [doi: 10.1002/jmv.29937] [Medline: 39323078]
18. Chen H, Hao J, Hu J, et al. Pretreatment HIV drug resistance and the molecular transmission network among HIV-positive individuals in China in 2022: multicenter observational study. JMIR Public Health Surveill 2023 Nov 17;9:e50894. [doi: 10.2196/50894] [Medline: 37976080]
19. Li M, Zhou J, Zhang K, et al. Characteristics of genotype, drug resistance, and molecular transmission network among newly diagnosed HIV-1 infections in Shenzhen, China. J Med Virol 2023 Jul;95(7):e28973. [doi: 10.1002/jmv.28973] [Medline: 37477806]
20. Xia H, Jin J, Ba H, et al. Genetic diversity and characteristics of drug resistance among treatment-naive people living with HIV in Xi'an, China. Drug Des Devel Ther 2023;17:1485-1494. [doi: 10.2147/DDDT.S406255] [Medline: 37220545]
21. Gao W, Zhou G, Li M, Wang P, Li J, Deng R. HIV drug resistance: analysis of viral genotypes and mutation loci in people living with HIV in Chongqing, China (2016–2023). AIDS Res Ther 2024;21(1):62. [doi: 10.1186/s12981-024-00646-z] [Medline: 39272106]
22. Song YX, Xin RL, Li ZC, et al. Prevalence of transmitted drug resistance among HIV-1 treatment-naive patients in Beijing. Epidemiol Infect 2018 Feb;146(3):339-344. [doi: 10.1017/S0950268817003016] [Medline: 29345606]
23. Gao L, Xia H, Zeng R, et al. Pre-treatment and acquired antiretroviral drug resistance among people living with HIV in Tianjin, China. HIV Med 2022 Mar;23 Suppl 1:84-94. [doi: 10.1111/hiv.13252] [Medline: 35293099]
24. Cao D, Xing H, Feng Y, et al. Molecular transmission network analysis reveals the challenge of HIV-1 in ageing patients in China: elderly people play a crucial role in the transmission of subtypes and high pretreatment drug resistance in developed Eastern China, 2019-2023. Virol J 2024 Aug 26;21(1):199. [doi: 10.1186/s12985-024-02455-2] [Medline: 39187869]
25. Zhang M, Ma Y, Wang Z, et al. Prevalence and transmission of pretreatment drug resistance in people living with HIV-1 in Shanghai China, 2017-2021. Virulence 2024 Dec;15(1):2373105. [doi: 10.1080/21505594.2024.2373105] [Medline: 38934465]
26. Liu M, He XQ, Deng RN, et al. Pretreatment drug resistance in people living with HIV: a large retrospective cohort study in Chongqing, China. HIV Med 2022 Mar;23 Suppl 1(Suppl 1):95-105. [doi: 10.1111/hiv.13253] [Medline: 35293098]
27. Minna Z, Hehe Z, Tielin N, et al. Divergent transmission dynamics and drug resistance evolution of HIV-1 CRF01_AE and CRF07_BC in Tianjin, China (2013-2022). Virol J 2025 May 8;22(1):137. [doi: 10.1186/s12985-025-02704-y] [Medline: 40340650]
28. Gibas KM, van den Berg P, Powell VE, Krakower DS. Drug resistance during HIV pre-exposure prophylaxis. Drugs 2019 Apr;79(6):609-619. [doi: 10.1007/s40265-019-01108-x] [Medline: 30963509]
29. Johnson KA, Chen MJ, Kohn R, et al. Acute HIV at the time of initiation of pre-exposure or post-exposure prophylaxis: impact on drug resistance and clinical outcomes. J Acquir Immune Defic Syndr 2021 Jun 1;87(2):818-825. [doi: 10.1097/QAI.0000000000002638] [Medline: 33512849]
30. Ochieng W, Suraratdecha C. HIV self-testing, PrEP, and drug resistance: some insights. Lancet HIV 2024 Mar;11(3):e134-e136. [doi: 10.1016/S2352-3018(23)00290-4] [Medline: 38301669]

XSL•FO
RenderX

31. Cox SN, Wu L, Wittenauer R, et al. Impact of HIV self-testing for oral pre-exposure prophylaxis scale-up on drug resistance and HIV outcomes in western Kenya: a modelling study. Lancet HIV 2024 Mar;11(3):e167-e175. [doi: 10.1016/S2352-3018(23)00268-0] [Medline: 38301668]

Abbreviations

3TC: lamivudine
aOR: adjusted odds ratio
ART: antiretroviral therapy
ARV: antiretroviral
CDC: Center for Disease Control and Prevention
COMET: Context-Based Modeling for Expeditious Typing
EFV: efavirenz
EPV: events per variable
FTC: emtricitabine
INSTI: integrase strand transfer inhibitor
NNRTI: nonnucleoside reverse-transcriptase inhibitor
NRTI: nucleoside reverse-transcriptase inhibitor
NVP: nevirapine
OR: odds ratio
PCR: polymerase chain reaction
PDR: pretreatment drug resistance
PEP: postexposure prophylaxis
PI: protease inhibitor
PrEP: pre-exposure prophylaxis
RT-PCR: reverse-transcriptase polymerase chain reaction
TDF: tenofovir
TDR: transmitted drug resistance

Edited by A Mavragani, T Sanchez; submitted 26.Mar.2026; peer-reviewed by D Yuan, E Tchiakpe; revised version received 11.Aug.2026; accepted 19.Aug.2026; published 17.Sep.2026.

Please cite as:
Cheng W, Fan Q, Zhang J, Jiang J, Yang J, Xu Y, Chen L, Chai C, Pan X
Genetic Subtypes, Pretreatment Drug Resistance, and Associated Factors Among Individuals Newly Diagnosed With HIV-1 in Zhejiang Province From 2022 to 2024: Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e96139
URL: https://publichealth.jmir.org/2026/1/e96139
doi:10.2196/96139

© Wei Cheng, Qin Fan, Jiafeng Zhang, Jun Jiang, Jiezhe Yang, Yun Xu, Lin Chen, Chengliang Chai, Xiaohong Pan. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 17.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Crimean-Congo Hemorrhagic Fever Disease Awareness Scale: Scale Development, Validity, and Reliability Study

Gülay Yilmazel[1*], PhD; Ayşe Çalmaz[2*], PhD; Emre Keleş[3*], PhD

[1]Application and Research Center for Occupational Diseases & Faculty of Health Sciences, Hitit Üniversitesi, Gazi St, Çorum, Türkiye

[2]Hitit University Iskilip MYO, Çorum

[3]Department of Public Health Nursing, Faculty of Health Sciences, Hitit University, Çorum

[*]all authors contributed equally

Corresponding Author:
Gülay Yilmazel, PhD
Application and Research Center for Occupational Diseases & Faculty of Health Sciences, Hitit Üniversitesi, Gazi St, Çorum, Türkiye

Abstract

Background: Crimean-Congo Hemorrhagic Fever (CCHF) is a life-threatening, tick-borne, viral disease with substantial public health importance, particularly in endemic regions. Because no licensed vaccine or specific antiviral treatment is currently available, improving public awareness remains a key strategy for preventing transmission and promoting early diagnosis. However, no standardized instrument has been available to comprehensively assess CCHF awareness among adults.

Objective: This study aimed to develop and psychometrically evaluate the Crimean-Congo Hemorrhagic Fever Disease Awareness Scale (CCHF-DAS) for use among adults.

Methods: This cross-sectional methodological study was conducted between March and June 2025 among 293 adults recruited from 2 family health centers in Türkiye. An initial 29-item scale was developed following a comprehensive literature review and expert evaluation using the Davis technique. A pilot study involving 45 individuals assessed item clarity and comprehensibility. Construct validity was examined using exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). EFA was performed to determine the underlying factor structure of the scale, and CFA was subsequently conducted to evaluate the fit of the proposed factor structure. Reliability was evaluated using Cronbach α, split-half reliability, composite reliability, average variance extracted (AVE), McDonald ω, and item-total correlations. The Harman single-factor test was performed to assess common method variance.

Results: Content validity indices indicated excellent expert agreement for all items (Item Content Validity Index=1.00). EFA resulted in a unidimensional 25-item structure explaining 35.66% of the total variance. The Kaiser-Meyer-Olkin value was 0.912, and the Bartlett test was significant (χ^2_4=3385.4; P<.001), supporting sample adequacy. CFA demonstrated acceptable model fit ($\chi^2_{4.15}$/df=4.15, root mean square error of approximation=0.075; standardized root mean square residual=0.072; goodness-of-fit index=0.95; adjusted goodness-of-fit index=0.94; comparative fit index=0.97). The scale demonstrated excellent internal consistency (Cronbach α=0.916; McDonald ω=0.931; composite reliability=0.931), although the AVE (0.356) was below the recommended threshold. Split-half reliability coefficients (Spearman-Brown=0.899; Guttman=0.898) further supported reliability. The Harman single-factor test indicated that common method bias was unlikely to be substantial.

Conclusions: The CCHF-DAS is the first validated instrument specifically designed to assess adult awareness of CCHF. The scale demonstrated satisfactory construct validity and excellent reliability, supporting its use in epidemiological research, public health surveillance, and evaluation of educational interventions. Further validation in larger, culturally diverse populations and independent samples is recommended.

(JMIR Public Health Surveill 2026;12:e93724) doi:10.2196/93724

KEYWORDS

awareness; scale development; psychometric validation; validity; reliability; public health; Crimean-Congo hemorrhagic fever

Introduction

Crimean-Congo Hemorrhagic Fever (CCHF) is a tick-borne, viral, zoonotic disease caused by an enveloped, single-stranded RNA virus belonging to the *Nairovirus* genus of the *Bunyaviridae* family. The virus is primarily transmitted through ticks of the *Hyalomma* genus and is responsible for severe viral hemorrhagic fever outbreaks, with reported case-fatality rates ranging from 10% to 40% worldwide [1]. Owing to its high mortality rate, broad geographic distribution, and potential for

XSL•FO
RenderX

epidemic spread, CCHF represents a major and persistent public health threat in endemic regions.

In Türkiye, CCHF first emerged as a public health concern in 2002, with confirmed diagnoses reported in 2003 following case clusters in Tokat Province. Since then, the disease has shown sustained endemicity, particularly in Central and Eastern Anatolia and the western and southern parts of the Black Sea Region. According to official data from the Ministry of Health of the Republic of Türkiye, 11,041 CCHF cases and 528 related deaths were reported between 2002 and 2018, corresponding to a case-fatality rate of 4.78% [2-5]. The disease is associated with a reported global case-fatality rate ranging from 10% to 40%, although mortality varies considerably across countries depending on factors such as health care infrastructure, surveillance systems, early diagnosis, and access to supportive treatment. This lower national fatality rate is thought to reflect improvements in disease surveillance, prompt diagnosis, experienced clinical management, and supportive care. Nevertheless, CCHF continues to represent a significant public health challenge because of its epidemic potential, wide geographic distribution, and lack of a licensed vaccine or specific antiviral therapy. Human infection occurs primarily through the bite or crushing of infected *Hyalomma* ticks, as well as through direct contact with the blood or bodily fluids of infected humans or animals [1,6,7]. A wide range of domestic and wild animals—including cattle, sheep, and goats—serve as amplifying hosts for the virus. Although these animals are typically asymptomatic, they can sustain viral replication, thereby facilitating virus circulation among ticks and increasing the risk of zoonotic transmission to humans through occupational or environmental exposure [2,8].

The epidemiology of CCHF is strongly influenced by ecological and environmental factors. Climate change, migratory bird patterns, and uncontrolled animal movement have been shown to contribute significantly to the expansion of tick habitats and the geographic spread of the virus [8,9]. These dynamics increase the likelihood of human–vector contact and complicate disease control efforts, particularly in rural and agricultural settings.

Following an incubation period of approximately 3 to 7 days, infected individuals typically present with nonspecific symptoms such as sudden-onset fever, headache, myalgia, arthralgia, nausea, vomiting, fatigue, and anorexia. In severe cases, the disease may progress to hemorrhagic manifestations involving the skin, mucous membranes, and internal organs, substantially increasing the risk of fatal outcomes [2,3].

Despite its high fatality rate and expanding geographic range, there is currently no specific antiviral treatment or licensed vaccine available for CCHF. For this reason, the World Health Organization has classified CCHF as a high-priority pathogen requiring urgent research and strengthened public health interventions [1,10]. In the absence of effective medical countermeasures, prevention strategies rely heavily on minimizing exposure risk through personal protective behaviors and public awareness initiatives [5].

Preventive measures recommended for individuals living in or traveling to endemic areas include wearing light-colored protective clothing, tucking trousers into socks, performing thorough body checks after potential tick exposure, and ensuring prompt and safe tick removal without direct hand contact [1,5]. Individuals are also advised to monitor for symptoms such as fever, headache, and malaise for up to 10 days following tick removal to facilitate early diagnosis and treatment [10].

Certain population groups are at disproportionately higher risk for CCHF, including individuals residing in endemic regions, agricultural and livestock workers, veterinarians, hunters, soldiers, campers, and health care workers exposed to infected patients or biological materials [1,4]. Despite these well-established risk factors and the disease's significant epidemic potential, CCHF remains relatively neglected in terms of systematic public awareness assessment and preventive education.

In the absence of an effective vaccine, raising awareness of CCHF risk factors and preventive behaviors constitutes the most critical strategy for reducing infection rates and disease-related mortality. Public awareness not only supports primary prevention but also facilitates early recognition and timely medical intervention, particularly during seasonal peaks of disease transmission. However, a review of national and international literature indicates that no validated measurement tool currently exists to assess adults' awareness of CCHF and its prevention.

Within this context, the present study aims to address this critical gap by developing a valid and reliable awareness scale to assess adults' knowledge, perceptions, and preventive behaviors related to CCHF at both national and international levels.

Methods

Study Design and Participants

This study was conducted using a cross-sectional methodological design aimed at developing and validating a disease-specific awareness scale. The study was reported in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines to ensure transparency and methodological rigor.

Study Setting and Participant Recruitment

The study was conducted in 2 government-affiliated family health centers located in the city center of Çorum, Türkiye. These centers provide comprehensive primary health care services to approximately 13,000 registered individuals from urban and peri-urban communities, including preventive, curative, maternal and child health, vaccination, and chronic disease follow-up services. Participant recruitment was carried out between March and June 2025 using a consecutive sampling approach. During the data collection period, all adults attending either family health center for any health care–related reason (eg, routine examinations, prescription renewal, vaccination, preventive health services, or follow-up appointments) were systematically approached by a trained researcher while waiting for their consultation. Individuals were first screened for eligibility according to the predefined inclusion and exclusion criteria. Eligible individuals received a standardized explanation of the study objectives and procedures, and those who agreed

XSL•FO
RenderX

to participate provided written informed consent before enrollment.

Recruitment was conducted on all working days throughout the study period to minimize selection bias. Each participant completed the questionnaire only once through a face-to-face interview conducted in a private area of the health center to ensure confidentiality and reduce external influence on responses. Recruitment continued consecutively until the predetermined target sample size was reached. Individuals who declined participation or did not meet the eligibility criteria were not replaced using selective sampling.

Ethical Considerations

Ethical approval for the study was obtained from the Hitit University Non-Interventional Research Ethics Committee (Approval 2024 - 22).

Informed Consent

All procedures were carried out in accordance with the principles of the Declaration of Helsinki. Written informed consent was obtained from all participants before enrollment. Before participation, eligible individuals received standardized verbal and written information regarding the study objectives, procedures, voluntary nature of participation, confidentiality of the collected data, and their right to withdraw from the study at any time without any consequences. Only participants who provided written informed consent were included in the study.

Privacy and Confidentiality

To protect participants' privacy, no personally identifiable information (eg, names, national identification numbers, telephone numbers, or addresses) was collected during the study. Each questionnaire was assigned a unique study identification code, and all data were deidentified prior to analysis. The dataset was accessible only to the research team and was stored on password-protected computers in accordance with institutional data security procedures. Study findings are reported only in aggregate form, ensuring that individual participants cannot be identified directly or indirectly.

Participant Compensation

Participants did not receive any financial compensation, incentives, gifts, or reimbursement for their participation in the study. Participation was entirely voluntary.

Data Collection Instruments

Data were collected using 2 instruments: a personal information form and the Crimean-Congo Hemorrhagic Fever Disease Awareness Scale (CCHF-DAS), which was developed by the researchers based on a comprehensive review of the relevant literature.

The personal information form consisted of 15 items designed to collect sociodemographic and health-related characteristics of the participants. The form included questions on age, gender, educational level, marital status, long-term place of residence, presence of chronic diseases, personal history of tick bites, tick exposure among relatives or close contacts, perceived occupational risk, and prior training or education related to CCHF. Participants were asked whether they had ever (during their lifetime) been bitten by a tick. Among those reporting a lifetime history of tick bites, information on the method of tick removal and hospital attendance and treatment was collected as lifetime history and was not linked to a specific tick bite episode. Participants were asked whether they had ever received any education or information regarding CCHF. This variable included both formal and informal educational activities, regardless of the setting. Participants who responded positively were additionally asked to identify the primary source of this information (health care professionals, media, or other sources).

Development Process of the CCHF-DAS

Overview

The CCHF-DAS was developed specifically for this study to assess adults' awareness of CCHF and preventive measures. The scale was originally developed in Turkish, and its English version is provided in Multimedia Appendix 1. The scale development process included item generation, content validity assessment, pilot testing, and psychometric evaluation.

Item Generation

Following the identification of the research objective, a comprehensive literature review was conducted to identify existing studies on CCHF-related knowledge, awareness, attitudes, and preventive practices. Searches were performed in the Web of Science, Scopus, PubMed, and ScienceDirect databases using the keywords "Crimean-Congo Hemorrhagic Fever," "vector-borne disease," "knowledge," "attitude," and "practice." As no disease-specific psychometric tool addressing CCHF awareness was identified, the development of a new measurement instrument was deemed necessary.

An initial pool of 29 items was generated based on the literature review. Items were structured using a 5-point Likert-type scale ranging from "Strongly agree" to "Strongly disagree." The initial item pool was developed to comprehensively represent 4 conceptual content domains of CCHF awareness—disease transmission, symptoms, risk factors, and preventive behaviors. However, these domains were conceptualized as complementary components of a single overarching construct, namely overall CCHF disease awareness, rather than as independent latent dimensions. Therefore, exploratory factor analysis (EFA) was conducted to determine whether the items reflected a unidimensional measurement structure.

Content Validity

Content validity of the draft CCHF-DAS was evaluated using expert review and the Davis technique. A total of 5 academic experts—2 from infectious diseases and 3 from public health—were consulted to assess the relevance, clarity, and cultural appropriateness of each item. Experts rated each item on a 4-point scale (1=not appropriate, 2=partially appropriate, 3=fairly appropriate, and 4=completely appropriate).

The Item Content Validity Index (I-CVI) was calculated for each item, with a threshold of ≥0.80 considered acceptable. All items achieved an I-CVI value of 1.00; therefore, no items were removed at this stage. Minor revisions were made based on expert feedback to improve clarity and wording.

XSL•FO
RenderX

Pilot Testing

The 29-item draft version of the CCHF-DAS was pilot tested with 45 individuals (25 women and 20 men) who shared similar characteristics with the target study population. Face-to-face interviews were conducted to assess item clarity, comprehensibility, and completion time. Participants were asked to review each item carefully and provide feedback regarding its meaning and ease of understanding. The average time required to complete the scale was recorded. Based on positive feedback and the absence of major comprehension issues, the scale development process proceeded to the main study. Participants involved in the pilot study were excluded from the final sample.

Sample Size

In scale development studies, it is recommended that the sample size be at least 10 times the number of items included in the scale. Accordingly, given the 29-item draft version of the CCHF-DAS, a minimum sample size of 290 participants was required [11]. Additionally, methodological guidelines suggest that a sample size of at least 200 participants is sufficient for conducting EFA and confirmatory factor analysis (CFA). Taking potential data loss into account, the target sample size was set at 300 participants, and data were ultimately collected from 293 individuals. Inclusion criteria were being aged 18 years or older; being registered at the participating family health centers; having no visual, hearing, speech, or diagnosed psychiatric disorders; being able to communicate in Turkish; and providing informed consent. Exclusion criteria were age less than 18 years; inability to communicate in Turkish; cognitive impairment; visual, hearing, or speech disorders that could interfere with questionnaire completion; and refusal to participate. Medical education and previous CCHF infection were not assessed and therefore were not used as exclusion criteria.

Data Analysis

Data were analyzed using SPSS (IBM Corp) for Windows version 22.0 and LISREL version 8.80. Descriptive statistics, including frequencies, percentages, minimum and maximum values, means, and SDs, were calculated to summarize the characteristics of the study sample. Content validity was assessed using the Davis technique. The suitability of the data for EFA was evaluated using the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett test of sphericity. Factor retention was initially evaluated using the Kaiser criterion (eigenvalues >1.0), examination of the scree plot, factor loading patterns, and conceptual interpretability of the extracted factors. Although the Kaiser criterion identified multiple factors with eigenvalues greater than 1.0, the final number of retained factors was determined by considering statistical evidence together with the theoretical framework of the scale. Preference was given to the most parsimonious and conceptually interpretable solution.

EFA was first performed to identify the underlying factor structure of the scale. Subsequently, CFA was conducted to evaluate the goodness-of-fit of the resulting measurement model. Although the use of an independent validation sample is recommended, EFA and CFA have also been applied to the same dataset in scale development studies when obtaining an additional sample is not feasible [12,13]. Factor retention was initially evaluated using the Kaiser criterion (eigenvalues >1.0), examination of the scree plot, factor loading patterns, and conceptual interpretability of the extracted factors. Although factors with eigenvalues greater than 1.0 were identified, the final number of retained factors was determined by considering both statistical evidence and theoretical coherence. Preference was given to the most parsimonious and conceptually interpretable solution. Model fit was evaluated using multiple goodness-of-fit indices, including the chi-square to *df* ratio (χ^2/df), goodness-of-fit index (GFI), adjusted goodness-of-fit index (AGFI), comparative fit index (CFI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR). The structural relationships between observed variables and latent constructs were illustrated using a path diagram. Internal consistency and reliability of the scale were assessed using Cronbach α coefficient, item-total correlation coefficients, split-half reliability, composite reliability (CR), average variance extracted (AVE), and McDonald ω coefficient. Because the final scale demonstrated a unidimensional structure, discriminant validity using the Fornell-Larcker criterion was not applicable, as this method requires at least 2 distinct latent constructs for comparison [14].

In this study, variance inflation factor was not reported because the present study was not based on a regression or PLS-SEM model, where variance inflation factor is routinely used to assess multicollinearity. To evaluate the potential influence of common method variance (CMV), the Harman single-factor test was performed [15-17].

Results

Demographic Characteristics of the Participants

The demographic characteristics of the participants are summarized in Table 1. The mean age of the participants was 38.76 (SD 12.92) years. More than half of the participants were female (173/293, 59.0%), and approximately one-third had completed high school education (99/293, 33.8%). The majority of participants were married (196/293, 66.8%) and resided in the city center (250/293, 85.3%). Regarding employment status, 41.3% (121/293) of the participants were employed by an employer. Most participants reported having no chronic disease (231/293, 78.8%). Tick bite, tick removal method, and hospitalization or treatment refer to participants' lifetime history. A history of tick exposure was reported by 16.0% (47/293) of the participants, while 34.5% (101/293) indicated that a relative or someone in their close environment had experienced a tick bite. Among participants who had been bitten by a tick in lifetime, 46.8% (22/47) reported having the tick removed at a health care facility, and 51.1% (24/47) received outpatient medical care following the tick bite. Participants were asked whether they had ever received any education or information regarding CCHF. This variable included formal or informal education received through health care professionals, media, community health education campaigns, or other sources. With respect to CCHF-related education, 72.7% (213/293) of the participants reported that they had not received any formal

XSL•FO
RenderX

training on CCHF. Among those who had received training, health care personnel were identified as the primary source of information (38/80, 47.5%). Additionally, nearly one-quarter of the participants (68/293, 23.2%) were employed in occupations considered to be at risk for CCHF.

XSL•FO
RenderX

Table . Demographic characteristics of the participants (N=293).

Characteristics	Value
Sex, n (%)	
Male	120 (41.0)
Female	173 (59.0)
Educational level, n (%)	
Illiterate	4 (1.4)
Literate	2 (0.7)
Primary school	57 (19.5)
Secondary school	47 (16.0)
High school	99 (33.8)
University	75 (25.6)
Postgraduate	9 (3.1)
Marital status, n (%)	
Married	196 (66.8)
Nonmarried	97 (33.1)
Place of residence, n (%)	
City center	250 (85.3)
District	20 (6.8)
Village or town	23 (7.8)
Occupation, n (%)	
Housewife	74 (25.3)
Working for someone else	121 (41.3)
Self-employed	48 (16.4)
Retired	8 (2.7)
Student	30 (10.2)
Unemployed	12 (4.1)
Chronic disease, n (%)	
Had	62 (21.2)
Had not	231 (78.8)
Tick bite on a relative, n (%)	
Yes	101 (34.5)
No	192 (65.5)
Bitten by a tick (lifetime), n (%)	
Yes	47 (16.0)
No	246 (84.0)
How a tick was removed (n=47), n (%)	
In the health facility	22 (46.8)
By relatives and friends	10 (21.3)
By oneself	15 (31.9)
Health care use following lifetime tick bite (n=47)[a], n (%)	
Yes hospitalized	1 (2.1)
Received outpatient treatment	24 (51.1)
Did not seek medical care	22 (46.8)

XSL•FO
RenderX

Receiving CCHF[b] training, n (%)	
Yes	80 (27.3)
No	213 (72.7)
Place of education (n=80), n (%)	
Health professional	38 (47.5)
Media	27 (33.8)
Other	15 (18.8)
Occupation with CCHF risk[c], n (%)	
Yes	68 (23.2)
No	225 (76.8)
Age (years; N=293), mean (SD; range)	38.76 (12.92; 13.00-93.00)

[a]Health care usage following lifetime tick bite: This variable was assessed only among participants who reported a lifetime history of tick bites (n=47). It reflects health care usage (outpatient treatment or hospitalization) following the tick bite and does not indicate hospitalization due to confirmed CCHF. The cause of hospitalization beyond the tick bite episode was not assessed.

[b]CCHF: Crimean-Congo Hemorrhagic Fever.

[c]Occupation with CCHF risk: Participants were asked whether they considered their own occupation to be associated with an increased risk of CCHF exposure (eg, agricultural workers, livestock breeders and farmers, veterinarians, slaughterhouse workers, forestry workers, and hunters). This variable reflects self-reported occupational risk perception and was not based on a predefined list of occupations.

Factor Analysis Findings for Construct Validity

Factor analysis was applied to determine the construct validity of the scale. KMO and Bartlett tests were conducted to assess the adequacy of the sample for EFA and the suitability of the data for factor analysis. The KMO value was determined as 0.898, and the Bartlett test result was χ^2_{300}=3659.2 (P<.001).

In line with the above findings, the principal components method was applied as an EFA to the 29-item CCHF-DAS. The factor analysis findings of the scale are presented in Table 2. In this analysis, the scale was examined under a single subdimension in accordance with the theoretical structure.

XSL•FO
RenderX

Table . Factor analysis findings for the Crimean-Congo Hemorrhagic Fever Disease Awareness Scale (CCHF-DAS) (29-Item structure).[a]

Item number	Items	Factor loading
1	CCHF[b] disease is seen common in the spring and summer months when ticks are active.	0.506
2	CCHF disease is transmitted via bitten infected ticks.	0.459
3	CCHF disease can be transmitted through direct contact with the blood and body fluids of farm animals carrying the CCHF virus.	0.639
4	The disease is caused by an infected tick attaching to an individual.	0.435
5	Shaking hands with someone who has CCHF may cause the disease[c]	-0.008
6	The lethality of the CCHF virus is high.	0.476
7	There is no proven treatment for the virus in CCHF disease.	0.628
8	There is no medical measure such as protective serum or antibiotics to prevent CCHF disease.[c]	0.230
9	CCHF disease is an occupational disease.	0.434
10	CCHF is transmitted from person to person through blood and body fluids.	0.618
11	CCHF virus only causes disease in humans.	0.584
12	CCHF virus can be carried by animals.[c]	0.238
13	Sudden onset of high fever is one of the important symptoms of CCHF disease.	0.715
14	Supportive treatment is given in CCHF disease.	0.577
15	Treatment for CCHF disease is based on symptoms.	0.565
16	The most effective way to prevent CCHF disease is personal protection.	0.678
17	Personal protective equipment should be used during contact with animals, as animals that may carry the CCHF virus do not show any signs of disease.	0.671
18	Light-colored clothing should be worn in risky areas to easily see ticks, which play an important role in CCHF transmission.	0.646
19	People coming from the field should scan their bodies for ticks.	0.763
20	There is no vaccine for CCHF disease.	0.550
21	Ticks should be removed from the place where they are attached to the body with appropriate materials and by experienced people.	0.716
22	Alcohol, oil, and similar products should not be poured on the tick.	0.593
23	The first period of CCHF disease resembles the flu.	0.598
24	In the advanced stages of CCHF disease, bleeding may occur from various parts of the body.	0.501
25	Crushing a tick because of the possibility of causing harm is a wrong practice to protect against CCHF disease.	0.553

XSL•FO
RenderX

Item number	Items	Factor loading
26	Symptoms of the disease may appear within 1 - 9 days.	0.487
27	The virus is not transmitted by consuming cooked infected animal meat.[c]	0.158
28	People working in agriculture and livestock are at risk of CCHF disease.	0.643
29	Health care workers working in areas where the disease is common are at risk of the disease.	0.705

[a]Total variance explained (%)=31.153.

[b]CCHF: Crimean-Congo Hemorrhagic Fever Disease.

[c]Those excluded (factor loading below 0.30).

As seen in Table 2, when the scale items were examined as a single subdimension, the factor loadings of items 5, 8, 12, and 27 were determined to be below 0.30. Therefore, it was decided to remove these 4 items from the scale and reexamine them. However, they could have higher factor loadings on the second or subsequent components which have not been assessed.

Because the number of items had changed, the KMO and Bartlet tests were readministered. The scale's KMO value was determined to be 0.912, indicating suitability for principal components analysis and adequacy of the sample size. Similarly, the Bartlett test results (χ^2_4=3385.4; P<.001) indicate that the data are correlated and suitable for factor analysis. Based on these findings, the principal components method was applied to the 25-item CCHF-DAS as an EFA. The factor analysis findings for the CCHF-DAS are presented in Table 3.

XSL•FO
RenderX

Table . Factor analysis findings for the Crimean-Congo Hemorrhagic Fever Disease Awareness Scale (CCHF-DAS) (25-item structure).[a]

Item number	Items	Factor loading
1	CCHF[b] disease is seen common in the spring and summer months when ticks are active.	0.508
2	CCHF disease is transmitted via bitten infected ticks.	0.460
3	CCHF disease can be transmitted through direct contact with the blood and body fluids of farm animals carrying the CCHF virus.	0.640
4	The disease is caused by an infected tick attaching to an individual.	0.437
6	The lethality of the CCHF virus is high.	0.474
7	There is no proven treatment for the virus in CCHF disease.	0.625
9	CCHF disease is an occupational disease.	0.429
10	CCHF is transmitted from person to person through blood and body fluids.	0.616
11	CCHF virus only causes disease in humans.	0.584
13	Sudden onset of high fever is one of the important symptoms of CCHF disease.	0.716
14	Supportive treatment is given in CCHF disease.	0.576
15	Treatment for CCHF disease is based on symptoms.	0.567
16	The most effective way to prevent CCHF disease is personal protection.	0.679
17	Personal protective equipment should be used during contact with animals, as animals that may carry the CCHF virus do not show any signs of disease.	0.671
18	Light-colored clothing should be worn in risky areas to easily see ticks, which play an important role in CCHF transmission.	0.649
19	People coming from the field should scan their bodies for ticks.	0.771
20	There is no vaccine for CCHF disease.	0.547
21	Ticks should be removed from the place where they are attached to the body with appropriate materials and by experienced people.	0.721
22	Alcohol, oil, and similar products should not be poured on the tick.	0.592
23	The first period of CCHF disease resembles the flu.	0.596
24	In the advanced stages of CCHF disease, bleeding may occur from various parts of the body.	0.499
25	Crushing a tick because of the possibility of causing harm is a wrong practice to protect against CCHF disease.	0.552
26	Symptoms of the disease may appear within 1 - 9 days.	0.484
28	People working in agriculture and livestock are at risk of CCHF disease.	0.647
29	Health care workers working in areas where the disease is common are at risk of the disease.	0.706

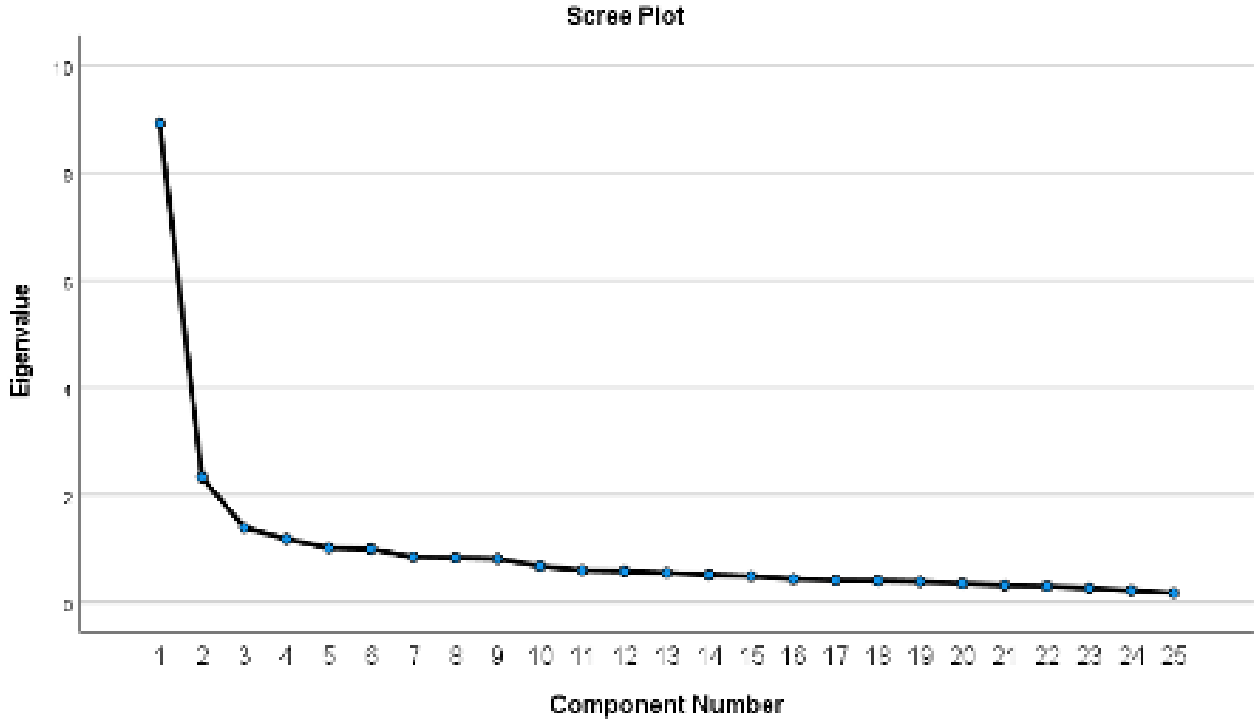
XSL•FO
RenderX

[a]Total variance explained (%)=35.657.

[b]CCHF: Crimean-Congo Hemorrhagic Fever Disease.

As seen in Table 3 and Figure 1, the CCHF-DAS conformed to the theoretical framework with a single-factor structure consisting of 25 items. Factor analysis revealed that all items had factor loadings above 0.30. The total variance explained for the scale was 35.157%. Therefore, no items were removed from the scale at this stage, and the single-dimension, 25-item structure was adopted.

Figure 1. Crimean-Congo Hemorrhagic Fever Disease Awareness Scale (CCHF-DAS) scree plot graph.



The CCHF-DAS scree plot graph is presented in Figure 1.

The eigenvalues of the CCHF-DAS in the single-factor structure are presented in Figure 1. After the EFA, structural equation modeling was established with CFA to obtain more precise findings.

CFA Findings

Table 4 presents the fit index values, normal and acceptable values determined for the CCHF-DAS.

Table . Values of the adaptation index for the Crimean-Congo Hemorrhagic Fever Disease Awareness Scale (CCHF-DAS; normal and acceptable values).

Index	Normal value	Acceptable value	Determined value
χ^2/SD	<2	<5	4.15
GFI[a]	>0.95	>0.90	0.95
AGFI[b]	>0.95	>0.90	0.94
CFI[c]	>0.95	>0.90	0.97
RMSEA[d]	<0.05	<0.08	0.075
SRMR[e]	<0.05	<0.08	0.072

[a]GFI: goodness-of-fit index.

[b]AGFI: adjusted goodness-of-fit index.

[c]CFI: comparative fit index.

[d]RMSEA: root mean square error of approximation.

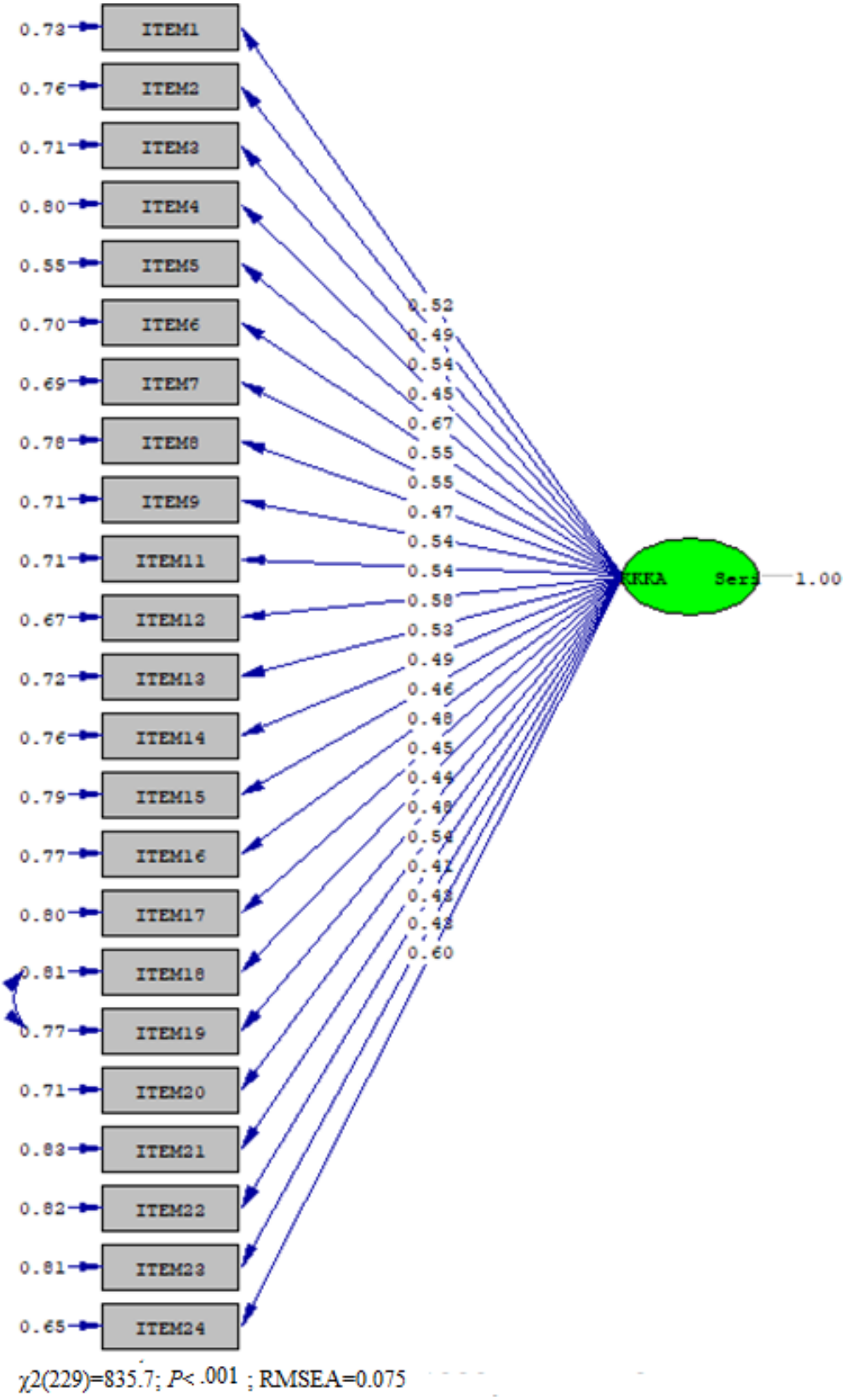
[e]SRMR: standardized root mean square residual.

XSL•FO
RenderX

As shown in Table 4, several indices were used to examine the fit of the CCHF-DAS model. Of these, the χ^2/SD value was determined to be 4.15, GFI 0.95, AGFI 0.94, CFI 0.97, RMSEA 0.075, and SRMR 0.072. Some of the relevant fit index values were deemed not to be within the desired range.

In Figure 2, factor loadings for the CCHF-DAS are presented as a PATH diagram.

Figure 2. PATH diagram of the Crimean-Congo Hemorrhagic Fever Disease Awareness Scale (CCHF-DAS). RMSEA: root mean square error of approximation.



XSL•FO
RenderX

As seen in Figure 2, the factor loadings of the model range from 0.41 to 0.67. The 2-tailed *t*-values of all items in the model are above 1.96 (7.84 - 12.74).

Internal Validity Findings

Table 5 presents the item means, item total correlations, and Cronbach α coefficients if the item is deleted for the CCHF-DAS.

XSL•FO
RenderX

Table . Item-total correlations and Cronbach α coefficients of the Crimean-Congo Hemorrhagic Fever Disease Awareness Scale (CCHF-DAS).[a]

Item number	Items	Value, n	Mean (SD)	Item-total correlation	If the item is deleted, α
1	CCHF disease is seen common in the spring and summer months when ticks are active.	293	4.29 (0.77)	0.440	0.915
2	CCHF disease is transmitted via bitten infected ticks.	293	4.09 (0.78)	0.397	0.915
3	CCHF disease can be transmitted through direct contact with the blood and body fluids of farm animals carrying the CCHF virus.	293	3.57 (1.28)	0.609	0.912
4	The disease is caused by an infected tick attaching to an individual.	293	4.04 (0.91)	0.381	0.916
6	The lethality of the CCHF virus is high.	293	4.10 (0.90)	0.420	0.915
7	There is no proven treatment for the virus in CCHF disease.	293	3.33 (1.29)	0.606	0.912
9	CCHF disease is an occupational disease.	293	3.39 (1.45)	0.369	0.918
10	CCHF is transmitted from person to person through blood and body fluids.	293	3.54 (1.40)	0.601	0.912
11	CCHF virus only causes disease in humans.	293	3.35 (1.41)	0.567	0.913
13	Sudden onset of high fever is one of the important symptoms of CCHF disease.	293	4.28 (0.70)	0.669	0.912
14	Supportive treatment is given in CCHF disease.	293	3.84 (0.86)	0.523	0.913
15	Treatment for CCHF disease is based on symptoms.	293	3.95 (0.84)	0.506	0.914
16	The most effective way to prevent CCHF disease is personal protection.	293	4.27 (0.78)	0.611	0.912
17	Personal protective equipment should be used during contact with animals, as animals that may carry the CCHF virus do not show any signs of disease.	293	4.20 (0.90)	0.607	0.912
18	Light-colored clothing should be worn in risky areas to easily see ticks, which play an important role in CCHF transmission.	293	4.31 (0.90)	0.579	0.912

XSL•FO
RenderX

Item number	Items	Value, n	Mean (SD)	Item-total correlation	If the item is deleted, α
19	People coming from the field should scan their bodies for ticks.	293	4.28 (1.11)	0.746	0.909
20	There is no vaccine for CCHF disease.	293	3.30 (1.29)	0.528	0.913
21	Ticks should be removed from the place where they are attached to the body with appropriate materials and by experienced people.	293	4.43 (0.82)	0.660	0.911
22	Alcohol, oil, and similar products should not be poured on the tick.	293	4.03 (1.14)	0.552	0.913
23	The first period of CCHF disease resembles the flu.	293	3.30 (1.08)	0.575	0.912
24	In the advanced stages of CCHF disease, bleeding may occur from various parts of the body.	293	3.59 (0.96)	0.453	0.914
25	Crushing a tick because of the possibility of causing harm is a wrong practice to protect against CCHF disease.	293	4.04 (0.97)	0.494	0.914
26	Symptoms of the disease may appear within 1 - 9 days.	293	3.73 (0.90)	0.429	0.915
28	People working in agriculture and livestock are at risk of CCHF disease.	293	4.27 (0.84)	0.571	0.913
29	Health care workers working in areas where the disease is common are at risk of the disease.	293	4.21 (0.84)	0.642	0.912

[a]Total Cronbach α=0.916.

As shown in Table 5, the total Cronbach α coefficient for the CCHF-DAS is 0.916. Total item correlation values for all items on the scale are above 0.30, and deleting any item does not cause a significant increase in the Cronbach α coefficient. Therefore, no items were removed from the scale at this stage.

For the internal consistency reliability coefficient of the CCHF-DAS, the scale was divided into 2 halves, and the consistency values for the 2 halves are given in Table 6.

XSL•FO
RenderX

Table . Split-half reliability values of the Crimean-Congo Hemorrhagic Fever Disease Awareness Scale (CCHF-DAS).

Split-half reliability	Values
Cronbach α	
First half	
Value	0.833
Number of items[a]	13
Second half	
Value	0.868
Number of items[b]	12
Total number of items	25
Correlation between the 2 halves	0.816
Spearman-Brown coefficient	
Equal length	0.899
Unequal length	0.899
Guttman Split-Half coefficient	0.898

[a]The items are Item1, Item2, Item3, Item4, Item5, Item6, Item7, Item8, Item9, Item10, Item11, Item12, Item13.

[b]The items are: Item14, Item15, Item16, Item17, Item18, Item19, Item20, Item21, Item22, Item23, Item24, Item25.

An examination of Table 6 revealed that the split-half reliability values for the internal consistency of the CCHF-DAS were found to be at acceptable levels. This value was 0.833 for the first half of the 13-item scale and 0.868 for the second half, consisting of 12 items. The results indicated that the correlation between the 2 halves was 0.816, the Spearman-Brown coefficient was 0.899, and the Guttman Split-Half coefficient was 0.898, indicating that the scale had good internal consistency reliability. These findings indicate that the CCHF-DAS has high internal consistency reliability.

Convergent and divergent validity analysis results are presented in Table 7.

Table . Convergent and divergent validity analysis results.

Subdimensions	CR[a]	AVE[b]	Omega
Crimean-Congo Hemorrhagic Fever Disease Awareness Scale	0.931	0.356	0.931
Expected values	≥0.70	≥0.50	≥0.70

[a]CR: composite reliability.

[b]AVE: average variance extracted.

In the convergent-divergent validity analysis, CR and AVE values were examined. CR values were 0.931 and ranged from acceptable to excellent reliability. AVE was 0.356 and below 0.50. Additionally, McDonald ω values were 0.931 and ranged from acceptable to excellent.

The distribution of minimum, maximum, and mean scores obtained from the CCHF-DAS are 37.00, 125.00, and 125, respectively.

Participants received a total score of 97.75 (SD 14.82) on the CCHF Awareness Scale.

Discussion

Interpretation of the Findings

The primary objective of this study was to develop and psychometrically evaluate a valid and reliable instrument for measuring adults' awareness of CCHF. This objective was achieved through a systematic scale development process involving content validation, pilot testing, EFA and CFA, and multiple reliability assessments. The final 25-item, unidimensional CCHF-DAS demonstrated satisfactory construct validity, excellent internal consistency, and acceptable model fit indices, supporting its use as a standardized instrument for assessing CCHF awareness in adult populations.

Principal Findings

Studies on CCHF in the literature generally consist of cross-sectional investigations evaluating knowledge, attitudes, and practices levels among various risk groups, such as the general public, livestock value chain actors, medical and pharmacy students, health care personnel, or slaughterhouse workers [18-24]. Accurately measuring societal awareness and risk perception in combating CCHF, which possesses an endemic character in Turkey and worldwide, is of vital importance for the development of preventive health policies. However, a review of the literature reveals a notable lack of a psychometrically robust instrument specifically measuring

XSL•FO
RenderX

CCHF awareness in the adult population. The CCHF-DAS developed in this study makes an important contribution to the literature by filling this methodological gap. The stages followed in the scale development process [25], data analysis techniques [26,27], and validation processes [12-15,28-31] were conducted in accordance with international scale development standards.

As a result of the EFA conducted to examine the construct validity of the scale, it was determined that the items had factor loadings above 0.30, and the scale exhibited a single-factor structure consisting of 25 items. The total variance explained by the single-factor structure was found to be 35.157%. Following the EFA, the fit indices (χ^2/SD=4.15; GFI=0.95; AGFI=0.94; CFI=0.97, RMSEA=0.075; and SRMR=0.072) obtained from the CFA established to validate the model were examined. Although some fit indices were at borderline values, they were generally observed to be within acceptable limits. It was determined that the factor loadings of the items in the model ranged from 0.41 to 0.67, and the *t*-values of all items were above 1.96 (ranging between 7.84 and 12.74).

The KMO value of 0.912 and a statistically significant Bartlett test of sphericity (P<.001) indicated that the sample size and data structure were highly suitable for factor analysis [11,25]. These results reflect strong inter-item correlations and support the adequacy of the dataset for psychometric evaluation.

Although items 5, 8, 12, and 27 were removed from the final scale because of low factor loadings, their weak association with the overall awareness construct may reflect inconsistent or insufficient knowledge regarding specific aspects of CCHF transmission, animal reservoirs, and prevention. Therefore, these items may still provide useful insight into community-level misconceptions and should be considered when planning targeted health education strategies.

Although the Kaiser criterion initially suggested more than one factor with eigenvalues greater than 1.0, factor retention was not determined solely on the basis of the eigenvalue criterion. The scree plot indicated a clear inflection after the first factor, suggesting the presence of 1 dominant underlying construct. Moreover, the CCHF-DAS was theoretically developed to assess overall disease awareness as a single latent construct, with items covering transmission, symptoms, prevention, and treatment as complementary components of the same construct rather than independent dimensions. Therefore, considering the theoretical framework, the scree plot, and the principle of parsimony, a unidimensional model was retained. This decision was further supported by the CFA, which demonstrated acceptable model fit indices and excellent internal consistency for the final 25-item scale [26,27].

In our study, the total Cronbach α coefficient calculated to evaluate the internal consistency reliability of the scale was found to be 0.916. The item-total correlations being above 0.30 and the removal of any item not causing a significant increase in the Cronbach α coefficient indicate that all items in the scale contribute homogeneously to the structure. Furthermore, within the scope of split-half reliability, the values were determined as 0.833 for the first half, 0.868 for the second half, with a correlation of 0.816 between the 2 halves, a Spearman-Brown coefficient of 0.899, and a Guttman Split-Half coefficient of 0.898, proving that the CCHF-DAS possesses a high level of internal consistency reliability. In the convergent and discriminant validity analysis, the CR value was found to be 0.931 (at an acceptable and excellent level), while the AVE value remained below 0.50 at 0.356. Similarly, McDonald Ω value was calculated as 0.931. Reliability analyses demonstrated excellent internal consistency of the CCHF-DAS. The Cronbach α coefficient for the overall scale was 0.916, exceeding the recommended threshold for high reliability [25]. In addition, split-half reliability coefficients and other internal consistency indices confirmed the stability and consistency of the scale items [14,28-31]. These findings suggest that the CCHF-DAS provides reliable measurements of disease awareness across items and respondents.

Although the AVE value (0.356) was below the recommended threshold of 0.50, the CR value of 0.931 indicated excellent construct reliability. As suggested by Fornell and Larcker [14], convergent validity may still be acceptable when CR exceeds 0.60. Moreover, Hair et al [26] recommend evaluating convergent validity using multiple complementary indicators, including factor loadings, CR, AVE, and overall model fit, rather than relying solely on AVE [26]. Nevertheless, the relatively low AVE should be considered a limitation, and future studies should further evaluate and refine the scale using independent samples.

Because the data were collected using self-report questionnaires administered at a single point in time, the potential influence of CMV was considered. The Harman single-factor test indicated that the first unrotated factor accounted for 35.66% of the total variance, which is below the commonly accepted threshold of 50%. This finding suggests that common method bias is unlikely to have substantially affected the observed relationships. Nevertheless, as with all self-report measures, the possibility of residual method bias cannot be completely excluded. Future studies should further minimize this potential limitation by incorporating multiple data sources, longitudinal designs, or procedural and statistical remedies for CMV.

Participants obtained a mean total score of 97.75 (SD 14.82) on the CCHF-DAS. This score indicates that while the target population possesses a certain level of awareness regarding CCHF, there remains an ongoing need for sustainable educational and intervention programs across the general public. Previous studies in the literature have examined individuals' knowledge, attitudes, and practices related to CCHF; however, these studies predominantly relied on author-generated questionnaires rather than standardized measurement instruments [18-24]. As a result, comparisons across studies and populations have been limited. To the best of our knowledge, the CCHF-DAS is the first validated and reliable scale that systematically assesses awareness of CCHF by addressing the causative agent, transmission routes, clinical manifestations, treatment, and preventive measures within a single measurement framework.

National and international studies have consistently reported insufficient levels of knowledge and preventive practices related to CCHF and have emphasized the need for comprehensive

XSL•FO
RenderX

strategies to control the disease [18-24]. The findings of the present study indicate that although participants demonstrated an overall acceptable level of awareness, there remains a substantial need to further enhance awareness, particularly given the ongoing disease burden, high mortality rates, and absence of a specific antiviral treatment or licensed vaccine. Assessing disease awareness using a validated instrument such as the CCHF-DAS may facilitate the identification of knowledge gaps and inform the development of targeted public health interventions aimed at promoting preventive behaviors and early health care seeking.

Limitations

While this study provides a robust and psychometrically sound instrument for assessing CCHF awareness, several limitations must be acknowledged to contextualize the findings appropriately:

- Geographic and sampling constraints: The study was conducted within a single geographic region, with participants recruited exclusively through family health centers. Consequently, the generalizability of the findings to broader populations, diverse demographic groups, or individuals residing in nonendemic or distinct socioeconomic settings remains limited. Future research should validate the CCHF-DAS across multicenter, nationwide, and culturally diverse cohorts.
- Cross-sectional design: The cross-sectional nature of the study precludes any causal inferences regarding the directionality of the relationships between influencing factors and CCHF awareness. Longitudinal designs are warranted to track how awareness evolves over time and to evaluate the predictive validity of the scale regarding actual protective behaviors.
- Absence of temporal reliability metrics: Because the scale was administered for the first time in this study, test-retest reliability and sensitivity to change over time could not be evaluated. Subsequent longitudinal investigations should assess the stability and responsiveness of the CCHF-DAS following targeted educational interventions.
- Methodological approach to factor analysis: Both EFA and CFA were performed using the same sample dataset. Although this practice appears in scale development literature, relying on a single sample increases the risk of overfitting. Future studies should test and confirm the factor structure using independent, newly drawn samples to establish stronger evidence of structural stability and external validity.

Conclusions

The CCHF-DAS, developed in this study, is a novel and psychometrically robust instrument capable of reliably and validly measuring CCHF awareness among adult populations, thereby filling a significant methodological gap in the literature. The findings demonstrate that the structural and internal consistency properties of the scale largely meet internationally accepted standards, confirming its utility for public health research. Beyond providing a psychometric measurement tool, the broader implications of this study offer critical insights for comprehensive public health policies and preventive medicine strategies:

- Epidemiological surveillance and risk stratification: The ability to monitor the awareness levels of adult populations residing in endemic regions using a standardized scale enables public health authorities to identify at-risk groups with greater precision. This facilitates the rational allocation of limited health resources, directing them toward the most vulnerable segments of the population.
- Evidence-based health education and intervention programs: The total mean scores obtained in this study and the detailed feedback provided by the scale items offer a concrete baseline to enhance the efficacy of community-based educational campaigns. Rather than relying on uniform and traditional informational strategies, the scale guides the design of dynamic intervention programs that directly target individuals' knowledge gaps and misconceptions.
- Interdisciplinary and sustainable control: Combating vector-borne zoonotic diseases such as CCHF requires coordinated efforts across medicine, veterinary science, environmental health, and the social sciences. The developed scale can serve as a common metric for evaluating the effectiveness of field studies conducted by various stakeholders, while also providing empirical data support for future epidemiological modeling.

In conclusion, the data generated through the CCHF-DAS should serve as a core component of strategic planning aimed at mitigating societal vulnerabilities, extending far beyond the mere measurement of individual awareness. Future longitudinal and independent sample validation studies across diverse geographical and cultural contexts will reinforce the universal validity of this scale and add significant value to global zoonotic disease management.

Public Health Implications

CCHF remains a significant public health challenge due to its high mortality rate, expanding geographic distribution, and the absence of a specific antiviral treatment or licensed vaccine. In this context, prevention and early detection are critically dependent on individual and community-level awareness. The CCHF-DAS provides a practical and evidence-based tool for assessing awareness levels and identifying knowledge gaps related to CCHF among adults.

The use of the CCHF-DAS in public health practice may support the planning, implementation, and evaluation of targeted educational and preventive interventions, particularly for high-risk groups such as agricultural workers, livestock handlers, health care workers, and individuals living in endemic regions. By enabling the systematic measurement of awareness, the scale can help guide resource allocation, tailor risk communication strategies, and monitor the effectiveness of awareness-raising programs over time. Ultimately, improving individual awareness of CCHF through informed public health interventions may contribute to reducing disease transmission, morbidity, and mortality at the population level.

XSL•FO
RenderX

Acknowledgments

No generative AI tools were used.

Funding

The authors declared no financial support was received for this work.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: GY, AÇ
Data curation: GY, AÇ
Formal analysis: GY, AÇ
Methodology: GY, AÇ
Project administration: GY, AÇ
Supervision: GY, AÇ
Writing – original draft: GY, AÇ, EK
Writing – review & editing: GY, AÇ

Conflicts of Interest

None declared.

Multimedia Appendix 1
Turkish version of the scale.
[DOCX File, 17 KB - publichealth_v12i1e93724_app1.docx]

References

1. Crimean-Congo haemorrhagic fever. World Health Organization. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/crimean-congo-haemorrhagic-fever [accessed 2026-08-26]
2. Fırat M, Hashemi MH, Dawod MAA. Crimean-Congo haemorrhagic fever (CCHF): an overview [Article in Turkish]. Karatekin Univ J Sci 2023;2(1):33-42 [FREE Full text]
3. Kara SS. Retrospective evaluation of children with Crimean-Congo hemorrhagic fever syndrome in the city of Erzurum [Article in Turkish]. Pamukkale Med J 2018;11(3):209-214. [doi: 10.31362/patd.399695]
4. Sayiner HS, Sahin MS, Biyik M, Selcuk MY, Aksoz S, Oznas O. Crimean-Congo haemorrhagic fever cases without travel history to endemic areas in Adıyaman. Klimik Dergisi 2017;30(3):142-145. [doi: 10.5152/kd.2017.34]
5. Kırım kongo kanamalı ateşi (hekimlere yönelik) [Report in Turkish]. : T.C. Sağlik Bakanliği; 2023 URL: https://hsgm.saglik.gov.tr/depo/birimler/zoonotik-ve-vektorel-hastaliklar-db/Dokumanlar/Sunumlar/KKKA_Sunum_Hekimlere_Yonelik_2023.pdf [accessed 2026-05-30]
6. Coşkun MV, Öztürk M, Özdemir YE. A healthcare worker's nıghtmare: delayed dıagnosıs of CCHF case [Article in Turkish]. Eskisehir Med J 2024;5(1):29-31 [FREE Full text]
7. Korkmaz D, Konya P, Demirtürk N. Investigation of the characteristics of Crimean Congo hemorrhagic fever cases reported in Afyonkarahisar province. Turkiye Parazitol Derg 2022 Sep 12;46(3):224-227. [doi: 10.4274/tpd.galenos.2022.14633] [Medline: 36094125]
8. Bernard C, Holzmuller P, Bah MT, et al. Systematic review on Crimean-Congo hemorrhagic fever enzootic cycle and factors favoring virus transmission: special focus on France, an apparently free-disease area in Europe. Front Vet Sci 2022;9:932304. [doi: 10.3389/fvets.2022.932304] [Medline: 35928117]
9. Leblebicioglu H, Ozaras R, Irmak H, Sencan I. Crimean-Congo hemorrhagic fever in Turkey: current status and future challenges. Antiviral Res 2016 Feb;126:21-34. [doi: 10.1016/j.antiviral.2015.12.003] [Medline: 26695860]
10. Ser O, Cetin H. The current situation of Crimean-Congo hemorrhagic fever [Article in Turkish]. TAF Prev Med Bull 2016;15(1):58. [doi: 10.5455/pmb.1-1437766801]
11. Tavsancil E. Measurement of Attitudes and Data Analysis With SPSS [Book in Turkish], 4th edition: Nobel Publishing Distribution; 2010. URL: https://www.nobelyayin.com/tutumlarin-olculmesi-ve-spss-ile-veri-analizi-2495.html [accessed 2026-09-01]
12. Brown TA. Confirmatory Factor Analysis for Applied Research, 2nd edition: Guilford Publications; 2015.
13. Kline RB. Principles and Practice of Structural Equation Modeling, 4th edition: Guilford Press; 2016.
14. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. J Mark Res 1981 Feb;18(1):39-50. [doi: 10.1177/002224378101800104]

XSL•FO
RenderX

15. Podsakoff PM, MacKenzie SB, Lee JY, Podsakoff NP. Common method biases in behavioral research: a critical review of the literature and recommended remedies. J Appl Psychol 2003 Oct;88(5):879-903. [doi: 10.1037/0021-9010.88.5.879] [Medline: 14516251]
16. Kock N. Common method bias in PLS-SEM: a full collinearity assessment approach. IJeC 2015;11(4):1-10. [doi: 10.4018/ijec.2015100101]
17. Harman D. A single factor test of common method variance. J Psychol 1967;35:359-378 [FREE Full text]
18. Ayebare D, Menya M, Mulyowa A, Muhwezi A, Tweyongyere R, Atim SA. Knowledge, attitudes, and practices of Crimean Congo hemorrhagic fever among livestock value chain actors in Kagadi district, Uganda. PLoS Negl Trop Dis 2023 Feb;17(2):e0011107. [doi: 10.1371/journal.pntd.0011107] [Medline: 36730376]
19. Ahmed A, Tanveer M, Saqlain M, Khan GM. Knowledge, perception and attitude about Crimean Congo Hemorrhagic Fever (CCHF) among medical and pharmacy students of Pakistan. BMC Public Health 2018 Dec 3;18(1):1333. [doi: 10.1186/s12889-018-6248-1] [Medline: 30509226]
20. Nejati J, Mohammadi M, Okati-Aliabad H. Knowledge, attitudes, and practices regarding Crimean-Congo hemorrhagic fever in a high-prevalence suburban community, southeast of Iran. Heliyon 2023 Jan 15;10(1):e23414. [doi: 10.1016/j.heliyon.2023.e23414] [Medline: 38148794]
21. Jamil H, Din MFU, Tahir MJ, et al. Knowledge, attitudes, and practices regarding Crimean-Congo hemorrhagic fever among general people: a cross-sectional study in Pakistan. PLoS Negl Trop Dis 2022 Dec;16(12):e0010988. [doi: 10.1371/journal.pntd.0010988] [Medline: 36480553]
22. Yilmaz R, Ozcetin M, Erkorkmaz U, Ozer S, Ekici F. Public knowledge and attitude toward Crimean Congo hemorrhagic fever in Tokat Turkey. Iran J Arthropod Borne Dis 2009;3(2):12-17. [Medline: 22808377]
23. Koculu S, Oncul A, Onal O, Yesilbag Z, Uzun N. Evaluation of knowledge of the healthcare personnel working in Giresun province regarding Crimean-Congo hemorrhagic fever before and after educational training. J Vector Borne Dis 2015 Jun;52(2):166-170. [Medline: 26119550]
24. Almorish MAW, Al-Sayaghi KM, Alshoabi SA, et al. Knowledge and attitudes regarding Crimean-Congo hemorrhagic fever among slaughterhouse workers in Sana'a and Dhamar cities-Yemen: a cross-sectional study. J Multidiscip Healthc 2024;17:3875-3886. [doi: 10.2147/JMDH.S470446] [Medline: 39155975]
25. Çapık C, Gözüm S, Aksayan S. Intercultural scale adaptation stages, language and culture adaptation: updated guideline. Florence Nightingale J Nurs 2018;26(3):199-210. [doi: 10.26650/FNJN397481]
26. Hair JF, Black WC, Babin BJ, Anderson RE. Multivariate Data Analysis, 7th edition: Pearson; 2010.
27. Costello AB, Osborne J. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. Practical Assess Res Eval 2005;10(1):7. [doi: 10.7275/jyj1-4868]
28. Çapık C. Use of confirmatory factor analysis in validity and reliability studies [Article in Turkish]. J Anatolia Nurs Health Sci 2014;17(3):196-205 [FREE Full text]
29. McDonald RP. Test Theory: A Unified Treatment: Lawrence Erlbaum Associates; 1999.
30. Dunn TJ, Baguley T, Brunsden V. From alpha to omega: a practical solution to the pervasive problem of internal consistency estimation. Br J Psychol 2014 Aug;105(3):399-412. [doi: 10.1111/bjop.12046] [Medline: 24844115]
31. Henseler J, Ringle CM, Sarstedt M. A new criterion for assessing discriminant validity in variance-based structural equation modeling. J Acad Mark Sci 2015 Jan;43(1):115-135. [doi: 10.1007/s11747-014-0403-8]

Abbreviations

AGFI: adjusted goodness-of-fit index
AVE: average variance extracted
CCHF: Crimean-Congo Hemorrhagic Fever
CCHF-DAS: Crimean-Congo Hemorrhagic Fever Disease Awareness Scale
CFA: confirmatory factor analysis
CFI: comparative fit index
CMV: common method variance
CR: composite reliability
EFA: exploratory factor analysis
GFI: goodness-of-fit index
I-CVI: Item Content Validity Index
KMO: Kaiser-Meyer-Olkin
RMSEA: root mean square error of approximation
SRMR: standardized root mean square residual
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 18.Feb.2026; peer-reviewed by E Lau, GUN Iskender, O Basol; revised version received 04.Aug.2026; accepted 24.Aug.2026; published 14.Sep.2026.

Please cite as:
Yilmazel G, Çalmaz A, Keleş E
Crimean-Congo Hemorrhagic Fever Disease Awareness Scale: Scale Development, Validity, and Reliability Study
JMIR Public Health Surveill 2026;12:e93724
URL: https://publichealth.jmir.org/2026/1/e93724
doi:10.2196/93724

© Gülay Yilmazel, Ayşe Çalmaz, Emre Keleş. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 14.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Antimicrobial Resistance Across Four Iraqi Governorates: A Multicenter Cross-Sectional Surveillance Analysis With Implications for Antimicrobial Stewardship Policy and WHO AWaRe–Aligned Public Health Strategies

Hemn Omer[1], MSc; Omeed Darweesh[2], PhD; Gulbahar Karim[3], PhD; Baref Rashid[4], PhD; Rozhgar Ahmed[2], MSc; Andrew Seaton[5], MD; Brian Godman[6,7,8], PhD; Amanj Kurdi[2,6,7], PhD

[1]College of Pharmacy, Al-Qalam University, Kirkuk, Kirkuk, Iraq

[2]College of Pharmacy, Al-Kitab University, Kirkuk 36015, Iraq

[3]Department of Basic Science, College of Nursing, University of Kirkuk, Kirkuk, Iraq

[4]Kurdistan Higher Council of Medical specialties, Erbil, Iraq

[5]Infection unit, Queen Elizabeth University Hospital, NHS Greater Glasgow and Clyde, Glasgow, Glasgow, United Kingdom

[6]Department of Public Health Pharmacy and Management, School of Pharmacy, Sefako Makgatho Health Sciences University, Pertoria, Gauteng, South Africa

[7]Strathclyde Institute of Pharmacy and Biomedical Sciences, University of Strathclyde, 161 Cathedral Street, Glasgow, Scotland, United Kingdom

[8]Antibiotic Policy Group, Institute for Infection and Immunity, City St. George's, University of London, London, United Kingdom

Corresponding Author:
Amanj Kurdi, PhD
College of Pharmacy, Al-Kitab University, Kirkuk 36015, Iraq

Abstract

Background: Antimicrobial resistance is a major public health threat globally, with disproportionate impact in low- and middle-income countries, where surveillance systems are often fragmented. In Iraq, nationally integrated multicenter resistance data to support stewardship policy and World Health Organization Access Watch Reserve (WHO AWaRe)–aligned interventions remain limited.

Objective: The objective of this study is to characterize multicenter antimicrobial resistance patterns across inpatient and outpatient laboratory settings in selected Iraqi governorates, and evaluate their implications for surveillance strengthening and antimicrobial stewardship policy.

Methods: We performed a multicenter cross-sectional surveillance analysis of routine microbiology data from four Iraqi governorates, representing approximately 10% of Iraq's population. Deduplicated isolates (first isolate per patient per specimen type) were included. Multidrug resistance was defined as non-susceptibility to at least one agent in three or more antimicrobial classes. Resistance proportions with 95% CI were calculated and compared by clinical setting and specimen type.

Results: Among 846 isolates, *Escherichia coli* (n=235/846, 27.8%), *Staphylococcus aureus* (n=151/846, 17.9%), *Pseudomonas aeruginosa* (n=65/846, 7.7%), and *Klebsiella pneumoniae* (n=59/846, 7.0%) predominated. Overall multidrug resistance prevalence was 57.3% (95% CI 53.9 - 60.6) and was significantly higher among inpatients (P=.04). Third-generation cephalosporin resistance was detected in 69.3% (106/153) of *Escherichia coli* and 44.4% (8/18) of *Klebsiella pneumoniae* isolates. Carbapenem resistance was present in 12.4% (20/161) of *Escherichia coli*, 11.1% (2/18) of *Klebsiella pneumoniae*, and 29.4% (5/17) of *Pseudomonas aeruginosa*. Methicillin resistance was identified in 89.8% (136/151) of *Staphylococcus aureus* isolates.

Conclusions: High multidrug resistance prevalence and substantial resistance among priority pathogens highlight urgent needs for strengthened surveillance integration, routine antibiogram reporting, and antimicrobial stewardship interventions aligned with WHO AWaRe targets. These findings provide multicenter subnational evidence to support surveillance strengthening, routine antibiogram reporting, and antimicrobial stewardship policy refinement in Iraq.

(JMIR Public Health Surveill 2026;12:e94436) doi:10.2196/94436

KEYWORDS

antimicrobial resistance; surveillance; multidrug resistance; WHO AWaRe; antimicrobial stewardship; public health policy; health systems strengthening.; World Health Organization Access Watch Reserve

XSL•FO
RenderX

Introduction

Antimicrobial resistance (AMR) is a major global public health challenge that threatens progress toward universal health coverage and sustainable development. According to the 2019 global burden of AMR analysis, bacterial AMR was estimated to be directly attributable to approximately 1.27 million deaths globally, with a further 4.95 million associated deaths, disproportionately affecting low- and middle-income countries (LMICs) where surveillance capacity is often limited [1-3]. In fragile and postconflict settings such as Iraq, structural challenges affecting laboratory capacity, antimicrobial regulation, and data integration may amplify resistance emergence and constrain effective surveillance and policy response [4,5].

Effective AMR response depends on robust, representative, and integrated surveillance systems capable of generating actionable data for clinicians, policymakers, and public health authorities. Surveillance data not only inform empiric treatment guidelines and antimicrobial stewardship (AMS) interventions, but also underpin national action planning, resource allocation, and monitoring of progress toward World Health Organization (WHO) and United Nations AMR targets. The WHO established the Global Antimicrobial Resistance and Use Surveillance System (GLASS) to standardize AMR reporting and strengthen national surveillance capacity [6]. However, while GLASS provides an essential global surveillance framework, important methodological and structural constraints affect interpretation of country-specific data, particularly in LMICs with incomplete laboratory coverage, nonstandardized testing panels, and limited denominator information [6]. In many participating countries, including Iraq, surveillance data are derived from selected sentinel laboratories without systematic linkage to population denominators, health care usage metrics, or standardized case definitions. Addressing these gaps requires complementary analyses integrating deduplicated, isolate-level data across multiple centers to improve operational relevance and local interpretability. In Iraq, available AMR data are largely derived from selected sentinel laboratories and may not fully capture national pathogen distribution, rural–urban variation, private-sector testing, or health care-associated infection dynamics [7,8]. This is important as the Middle East has been identified as a region with a high and growing burden of AMR, driven by a combination of health system fragility, regulatory challenges, and heterogeneous laboratory infrastructures [5]. Iraq represents a critical case study for AMR surveillance strengthening in fragile health systems. Decades of conflict, economic instability, and infrastructure disruption have affected laboratory networks, data integration systems, infection prevention capacity, and antimicrobial governance frameworks [9]. While Iraq participates in GLASS and has developed a national action plan on AMR (2018 - 2022), implementation challenges have been documented, including fragmented surveillance systems, limited routine data integration, and variable enforcement of antimicrobial regulations [9,10]. This is similar to other LMICs that experience difficulties with meeting national action plan goals due to concerns with available resources as well as personnel [11],

Published studies from Iraq consistently report high levels of resistance among key bacterial pathogens. Multidrug-resistant (MDR) *Acinetobacter spp*., *extended-spectrum beta-lactamase* (ESBL)-producing *Escherichia coli*, and *Klebsiella pneumoniae*, *carbapenem-resistant Pseudomonas aeruginosa*, and *methicillin-resistant Staphylococcus aureus (MRSA*) have been described across multiple hospital settings [4,8,12-15]. Intensive care unit-based studies have further demonstrated discordance between empiric antibiotic prescribing and susceptibility results, reflecting both diagnostic constraints and evolving resistance profiles [4,7,16]. Community-derived isolates have also shown increasing resistance to commonly used antibiotics, raising concerns regarding first-line treatment effectiveness [4,9].

Despite this growing body of evidence, AMR data in Iraq remain fragmented across provinces, institutions, and study designs, with limited integration into a coordinated national surveillance platform. To date, no multicenter analysis has synthesized deduplicated hospital and community resistance data across major bacterial pathogens while explicitly contextualizing findings within GLASS reporting and WHO Access Watch Reserve (AWaRe) stewardship frameworks. Iraq occupies a strategic geographic position with substantial population mobility, cross-border health care usage, regional displacement, and international travel. Resistant organisms emerging in fragile or conflict-affected settings may disseminate regionally and globally through migration, cross-border health care usage, and population displacement, reinforcing the importance of strengthening surveillance systems beyond national boundaries [17]. In addition, patterns observed in Iraq reflect challenges common to many LMIC and postconflict health systems, including limited diagnostic infrastructure, empiric prescribing pressures, and uneven surveillance integration. Consequently, characterizing current AMR patterns in Iraq contributes to the global understanding of AMR epidemiology in fragile settings and provides transferable lessons for surveillance strengthening and empiric therapy optimization internationally. In settings where empiric therapy is frequently initiated prior to culture results, absence of consolidated national resistance data may contribute to suboptimal treatment selection and further resistance selection pressure [7,18,19]. This gap can be partly addressed by conducting a systematic national synthesis of AMR patterns in Iraq. Such an analysis can clarify pathogen-specific resistance distributions, identify priority resistance threats, contextualize Iraqi data within global estimates, and provide an evidence base to inform empiric treatment recommendations and surveillance strengthening. Consequently, this study aims to generate a multicenter, deduplicated synthesis of bacterial AMR patterns across inpatient and outpatient laboratory settings in selected Iraqi governorates in Iraq and to evaluate their implications for strengthening surveillance integration, stewardship policy reform, and WHO AWaRe–aligned public health strategies.

Methods

Study Design and Setting

This study was designed as a multicenter cross-sectional surveillance analysis of routinely collected microbiology

XSL•FO
RenderX

laboratory data from four Iraqi governorates: Kirkuk, Tikrit, Sulaymaniyah, and Mosul. The primary objective was descriptive surveillance characterization of resistance patterns across participating laboratories rather than inferential epidemiological estimation or causal assessment of resistance determinants. Accordingly, the findings should be interpreted within the context of routine operational surveillance data generated through clinical diagnostic practice. These provinces were purposively selected to provide geographically and demographically diverse surveillance coverage; however, they were not intended to constitute a nationally representative sampling framework. Collectively, the included catchment areas cover approximately 10% of Iraq's estimated population of 47 million. The inclusion of multiple provinces was intended to overcome the limitations of prior Iraqi studies [12,14,15,19]; which were often restricted to single cities, and/or pathogens. In each governorate, data were obtained from the principal tertiary public hospital laboratory, which functions as the primary referral microbiology center for surrounding districts. These laboratories constitute major public-sector diagnostic nodes within their respective regions, although they do not encompass private laboratories or smaller peripheral facilities. Data were collected between July and December 2025 and included isolates from both inpatient and outpatient settings. Outpatient status should not be interpreted as equivalent to community-acquired infection, as information on prior health care exposure, referral history, recent hospitalization, or previous antibiotic use was not available within the dataset. Microbiological testing was ordered according to routine clinician-driven practice rather than through a predefined surveillance sampling protocol. Testing is provided free of charge within participating public hospitals. As denominator data on total admissions, outpatient visits, and infection episodes were unavailable, findings represent laboratory-confirmed resistance patterns among tested patients rather than population incidence or prevalence estimates. Because testing followed routine clinician-directed practice rather than structured sampling, the dataset may preferentially capture clinically complex, severe, treatment-refractory, or diagnostically investigated cases. Consequently, the findings should be interpreted as operational surveillance observations, and the scope of inference is restricted to describing resistance distributions within participating laboratory settings rather than deriving nationally representative epidemiological estimates. This analysis therefore reflects operational surveillance data derived from routine diagnostic activity rather than structured sentinel sampling. The study was conducted and reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology

guidance for cross-sectional observational studies [20]; (Checklist 1) and aligns with WHO GLASS methodological principles for isolate-based AMR surveillance [21].

Study Population and Eligibility Criteria

All patients of any age or sex with at least one positive bacterial culture during the study period were eligible for inclusion. Extracted variables included patient age, sex, inpatient or outpatient status, governorate, specimen type, bacterial species identification, and antimicrobial susceptibility testing (AST) results. Anonymized laboratory records were extracted retrospectively from the electronic laboratory information systems of the principal tertiary public hospital laboratory in each participating governorate using a standardized data extraction template. Data extraction was undertaken locally by designated laboratory personnel, after which records were anonymized, quality checked for completeness and consistency, deduplicated in accordance with WHO GLASS guidance, and combined into a single dataset for analysis. To minimize duplication and align with international surveillance standards, only the first isolate per patient per specimen type during the study period was included in the analysis. Repeat isolates of the same organism from the same patient and specimen source were excluded, consistent with WHO GLASS deduplication guidance [21]. Negative cultures, fungal-only isolates, and records with incomplete susceptibility or missing key demographic variables were excluded. Unexpected or discrepant susceptibility profiles underwent internal laboratory review and repeat testing using disk diffusion and/or VITEK 2 (bioMérieux) confirmation where applicable. Isolates lacking essential susceptibility information required for resistance phenotype assignment, including MDR classification, were excluded from final analysis. Data were anonymized prior to aggregation and analysis to ensure compliance with ethical and data governance standards.

Specimen Collection and Microbiological Identification

Specimens were obtained as part of routine clinical care in accordance with standard clinical microbiology practice. Specimen types included urine, blood, wound swabs, vaginal swabs, throat swabs, stool samples, respiratory tract specimens, and sterile body fluids such as cerebrospinal, pleural, synovial, and ascitic fluids. Blood cultures were processed using automated systems including BACT or ALERT 3D (bioMérieux) and BD BACTEC FX40 (Becton, Dickinson and Company). Other specimens were inoculated onto appropriate selective and differential media including blood agar, chocolate agar, and Brilliance UTI Clarity agar (Oxoid). Primary bacterial identification was performed locally using routine microbiological workflows combining conventional phenotypic and biochemical methods with automated platforms according to laboratory practice. Standard biochemical testing included Gram staining, catalase, and coagulase testing for Gram-positive cocci, and oxidase, triple sugar iron, citrate, and related biochemical assays for Gram-negative bacilli. The VITEK 2 Compact automated system was used for automated organism identification and susceptibility confirmation, particularly for selected MDR, complex, or diagnostically challenging isolates including *P. aeruginosa* and *K. pneumonia*. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) [22]; was not available across all participating centers but was used centrally at the Infectious Diseases Control Center in Baghdad as a confirmatory reference method for isolates with atypical biochemical profiles, unusual clinical contexts, or suspected novel resistance phenotypes.

AST

AST was performed using Kirby–Bauer disk diffusion, VITEK 2 automated testing, and/or minimum inhibitory concentration

XSL•FO
RenderX

(MIC)-based methods depending on organism type, isolate complexity, and institutional laboratory protocol. Kirby–Bauer disk diffusion represented the primary AST method across participating laboratories, while VITEK 2 and MIC-based approaches were used for selected isolates requiring confirmatory susceptibility testing, particularly suspected MDR organisms and selected *P. aeruginosa* and *K. pneumonia* isolates. AST interpretation was standardized across centers using Clinical and Laboratory Standards Institute (CLSI) M100, 35th edition (2025) criteria [23]. All susceptibility categories and MIC breakpoints were interpreted according to the 2025 standards to support intercenter comparability. Although routine antibiotic panels varied modestly across centers because of organism-specific testing practices, reagent availability, and local laboratory capacity, laboratories maintained core pathogen-specific antibiotic panels to support consistent classification of multidrug resistance. Routine quality control procedures were implemented using standard reference strains, including *E. coli* American Type Culture Collection 25922 and *S. aureus* American Type Culture Collection 25923, tested monthly and when new reagent or disk batches were introduced, in accordance with CLSI M100-2025 guidance. For MRSA classification, cefoxitin disk diffusion (30 µg) was used as the primary standardized screening method across participating centers in accordance with CLSI recommendations. Although oxacillin was used in a limited number of local workflows, all isolates included in the final MRSA analysis were validated using cefoxitin criteria to ensure uniform classification. Carbapenem resistance was defined as nonsusceptibility to at least one tested carbapenem agent (imipenem, meropenem, and/or ertapenem) according to CLSI M100-2025 criteria. To mitigate interlaboratory heterogeneity, interpretive standardization was maintained through uniform application of CLSI M100-2025 criteria across all testing platforms. Discrepant or unusual resistance phenotypes underwent confirmatory assessment through advanced methods at the Infectious Diseases Control Center reference laboratory where required. Glycopeptide susceptibility findings, including vancomycin non-susceptibility among S. aureus, should be interpreted within the context of routine multicenter laboratory surveillance data, recognizing that confirmatory MIC testing and reference laboratory validation were not uniformly available for all isolates. These standards provide internationally accepted breakpoints for categorizing isolates as susceptible, intermediate, or resistant. For surveillance consistency and conservative public health estimation, isolates categorized as intermediate were grouped with resistant isolates, consistent with WHO GLASS analytical methodology [21]. MDR was defined as non-susceptibility to at least one agent in three or more antimicrobial classes, in accordance with the internationally standardized definitions proposed by Magiorakos et al [24]. Terminology throughout the manuscript was aligned with WHO GLASS reporting conventions. The term “non-susceptible” was used for AST outcomes where appropriate, while specific resistance terminology was retained for recognized resistance phenotypes such as MRSA, carbapenem-resistant organisms, and intrinsic resistance classifications.

Statistical Analysis

Descriptive surveillance analytics were applied to summarize demographic characteristics, specimen distribution, pathogen prevalence, and AMR proportions. Given the surveillance-oriented design, analyses were intentionally descriptive and exploratory, aiming to characterize operational resistance patterns across clinical settings and specimen types rather than test causal hypotheses or estimate independent risk relationships. Survey weighting was not applied because the dataset comprised all eligible deduplicated isolates obtained through routine laboratory surveillance from participating centers rather than a probability-based sample of the Iraqi population. Consequently, no sampling probabilities or survey weights were available or appropriate, and all analyses were performed using unweighted data. Although selected demographic and laboratory variables were available, the dataset lacked several key determinants required for robust adjusted modeling, including prior antibiotic exposure, recent hospitalization, health care contact history, comorbidity burden, invasive procedures, intensive care unit admission, infection severity, and prior microbiological history. Consequently, multivariable modeling was not undertaken because adjusted estimates based on incompletely measured resistance determinants could be vulnerable to substantial residual confounding and potentially misleading causal interpretation. Resistance rates were calculated as the proportion of resistant isolates among tested isolates for each pathogen–antibiotic combination, consistent with WHO GLASS reporting methodology [21]. Proportions were reported with 95% CI to quantify statistical uncertainty. Comparisons between inpatient and outpatient populations, across governorates, and between specimen types were conducted using χ^2 or Fisher exact tests as appropriate. These subgroup comparisons were intended as descriptive exploratory analyses to characterize variation across clinical contexts rather than to identify independent predictors of resistance. Accordingly, multivariable modeling and formal adjustment for multiple comparisons were not undertaken because they were outside the primary surveillance objectives of this study. Records with missing key demographic variables or incomplete susceptibility information required for resistance phenotype classification were excluded during data cleaning, as described in the eligibility criteria. Analyses were therefore performed using complete available data for each relevant pathogen–antibiotic combination, resulting in variable denominators across selected comparisons. A two-sided P value <.05 was considered statistically significant. Statistical analyses were performed using SPSS version 28.0 (IBM Corp).

Ethical Considerations

Ethical approval was obtained from the Institutional Review Board of Al-Kitab University (approval number: 6/557). As the study involved retrospective analysis of anonymized laboratory data without direct patient contact, the requirement for informed consent was waived. All procedures were conducted in accordance with the Declaration of Helsinki and applicable national ethical standards [25]. Data were handled in accordance with institutional data governance procedures, and no identifiable patient information was retained within the analytical dataset.

XSL•FO
RenderX

Results

Study Population and Specimen Characteristics

A total of 846 deduplicated bacterial isolates were included in this multicenter surveillance dataset. Urine specimens accounted for over half of isolates (n=449, 53.0%), followed by vaginal swabs (119/846, 14.0%), wound swabs (95/846, 11.2%), and blood cultures (86/846, 10.1%) (Table 1). The predominance of urine specimens is consistent with routine diagnostic testing patterns and indicates that urinary tract infections constitute a major driver of microbiological sampling within participating centers. Two-thirds of isolates were obtained from female patients (559/846, 66.0%), consistent with the high proportion of urinary and vaginal specimens. The largest age group represented was 21 - 30 years (194/846, 22.9%). Most isolates were derived from outpatient settings (535/846, 63.2%), indicating that a substantial proportion of detected resistance was identified among isolates obtained through outpatient clinical pathways. However, outpatient status does not necessarily imply community-acquired infection because health care exposure history was unavailable. Geographically, isolates were distributed across four governorates, with Kirkuk contributing 43.9% (372/846) and Tikrit 25.1% (213/846), reflecting laboratory catchment activity rather than population-weighted sampling.

Table . Demographic and clinical characteristics of unique patients (N=846) and specimen distribution of bacterial isolates across four Iraqi governorates[a].

Variables and categories	Total, N (%)	Urine, n (%)	Vaginal swab, n (%)	Wound swab n (%)	Blood, n (%)	Throat swab, n (%)	Stool, n (%)
Total							
Total patients	846 (100)	449 (53.0)	119 (14.0)	95 (11.2)	86 (10.1)	68 (8.0)	29 (3.4)
Sex							
Male	287 (33.9)	154 (53.6)	0 (0)	51 (17.7)	43 (14.9)	27 (9.4)	12 (4.1)
Female	559 (66.1)	295 (52.7)	119 (21.2)	44 (7.8)	43 (7.6)	41 (7.3)	17 (3.0)
Age, years							
0 - 2	36 (4.2)	16 (44.4)	0 (0)	6 (16.6)	10 (27.7)	3 (8.3)	1 (2.7)
>2 - 12	57 (6.7)	31 (54.3)	0 (0)	7 (12.2)	13 (22.8)	1 (1.7)	5 (8.7)
13 - 20	103 (12.1)	47 (45.6)	11 (10.6)	14 (13.5)	8 (7.7)	17 (16.5)	6 (5.8)
21 - 30	194 (22.9)	109 (56.1)	34 (17.5)	22 (11.3)	18 (9.2)	6 (3.0)	5 (2.5)
31 - 40	173 (20.4)	87 (50.2)	38 (21.9)	18 (10.4)	10 (5.7)	14 (8.0)	6 (3.4)
41 - 50	98 (11.5)	58 (59.1)	19 (19.3)	8 (8.1)	6 (6.1)	5 (5.1)	2 (2.0)
51 - 60	68 (8.0)	30 (44.1)	10 (14.7)	5 (7.3)	8 (11.7)	14 (20.6)	1 (1.4)
>60	117 (13.8)	71 (60.7)	7 (5.9)	15 (12.8)	13 (11.1)	8 (6.8)	3 (2.5)
Patient status							
Outpatient	535 (63.2)	307 (57.3)	94 (17.5)	56 (10.4)	13 (2.4)	48 (8.9)	17 (3.1)
Inpatient	311 (36.7)	142 (45.6)	25 (8.0)	39 (12.5)	73 (23.4)	20 (6.4)	12 (3.8)
Governorate distribution							
Kirkuk	372 (43.9)	196 (52.7)	32 (8.6)	37 (9.9)	70 (18.8)	22 (5.9)	15 (4.0)
Tikrit	213 (25.1)	115 (53.9)	54 (25.3)	15 (7.0)	0 (0)	19 (8.9)	10 (4.6)
Sulaymaniyah	138 (16.3)	63 (45.6)	25 (18.1)	19 (13.7)	14 (10.1)	17 (12.3)	0 (0.0)
Mosul	123 (14.5)	75 (60.9)	8 (6.5)	24 (19.5)	2 (1.6)	10 (8.1)	4 (3.2)

[a]Percentages in the total column are calculated relative to all included patients. Percentages presented across specimen columns are row percentages showing the distribution of each demographic or clinical subgroup across specimen categories; therefore, percentages within rows may sum to approximately 100% owing to rounding.

Distribution of Bacterial Pathogens

Among the 846 isolates, more than 20 bacterial species were identified. *E. coli* was the most frequently isolated pathogen (n=235, 27.78%), followed by *S. aureus* (n=151, 17.85%), *P. aeruginosa* (n=65, 7.68%), and *K. pneumonia* (n=59, 6.97%). Together, these four pathogens accounted for more than 60% (510/846) of all isolates. Other notable pathogens included *Enterococcus faecalis* (n/N43/846, 5.08%), *Staphylococcus haemolyticus* (41/846, 4.85%), *Acinetobacter baumannii* (36/846, 4.26%), and *Streptococcus pyogenes* (27/846, 3.19%) (Figure 1). The distribution of the top 16 most common pathogens demonstrates the predominance of Gram-negative organisms, particularly *E. coli*. Comparison between inpatient and outpatient populations showed variation in pathogen distribution (Figure 2). *E. coli* was more frequently identified

XSL•FO
RenderX

among outpatients, whereas *Acinetobacter baumannii* and *Staphylococcus haemolyticus* were proportionally more common among inpatients.

Figure 1. Distribution of the top 16 bacterial pathogens isolated from clinical specimens in four Iraqi Governorates (N=846).

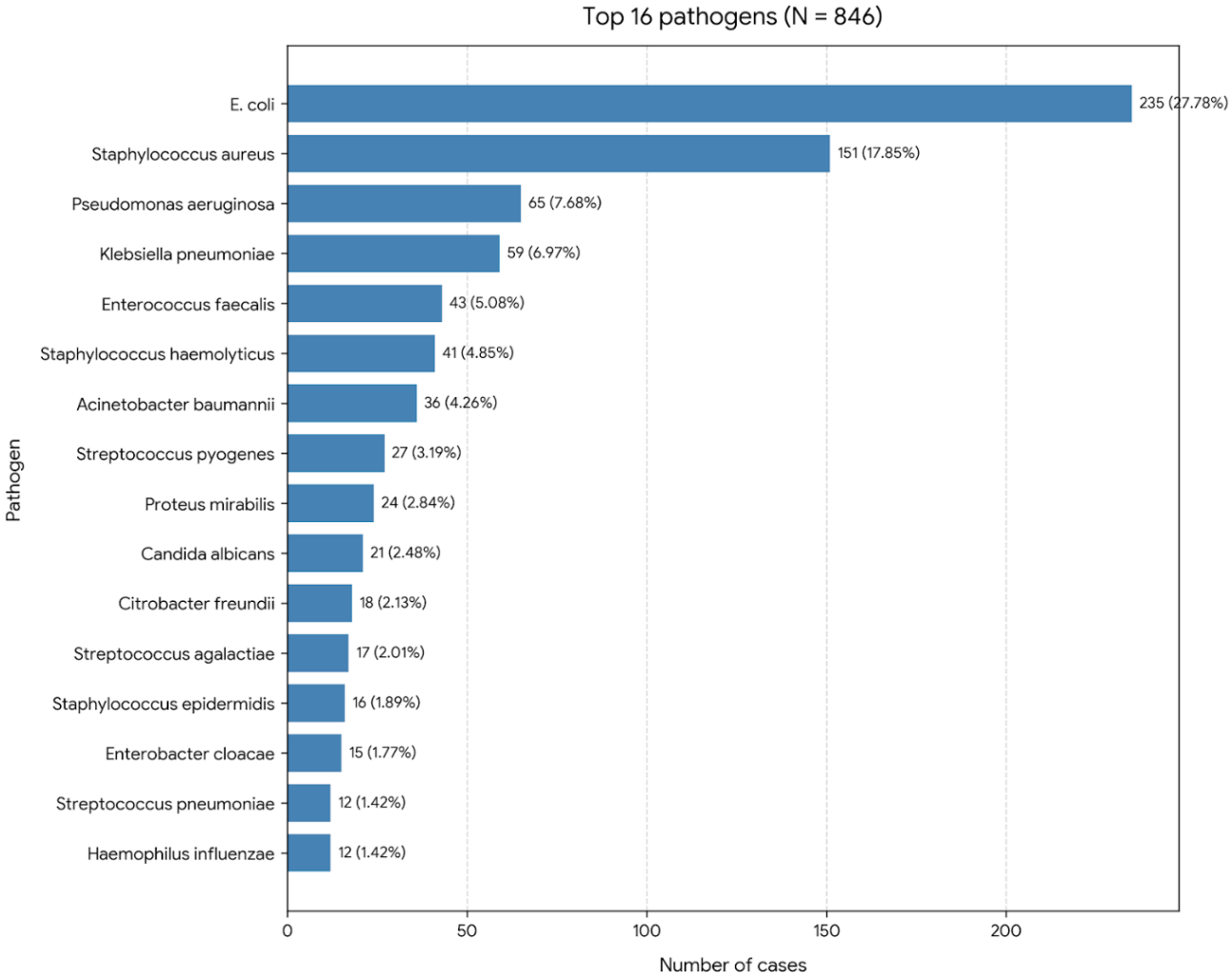
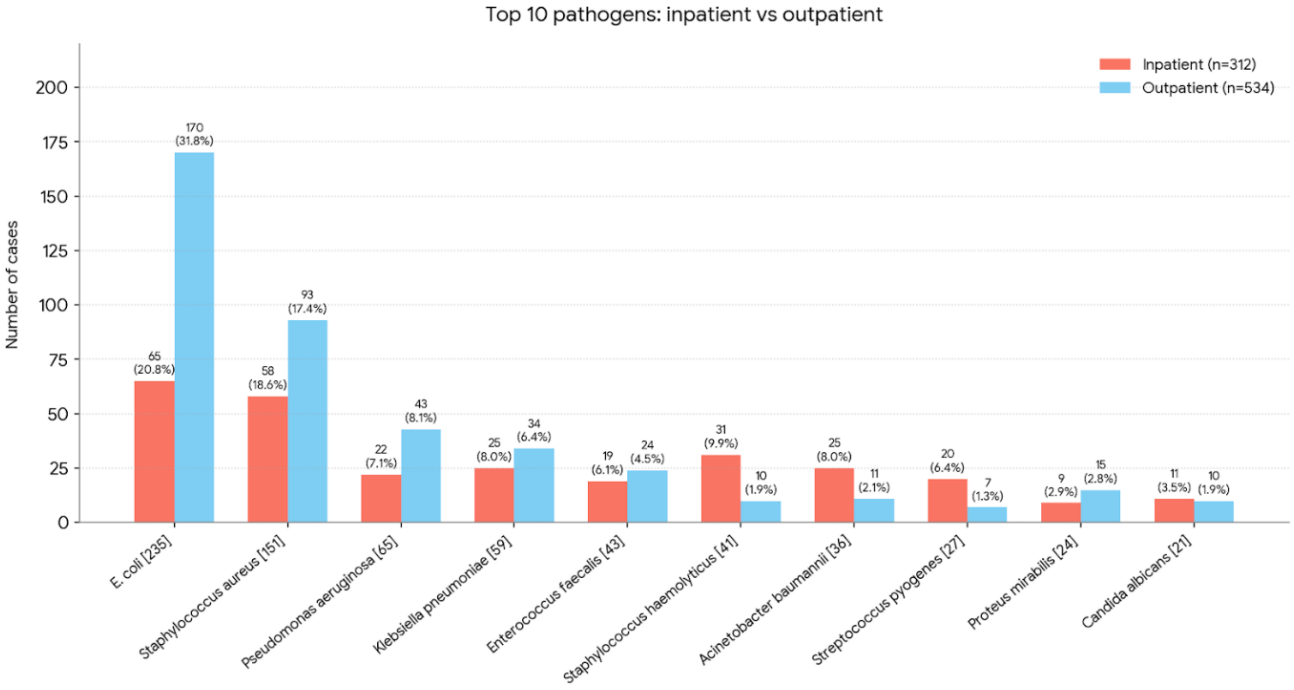


Figure 2. Distribution of the top 10 bacterial pathogens by clinical setting (Inpatient vs outpatient) in four Iraqi Governorates.

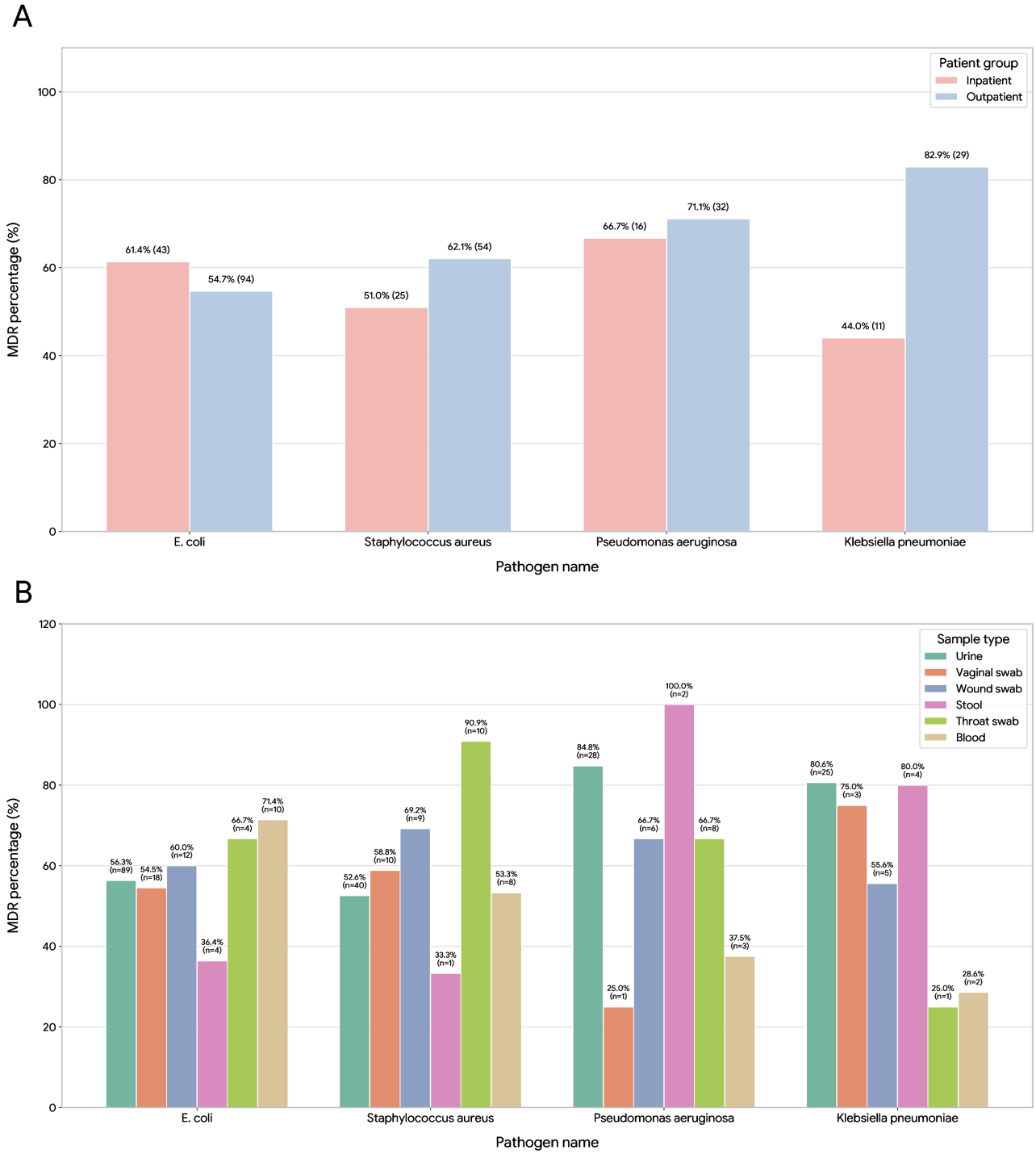


XSL•FO
RenderX

Composite AMR Indicators and MDR

The overall prevalence of MDR, defined as resistance to at least one agent in three or more antimicrobial classes, was 57.3% (485/846, 95% CI 53.9 - 60.6). MDR was significantly more frequent among inpatients than outpatients (62.7%, 195/311, 95% CI 57.2 - 68.1 vs 54.2% 290/535, 95% CI 49.9 - 58.4; *P*=.04), which is consistent with a greater concentration of resistant organisms within hospital settings. However, these comparisons reflect descriptive variation across clinical settings and should not be interpreted as evidence of independent associations because adjustment for important patient- and health care-level confounders was not feasible within the available surveillance dataset. When analyzed at the pathogen level, however, differences in MDR prevalence between inpatient and outpatient isolates were not statistically significant for *E. coli* (*P*=.39), *S. aureus* (*P*=.28), or *P. aeruginosa* (*P*=.79). In contrast, MDR prevalence differed significantly for *K. pneumonia* between clinical settings (*P*=.002, Fisher exact test), suggesting pathogen-specific variation in health care-associated resistance dynamics (Figure 3A).

Figure 3. Multidrug resistance prevalence by (A) clinical setting and (B) specimen type among major bacterial pathogens in Iraq. MDR was defined as resistance to at least one antimicrobial agent in three or more antimicrobial classes. MDR: multidrug resistant;



XSL•FO
RenderX

Among the most common pathogens, MDR prevalence was highest in *P. aeruginosa* (71.6%, 47/65, 95% CI 59.6 - 81.2), followed by *K. pneumonia* (66.1%, 39/59, 95% CI 53.3 - 76.8), *S. aureus* (58.3%, 88/151, 95% CI 50.3 - 65.9), and *E. coli* (56.6%, 133/235, 95% CI 50.2 - 62.8). Carbapenem resistance, defined as resistance to at least one carbapenem (imipenem and/or meropenem) among isolates tested for carbapenems, was observed in 12.4% of *E. coli* (20/161, 95% CI 8.2 - 18.3), 11.1% of *K. pneumoniae* (2/18, 95% CI 3.1 - 32.8), and 29.4% of *P. aeruginosa* (5/17, 95% CI 13.3 - 53.1). Although denominators varied according to testing availability, carbapenem resistance was consistently present across key Gram-negative pathogens. Third-generation cephalosporin resistance, used as a surveillance proxy indicator for probable ESBL activity rather than confirmed ESBL production and defined as resistance to at least one of ceftriaxone, ceftazidime, or cefotaxime among isolates tested for these agents, was detected in 69.3% of *E. coli* (106/153, 95% CI 61.6 - 76.0) and 44.4% of *K. pneumoniae* (8/18, 95% CI 24.6 - 66.3). At the genus level, 70.0% of *Klebsiella* spp. isolates tested (21/30, 95% CI 52.1 - 83.3) demonstrated third-generation cephalosporin resistance. MRSA, defined by oxacillin resistance, accounted for 89.8% of *S. aureus* isolates (136/151, 95% CI 83.9 - 93.6), reflecting a markedly elevated burden of methicillin resistance.

Specimen-Specific Multidrug Resistance (MDR) Variation

Analysis of MDR prevalence by specimen type demonstrated significant variation for selected pathogens. Among *K. pneumonia*, urinary isolates exhibited substantially higher MDR prevalence compared with blood isolates (80.6% vs 28.6%; P=.01) and throat swab isolates (80.6% vs 25.0%; P=.04). Similarly, *P. aeruginosa* showed significantly higher MDR prevalence in urine compared with blood (84.8% vs 37.5%; P=.01) and vaginal swabs (84.8% vs 25.0%; P=.02). For *S. aureus*, throat swab isolates demonstrated significantly higher MDR prevalence than urinary isolates (90.9% vs 52.6%; P=.02). No statistically significant specimen-level differences were observed for *E. coli* (all P>.05) (Figure 3B). These findings indicate that resistance burden varies not only by pathogen and clinical setting but also by anatomical source, with urinary isolates of certain Gram-negative pathogens demonstrating particularly elevated MDR prevalence.

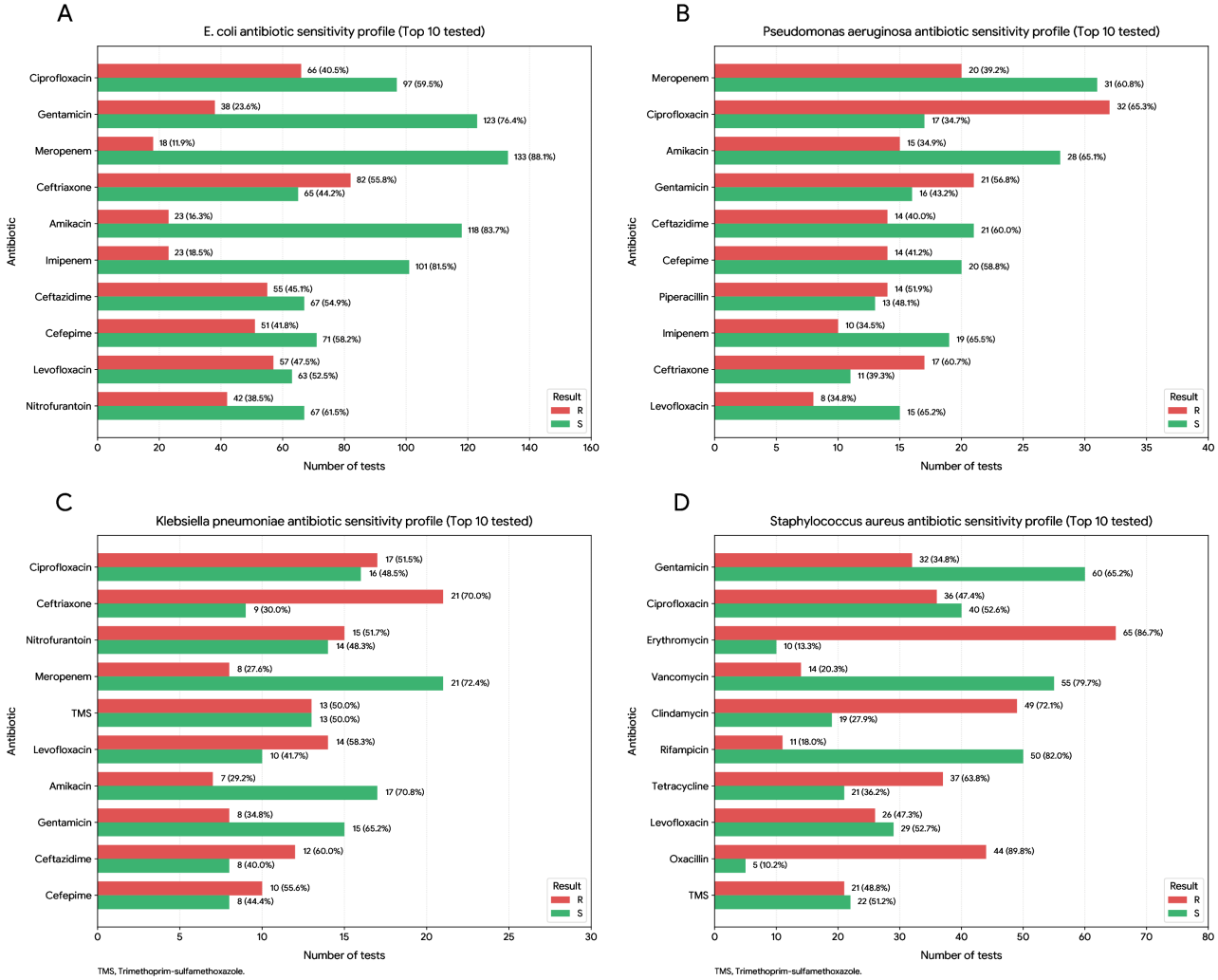
AMR Patterns Among Major Gram-Negative Pathogens

E. Coli

Resistance rates among *E. coli* isolates varied substantially across antibiotic classes (Figure 4A). High resistance was observed to ceftriaxone (82/147, 55.8%), ceftazidime (55/122, 45.1%), and ciprofloxacin (66/163, 40.5%). Resistance to levofloxacin was 47.5% (57/120), and to cefepime 41.8% (51/122). Nitrofurantoin resistance was 38.5% (42/109). Lower resistance rates were observed for carbapenems and aminoglycosides. Resistance to meropenem was 11.9% (18/151), and to imipenem 18.5% (23/124). Amikacin resistance was 16.3% (23/141), and gentamicin resistance 23.6% (38/161).

XSL•FO
RenderX

Figure 4. Antimicrobial susceptibility profiles of major bacterial pathogens isolated in four Iraqi governorates. (A) Escherichia coli; (B) *Pseudomonas aeruginosa*; (C) Klebsiella pneumoniae; (D) Staphylococcus aureus. Bars represent proportions of resistant (R) and susceptible (S) isolates among tested samples for each pathogen–antibiotic combination. Denominators varied according to organism-specific testing panels, antibiotic availability, and laboratory testing practices across participating centers. R: resistant; S: susceptible.



P. Aeruginosa

Among *P. aeruginosa* isolates, resistance to ciprofloxacin was 65.3% (32/49), representing the highest resistance observed in this species (Figure 4B). Gentamicin resistance was 56.8% (21/37), and piperacillin resistance 51.9% (14/27). Ceftriaxone nonsusceptibility was observed in 60.7% of isolates; however, because *P. aeruginosa* is intrinsically nonsusceptible to ceftriaxone, this finding is not clinically informative as a marker of acquired resistance or therapeutic option loss. Carbapenem resistance was notable, with 39.2% (20/51) resistance to meropenem and 34.5% (10/29) to imipenem. Amikacin resistance was 34.9% (15/36). Resistance to ceftazidime and cefepime was 40.0% (14/35) and 41.2% (14/34), respectively (Figure 4B).

K. Pneumoniae

K. pneumoniae demonstrated high resistance to ceftriaxone (21/30, 70.0%) and ceftazidime (12/20, 60.0%). Resistance to ciprofloxacin was 51.5% (17/33), and to levofloxacin 58.3% (14/24). Nitrofurantoin resistance was 51.7% (15/29). Carbapenem resistance was lower compared with third-generation cephalosporins, with 27.6% (8/29) resistance to meropenem. Aminoglycoside resistance was 29.2% (7/24) for amikacin and 34.8% (8/23) for gentamicin (Figure 3C).

AMR Patterns Among S. Aureus

Among *S. aureus* isolates, oxacillin resistance was 89.8% (44/49), consistent with a high prevalence of methicillin resistance. Erythromycin resistance was 86.7% (65/75), and clindamycin resistance 72.1% (49/68). Tetracycline resistance was 63.8% (37/58). Resistance to ciprofloxacin and levofloxacin was 47.4% (36/76) and 47.3% (26/55), respectively. Trimethoprim–sulfamethoxazole resistance was 48.8% (21/43). Gentamicin resistance was 34.8% (32/92) (Figure 4D). Vancomycin non-susceptibility was observed in 20.3% (14/69) of tested *S. aureus* isolates; however, this finding should be interpreted with considerable caution given the global rarity of confirmed vancomycin-resistant *S. aureus* (VRSA). Possible contributing explanations include testing methodology variability, misclassification of intermediate isolates as resistant, instrument- or platform-related limitations, lack of uniform confirmatory MIC testing, or potential data recording inconsistencies

XSL•FO
RenderX

Discussion

Key Findings

This multicenter surveillance analysis addressed the study objective of characterizing AMR patterns across hospital and outpatient laboratory settings in Iraq and evaluating their implications for surveillance strengthening and AMS policy. The findings demonstrated a high overall prevalence of multidrug resistance (485/846, 57.3%), substantial third-generation cephalosporin resistance among Enterobacterales, elevated carbapenem resistance among nonfermenters, and high methicillin resistance among *S. aureus*. The elevated third-generation cephalosporin resistance observed among Enterobacterales is consistent with a substantial burden of ESBL-probable organisms; however, these findings should be interpreted cautiously because phenotypic confirmatory testing and molecular characterization were not performed. Consequently, the results reflect surveillance indicators compatible with probable ESBL activity rather than microbiologically confirmed ESBL prevalence. Beyond the microbiological burden, these results underscore the urgent need for strengthened surveillance coordination, routine antibiogram reporting, and improved integration of laboratory data within broader AMR monitoring systems aligned with WHO GLASS and AWaRe frameworks. MDR prevalence was higher among inpatient isolates (195/311, 62.7%) than outpatient isolates (290/535, 54.2%), although resistance burden remained substantial across both health care settings, with resistant organisms observed across both inpatient and outpatient laboratory populations. However, outpatient isolates should not be interpreted as synonymous with community-acquired infection because prior health care exposure could not be assessed. Although MDR prevalence was higher among inpatients, resistance remained substantial across settings. Interpretation of subgroup differences warrants caution. Although higher MDR prevalence was observed among inpatient isolates and selected specimen categories, the available surveillance dataset did not include several important resistance determinants required for robust adjusted modeling, including previous antibiotic exposure, recent hospitalization, health care contact history, comorbidity burden, invasive procedures, intensive care unit admission, infection severity, or prior microbiological findings. Consequently, observed differences may partly reflect unmeasured variation in patient complexity, clinical indication, health care exposure, or testing practices rather than independent effects of clinical setting or specimen source alone. The present analysis therefore prioritized descriptive surveillance characterization over inferential modeling, consistent with the study's operational surveillance objectives. These findings should be interpreted within Iraq's broader AMR and surveillance landscape. Institute for Health Metrics and Evaluation (IHME) modeling estimated approximately 3450 AMR-attributable deaths in Iraq in 2021, with carbapenem-resistant Gram-negative organisms and MRSA contributing substantially to mortality. While such estimates rely on aggregated datasets, this study provides complementary operational laboratory data with context-specific resistance proportions relevant to stewardship planning and health system response. The observed high MRSA prevalence and substantial carbapenem resistance among *P. aeruginosa* and Enterobacterales are consistent with national mortality patterns. Recent Iraqi surveillance studies similarly report sustained high ESBL prevalence, increasing carbapenem resistance, and substantial resistance to third-generation cephalosporins and fluoroquinolones across multiple governorates and health care settings [8,18]. Regarding ceftriaxone findings in *P. aeruginosa*, it is important to note that *P. aeruginosa* is intrinsically nonsusceptible to most third-generation cephalosporins, including ceftriaxone, due to low outer membrane permeability, chromosomal AmpC β-lactamase production, and efflux mechanisms [26]. Accordingly, the ceftriaxone result is presented for completeness of the routine laboratory dataset but should not be interpreted as a clinically meaningful surveillance marker of acquired resistance or loss of a recommended therapeutic option, since ceftriaxone is not routinely indicated for *P. aeruginosa* infections [27].

Interpretations, Implications, and Comparisons of the Study Findings in the Context of Existing Literature

Comparison With Global and Regional Surveillance

The WHO GLASS 2025 report includes Iraqi data from sentinel laboratories and reports resistance patterns broadly consistent with our findings. However, GLASS outputs are aggregated and adjusted for surveillance coverage, whereas this study provides deduplicated, center-level isolate data with inpatient–outpatient and specimen-level stratification, allowing greater insight into within-country and clinical-context variation [6]. Thus, this multicenter analysis complements rather than duplicates national reporting by enhancing granularity and operational relevance for stewardship and guideline refinement. Gram-negative pathogens dominated the resistance landscape, consistent with global and regional evidence identifying Enterobacterales and nonfermenters as priority AMR threats, particularly in Middle Eastern, postconflict, and resource-constrained settings [1,5,6]. Fragmented laboratory systems and inconsistent antimicrobial regulation, recognized drivers of AMR expansion in Iraq and neighboring countries, may contribute to these patterns [5]. The high prevalence of third-generation cephalosporin resistance in *E. coli* (106/153, 69.3%) and *K. pneumoniae* (8/18, 44.4% - 21/30, 70.0%) is particularly concerning and suggests a substantial burden of ESBL-probable Enterobacterales. However, because phenotypic ESBL confirmation and molecular testing were not undertaken, these observations should be interpreted as surveillance indicators of probable resistance mechanisms rather than confirmed ESBL prevalence estimates. GLASS 2025 reports persistent and, in some regions, rising resistance to third-generation cephalosporins among bloodstream and urinary isolates of *E. coli* and *K. pneumoniae*, especially in LMIC settings [6]. Our findings place Iraq toward the upper range of these global estimates and, importantly, demonstrate statistically significant inpatient–outpatient differences for *K. pneumoniae* and specimen-level heterogeneity for selected pathogens. Such localized variation has direct implications for empiric therapy optimization and AMS prioritization but cannot be discerned from national-level GLASS outputs alone. Rather than

XSL•FO
RenderX

duplicating GLASS, this multicenter analysis complements national surveillance by providing operationally relevant, context-specific insights into resistance distribution across health care settings within Iraq.

Moreover, carbapenem resistance, although lower than cephalosporin resistance, was consistently present, particularly in *P. aeruginosa* (5/17, 29.4%). Carbapenem-resistant nonfermenters are recognized by WHO as critical priority pathogens [28], and even moderate resistance levels can restrict empiric therapy and increase mortality risk in severe infection [1,29]. The very high MRSA prevalence (136/151, 89.8%) exceeds many international benchmarks; rates above 50% are commonly considered markers of entrenched health care-associated transmission and IPC challenges [6]. Despite the overall resistance burden, some susceptibility patterns suggest limited stewardship opportunities. Among Gram-negative pathogens, amikacin showed lower resistance than third-generation cephalosporins and fluoroquinolones, indicating potential utility within empiric combination regimens for suspected MDR Gram-negative infections, pending susceptibility results. Among *S. aureus*, comparatively lower resistance to trimethoprim–sulfamethoxazole and rifampicin suggests that selected oral agents may retain utility in some culture-guided community or step-down treatment settings. Nevertheless, the high MRSA burden substantially limits empiric reliance on beta-lactams for suspected staphylococcal infections. Overall, these findings emphasize the importance of timely microbiological diagnostics to optimize therapy and minimize unnecessary escalation to watch or reserve antibiotics within the WHO AWaRe framework.

In addition, the observed vancomycin nonsusceptibility rate among *S. aureus* isolates warrants careful interpretation within a surveillance quality assurance context. While confirmed VRSA remains rare globally [6,30], a reported nonsusceptibility proportion of this magnitude would be expected to attract substantial microbiological scrutiny. Several methodological explanations should therefore be considered, including testing methodology variability, breakpoint interpretation differences, misclassification of intermediate isolates as resistant, instrument-related limitations, heterogeneous laboratory workflows, lack of universal confirmatory MIC testing, and possible data entry or transcription inconsistencies. Elevated phenotypic resistance rates have nevertheless been reported in several LMIC settings [27], often reflecting laboratory methodology limitations, heterogeneous vancomycin-intermediate *Staphylococcus aureus* (VISA) phenotypes, or high glycopeptide selection pressure rather than fully confirmed vancomycin-resistant strains. If true VRSA were present at this magnitude, the implications would be substantial, including restriction of last-line glycopeptide therapy, increased reliance on linezolid or daptomycin, and urgent reinforcement of infection prevention and control measures [31]. Consequently, these findings should not be interpreted as definitive evidence of confirmed VRSA prevalence but rather as a signal warranting confirmatory MIC-based testing, strengthened microbiological quality assurance, and further prospective surveillance investigation. Similar variability in phenotypic resistance reporting has been observed in recent Iraqi surveillance reports [32].

Implications for Iraqi National Antimicrobial Guidelines (INAG)

The Iraqi National Antimicrobial Guidelines (INAG) emphasize rational prescribing, culture-guided therapy, short treatment duration, and avoidance of unnecessary broad-spectrum antibiotics [33]. However, effective guideline implementation depends on timely, representative resistance data. The resistance patterns identified in this surveillance analysis suggest that periodic revision of INAG should be explicitly linked to structured national antibiogram reporting cycles. Embedding surveillance–guideline feedback mechanisms within policy frameworks would enhance responsiveness to evolving resistance patterns. The guidelines explicitly recommend using narrow-spectrum agents whenever effective and caution against broad-spectrum escalation unless clinically justified [33]. There are similar sentiments in the WHO AWaRe guidance following concerns with the robustness of national antibiotic guidelines from many LMICs [34,35]. However, the resistance patterns observed in this study have several potentially actionable implications for empiric prescribing policy and periodic INAG revision. High third-generation cephalosporin resistance among Enterobacterales may reduce the reliability of ceftriaxone-based empiric regimens, particularly for severe urinary tract, bloodstream, and intraabdominal infections where *E. coli* and *K. pneumoniae* are common pathogens. Elevated fluoroquinolone resistance may similarly limit commonly used oral step-down and outpatient treatment strategies. Conversely, comparatively lower resistance to amikacin across major Gram-negative pathogens suggests a possible role within carefully supervised empiric combination approaches for suspected MDR Gram-negative infections, pending culture and susceptibility results. The substantial MRSA burden further supports review of empiric antistaphylococcal coverage strategies in selected hospital-associated, surgical, and skin or soft tissue infection contexts. The INAG emphasizes that empiric therapy should be informed by local antibiogram data. These findings support a more formal surveillance-to-guideline feedback mechanism whereby routine regional antibiograms directly inform scheduled INAG revision cycles, syndrome-specific empiric recommendations, AWaRe stewardship priorities, and structured 48 - 72-hour antimicrobial review and de-escalation processes. Our findings underscore the urgency of updating national and institutional antibiograms regularly and integrating resistance surveillance into guideline revision cycles, which is currently not happening despite growing AMR rates. Without periodic updates reflecting current resistance patterns, empiric recommendations risk being misaligned with microbiological reality. Importantly, INAG also advises avoiding unnecessary macrolides, tetracyclines, quinolones, and aminoglycosides in pregnancy and pediatric populations. Rising resistance to these agents may further complicate safe empiric therapy selection in vulnerable groups.

Alignment With WHO AWaRe Framework

The WHO AWaRe classification categorizes antibiotics into access, watch, and reserve groups to optimize stewardship and

XSL•FO
RenderX

preserve last-line agents, which typically include watch and reserve antibiotics [36,37]. The WHO recommended that at least 60% of national antibiotic consumption should derive from the access group to help limit the use of watch and reserve antibiotics, particularly in LMICs, where the use of broader-spectrum agents may contribute to increased resistance selection pressure [37]. The United Nations General Assembly (UN-GA) recently increased this to 70% access antibiotics across sectors to reduce AMR [38,39]. In this study, resistance levels were particularly high for several watch-group antibiotics, including fluoroquinolones and third-generation cephalosporins, patterns consistent with selective pressure reported across LMICs [29,39,40]. Carbapenems remain critical for treating ESBL-producing Enterobacterales, and the presence of carbapenem resistance signals potential erosion of these last-line therapies. WHO AWaRe guidance emphasizes restricting watch and reserve antibiotics to appropriate indications and reinforcing microbiological testing before escalation [41]. Our findings therefore reinforce three key AWaRe-aligned priorities for Iraq, including strengthening Access antibiotic reliability through surveillance, reducing unnecessary watch antibiotic use, expanding AWaRe-based usage monitoring to inform AMS activities [42], and preserving carbapenems through formal stewardship and preauthorization systems, particularly given concerns regarding hospital antibiotic prescribing in Iraq [43].

The higher MDR prevalence among inpatients compared with outpatients is consistent with the recognized association between health care exposure to antibiotics and resistant infection risk reported in previous literature [29]. However, the descriptive design of this study does not permit causal attribution or assessment of independent risk factors. Hospital environments concentrate antibiotic use, invasive procedures, and vulnerable patients, all of which amplify resistance selection. However, the presence of substantial resistance among outpatient isolates demonstrates that resistant organisms were identified beyond inpatient settings, with an increasing focus on community antibiotic use in recent years given their appreciable prescribing in this sector versus hospital sectors [44,45]; however, outpatient isolates may include patients with previous health care contact, referral pathways, or other unmeasured health care exposures. Concerns with community dispensing of antibiotics in Iraq are not helped by high rates of their dispensing without a prescription [7,45-47], which may contribute to antimicrobial selection pressure. In addition, concerns with current nurse prescribing in ambulatory care clinics [46]. Both situations are not helped by concerns with current patient knowledge regarding antibiotics, AMR, and AMS in Iraq [45]. Consequently, addressing AMR in Iraq therefore requires a dual focus: strengthening hospital stewardship activities while also improving prescribing and dispensing practices in ambulatory care.

The observed specimen-level heterogeneity further refines the epidemiological interpretation. Urinary isolates of *K. pneumoniae* and *P. aeruginosa* demonstrated significantly higher MDR prevalence than bloodstream or mucosal isolates. This may be consistent with several possible explanations, including differential antimicrobial exposure, infection complexity, or clinical testing patterns associated with urinary tract infections; however, causal mechanisms cannot be determined from the available surveillance data. In contrast, *E. coli* showed no statistically significant variation in MDR across specimen types, suggesting broader dissemination of resistance traits across clinical syndromes. Combined with this, there is a need for rapid diagnostic testing expansion, risk-stratified empiric algorithms, and de-escalation protocols once susceptibility results are available. INAG explicitly states that antibiotic therapy should be reviewed continuously for adjustment or discontinuation [33]. Embedding structured 48‑72-hour review checkpoints into routine clinic practice, especially in hospitals, may mitigate unnecessary prolonged broad-spectrum exposure. The high MRSA prevalence also suggests a need to review empiric coverage for suspected staphylococcal infections, particularly in hospitalized patients. However, expanding glycopeptide or linezolid use without stewardship oversight would be counterproductive.

Policy and Surveillance Implications

The modest number of isolates relative to population catchment size highlights structural limitations in routine AMR surveillance in Iraq. Because testing was based on routine clinician-directed diagnostic practice rather than systematic sampling, the dataset may be enriched for clinically complex, severe, treatment-refractory, or health care-exposed cases. This potential selection bias should be considered when interpreting resistance frequencies and limits direct extrapolation to population-level prevalence or nationally representative pathogen distributions. Current surveillance remains largely clinician-driven and laboratory-based, with limited denominator integration or standardized sampling. Strengthening surveillance will require integrated laboratory information systems, standardized reporting frameworks, improved diagnostic capacity, and denominator-informed metrics linked to health care usage. More systematic diagnostic sampling across common infection syndromes could generate surveillance data that are less biased toward resistant phenotypes and more representative of routine clinical microbiology, thereby improving empiric therapy guidance and preservation of effective Access antibiotics.

The Iraqi National Action Plan on AMR (2018‑2022) emphasized strengthening surveillance and rational prescribing [9]. While this study contributes multicenter evidence, it also highlights persistent resistance challenges and the need for unified laboratory reporting aligned with WHO GLASS standards [21]. Although GLASS provides essential surveillance infrastructure, sentinel reporting may underrepresent subnational variability [6]. Regular antibiogram reporting, interoperable data systems, and coordinated AWaRe-focused education and quality indicators could support more evidence-based prescribing and progress toward UN General Assembly AMR targets [48].

Limitations

This multicenter analysis incorporated deduplicated microbiological data from four Iraqi governorates using internationally recognized GLASS-aligned methodological principles and standardized multidrug resistance definitions, improving comparability with regional and global surveillance reports [21,24]. However, several limitations warrant careful

XSL•FO
RenderX

consideration. First, the study included selected governorates representing approximately 10% of the Iraqi population and should not be interpreted as nationally representative. Data were generated through routine clinician-directed microbiological testing without systematic sampling frameworks or denominator information. Consequently, the dataset may be biased toward patients with more severe illness, treatment failure, health care exposure, or increased clinical suspicion of resistant infection. These factors may influence observed resistance proportions and constrain generalizability beyond participating laboratory settings. The study did not have access to denominator data on total hospital admissions, outpatient consultations, infection episodes, or total specimens processed; therefore, resistance proportions cannot be interpreted as population incidence or prevalence estimates. Although the study included geographically distinct governorates and major referral laboratories, the sampling strategy was not designed to produce nationally representative AMR estimates. Findings therefore reflect resistance patterns within selected public-sector referral settings rather than the entire Iraqi health care system. Resistance estimates may differ in non-participating governorates, private-sector laboratories, smaller health care facilities, or rural settings with different diagnostic access and prescribing practices. The study categorized isolates by inpatient versus outpatient laboratory setting; however, outpatient status should not be interpreted as equivalent to community-acquired infection. The dataset did not include information on prior hospitalization, health care contact, referral status, invasive procedures, or recent antibiotic exposure. Consequently, comparisons between inpatient and outpatient isolates reflect clinical care setting rather than definitive health care-associated versus community-acquired infection epidemiology. Second, data were derived from the principal public tertiary hospital laboratory in each city. While these centers function as major referral facilities, they do not capture all microbiological testing conducted in smaller hospitals, private laboratories, or rural health facilities. Consequently, the findings may not fully represent resistance patterns across the entire health care system. Third, culture and susceptibility testing was ordered at clinician discretion. Although testing is provided free of charge in the included public hospitals, it is plausible that testing was more frequently performed in patients with severe disease, hospitalization, treatment failure, or diagnostic uncertainty. Such selective testing may enrich the dataset for more resistant organisms and potentially overestimate resistance prevalence relative to all infections occurring in the community. This limitation is inherent to many laboratory-based AMR surveillance systems internationally [21,29]. Laboratory testing panels were not completely uniform across centers, resulting in variable denominators for some antibiotic classes. Phenotypic confirmatory testing and molecular characterization of ESBL production were not performed. Accordingly, third-generation cephalosporin resistance findings should be interpreted as proxy indicators of probable ESBL activity rather than definitive measures of confirmed ESBL prevalence. As a cross-sectional laboratory-based analysis, causal inferences regarding drivers of resistance cannot be drawn. Although subgroup comparisons were undertaken by clinical setting, specimen type, and governorate, the dataset lacked several critical patient- and system-level variables necessary for robust multivariable adjustment, such as prior antimicrobial exposure and recent hospitalization. Consequently, observed differences should be interpreted as descriptive surveillance findings rather than adjusted estimates of independent resistance determinants. Future studies using prospectively collected clinical datasets or linked health system data would be valuable to enable more robust multivariable analyses of resistance risk factors. Furthermore, the analytical approach was intentionally descriptive and surveillance-oriented. Although subgroup comparisons by clinical setting and specimen type were undertaken to characterize variation in resistance patterns, the study did not use surveillance-oriented or adjustment for multiple comparisons because these analyses were outside the primary objectives of this multicenter surveillance synthesis. Accordingly, subgroup findings should be interpreted as descriptive rather than adjusted inferential estimates.

Conclusions

This multicenter surveillance analysis highlights substantial multidrug resistance and clinically important resistance among priority bacterial pathogens across four Iraqi governorates. Beyond the observed microbiological burden, the findings illustrate important structural gaps in routine AMR surveillance, including limited denominator integration, heterogeneous laboratory reporting, and dependence on clinician-directed testing practices. Strengthening integrated laboratory information systems, standardized antibiogram reporting, and linkage between surveillance outputs and antimicrobial guideline revision processes will be essential for improving stewardship effectiveness and informing evidence-based AMR policy in Iraq. More broadly, the study demonstrates the feasibility and value of multicenter laboratory data aggregation in fragile and resource-constrained settings, providing potentially transferable lessons for AMR surveillance strengthening in other LMIC and postconflict health systems. While findings are not intended to provide population-level national estimates, they offer operationally relevant surveillance evidence to inform laboratory strengthening, stewardship planning, and future development of integrated AMR monitoring systems in Iraq.

Acknowledgments

During manuscript preparation, generative artificial intelligence (AI) tools were used to assist with language editing, structural refinement, and formatting support. All scientific content, interpretation of findings, methodological decisions, and conclusions were critically reviewed, verified, and approved by the authors. No generative AI tools were used for data analysis, statistical computation, or autonomous interpretation of study findings. Authors remain fully responsible for the accuracy, integrity, and originality of the manuscript content.

XSL•FO
RenderX

Funding

The authors declared no financial support was received for this work.

Data Availability

The dataset analyzed in this study consists of de-identified routine clinical microbiology laboratory records collected from participating tertiary public hospitals in four Iraqi governorates. Due to institutional data governance policies and ethical restrictions related to patient-level health data, the full raw dataset cannot be made publicly available. Aggregated data supporting the findings of this study are presented within the manuscript and its supplementary materials. De-identified analytical datasets may be made available from the corresponding author upon reasonable request, subject to approval from the participating institutions and compliance with applicable data protection and ethical regulations.

Authors' Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work. AK and OD are joint co-corresponding authors.

Conflicts of Interest

None declared.

Checklist 1
STROBE checklist.
[PDF File, 121 KB - publichealth_v12i1e94436_app1.pdf]

References

1. Murray C, Ikuta KS, Sharara F. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. Lancet 2022 Feb 12;399(10325):629-655. [doi: 10.1016/S0140-6736(21)02724-0] [Medline: 35065702]
2. Patra M, Gupta AK, Kumar D, Kumar B. Antimicrobial resistance: a rising global threat to public health. Infect Drug Resist 2025 Oct;18:5419-5437. [doi: 10.2147/IDR.S530557] [Medline: 41158783]
3. Ho CS, Wong CTH, Aung TT, et al. Antimicrobial resistance: a concise update. Lancet Microbe 2025 Jan;6(1):100947. [doi: 10.1016/j.lanmic.2024.07.010] [Medline: 39305919]
4. The burden of antimicrobial resistance (AMR) in Iraq. Institute for Health Metrics and Evaluation (IHME).: University of Washington, Global Research on AntiMicrobial resistance (GRAM); 2023. URL: https://www.healthdata.org/sites/default/files/2023-09/Iraq.pdf [accessed 2026-07-27]
5. Mathu R, Diago-Navarro E, Lynch E, et al. Antibiotic resistance in the Middle East and Southern Asia: a systematic review and meta-analysis. JAC Antimicrob Resist 2025 Feb;7(1):dlaf010. [doi: 10.1093/jacamr/dlaf010] [Medline: 39973906]
6. World Health Organization. Global antibiotic resistance surveillance report 2025 report 2025. : World Health Organization; 2025 URL: https://www.who.int/publications/i/item/9789240116337 [accessed 2026-07-31]
7. Al-Jumaili AA, Ahmed KK. A review of antibiotic misuse and bacterial resistance in Iraq. East Mediterr Health J 2024 Nov 3;30(10):663-670. [doi: 10.26719/2024.30.10.663] [Medline: 39574365]
8. Al-Fahad DK, Alpofead JA, Chawsheen MA, Al-Naqshbandi AA, Abas AT. Surveillance of antimicrobial resistance in Iraq: a comprehensive data collection approach. Aro 2024;12(2):179-193. [doi: 10.14500/aro.11689]
9. National action plan on antimicrobial resistance in Iraq (2018–2022). Ministry of Health/ Environment and Ministry of Agriculture. 2018. URL: https://cdn.who.int/media/docs/default-source/antimicrobial-resistance/amr-spc-npm/nap-library/national-action-plan-iraq.pdf?sfvrsn=8a7c49b1_1&download=true [accessed 2026-07-27]
10. Yassen AO, Bilal AD, Ismael BS, Al-Ani AHH. Perceptions and practices of community pharmacists towards antimicrobial stewardship program in Erbil City, Iraq: a cross-sectional study: community pharmacists towards antimicrobial stewardship program. Zanco J Med Sci 2025 Oct 30;Special(October 2025):423-434. [doi: 10.15218/zjms.2025.045]
11. Charani E, Mendelson M, Pallett SJC, et al. An analysis of existing national action plans for antimicrobial resistance-gaps and opportunities in strategies optimising antibiotic use in human populations. Lancet Glob Health 2023 Mar;11(3):e466-e474. [doi: 10.1016/S2214-109X(23)00019-0] [Medline: 36739875]
12. Mohammed S, Ahmed M, Karem K. Incidence of multi-drug resistant escherichia coli isolates from blood and urine in Kerbala, Iraq. J Kerbala Univ 2014;10(2):222-227 [FREE Full text]
13. Antimicrobial resistance TrACSS Iraq 2022 country profile. World Health Organization. 2022. URL: https://www.who.int/publications/m/item/Antimicrobial-resistance-tracss-irq-2022-country-profile [accessed 2026-07-27]
14. Abdullah IM. Multiple drugs resistance among urinary tract infection patients in Duhok City –Kurdistan region –Iraq. Duhok Med J 2019 May 31;13(1):22-31 [FREE Full text] [doi: 10.31386/dmj.2019.13.1.3]

XSL•FO
RenderX

15. Al-Derzi N. Pattern of resistance to pseudomonas infection in the north of Iraq: emphasis on the potential role of a combination antibiogram. Iraqi J Commun Med 2012;11:193-198 [FREE Full text]
16. M. Al-Nema Z. Trend of antibiotics usage in the intensive care unit in the medical city in Baghdad. Iraqi J Pharm Sci 2016;25(1):50-58. [doi: 10.31351/vol25iss1pp50-58]
17. Abou Fayad A, Rizk A, El Sayed S, et al. Antimicrobial resistance and the Iraq wars: armed conflict as an underinvestigated pathway with growing significance. BMJ Glob Health 2023 Jan;7(Suppl 8):e010863. [doi: 10.1136/bmjgh-2022-010863] [Medline: 36781284]
18. Hashim HT, Hashim AT, Ali HT, et al. Prevalence and pattern of antibiotic use and resistance among Iraqi patients: a cross-sectional study. Afr Health Sci 2024 Sep;24(3):47-57. [doi: 10.4314/ahs.v24i3.7] [Medline: 40777977]
19. Salim KS, Alsabah AS, Taghi HS, et al. Misuse of antibiotics in Iraq: a review of Iraqi published studies. Al-Mustansiriyah J Pharm Sci 2021;21(2):15-20. [doi: 10.32947/ajps.v21i2.805]
20. Elm EV, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ 2007 Oct 20;335(7624):806-808. [doi: 10.1136/bmj.39335.541782.AD]
21. GLASS manual for early implementation. World Health Organization. 2017. URL: https://www.who.int/publications/i/item/9789241549400 [accessed 2026-07-27]
22. Clark AE, Kaleta EJ, Arora A, Wolk DM. Matrix-assisted laser desorption ionization-time of flight mass spectrometry: a fundamental shift in the routine practice of clinical microbiology. Clin Microbiol Rev 2013 Jul;26(3):547-603. [doi: 10.1128/CMR.00072-12] [Medline: 23824373]
23. CLSI M100 performance standards for antimicrobial susceptibility testing. CLSI. URL: https://clsi.org/shop/standards/m100/ [accessed 2026-07-31]
24. Magiorakos AP, Srinivasan A, Carey RB, et al. Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. Clin Microbiol Infect 2012 Mar;18(3):268-281. [doi: 10.1111/j.1469-0691.2011.03570.x] [Medline: 21793988]
25. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
26. Pang Z, Raudonis R, Glick BR, Lin TJ, Cheng Z. Antibiotic resistance in pseudomonas aeruginosa: mechanisms and alternative therapeutic strategies. Biotechnol Adv 2019;37(1):177-192. [doi: 10.1016/j.biotechadv.2018.11.013] [Medline: 30500353]
27. Gelaw LY, Bitew AA, Gashey EM, Ademe MN. Ceftriaxone resistance among patients at GAMBY teaching general hospital. Sci Rep 2022 Jul 14;12(1):12000. [doi: 10.1038/s41598-022-16132-3] [Medline: 35835837]
28. Global action plan on antimicrobial resistance. World Health Organization. 2015. URL: https://www.who.int/publications/i/item/9789241509763 [accessed 2026-07-27]
29. Holmes AH, Moore LSP, Sundsfjord A, et al. Understanding the mechanisms and drivers of antimicrobial resistance. Lancet 2016 Jan 9;387(10014):176-187. [doi: 10.1016/S0140-6736(15)00473-0] [Medline: 26603922]
30. McGuinness WA, Malachowa N, DeLeo FR. Vancomycin resistance in Staphylococcus aureus. Yale J Biol Med 2017;90(2):269. [Medline: 28656013]
31. Chang S, Sievert DM, Hageman JC, et al. Infection with vancomycin-resistant Staphylococcus aureus containing the vanA resistance gene. N Engl J Med 2003 Apr 3;348(14):1342-1347. [doi: 10.1056/NEJMoa025025] [Medline: 12672861]
32. Alkhulaifi ZM, Mohammed KA, Alkhulaifi ZM, Mohammed KA. The prevalence of cephalosporins resistance in pseudomonas aeruginosa isolated from clinical specimens in Basra, Iraq. Univ Thi-Qar J Sci 2023;10(1(SI):150-153. [doi: 10.32792/utq/utjsci/v10i1(SI).1010]
33. Al-Jawadi H, Alwiswasi N, Hasan A. The Iraqi National Antimicrobial Guidelines (INAG). Ministry of Health and Environment. 2020. URL: https://www.iraqipharm.com/wp-content/uploads/2020/06/INAG-After-Protect.pdf [accessed 2026-07-27]
34. Jamil E, Saleem Z, Godman B, et al. Global variation in antibiotic prescribing guidelines and the implications for decreasing AMR in the future. Front Pharmacol 2025;16:1600787. [doi: 10.3389/fphar.2025.1600787] [Medline: 40852607]
35. Sharland M, Zanichelli V, Ombajo LA, et al. The WHO essential medicines list AWaRe book: from a list to a quality improvement system. Clin Microbiol Infect 2022 Dec;28(12):1533-1535. [doi: 10.1016/j.cmi.2022.08.009] [Medline: 36007869]
36. Zanichelli V, Sharland M, Cappello B, et al. The WHO AWaRe (Access, Watch, Reserve) antibiotic book and prevention of antimicrobial resistance. Bull World Health Org 2023 Apr 1;101(4):290-296. [doi: 10.2471/BLT.22.288614]
37. Sharland M, Gandra S, Huttner B, et al. Encouraging AWaRe-ness and discouraging inappropriate antibiotic use-the new 2019 Essential Medicines List becomes a global antibiotic stewardship tool. Lancet Infect Dis 2019 Dec;19(12):1278-1280. [doi: 10.1016/S1473-3099(19)30532-8] [Medline: 31782385]
38. Political declaration of the high-level meeting on antimicrobial resistance. United Nations. 2024. URL: https://www.un.org/pga/wp-content/uploads/sites/108/2024/09/FINAL-Text-AMR-to-PGA.pdf [accessed 2026-07-27]

XSL•FO
RenderX

39. Sulis G, Sayood S, Katukoori S, et al. Exposure to World Health Organization's AWaRe antibiotics and isolation of multidrug resistant bacteria: a systematic review and meta-analysis. Clin Microbiol Infect 2022 Sep;28(9):1193-1202. [doi: 10.1016/j.cmi.2022.03.014] [Medline: 35339675]
40. Klein EY, Milkowska-Shibata M, Tseng KK, et al. Assessment of WHO antibiotic consumption and access targets in 76 countries, 2000-15: an analysis of pharmaceutical sales data. Lancet Infect Dis 2021 Jan;21(1):107-115. [doi: 10.1016/S1473-3099(20)30332-7] [Medline: 32717205]
41. Sharland M, Cappello B, Ombajo LA, et al. The WHO AWaRe Antibiotic Book: providing guidance on optimal use and informing policy. Lancet Infect Dis 2022 Nov;22(11):1528-1530. [doi: 10.1016/S1473-3099(22)00683-1] [Medline: 36309019]
42. Saleem Z, Sheikh S, Godman B, et al. Increasing the use of the WHO AWaRe system in antibiotic surveillance and stewardship programmes in low- and middle-income countries. JAC Antimicrob Resist 2025 Apr;7(2):dlaf031. [doi: 10.1093/jacamr/dlaf031] [Medline: 40110554]
43. Kurdi A, Hasan AJ, Baker KI, et al. A multicentre point prevalence survey of hospital antibiotic prescribing and quality indices in the Kurdistan regional government of Northern Iraq: the need for urgent action. Expert Rev Anti Infect Ther 2021 Jun;19(6):805-814. [doi: 10.1080/14787210.2021.1834852] [Medline: 33034234]
44. Saleem Z, Moore CE, Kalungia AC, et al. Status and implications of the knowledge, attitudes and practices towards AWaRe antibiotic use, resistance and stewardship among low- and middle-income countries. JAC Antimicrob Resist 2025 Apr;7(2):dlaf033. [doi: 10.1093/jacamr/dlaf033] [Medline: 40134815]
45. Darweesh O, Kurdi A, Merkhan M, et al. Knowledge, attitudes, and practices of Iraqi parents regarding antibiotic use in children and the implications. Antibiotics (Basel) 2025 Apr 3;14(4):376. [doi: 10.3390/antibiotics14040376] [Medline: 40298563]
46. Kurdi A, Hasan AJ, Hassan Ii AJ, et al. A simulated client study on non-prescription antibiotic sales and counselling practices in Iraqi nurse-led clinics: implications for antimicrobial resistance. Infect Drug Resist 2025 Oct;18:5599-5616. [doi: 10.2147/IDR.S533590] [Medline: 41189960]
47. Alagha HZ. Dispensing antibiotics without prescription in the Arab world: a narrative review of the prevalence, appropriateness, facilitators and preventive interventions. Bull Pharm Sci Assiut Univ 2023 Jun 1;46(1):399-420. [doi: 10.21608/bfsa.2023.301131]
48. Cook A, Kalungia AC, Ubaka CM, et al. Developing AWaRe-ness in primary care across low- and middle-income countries - a vital challenge for antibiotic stewardship programs. Expert Rev Anti Infect Ther 2026 Jan;24(1):1-5. [doi: 10.1080/14787210.2026.2628047] [Medline: 41652914]

Abbreviations

AMR: antimicrobial resistance
AMS: antimicrobial stewardship
AST: antimicrobial susceptibility testing
AWaRe: Access Watch Reserve
ESBL: extended-spectrum beta-lactamase
GLASS: Global antimicrobial resistance and use surveillance system
LMIC: low-and middle-income countries
MDR: multidrug-resistant
MRSA: methicillin-resistant Staphylococcus aureus
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 01.Mar.2026; peer-reviewed by D Iryawati, G Mboowa, H Mwanja; revised version received 03.Jul.2026; accepted 19.Jul.2026; published 20.Aug.2026.

Please cite as:
Omer H, Darweesh O, Karim G, Rashid B, Ahmed R, Seaton A, Godman B, Kurdi A
Antimicrobial Resistance Across Four Iraqi Governorates: A Multicenter Cross-Sectional Surveillance Analysis With Implications for Antimicrobial Stewardship Policy and WHO AWaRe–Aligned Public Health Strategies
JMIR Public Health Surveill 2026;12:e94436
URL: https://publichealth.jmir.org/2026/1/e94436
doi:10.2196/94436

© Hemn Omer, Omeed Darweesh, Gulbahar Karim, Baref Rashid, Rozhgar Ahmed, Andrew Seaton, Brian Godman, Amanj Kurdi. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 20.Aug.2026. This is an

XSL•FO
RenderX

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

XSL•FO
RenderX

Adolescent Self-Reported Recovery for Substance Use in Illinois: Statewide Representative Epidemiological Study

Douglas C Smith, PhD; Crystal A Reinhart, PhD; Shahana Begum, MS; Tarjani Bhatt, MPH

Center for Prevention Research and Development (CPRD), School of Social Work, University of Illinois Urbana-Champaign, 100 Trade Centre Drive, Suite 300, Champaign, IL, United States

Corresponding Author:
Douglas C Smith, PhD
Center for Prevention Research and Development (CPRD), School of Social Work, University of Illinois Urbana-Champaign, 100 Trade Centre Drive, Suite 300, Champaign, IL, United States

Abstract

Background: Although recovery is a central tenet of the US substance use disorder service delivery system, empirical research on youth recovery remains limited and underdeveloped. Notably, no population-based representative surveys, either in the United States or internationally, currently assess recovery status among secondary school–aged youth (aged 14-18 years). Consequently, little is known about how many youth identify as being in recovery or about their characteristics and needs.

Objective: This study presents the first statewide representative estimate of adolescent self-reported recovery (ASR), derived from a large Midwestern state in the United States.

Methods: We used data from the 2024 Illinois Youth Survey, a weighted, statewide representative survey of students from 8th, 10th, and 12th grades across Illinois. We examined the prevalence of ASR with a widely used single-item question, “Do you consider yourself to be in recovery?” The question was presented after an instruction directing students to consider only substance use when responding. We estimated the prevalence of ASR and conducted descriptive analyses to characterize this group.

Results: Among the 6871 participating students from the 10th and 12th grades, the prevalence of ASR was 3.3% (95% CI 2.6%-4.1%). Among participants with ASR, 51.1% (118/231) were female, 39% (90/231) identified as Latino or Latina, 38.1% (88/231) identified as White, and 13% (30/231) identified as Black or African American. The average age of participants with ASR was 16.5 (SD 1.14) years. Participants with ASR were demographically diverse, and a little over half received free or reduced-price lunch.

Conclusions: Findings suggest that financial recovery capital may be particularly important for participants with ASR. This study provides the first population-based estimate of the prevalence of ASR and underscores the importance of including recovery status in large-scale surveys to inform and strengthen recovery support systems.

(JMIR Public Health Surveill 2026;12:e82792) doi:10.2196/82792

KEYWORDS

recovery; adolescents; substance use disorder; epidemiology; recovery capital

Introduction

Recovery is undoubtedly an organizing principle of the US substance use disorder (SUD) service delivery system. For instance, the US Surgeon General’s report has a dedicated chapter on recovery [1], and federal and state agencies have the word “recovery” appearing in their organization names (eg, the Substance Abuse and Mental Health Services Administration houses an Office of Recovery). Additionally, nationally representative epidemiological studies have measured substance use recovery among adults. On the basis of a study that used data from the National Survey of Drug Use and Health, about 8% of US adults consider themselves to be in substance use recovery [2].

Focusing on recovery as an organizing concept for care is part of a broader effort to frame SUD—a clinically diagnosable condition with substance use—as a treatable health condition. Additionally, the recovery movement seeks to destigmatize the millions of Americans who have recovered from SUD. Finally, there has been increasing attention on what supports are required to both initiate and maintain recovery [1].

Major advances have been made toward these goals. Historically, public stigma toward addiction and misperceptions about treatment effectiveness may have clouded health professionals’ willingness to treat patients with addiction. However, since the early 2000s, researchers have shown that treatment outcomes for people with SUD are similar to, and sometimes better than, those for other chronic health care conditions [3]. Major initiatives such as the implementation of the screening, brief intervention, and referral to treatment model

XSL•FO
RenderX

in primary care have also moved SUD treatment into mainstream health care and improved SUD case identification [4]. However, the opioid crisis in the United States revealed a fissure in health care provision for SUD, with very low rates of efficacious medications being prescribed for persons with opioid use disorders [5].

Regarding stigma reduction, researchers have focused on identifying stigmatizing terms [6,7] that can be avoided by professionals. One clear term to avoid is "substance abuse," which has been associated with professionals viewing persons with SUD as in need of punitive responses [7] and as dangerous [6]. Because of findings like these, some state agencies have removed "substance abuse" from their agency names. Another focus of destigmatizing recovery involves identifying potential workplace discrimination. For example, one national survey found that people remain leery of hiring individuals even when they are in recovery [8], signaling that additional work must be done to frame SUD as a treatable health condition.

Finally, there are many efforts to improve the infrastructure to maintain recovery once it is achieved. Two examples include the development of recovery community organizations and collegiate recovery programs, which both provide support for recovery maintenance [9,10]. While the former are community centers run by people in recovery through many programs (ie, antistigma events, support groups, and training of peer recovery specialists), the latter are specialized programs in universities designed to provide social support to college students (ie, support groups and sober tailgates). Additionally, many health care systems have integrated persons with lived experience into paraprofessional roles to help support those entering recovery [11]. All these efforts are part of an overall movement toward creating a recovery-oriented system of care that prizes lived experience, provides recovery support, and destigmatizes people with SUD.

However, with few exceptions, what is known about the pathways into and benefits of recovery is almost entirely based on adult research. Some progress has been made in adapting adult-focused models of recovery support for adolescents. For example, recovery high schools congregate adolescents who have entered recovery and could benefit from intense, structured social support [12,13]. Another model of adolescent recovery support is the alternative peer group model, which emphasizes sober recreation, peer support, and linkage to services [14]. While promising, both of these models are not widely available, which is likely driven by the fact that there are no nationally representative epidemiological data on the number of youth in recovery who may benefit from such supports [1].

The purpose of this study is to obtain a more accurate estimate of the prevalence of adolescent self-reported recovery (ASR) than that reported in prior studies. This could help policymakers better assess the need for structured recovery support for younger adolescents. This study used a weighted random sample of students from 10th and 12th grade in a Midwestern US state, which improves upon prior estimates. For example, to our knowledge, only 1 prior study estimated ASR, but it was not representative of the state from which the sample was drawn. It is estimated that 4.9% of 10th graders and 4.6% of 12th graders (roughly aged 16-18 years) may consider themselves to be in recovery [15]. In this study, we address a major weakness of the prior study by using 2024 data from the same statewide survey, which is representative. Thus, this study is the first statewide representative estimate of ASR to be conducted.

Methods

Ethical Considerations

Human subjects approval for the primary study was obtained from the authors' institutional review board (University of Illinois; IRB24-0832), which included authorization to analyze and disseminate aggregated findings for secondary analyses such as this study. This study was approved for a waiver of parental consent, with parental notification and opt-out procedures in place. Schools were instructed not to survey students whose parents had opted them out of the survey and to provide such students with an alternative activity. However, the research team did not receive data on the number of students whose parents declined their participation in the survey.

Study Description

The data used in this study come from the 2024 Illinois Youth Survey (IYS), a biennial survey administered to students in grades 8, 10, and 12. Participation was open to all schools in Illinois. Schools were recruited through a variety of mechanisms, including direct emails to principals and other school contacts from previous survey administrations, telephone calls, online newsletters and press releases, and conference exhibits (eg, the Illinois Principals' Association conference). Those who opted to participate registered online and were sent instructions on distributing parent notifications, instructions on administering the survey across grades at the school, teacher or proctor instructions to be read aloud to students on survey day, and login information for the online survey. Schools administered the survey on a date of their choosing between January 2024 and June 2024, and students completed the survey on school-issued devices during scheduled classroom periods. Some schools, especially larger schools, are surveyed over a few days or a week.

Data Validation

Multiple procedures were applied to ensure internal consistency and data quality [16]. Internal consistency was evaluated by cross-validating related items (eg, past 30-day vs past-year substance use). For instance, respondents who reported alcohol use in the past month but not in the past year were classified as inconsistent and the corresponding survey items were recoded as missing. Following these checks, 12.4% of surveys were invalidated. Surveys were removed if respondents (1) endorsed the use of a fictitious drug (1%), (2) completed fewer than 40% of survey items (9%), or (3) selected 2 of the lowest honesty options (2.8%). Some surveys met more than one exclusion criterion, resulting in overlap.

Sampling Procedures and School Participation

While participation in the IYS was open to all schools, the data used for this study were gathered from schools randomly

XSL•FO
RenderX

selected to comprise a statewide representative sample (refer to sampling procedures explained in subsequent sections). The sampling frame excluded very small schools (fewer than 16 students per grade) and all specialized institutions, such as alternative and charter schools, which may differ from traditional public schools in student composition.

During the 2024 administration cycle, a total of 10,490 surveys were collected from 154 sampled schools, including 74 middle schools and 80 high schools. Student participation rates varied across schools, with median participation rates of 81.5% (IQR 71.9%-86.5%) in middle schools and 55.3% (IQR 34.5%-69.2%) in high schools. This is consistent with participation rates in recent nationally representative surveys, which have dropped since the COVID-19 pandemic [17]. Participating schools included 60 from suburban Chicago, 29 from Chicago, 34 from other urban areas, and 31 from rural Illinois. To ensure adequate representation across grade-by-region strata, schools were oversampled at recruitment. Consequently, the target number of participating schools was achieved in most strata, and all participating schools were retained for analysis. Sampling weights were applied to account for differential probabilities of selection and participation.

The IYS uses a 2-stage, stratified cluster sampling design to generate a statewide representative sample of Illinois public schools serving students from 8th, 10th, and 12th grades. Schools were stratified by geographic region (Chicago, suburban Chicago, other urban or suburban counties, and rural counties) and grade level (8th, 10th, and 12th), yielding 12 strata to account for variation in student distribution across the state. In the first stage, schools were randomly selected within each stratum using a probability-proportional-to-size approach, accounting for differences in grade-level enrollment. In the second stage, all eligible students within selected schools were invited to participate, rather than sampling classrooms, to reduce logistical constraints and minimize within-class clustering. Within each participating school and grade, up to 50 completed surveys were randomly selected; schools with fewer than 50 respondents contributed all completed surveys [18-20]. This approach ensured equal selection probabilities across grades.

Survey weights were constructed to account for differential selection probabilities at the school and student levels. An additional absenteeism-adjustment factor was applied, giving greater weight to students with higher levels of school absence due to the known association between absenteeism and substance use [21]. Final weights were normalized to preserve the original sample size. When used with the specified survey design (including stratification), they yield estimates representative of the statewide public school population.

The 2024 IYS weighted dataset consisted of 10,490 students from 154 sampled schools. For this study, we focused on the subset of students (n=6986) from 10th and 12th grades who were administered the recovery status item. This study provides the first population-based estimate of the prevalence of ASR in Illinois. We hypothesized that the prevalence of ASR would be lower than that reported for adults in recovery.

Measures

The IYS functions primarily as a prevention survey, with most items adapted from the Communities That Care Survey [22]. It includes a wide range of demographic measures, risk factors for initiating substance use, detailed substance use and consequence items, and a variety of other health and safety measures.

The primary variable of interest in this study was self-identified recovery status. Students were asked a single yes-or-no question, “Do you consider yourself to be in recovery?”, which has been used in prior adult research [2,19]. To avoid confusion with mental health recovery, this item appeared within a section clearly labeled: “The following questions are about recovery from substance use.” As done in other adult and adolescent research, immediately before the recovery item, students were asked, “Besides nicotine, did you used to have a problem with drugs or alcohol, but no longer do?” [2,15,19]. An affirmative response to this item is indicative of problem resolution but is not considered indicative of being in recovery.

Analyses

All analyses accounted for the complex sampling design of the IYS and were conducted in both SPSS (version 31.0; IBM Corp) and Stata (version 19.1; StataCorp). Descriptive statistics, including prevalence estimates and 95% CIs, were calculated using the survey design and sampling weights to obtain design-adjusted SEs.

Missing data were minimal for demographic variables (<1.5%) and for the recovery item (3.5%). To evaluate the missing data mechanism, we conducted the Little Missing Completely at Random test on analytic variables, which did not indicate evidence against missing completely at random (χ^2_{50}=63.94; P=.09). However, given the strong conceptual and empirical links between recovery status and related substance use behaviors, multiple imputations were conducted under the more general missing-at-random assumption. The imputation model incorporated a broad set of alcohol-related and marijuana-related measures to improve the plausibility of missing at random and the quality of imputed recovery estimates.

Recovery prevalence estimates were calculated overall and by grade level. Descriptive characteristics of youth who identified as being in recovery are also presented.

Analytic Sample

Among the 6986 students from 10th and 12th grades, we applied a consistency check to ensure the validity of self-reported recovery status. Students (115/6986, 1.6%) who indicated never having used substances on the age-of-first-use item but simultaneously endorsed being in recovery were classified as inconsistent and excluded from the analysis. This step was taken to improve the accuracy of the recovery prevalence estimates. After applying this exclusion, the final analytic, unweighted sample included 6871 students.

XSL•FO
RenderX

Results

Weighted Prevalence of Recovery by Grade

The weighted prevalence of youth who self-identified as being in recovery is presented in Table 1. Overall, 3.3% (231/6909; 95% CI 2.6%-4.1%) of students from 10th and 12th grades reported being in recovery, while 96.6% (6678/6909) reported not being in recovery. When examined by grade level, recovery prevalence was 3.3% (120/3592; 95% CI 2.4%-4.2%) among 10th graders and 3.3% (111/3317; 95% CI 2.4%-4.5%) among 12th graders. In both grades, more than 96% of students reported not being in recovery.

Table . Weighted prevalence of adolescent self-reported recovery (ASR) (N=6871).

Grade	Participants with ASR, n (%)	95% CI[a]	Adolescents not reporting recovery, n (%)
10th	120 (3.3)	2.4-4.2	3472 (96.6)
12th	111 (3.3)	2.4-4.5	3206 (96.6)

[a]95% CIs correspond to the prevalence estimates of ASR and are based on pooled results from multiple imputed analyses.

Weighted Prevalence of Recovery by Demographic Groups and Characteristics

Table 2 presents the weighted prevalence of ASR by demographic factors and characteristics of the participants with ASR. Among the 231 participants with ASR, the mean age was 16.5 (SD 1.14) years. Recovery prevalence was higher among female participants (n=118, 3.8%) than among male participants (n=110, 2.8%), with a small proportion preferring not to answer.

Table . Weighted prevalence and characteristics of adolescent self-reported recovery (ASR) (n=231)[a].

Variables	Participants with ASR, n (%)	Prevalence of ASR, % (95% CI)
Sex		
Female	118 (51.1)	3.8 (2.8-4.8)
Male	110 (47.6)	2.8 (2.0-3.5)
Prefer not to answer	3 (1.3)	—[b]
Race		
White	88 (38.1)	2.2 (1.5-3.0)
Black or African American	30 (13.0)	4.2 (1.8-6.5)
Latino or Latina	90 (39.0)	5.3 (3.5-7.2)
Asian American	10 (4.3)	—
Other	13 (5.6)	—
Grade		
10th	120 (52.1)	3.3 (2.4-4.2)
12th	111 (47.9)	3.3 (2.4-4.5)
Region		
Suburban Chicago	123 (53.2)	3.2 (2.5-4.0)
Chicago	40 (17.3)	4.3 (1.7-6.8)
Other urban or suburban areas	33 (14.3)	2.3 (1.2-3.5)
Rural areas	35 (15.2)	4.9 (2.9-6.9)
Free or reduced-price lunch		
Free lunch	116 (50.2)	4.3 (3.2-5.5)
Reduced-price lunch	13 (5.6)	3.4 (1.3-5.4)
Neither	102 (44.2)	2.4 (1.5-3.3)

[a]Mean age for participants with ASR was 16.5 (SD 1.14) years.

[b]Prevalence and CIs are not shown for cells with small sample sizes due to limited precision.

By race, Latino or Latina youth had the highest prevalence of ASR (5.3%), followed by Black or African American youth (4.2%) and White youth (2.2%). Sample sizes for Asian American and other racial groups were too small to support

XSL•FO
RenderX

reliable estimates. Although White youth had a lower prevalence of ASR, they accounted for a substantial share of the total ASR population (88/231, 38.1%) due to their larger representation in the overall student population. In contrast, Latino or Latina youth, who had the highest prevalence of ASR (5.3%), comprised a disproportionately large share of the ASR population (90/231, 39%) relative to their representation in the student population.

The prevalence of ASR was comparable across grades, with estimates of 3.3% among 10th graders and 3.3% among 12th graders, and the proportion of participants with ASR was approximately evenly split between the two grades (120/231, 52.1% and 111/231, 47.9%, respectively).

Regional differences in prevalence of ASR were observed, with the highest prevalence in rural areas (4.9%), followed by Chicago (4.3%), suburban Chicago (3.2%), and other urban or suburban areas (2.3%). Among all participants with ASR, most were from suburban Chicago (123/231, 53.2%).

Regarding socioeconomic status, as indicated by free or reduced-price lunch status, the highest prevalence of ASR was among those receiving free lunch (4.3%), followed by those receiving reduced-price lunch (3.4%). The prevalence of ASR was lowest among youth not receiving any lunch assistance (2.4%). Approximately half (116/231, 50.2%) of all participants with ASR received free lunch.

Discussion

Principal Findings

In this Midwestern US state, 3.3% of adolescents self-reported being in recovery. Based on the total number of students in the 10th and 12th grades in Illinois, that equates to about 4375 tenth graders and 4455 twelfth graders. Comparably, about 4% to 8% of adults in the United States indicate that they are in recovery [2,23]. This is consistent with our hypothesis that adolescent self-reported recovery would be lower than estimates for adults, owing to fewer years of disordered use and fewer opportunities to make lifestyle changes associated with being in recovery.

Furthermore, we note that the overall prevalence rate of ASR here (3.3%) is lower than that previously reported for adolescents (4.3%) in the same state [15]. This could be because this study used a statewide representative sample, whereas the prior sample did not. Additionally, this study used more stringent data screening, excluding students who reported “never using” substances on items about age of onset but also reported being in recovery. In this study, this exclusion criterion lowered the overall prevalence estimate from 5.1% to 3.3%, a considerable difference. This finding highlights the importance for researchers measuring prevalence of ASR to incorporate validity checks.

On the basis of this and other work, we urge national epidemiological studies to start measuring youth recovery. First, these estimates will help improve the planning and development of recovery support services for adolescents. For example, recovery high schools, specially designed to support youth in recovery, are only available in some areas. These findings may support the plausibility of establishing a recovery high school in Chicago or its suburbs. Second, compared to small clinic-based studies of adolescent recovery, studies based on large epidemiological designs can capture a wide variety of recovery pathways, including nonabstinent recovery. There are many definitions of recovery [1]. Regrettably, we know little about nonabstinent youth in recovery [24], although small studies suggest that many youth are interested in this pathway [25].

Third, asking about youth recovery on epidemiological surveys may lead to new insights through the identification of recovery hot spots; that is, it would enable researchers to evaluate whether systems-level initiatives in recovery promotion (ie, recovery community organizations) are working. Emerging research has shown that county-level recovery ecosystems, measured by the presence of treatment, support groups, and recovery-friendly policies, predict ASR indirectly through parent and peer disapproval [26]. Given the need for community recovery capital [27] to support youth recovery, these findings support the validity of the item we used here and in prior research to measure recovery [15].

Finally, youth recovery may be a critical prevention indicator in the future. Although recovery is generally thought of as a treatment outcome in existing patient placement criteria for SUDs, youth recovery could also be conceptualized as a prevention indicator in the broader health context; that is, participants with ASR could reasonably be expected to experience better health outcomes in adulthood due to lifestyle choices associated with earlier recovery.

Limitations

The following limitations apply to this study. First, the data were self-reported by adolescents, and it is unclear whether all students fully grasped the concept of recovery, despite the introductory heading clarifying that questions referred to substance use. While the survey contractor did not inquire whether schools used monitoring of school devices (ie, GoGuardian), schools were instructed to provide a private space for students to participate and were informed that the survey was anonymous. Research shows that group-administered computerized surveys can reduce reporting of socially undesirable topics due to concerns about confidentiality, even with anonymity assurances [28]. Nevertheless, it is possible that surveillance software used by schools resulted in underreporting. Future studies should track and transparently report whether schools use monitoring software. This would enable better sensitivity testing of its impact on reporting sensitive information.

Second, estimates for 12th graders in the Chicago strata may be less reliable due to small sample sizes at randomized schools and lower student participation. However, as the majority of students in Illinois (223,547/406,719, 55%) live in suburban school districts, the slightly higher recovery estimates in Chicago may only result in a slight upward bias in the statewide prevalence rate. Third, as this survey only included students who were present in school on the days of survey administration, it is a limitation that it does not include students who were absent or had left school. We do not know how this would affect ASR estimates. While absenteeism would affect substance use

XSL•FO
RenderX

estimates [21], participants with ASR may experience improvements in school attendance. Future research should clarify this phenomenon. Fourth, these findings are specific to Illinois public school 10th-grade and 12th-grade students and should not be extrapolated to other states or countries. Future research is needed to assess whether youth recovery prevalence differs nationally or internationally.

Fifth, because the data source is a prevention survey, some variables of interest are unavailable (eg, recovery supports and SUD diagnoses). However, we note that SUD screening scores were elevated among participants with ASR in our prior research [15], possibly supporting the validity of our single-item measure of recovery for adolescents. Finally, the statewide representative sample excluded small schools or specialized institutions, including alternative and charter schools. In sensitivity analyses, we found that alternative schools (in our nonweighted sample) had a higher prevalence of ASR. The exclusions of both absent students and students from alternative schools necessitate additional research on marginalized students. Such students are often less likely to attend school and are also overrepresented in alternative schools [29,30]. However, the overall impact on the state prevalence of the ASR estimate would be low because there are so few alternative schools. Nevertheless, although the data are representative of mainstream public schools, the prevalence estimates reported here may not fully generalize to all youth.

Conclusions

Recovery is an organizing concept of the US SUD service delivery system, which is why it is peculiar that there is no national estimate of youth recovery. This statewide representative survey of a large Midwestern US state found that 3% of youth in 10th and 12th grades considered themselves to be in recovery. We strongly encourage the adoption of a recovery status item on national and international epidemiological studies involving adolescents under the age of majority. Doing so may lead to better recovery support planning, a better understanding of nonabstinent recovery pathways, more research on recovery hot spots, and improved understanding of whether youth recovery functions as both a treatment outcome and a protective factor for future health status.

Funding

This study was funded by the Illinois Department of Human Services, Division of Behavioral Health and Recovery (Contract 43CDZ03292), through a grant from the US Substance Abuse and Mental Health Services Administration.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: DCS
Data curation: CAR
Formal analysis: SB, TB
Funding acquisition: DCS
Supervision: DCS
Writing—original draft: DCS, CAR
Writing—review and editing: DCS, CAR, SB, TB

Conflicts of Interest

None declared.

References

1. Facing addiction in America: the Surgeon General's report on alcohol, drugs, and health. : U.S. Department of Health & Human Services; 2016 URL: https://www.hhs.gov/sites/default/files/facing-addiction-in-america-surgeon-generals-report.pdf [accessed 2026-04-01]
2. Jones CM, Noonan RK, Compton WM. Prevalence and correlates of ever having a substance use problem and substance use recovery status among adults in the United States, 2018. Drug Alcohol Depend 2020 Sep 1;214:108169. [doi: 10.1016/j.drugalcdep.2020.108169] [Medline: 32682218]
3. McLellan AT, Lewis DC, O'Brien CP, Kleber HD. Drug dependence, a chronic medical illness: implications for treatment, insurance, and outcomes evaluation. JAMA 2000 Oct 4;284(13):1689-1695. [doi: 10.1001/jama.284.13.1689] [Medline: 11015800]
4. Moberg DP, Paltzer J. Clinical recognition of substance use disorders in Medicaid primary care associated with universal Screening, Brief Intervention and Referral to Treatment (SBIRT). J Stud Alcohol Drugs 2021 Nov;82(6):700-709. [doi: 10.15288/jsad.2021.82.700] [Medline: 34762029]
5. Mauro PM, Gutkind S, Annunziato EM, Samples H. Use of medication for opioid use disorder among US adolescents and adults with need for opioid treatment, 2019. JAMA Netw Open 2022 Mar 1;5(3):e223821. [doi: 10.1001/jamanetworkopen.2022.3821] [Medline: 35319762]

XSL•FO
RenderX

6. Lee A, Smith DC, Campbell CC. Impact of substance use terminology on MSW students: a randomized vignette study. Res Soc Work Pract 2025. [doi: 10.1177/10497315251338224]
7. Kelly JF, Westerhoff CM. Does it matter how we refer to individuals with substance-related conditions? A randomized study of two commonly used terms. Int J Drug Policy 2010 May;21(3):202-207. [doi: 10.1016/j.drugpo.2009.10.010] [Medline: 20005692]
8. Shatterproof addiction stigma index report: a new way forward. : Shatterproof; 2024 URL: https://www.shatterproof.org/sites/default/files/2024-09/Final%20SASI%202024.pdf [accessed 2026-04-01]
9. Kelly JF, O'Connor JB, Klein MR, et al. One-stop shopping for recovery? A systematic review of the evidence on U.S. recovery community centers. J Subst Use Addict Treat 2025 Aug;175:209723. [doi: 10.1016/j.josat.2025.209723] [Medline: 40447107]
10. Hennessy EA, Nichols LM, Brown TB, Tanner-Smith EE. Advancing the science of evaluating collegiate recovery program processes and outcomes: a recovery capital perspective. Eval Program Plann 2022 Apr;91:102057. [doi: 10.1016/j.evalprogplan.2022.102057] [Medline: 35217288]
11. Gormley MA, Pericot-Valverde I, Diaz L, et al. Effectiveness of peer recovery support services on stages of the opioid use disorder treatment cascade: a systematic review. Drug Alcohol Depend 2021 Dec 1;229(Pt B):109123. [doi: 10.1016/j.drugalcdep.2021.109123] [Medline: 34700201]
12. Finch AJ. Salvaging a Teenage Wasteland: Origins of the Recovery High School Movement: Oxford University Press; 2024. [doi: 10.1093/oso/9780190645502.001.0001]
13. Jurinsky J, George SS, Riccio AA, Finch AJ, Kelly JF, Hennessy EA. Recovery capital, the recovery ecosystem, and substance use among recovery high school students. Front Adolesc Med 2025;3:1567567. [doi: 10.3389/fradm.2025.1567567] [Medline: 41646431]
14. Akkala S, Mehta U, Nikhil Pavaskar R, et al. Alternative peer groups (APGs) in adolescent substance use disorder recovery: an insight into services and the need for robust infrastructure. J Subst Use 2025;31(1):117-123. [doi: 10.1080/14659891.2025.2474933]
15. Smith DC, Reinhart CA, Begum S, et al. Coming of age in recovery: the prevalence and correlates of substance use recovery status among adolescents and emerging adults. PLoS One 2023;18(12):e0295330. [doi: 10.1371/journal.pone.0295330] [Medline: 38113212]
16. Kosgolla JV, Smith DC, Begum S, Reinhart CA. Assessing the self-reported honesty threshold in adolescent epidemiological research: comparing supervised machine learning and inferential statistical techniques. BMC Med Res Methodol 2023 Sep 21;23(1):210. [doi: 10.1186/s12874-023-02035-y] [Medline: 37735353]
17. Santelli JS, Malden DE, Moss RA, Finkelstein M, Lindberg LD. Age bias, missing data, and declining response rates in the national Youth Risk Behavior Survey and their influence on estimates of trends in adolescent sexual experience, 2011–2023. Sex Res Soc Policy 2025. [doi: 10.1007/s13178-025-01223-3]
18. Lohr SL. Sampling: Design and Analysis, 3rd edition: Chapman and Hall/CRC; 2021. [doi: 10.1201/9780429298899]
19. Heeringa SG, West BT, Berglund PA. Applied Survey Data Analysis, 2nd edition: Chapman & Hall/CRC; 2017. [doi: 10.1201/9781315153278]
20. Levy PS, Lemeshow S. Sampling of Populations: Methods and Applications, 4th edition: John Wiley & Sons; 2013.
21. Gakh M, Coughenour C, Assoumou BO, Vanderstelt M. The relationship between school absenteeism and substance use: an integrative literature review. Subst Use Misuse 2020;55(3):491-502. [doi: 10.1080/10826084.2019.1686021] [Medline: 31805820]
22. Arthur MW, Hawkins JD, Pollard JA, Catalano RF, Baglioni AJJ. Measuring risk and protective factors for substance use, delinquency, and other adolescent problem behaviors. The Communities That Care Youth Survey. Eval Rev 2002 Dec;26(6):575-601. [doi: 10.1177/0193841X0202600601] [Medline: 12465571]
23. Kelly JF, Abry AW, Milligan CM, Bergman BG, Hoeppner BB. On being "in recovery": a national study of prevalence and correlates of adopting or not adopting a recovery identity among individuals resolving drug and alcohol problems. Psychol Addict Behav 2018 Sep;32(6):595-604. [doi: 10.1037/adb0000386] [Medline: 30070538]
24. Finch AJ, Jurinsky J, Anderson BM. Recovery and youth: an integrative review. Alcohol Res 2020;40(3):06. [doi: 10.35946/arcr.v40.3.06] [Medline: 33344101]
25. Smith DC, Evans JM, Reinhart CA, Taylor SE, Begum S, Jenkins KV. Adolescent perceptions of substance use problem resolution and recovery. Alcohol Treat Q 2024;42(2):168-178. [doi: 10.1080/07347324.2023.2292028]
26. Kosgolla J, Smith D, Lee A, Reinhart C, Jurinsky J. Community support and youth recovery: the mediating role of peer and parental disapproval. Front Adolesc Med 2025;3:1602279. [doi: 10.3389/fradm.2025.1602279]
27. Hennessy EA, Cristello JV, Kelly JF. RCAM: a proposed model of recovery capital for adolescents. Addict Res Theory 2019;27(5):429-436. [doi: 10.1080/16066359.2018.1540694]
28. Gnambs T, Kaspar K. Disclosure of sensitive behaviors across self-administered survey modes: a meta-analysis. Behav Res Methods 2015 Dec;47(4):1237-1259. [doi: 10.3758/s13428-014-0533-4] [Medline: 25410404]
29. Kearney CA, Childs J, Burke S. Social forces, social justice, and school attendance problems in youth. Contemp Sch Psychol 2023;27:136-151. [doi: 10.1007/s40688-022-00425-5]

XSL•FO
RenderX

30. Kho A, Rabovsky S. The students alternative schools serve. Urban Institute. 2022. URL: https://www.urban.org/sites/default/files/2022-07/2022%2007%2024_Alternative%20Schools_finalized.pdf [accessed 2026-04-01]

Abbreviations

ASR: adolescent self-reported recovery
IYS: Illinois Youth Survey
SUD: substance use disorder

Edited by A Mavragani, T Sanchez; submitted 21.Aug.2025; peer-reviewed by E Sahker, K Cummins; revised version received 13.Mar.2026; accepted 17.Mar.2026; published 29.Apr.2026.

Please cite as:
Smith DC, Reinhart CA, Begum S, Bhatt T
Adolescent Self-Reported Recovery for Substance Use in Illinois: Statewide Representative Epidemiological Study
JMIR Public Health Surveill 2026;12:e82792
URL: https://publichealth.jmir.org/2026/1/e82792
doi:10.2196/82792

© Douglas C Smith, Crystal A Reinhart, Shahana Begum, Tarjani Bhatt. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 29.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Determinants of Willingness to Receive Health Information From Neighborhood Food and Beauty Establishments: Cross-Sectional Study

Seema Aithal[1], MBBS, MPH; Huiling Guo[1], PhD; Zoe Lara-Jane Hildon[2], PhD; May O Lwin[3], PhD; Angela Chow[1], MBBS, PhD

[1]Department of Epidemiology and Preventive Medicine, Tan Tock Seng Hospital, 11 Jalan Tan Tock Seng, Singapore, Singapore

[2]Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore

[3]Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore, Singapore

Corresponding Author:
Angela Chow, MBBS, PhD
Department of Epidemiology and Preventive Medicine, Tan Tock Seng Hospital, 11 Jalan Tan Tock Seng, Singapore, Singapore

Abstract

Background: Although health care providers are the most trusted sources of health information, service establishments within communities represent important, yet underused, sources of health information. Specifically, food and beauty establishments can act as alternative settings for health communication, facilitating broader engagement with the general population.

Objective: This study examined factors associated with willingness to receive health information from these non–health care service establishments among community-dwelling adults in Singapore.

Methods: A cross-sectional survey was conducted among residents in 2 neighborhoods in central Singapore between November 2024 and April 2025. Data on sociodemographic characteristics, trust in information from health care and non–health care services, and willingness to receive health information were collected anonymously. The primary outcome was willingness to receive health information from non–health care services (yes or no), assessed among respondents with no prior exposure to health information from such services. Multivariable logistic regression was used to identify factors independently associated with willingness to receive health information from non–health care services.

Results: Among the 403 respondents, most were aged ≥50 years (n=223, 55.3%), female (n=219, 54.3%), Chinese (n=350, 86.9%), and highly educated (n=302, 74.9%). Of the 339 respondents without prior exposure to health information from non–health care services, approximately one-third (n=106, 31.3%) reported that they were willing to receive health information in the future. In adjusted analysis, greater trust in health information (adjusted odds ratio [AOR] 3.71, 95% CI 1.50-9.19) and high health information orientation (AOR 1.89, 95% CI 1.11-3.21) were associated with increased willingness to receive health information from non–health care services. Trust in health information was positively associated with willingness among those aged 21 to 34 years (AOR 4.96, 95% CI 1.35-18.30), those aged 35 to 49 years (AOR 8.02, 95% CI 2.62-24.59), and male respondents (AOR 6.22, 95% CI 2.79-13.89) to receive health information from these sources, but not among those aged ≥50 years (AOR 1.92, 95% CI 0.92-4.02) or female respondents (AOR 1.85, 95% CI 0.87-3.96).

Conclusions: Nearly one-third of community-dwelling adults expressed willingness to receive health information from non–health care (food and beauty) services, highlighting the potential for leveraging these establishments as alternative health communication channels. Willingness was positively associated with higher health information orientation and greater trust. Additionally, trust in non–health care (food and beauty) services was associated with higher odds of willingness to receive health information among those aged 21 to 49 years and male respondents. This suggests the need for tailored trust-building strategies to strengthen engagement through such alternative channels.

(JMIR Public Health Surveill 2026;12:e86435) doi:10.2196/86435

KEYWORDS

health promotion; health information; trust; willingness; non–health care services

Introduction

According to the World Health Organization, health promotion is defined as the process of enabling people to increase control over and improve their health by fostering activities that enhance the biological, social, psychological, environmental, and economic conditions necessary for achieving optimal health [1]. Health education is a central component of health promotion and empowers populations to adopt healthy behaviors [2].

XSL•FO
RenderX

Importantly, Sørensen et al [3] highlight that access to health information represents the essential first step before individuals can understand, evaluate, and apply such information to adopt desirable health behaviors.

Although health care providers represent the primary trusted information sources [4,5], equitable access to health information continues to pose significant challenges, especially for specific subpopulations and individuals with unique biopsychosocial needs [6]. Individuals who turn first to their health care providers for health information represent a unique population and mainly comprise older adults (aged >65 years), those in poor health, those with lower education, those having health insurance, and non–internet users [7]. This underscores the need to explore nontraditional approaches for disseminating health information across the wider community, including service establishments. Such approaches may be especially relevant for younger adults, who typically have fewer health care encounters owing to lower chronic disease burden, thereby emphasizing the importance of community-based social networks and touchpoints for engaging this demographic [8].

The African American barbershop health initiative in the United States illustrates the successful transformation of non–health care community spaces into health promotion venues by tapping into existing social networks and trusted relationships with barbers serving as community health advocates [9,10]. The initiative's success stemmed from recognizing barbershops as natural gathering spaces where health messages could be effectively conveyed in familiar, comfortable settings. This approach has been efficacious across multiple health areas, including outreach for infectious diseases (HIV and COVID-19), chronic disease management (diabetes and hypertension), cancer screening, and mental health promotion [11-18]. Furthermore, this community-based health promotion approach has proven remarkable adaptability across varied cultural and health care contexts internationally [19-25], attesting to the viability of engaging non–health care services for health promotion and education outreach.

Food and beauty establishments can serve as alternative venues for disseminating health information and expanding outreach to the general population. Beauty establishments, including barbershops and beauty salons, have been heavily studied, but little is known about the utility of other non–health care establishments in the community for disseminating health information at scale. In Singapore, where an estimated 60% of residents dine out regularly [26], hawker centers, food courts, and coffee shops often serve as "community dining rooms" that cater to all ages, classes, and schedules [27]. Programs promoting healthier diets and hawker center etiquette, mainly targeting patrons of these food establishments, have been successfully implemented [28,29]. Moreover, hawkers often recognize their regular patrons and keep up with their patrons' lives [30]. The potential of neighborhood establishments with high patronage, specifically food and beauty establishments, to serve as conduits for public health messaging warrants consideration. Such venues may be especially valuable for reaching out to younger demographics with infrequent health care contact while tapping into existing social connections formed in these community spaces.

Trust constitutes an important prerequisite for individuals' willingness to receive health information from non–health care sources. Prior research has shown that trust reduces perceived risk [31], implying that individuals with greater trust in the information provider are less likely to view the information as inappropriate or potentially harmful compared with those who have lower trust. Moreover, trust has also been shown to positively influence participant compliance and cooperation [32]. For example, a Dutch study [33] investigating the association between trust in health insurers and willingness to receive health care advice from these insurers found that higher levels of trust were significantly associated with increased willingness to receive health care advice. Against this backdrop, this study aims to explore the willingness of community-dwelling adults in Singapore to receive health information from non–health care services, specifically food and beauty establishments located within residential neighborhoods, and identify the factors that influence it.

Methods

Study Design and Setting

A cross-sectional survey was conducted in 2 neighborhoods located in central Singapore between November 2024 and April 2025. The sampling frame comprised all residential units located within the 2 selected neighborhoods, identified from the Housing and Development Board residential database and verified against existing physical mailboxes. After excluding units located above commercial premises, a survey package consisting of a study information sheet (appended with a URL or QR code for access to the online survey) and a step-by-step instructional infographic guide was delivered to all eligible residential units by postal mail. Each household received a participant information sheet containing a unique identifier to detect duplicate submissions and facilitate participant reimbursement.

The online self-administered survey included predominantly close-ended questions with a predefined set of responses. Participants were required to complete all questions in the survey to ensure valid submission and minimize item nonresponse. Only 1 resident per household, aged ≥21 years, was eligible to participate in the study. Initial recruitment involved invitations to 9913 households within the study area. To boost participation, reminder letters were sent to 500 randomly selected households from the list of remaining addresses that had not yet responded to the initial survey invitation. Subsequently, research assistants conducted door-to-door follow-up at these 500 household addresses to achieve the target minimum sample size of 385 participants. Assistance was provided to older participants who expressed interest but faced difficulties with self-completion, through interviewer administration.

Ethical Considerations

The study was approved by the National Healthcare Group Domain Specific Review Board of Singapore (2023/00118). Informed consent was waived, and participants provided implied consent by voluntarily completing and submitting the online survey. Participants who completed the survey received S$5 (US $3.93) as reimbursement for their time. No personal

XSL•FO
RenderX

identifiers were collected in the survey, and deidentified data were used for analysis.

Survey Instrument

The survey questionnaire collected comprehensive sociodemographic information, including age, sex, ethnicity, marital status, highest educational qualification, employment status, and duration of residency in the neighborhood (<10 years or ≥10 years). Questions on self-reported health status (very good and good vs fair, poor, and very poor), health information orientation, and health literacy were also included. Health information orientation [34], which describes an individual's inherent interest in actively seeking relevant health information to make informed decisions about their health, was assessed using 8 items measured on a 5-point Likert scale (1=strongly disagree to 5=strongly agree), with scores ≥32 (≥80% of the maximum score of 40) indicating a high level of health orientation. Health literacy [35] was evaluated using a 16-item European Health Literacy Survey Questionnaire Short form, where responses were dichotomized (very difficult or difficult=0 and easy or very easy=1), and a total score of ≥13 was classified as an adequate level of health literacy.

The survey also incorporated questions on whether participants passively received health information and actively sought health information from a range of sources, including health care professionals, staff from community care centers, social contacts (family members, friends, and neighbors), social influencers or bloggers, and staff from food and beauty establishments. Data were also collected on residents' passive receipt of information across 5 different types of health topics: health screening, lifestyle behaviors, chronic disease management, vaccination, and antibiotic use guidance. Trust was operationalized as overall trust in health information received from various sources, including health care clinics and non–health care (food and beauty) establishments. It was assessed using a 5-point Likert scale (1=never, 2=somewhat, 3=moderately, 4=a lot, and 5=completely), and participants who selected "moderately," "a lot," or "completely" were deemed to have trust in health information from health care and non–health care services. The questionnaire was developed in English and translated into 2 other local languages: Mandarin and Malay (the survey questionnaire is available in Multimedia Appendix 1).

Data Analysis

The primary outcome was future willingness to receive health information from non–health care services (yes or no), evaluated exclusively among those who had never previously received health information from such services. In accordance with the study hypothesis, non–health care services were operationally defined as food and beauty establishments. Participants who reported prior receipt were excluded to avoid potential confounding, as previous exposure could reasonably influence both their trust in health information from these services and their willingness to receive health information in the future.

Descriptive analysis involved categorical variables expressed as proportions and continuous variables as means with SDs. Chi-square test was used to compare differences between the 2 outcome groups. First, we conducted univariate analysis to inform variable selection for subsequent multivariable analysis. Covariates were identified a priori through a review of the literature and selected based on their relevance as determined by the research team. The initial multivariable logistic regression model adjusted for core sociodemographic variables (age group, gender, and ethnic group), along with variables that were significant or showed marginal significance (P<.20) in univariate analysis. Next, potential interactions between trust and other covariates were systematically explored based on theoretical considerations. Significant interaction terms were sequentially added in the subsequent models. Multicollinearity was assessed using the variance inflation factor. The final regression model was selected based on goodness-of-fit criteria, including Akaike information criteria and log likelihood ratio tests (Table S1 in Multimedia Appendix 1). Additional analyses explored effect modification by age and gender. Statistical significance was defined as P <.05, with all analyses conducted using Stata (version 18.0; StataCorp LLC).

Results

In total, 406 survey responses were collected from 9913 households, giving an overall response rate of 4.1%. After data cleaning, 403 responses were included in the final analysis sample, with 3 (0.01%) excluded due to duplicate responses from the same households.

Cohort Characteristics

Among the 403 respondents, 54.3% (n=219) were female, with a mean age of 50.9 (SD 15) years. According to the 2025 population census for the 2 study neighborhoods, the demographic profile of respondents was similar to the underlying population in terms of sex (male participants: n=20,310, 46.7% vs female participants: n=23,200, 53.3%; P=.68) and age distribution (21 - 34 years: n=7910, 18.2%; 35 - 49 years: n=10,380, 23.9%; ≥50 years: n=25,220, 58%; P=.06). Most respondents were Chinese (n=350, 86.9%), highly educated (n=302, 74.9%), and currently employed (n=270, 67%). Approximately two-thirds (n=262, 65%) were long-term residents with ≥10 in their neighborhoods. Roughly two-thirds of respondents (n=258, 64%) showed high health information orientation and 71.5% (n=288) displayed an adequate level of health literacy. Self-reported health status was fair, poor, or very poor among 41.4% (n=167) of respondents (Table 1). The proportion reporting fair, poor, or very poor health was 35.9% (n=23) among younger adults, 39.7% (n=46) among those aged 35 to 49 years, and 44% (n=98) among those aged ≥50 years.

XSL•FO
RenderX

Table . Sociodemographic characteristics of respondents from 2 neighborhoods in central Singapore (N=403).

Variables	Respondents
Age (years), mean (SD)	50.9 (15.0)
Age group (years), n (%)	
21 - 34	64 (15.9)
35 - 49	116 (28.8)
≥50	223 (55.3)
Sex, n (%)	
Male	184 (45.7)
Female	219 (54.3)
Ethnic group, n (%)	
Chinese	350 (86.8)
Non-Chinese	53 (13.2)
Education, n (%)	
Lower education (postsecondary education and below)	101 (25.1)
Higher education (diploma and above)	302 (74.9)
Marital status, n (%)	
Never married, widowed, divorced, or separated	265 (65.8)
Married	138 (34.2)
Employment status, n (%)	
Not currently working	133 (33.0)
Currently working	270 (67.0)
Duration of residency in the neighborhood (years), n (%)	
<10	141 (35.0)
≥10	262 (65.0)
Self-reported health status, n (%)	
Fair, poor, and very poor	167 (41.4)
Good and very good	236 (58.6)
Health information orientation[a], n (%)	
Low level	145 (36.0)
High level	258 (64.0)
Health literacy[b], n (%)	
Inadequate	115 (28.5)
Adequate	288 (71.5)
Trust in health information from health care sources, n (%)	
No	18 (4.5)
Yes	385 (95.5)
Trust in health information from non–health care (food and beauty) services, n (%)	
No	298 (74.0)
Yes	105 (26.0)
Ever received health information from health care sources, n (%)	
No	72 (17.9)
Yes	331 (82.1)

XSL•FO
RenderX

Variables	Respondents
Ever received health information from non–health care (food and beauty) services, n (%)	
No	339 (84.1)
Yes	64 (15.9)

[a]High health information orientation was defined as a score of ≥80% of the total score of 40 (cutoff ≥32).

[b]Adequate health literacy was defined using the European Health Literacy Survey Questionnaire–Short form (HLS-EU-Q16) standard cutoff score of ≥13.

Receipt of and Trust in Health Information

Table 2 shows the proportion of respondents who received information both actively and passively stratified by type of information source. Approximately four-fifths of respondents (n=331, 82.1%) had passively received health information from health care sources, but only 11.9% (n=48) and 11.7% (n=47) of respondents indicated prior receipt of health information from food and beauty establishments, respectively. Overall, 16% (64/403) of respondents reported previous receipt of health information from non–health care (food and beauty) services. Information on healthy lifestyle behaviors was the most commonly received health topic from non–health care (food and beauty) services (n=33, 8.2% to n=35, 8.7%) among the different health topics examined. Nearly three-quarters of respondents (n=313, 77.7%) reported having actively sought information from health care clinics for health-related matters, but a lower proportion of respondents had obtained health information from non–health care services (n=14, 3.5% from food establishments and n=16, 4% from beauty establishments). Overall, 95.5% (n=385) of the respondents expressed trust in health information from health care sources as compared to 23.1% (n=93) for food establishments and 20.6% (n=83) for beauty establishments (Table 2).

Table . Proportion of respondents who passively received, actively sought, and trusted health information, by source type (N=403).

Type of information source	Health care establishments, n (%)	Food establishments, n (%)	Beauty establishments, n (%)	Community care centers, n (%)	Social contacts, n (%)	Social influencers or bloggers, n (%)
Proportion who passively received information on health-related matters	331 (82.1)	48 (11.9)	47 (11.7)	123 (30.5)	189 (46.9)	38 (9.4)
Proportion who passively received information on						
Healthy lifestyle behaviors	260 (64.5)	35 (8.7)	33 (8.2)	81 (20.1)	162 (40.2)	32 (7.9)
Health screening	272 (67.5)	25 (6.2)	23 (5.7)	90 (22.3)	139 (34.5)	17 (4.2)
Appropriate use of antibiotics	260 (64.5)	23 (5.7)	18 (4.5)	70 (17.4)	102 (25.3)	16 (4.0)
Vaccination	253 (62.8)	27 (6.7)	22 (5.5)	80 (19.9)	120 (29.8)	16 (4.0)
Management of chronic diseases	244 (60.5)	26 (6.4)	19 (4.7)	82 (20.3)	128 (31.8)	14 (3.5)
Proportion who actively sought information on health-related matters	313 (77.7)	14 (3.5)	16 (4.0)	49 (12.2)	50 (12.4)	5 (1.2)
Proportion who trusted information on health-related matters	385 (95.5)	93 (23.1)	83 (20.6)	278 (69.0)	179 (44.4)	51 (12.6)

Willingness to Receive Health Information From Non–Health Care Services

Among the 339 respondents who had not previously received any health information from non–health care (food and beauty) services, nearly one-third (n=106, 31.3%) indicated their willingness to receive such information in the future. Respondents expressed the highest preference for receiving information on chronic disease management from food (99/377, 26.3%) and beauty (91/384, 23.7%) establishments, compared with other health topics they were willing to receive in the future (Table S2 in Multimedia Appendix 1).

A higher proportion of respondents who were willing to receive health information from non–health care (food and beauty) services exhibited high levels of health information orientation (74/106, 69.8% vs 133/233, 57.1%; *P*=.03) and trusted the health

XSL•FO
RenderX

information from non–health care services (40/106, 37.7% vs 36/233, 15.5%; $P<.001$) compared to those who were unwilling (Table 3). No significant differences were observed between willing and unwilling respondents with regard to age, sex, ethnicity, or education levels.

XSL•FO
RenderX

Table . Logistic regression analysis of factors associated with willingness to receive health information from non–health care (food and beauty) services among respondents across 2 neighborhoods in central Singapore (N=339).

Variables	Willingness to receive health information from non–health care services		Univariate model		Multivariable model	
	Yes (n=106)	No (n=233)	Crude OR[a] (95% CI)	*P* value[b]	Adjusted OR (95% CI)	*P* value[b]
Age (years), mean (SD)	48.2 (14.6)	51.6 (14.9)	__[c]	—	—	—
Age group (years), n (%)						
21 - 34	22 (20.8)	32 (13.7)	1.75 (0.93-3.28)	.08	1.30 (0.57-2.99)	.53
35 - 49	32 (30.2)	69 (29.6)	1.18 (0.69-1.99)	.54	1.03 (0.53 - 2.03)	.92
≥50	52 (49.0)	132 (56.7)	Reference		Reference	
Sex, n (%)						
Male	51 (48.1)	114 (48.9)	Reference		Reference	
Female	55 (51.9)	119 (51.1)	1.03 (0.65 - 1.64)	.89	1.41 (0.78 - 2.53)	.25
Ethnic group, n (%)						
Chinese	91 (85.8)	209 (89.7)	Reference		Reference	
Non-Chinese	15 (14.2)	24 (10.3)	1.44 (0.72 - 2.86)	.30	1.45 (0.68 - 3.07)	.33
Education, n (%)						
Lower education (postsecondary education and below)	31 (29.2)	54 (23.2)	Reference		—	
Higher education (diploma and above)	75 (70.8)	179 (76.8)	0.73 (0.44 - 1.22)	.23	—	—
Marital status, n (%)						
Never married, widowed, divorced, or separated	66 (62.3)	157 (67.4)	Reference		—	
Married	40 (37.7)	76 (32.6)	1.25 (0.78 - 2.02)	.36	—	—
Employment status, n (%)						
Not currently working	35 (33.0)	80 (34.3)	Reference		—	
Currently working	71 (67.0)	153 (65.7)	1.06 (0.65 - 1.73)	.81	—	—
Duration of residency in the neighborhood (years), n (%)						
<10	33 (31.1)	90 (38.6)	Reference		Reference	
≥10	73 (68.9)	143 (61.4)	1.39 (0.85 - 2.27)	.18	1.72 (0.98 - 3.03)	.06
Self-reported health status, n (%)						
Fair, poor, or very poor	44 (41.5)	104 (44.6)	Reference		—	
Good or very good	62 (58.5)	129 (55.4)	1.14 (0.71 - 1.81)	.59	—	—
Trust in health information from non–health care (food and beauty) services, n (%)						
No	66 (62.3)	197 (84.6)	Reference		Reference	
Yes	40 (37.7)	36 (15.4)	3.32 (1.95 - 5.63)	*<.001*	3.71 (1.50 - 9.19)	*.005*
Health information orientation[d], n (%)						
Low level	32 (30.2)	100 (42.9)	Reference		Reference	

XSL•FO
RenderX

Variables	Willingness to receive health information from non–health care services		Univariate model		Multivariable model	
	Yes (n=106)	No (n=233)	Crude OR[a] (95% CI)	*P* value[b]	Adjusted OR (95% CI)	*P* value[b]
High level	74 (69.8)	133 (57.1)	1.74 (1.07 - 2.84)	*.03*	1.89 (1.11 - 3.21)	*.02*
Health literacy[e], n (%)						
Inadequate	26 (24.5)	69 (29.6)	Reference		—	
Adequate	80 (75.5)	164 (70.4)	1.29 (0.77 - 2.19)	.33	—	—
Interaction between trust in health information from non–health care (food and beauty) services and respondents aged 21 to 34 years	—	—	—	—	3.43 (0.72 - 16.37)	.12
Interaction between trust in health information from non–health care (food and beauty) services and respondents aged 35 to 49 years	—	—	—	—	5.42 (1.31 - 22.40)	*.02*
Interaction between trust in health information from non–health care (food and beauty) services and female respondents	—	—	—	—	0.23 (0.07 - 0.74)	*.01*

[a]OR: odds ratio.

[b]Italicized values indicate statistical significance of *P*<.05.

[c]Not applicable.

[d]High health information orientation was defined as a score ≥80% of the total score of 40 (cutoff ≥32).

[e]Adequate health literacy was defined using the European Health Literacy Survey Questionnaire - Short form (HLS-EU-Q16) with a cutoff score of ≥13.

Determinants of Willingness to Receive Health Information From Non–Health Care Sources

Health information orientation (*P*=.03) and trust (*P*<.001) were the significant predictors of willingness to receive health information in the future from non–health care services such as food and beauty establishments. Compared to those with low health information orientation, respondents with high health information orientation were associated with greater willingness (adjusted odds ratio [AOR] 1.89, 95% CI 1.11 - 3.21) to receive health information from non–health care (food and beauty) services. Respondents who trusted health information from non–health care (food and beauty) services were associated with higher odds of willingness to receive health information from such services in the future (AOR 3.71, 95% CI 1.50 - 9.19) compared to those who did not trust health information from non–health care (food and beauty) services (Table 3).

Age and sex demonstrated significant interaction effects with trust in health information from non–health care (food and beauty) services. Among respondents aged 21 to 34 years and 35 to 49 years, trust in health information from non–health care (food and beauty) services was associated with nearly 5-fold (AOR 4.96, 95% CI 1.35-18.30) and 8-fold (AOR 8.02, 95% CI 2.62-24.59) higher odds of willingness to receive health information, respectively, after adjusting for sex, ethnicity, long-term residency, and health information orientation. However, this association was not significant among respondents aged ≥50 years (AOR 1.92, 95% CI 0.92-4.02; Table 4). Similarly, sex-stratified analysis showed that trust in health information from non–health care (food and beauty) services was associated with 6-fold higher odds of willingness to receive health information among male respondents (AOR 6.22, 95% CI 2.79-13.89), but not among female respondents (AOR 1.85, 95% CI 0.87-3.96), after adjusting for age, ethnicity, long-term residency, and health information orientation (Table 5).

XSL•FO
RenderX

Table . Association of willingness to receive health information and trust in health information from non–health care (food and beauty) services, according to age group (N=339).

Willingness to receive health information from non–health care (food and beauty) services	Respondents aged 21 to 34 years		Respondents aged 35 to 49 years		Respondents aged ≥50 years	
	Odds ratio (95% CI)	*P* value[a,b]	Odds ratio (95% CI)	*P* value[a,b]	Odds ratio (95% CI)	*P* value[a,b]
Unadjusted analysis						
Trust in health information from non-healthcare services (yes)	4.50 (1.26-16.04)	*.02*	7.18 (2.40 - 21.47)	*<.001*	2.08 (1.01 - 4.29)	*.05*
Adjusted analysis[c]						
Trust in health information from non-healthcare services (yes)	4.96 (1.35-18.30)	*.02*	8.02 (2.62 - 24.59)	*<.001*	1.92 (0.92 - 4.02)	.08

[a]Multiplicative scale.

[b]Italicized values indicate statistical significance of *P*<.05.

[c]Adjusted for gender, ethnicity, duration of residency in the neighborhood, and health information orientation.

Table . Association of willingness to receive health information and trust in health information from non–health care (food and beauty) services, according to sex (N=339).

Willingness to receive health information from non–health care services	Male respondents		Female respondents	
	Odds ratio (95% CI)	*P* value[a,b]	Odds ratio (95% CI)	*P* value[a,]
Unadjusted analysis				
Trust in health information from non-healthcare services (yes)	6.35 (2.90-13.91)	*<.001*	1.81 (0.86-3.80)	.12
Adjusted analysis[c]				
Trust in health information from non-healthcare services (yes)	6.22 (2.79-13.89)	*<.001*	1.85 (0.87-3.96)	.11

[a]Multiplicative scale.

[b]Italicized values indicate statistical significance of *P*<.05.

[c]Adjusted for age, ethnicity, duration of residency in the neighborhood, and health information orientation.

Discussion

Principal Findings

This study set out to examine the potential of non–health care service establishments within communities that function as routine, everyday touchpoints for residents and offer place-based opportunities for health communications. Our findings offer valuable insights into the willingness of community-dwelling adults to accept neighborhood non–health care (food and beauty) services as potential health information sources. Of interest, 16% (64/403) of the respondents reported having received health information passively from either food or beauty establishments in their neighborhoods. Among these, information on healthy lifestyle behaviors constituted the largest proportion , compared with health screening, appropriate antibiotic use, vaccination, and chronic disease management. However, nearly a quarter of respondents (83/385, 21.6% to 99/377, 26.3%) with no prior exposure to health information from such sources indicated their willingness to receive information on various health topics in the future. Thus, to harness this potential resource more effectively, it is imperative to first explore the factors that influence the involvement of food and beauty establishments in disseminating health information.

Of the 339 (84.1%) respondents who had never received health information from food and beauty establishments, nearly one-third (n=106, 31.3%) expressed willingness to receive such

XSL•FO
RenderX

information from these sources in the future. This finding highlights a meaningful opportunity and suggests that food and beauty establishments may offer some untapped potential as alternative settings for health promotion. By providing health information to receptive audiences in informal, everyday settings, these establishments could complement the traditional health care communication strategies and help broaden the reach of public health messaging.

The results also revealed that higher health information orientation and greater trust were positively associated with an individual's willingness to receive health information from non–health care (food and beauty) services. It is possible that individuals with high health information orientation are significantly more health conscious and more likely to seek health information from diverse sources compared to those with low health information orientation, who are less inclined to seek health information [36]. Such individuals often supplement the information provided by health care professionals with additional sources [37]. Furthermore, it is suggested that people who actively seek and value health information tend to be more engaged in their communities and more likely to participate in community health activities [38]. Thus, leveraging these individuals as health ambassadors may help to engage peers, build trust, and boost community participation in preventive health interventions.

In line with previous research, we found that the vast majority (n=385, 95.5%) of respondents trusted health information from health care establishments, whereas only a small proportion (n=83, 20.6% to n=93, 23.1%) trusted health information from food and beauty establishments. Prior literature indicated that expertise, accuracy, and credibility or reliability were the most important qualities for trusting information sources [39]. Trust in health care staff is rooted in their academic qualifications, medical knowledge, adherence to professional standards, and high level of authority [40]. Conversely, trust in non–health care sources such as family and friends relies on repeated interactions built over time through familiar relationships and is valued primarily for personal relevance, reassurance, and emotional support [41,42]. Social connections, based on familiarity, often exert significant influence on individual decision-making processes, especially regarding health care decisions [43]. Fostering trust may be an important component of strategies to enhance willingness, and future studies could investigate how trust operates within familiar social networks and how novel interpersonal sources of health information are used and perceived.

Regarding trust in health information from non–health care services, previous research suggests that disseminated messages should align with, and ideally be endorsed by, trusted health care authorities such as the Health Promotion Board and other relevant institutions [4]. This alignment is especially important in the current digital landscape, which is characterized by widespread misinformation amid vast amounts of online content. Consistent and endorsed health messaging could significantly enhance the perceived credibility of non–health care establishments as emerging health advocates within the community. In addition, leveraging the trust fostered through repeated interactions and social relationships between patrons and establishments [43,44] may further enhance the public's willingness to receive health information from these establishments. This underscores the potential value of engaging food and beauty establishments with high footfall for community health promotion efforts. Given the observed age and gender differences in trust, future research is warranted to explore the mechanisms that foster trust, particularly among younger and male respondents in relation to food and beauty establishments. Such insights could guide strategies for effectively using these venues as community-based health information sources.

Our study has several strengths. To our knowledge, it is the first study to explore the feasibility of using non–health care community services—specifically food and beauty establishments—as channels for public health messaging in a Southeast Asian context, thereby contributing to the limited research in this area. The findings offer useful insights for developing community-based interventions that leverage these non–health care service providers as alternative health communication channels. In particular, the role of non–health care service providers, especially food establishments, extends beyond local relevance and has broader public health implications, given the increasing prevalence of eating-out cultures in both Western and Southeast Asian societies [45]. As such, the insights generated may be applicable to other urban and sociocultural settings. Finally, the use of validated measurement scales for health information orientation and health literacy strengthens the methodological rigor and enables comparison with existing literature.

However, several limitations warrant consideration. First, the cross-sectional design restricts causal inference, thus limiting findings to associational relationships. The possibility of reverse causation or unmeasured confounding also cannot be ruled out due to the cross-sectional nature of our study. Second, although the survey response rate was low, comparisons with available census data from the 2 study neighborhoods on sex and age suggest that selection bias (if any) is likely to be minimal. Third, although trust is context dependent, we did not collect data on the factors that help to build rapport or trust, such as the nature of the interactions between patrons and establishments, or on the alignment of messages with the preferences of the different target audiences. These factors may influence how trust is formed and maintained in real-world settings. Finally, trust in health information from various sources was measured using a single 5-point Likert scale question, without exploring the different dimensions of trust. Future studies should investigate the complex interplay between cognitive and behavioral components of trust across different health information sources, while also considering diverse health conditions and disease severity in naturalistic settings.

Conclusions

Our findings highlight the potential for using non–health care services, particularly food and beauty establishments located within residential neighborhoods, as channels for health communication, with nearly 1-in-3 community-dwelling adults expressing willingness to receive health information from these sources. Individuals exhibiting higher health information orientation and greater trust were associated with increased

XSL•FO
RenderX

willingness. Fostering trust could further improve their willingness to receive health information disseminated through such establishments. Additionally, trust in non–health care services was associated with greater willingness among adults aged 21 to 49 years and among male respondents, but not among those aged ≥50 years or female respondents. Further research is warranted to explore the age- and gender-specific approaches that strengthen trust and subsequently enhance willingness to receive health information from non–health care service establishments.

Acknowledgments

The authors thank all residents from the Cheng San and Teck Ghee neighborhoods who participated in this study. The authors also acknowledge the hard work of the research assistants, Mr Ayden Nafi, Mr Gabriel Toh, Mr Jedrek Wang, and Ms Eryn Lim, who assisted with survey administration and participant recruitment.

Funding

This study was supported by the National Medical Research Council's Clinician Scientist Award (MOH-000150) from the Ministry of Health, Singapore. The funders had no role in the study design, data collection, or interpretation of the work for publication.

Data Availability

The datasets generated and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: AC, HG, MOL, ZH
Data curation: AC, HG
Formal analysis: SA, HG, AC
Funding acquisition: AC
Investigation: AC, HG
Methodology: AC, HG, MOL, and ZH
Visualization: SA
Writing—original draft: SA
Writing—review and editing: SA, HG, ZH, MOL, AC

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables and survey questionnaire used in the study.
[PDF File, 309 KB - publichealth_v12i1e86435_app1.pdf]

References

1. Ottawa Charter for Health Promotion. World Health Organization. 1986. URL: https://iris.who.int/server/api/core/bitstreams/b170c0c7-efbb-4b85-8191-cc67822e7cf0/content [accessed 2025-07-31]
2. Health promotion and health protection. Australian Institute of Health and Welfare. 2024. URL: https://www.aihw.gov.au/reports/australias-health/health-promotion [accessed 2025-07-31]
3. Sørensen K, Van den Broucke S, Fullam J, et al. Health literacy and public health: a systematic review and integration of definitions and models. BMC Public Health 2012 Jan 25;12:80. [doi: 10.1186/1471-2458-12-80] [Medline: 22276600]
4. Guo H, Hildon ZJL, Chow A. "Antibiotics are for everyone, our past and our future generations, right? If antibiotics are dead, we will be in big trouble": building on community values for public engagement on appropriate use of antibiotics in Singapore. Front Public Health 2022 Sep 30;10:1001282. [doi: 10.3389/fpubh.2022.1001282]
5. Hesse BW, Nelson DE, Kreps GL, et al. Trust and sources of health information: the impact of the internet and its implications for health care providers: findings from the first Health Information National Trends Survey. Arch Intern Med 2005;165(22):2618-2624. [doi: 10.1001/archinte.165.22.2618] [Medline: 16344419]
6. The Lancet Global Health. Contemporary challenges to health information for all. Lancet Glob Health 2024 Sep;12(9):e1370. [doi: 10.1016/S2214-109X(24)00336-X] [Medline: 39151961]
7. Swoboda CM, Van Hulle JM, McAlearney AS, Huerta TR. Odds of talking to healthcare providers as the initial source of healthcare information: updated cross-sectional results from the Health Information National Trends Survey (HINTS). BMC Fam Pract 2018 Aug 29;19(1):146. [doi: 10.1186/s12875-018-0805-7] [Medline: 30157770]

XSL•FO
RenderX

8. Yan S, Kwan YH, Thumboo J, Low LL. Characteristics and health care utilization of different segments of a multiethnic Asian population in Singapore. JAMA Netw Open 2019 Sep 4;2(9):e1910878. [doi: 10.1001/jamanetworkopen.2019.10878] [Medline: 31490539]
9. Palmer KNB, Rivers PS, Melton FL, et al. Health promotion interventions for African Americans delivered in U.S. barbershops and hair salons- a systematic review. BMC Public Health 2021 Aug 16;21(1):1553. [doi: 10.1186/s12889-021-11584-0] [Medline: 34399723]
10. Wippold GM, Frary SG, Garcia KA, Wilson DK. Implementing barbershop-based health-promotion interventions for Black men: a systematic scoping review. Health Psychol 2023 Jul;42(7):435-447. [doi: 10.1037/hea0001294] [Medline: 37227823]
11. Baker JL, Brawner B, Cederbaum JA, et al. Barbershops as venues to assess and intervene in HIV/STI risk among young, heterosexual African American men. Am J Mens Health 2012 Sep;6(5):368-382. [doi: 10.1177/1557988312437239] [Medline: 22398991]
12. Hill K, Haynes N, Momplaisir F. Partnering with barbershops and salons to engage vulnerable communities during the COVID-19 pandemic. Open Forum Infect Dis 2021;8(7):ofab191. [doi: 10.1093/ofid/ofab191] [Medline: 34316501]
13. Kaneri P, Lima do Vale M, Harding S, Molokhia M. A scoping review of the evidence available for the use of salons as health promotion environments, for the prevention and management of non-communicable diseases in women from different ethnic backgrounds. Front Public Health 2023;11:1161645. [doi: 10.3389/fpubh.2023.1161645] [Medline: 37529433]
14. Victor RG, Blyler CA, Li N, et al. Sustainability of blood pressure reduction in Black barbershops. Circulation 2019 Jan 2;139(1):10-19. [doi: 10.1161/CIRCULATIONAHA.118.038165] [Medline: 30592662]
15. Signorello LB, Schlundt DG, Cohen SS, et al. Comparing diabetes prevalence between African Americans and Whites of similar socioeconomic status. Am J Public Health 2007 Dec;97(12):2260-2267. [doi: 10.2105/AJPH.2006.094482] [Medline: 17971557]
16. Luque JS, Rivers BM, Gwede CK, Kambon M, Green BL, Meade CD. Barbershop communications on prostate cancer screening using barber health advisers. Am J Mens Health 2011 Mar;5(2):129-139. [doi: 10.1177/1557988310365167] [Medline: 20413392]
17. Rogers CR, Okuyemi K, Paskett ED, et al. Study protocol for developing #CuttingCRC: a barbershop-based trial on masculinity barriers to care and colorectal cancer screening uptake among African-American men using an exploratory sequential mixed-methods design. BMJ Open 2019 Jul 24;9(7):e030000. [doi: 10.1136/bmjopen-2019-030000] [Medline: 31345981]
18. Hammond WP. Taking it like a man: masculine role norms as moderators of the racial discrimination-depressive symptoms association among African American men. Am J Public Health 2012 May;102 Suppl 2(Suppl 2):S232-S241. [doi: 10.2105/AJPH.2011.300485] [Medline: 22401515]
19. Ogborn G, Bowden-Howe C, Burd P, Kleijn M, Michelson D. Barbershops as a setting for supporting men's mental health during the COVID-19 pandemic: a qualitative study from the UK. BJPsych Open 2022 Jun 27;8(4):e118. [doi: 10.1192/bjo.2022.520] [Medline: 35757926]
20. Zaman M, de Vale ML, Coultas C, et al. Factors affecting the delivery of community-based salon interventions to prevent cardiovascular disease and breast cancer among ethnically diverse women in south London: a concept-mapping approach. Lancet 2023 Nov;402 Suppl 1:S96. [doi: 10.1016/S0140-6736(23)02148-7] [Medline: 37997143]
21. McCann H, Myers K. Addressing the silence: utilising salon workers to respond to family violence. J Sociol (Melb) 2023 Mar;59(1):104-119. [doi: 10.1177/14407833211031005]
22. Boivin JM, Risse J, Laurière E, Burnier M. Screening for hypertension at the hairdresser: a feasibility study in France and Morocco. Blood Press 2020 Aug;29(4):202-208. [doi: 10.1080/08037051.2020.1728189] [Medline: 32054323]
23. Chiaborelli M, Kopeka M, Sekhesa P, et al. Hair salons as a promising space to provide HIV and sexual and reproductive health services for young women in Lesotho: a citizen scientist mixed-methods study. BMC Public Health 2025 Jul 2;25(1):2202. [doi: 10.1186/s12889-025-23435-3] [Medline: 40604538]
24. Makabe S, Fujiwara K, Kume Y, et al. Evaluating a frailty education program implemented through barbershops/salons in Japan: a preliminary study. SN Soc Sci 2022;2(5):55. [doi: 10.1007/s43545-022-00361-4] [Medline: 35499069]
25. Horinishi Y, Horinishi R, Sano C, Ohta R. The role of beauty salons in community health promotion: exploring health engagement and social connectivity in Japanese hair salons. Cureus 2025 Feb;17(2):e79511. [doi: 10.7759/cureus.79511] [Medline: 40135027]
26. National nutrition survey 2010 report. : Health Promotion Board URL: https://data.gov.sg/datasets/d_091a7ef168518786f5dd83d9f56a25a0/view [accessed 2026-04-24]
27. Radomskaya V, Bhati AS. Hawker centres: a social space approach to promoting community wellbeing. Urban Plan 2022;7(4). [doi: 10.17645/up.v7i4.5658]
28. Healthier Dining Programme. Health Promotion Board Singapore. URL: https://www.hpb.gov.sg/healthy-living/food-beverage/healthier-dining-programme [accessed 2025-07-31]
29. Clean Tables Campaign. National Environmental Agency. URL: https://www.nea.gov.sg/our-services/public-cleanliness/clean-tables/clean-tables-campaign [accessed 2025-07-31]

XSL•FO
RenderX

30. Raguraman A. Hawkers, postmen, shopkeepers to keep an eye on the elderly in Cheng San-Seletar. The Straits Times. 2023 Jan 8. URL: https://www.straitstimes.com/singapore/hawkers-postmen-shopkeepers-to-keep-an-eye-on-the-elderly-in-cheng-san-seletar [accessed 2025-05-31]
31. Siegrist M, Gutscher H, Earle TC. Perception of risk: the influence of general trust, and general confidence. J Risk Res 2005 Mar;8(2):145-156. [doi: 10.1080/1366987032000105315]
32. Yamagishi T, Cook KS. Generalized exchange and social dilemmas. Soc Psychol Q 1993 Dec;56(4):235-248. [doi: 10.2307/2786661]
33. van der Hulst FJP, Brabers AEM, de Jong JD. The relation between trust and the willingness of enrollees to receive healthcare advice from their health insurer. BMC Health Serv Res 2023 Jan 18;23(1):52. [doi: 10.1186/s12913-022-09016-9] [Medline: 36653840]
34. Dutta-Bergman MJ. Primary sources of health information: comparisons in the domain of health attitudes, health cognitions, and health behaviors. Health Commun 2004;16(3):273-288. [doi: 10.1207/S15327027HC1603_1] [Medline: 15265751]
35. Pelikan J, Ganahl K, Van den Broucke S, Sørensen K. Measuring health literacy in Europe: introducing the European Health Literacy Survey Questionnaire (HLS-EU-Q). In: Okan O, Bauer U, Pinheiro P, Levin-Zamir D, Sørensen K, editors. International Handbook of Health Literacy: Research, Practice and Policy across the Life-Span: Policy Press; 2019:115-138.
36. Guo H, Lim HY, Chow A. Health information orientation profiles and their association with knowledge of antibiotic use in a population with good internet access: a cross-sectional study. Antibiotics (Basel) 2022 Jun 4;11(6):769. [doi: 10.3390/antibiotics11060769] [Medline: 35740175]
37. Cutilli CC. Seeking health information: what sources do your patients use? Orthop Nurs 2010;29(3):214-219. [doi: 10.1097/NOR.0b013e3181db5471] [Medline: 20505493]
38. Basu A, Dutta MJ. The relationship between health information seeking and community participation: the roles of health information orientation and efficacy. Health Commun 2008;23(1):70-79. [doi: 10.1080/10410230701807121] [Medline: 18443994]
39. Sillence E, Briggs P, Harris PR, Fishwick L. How do patients evaluate and make use of online health information? Soc Sci Med 2007 May;64(9):1853-1862. [doi: 10.1016/j.socscimed.2007.01.012] [Medline: 17328998]
40. Birkhäuer J, Gaab J, Kossowsky J, et al. Trust in the health care professional and health outcome: a meta-analysis. PLoS One 2017;12(2):e0170988. [doi: 10.1371/journal.pone.0170988] [Medline: 28170443]
41. Cropanzano R, Mitchell MS. Social exchange theory: an interdisciplinary review. J Manage 2005 Dec;31(6):874-900. [doi: 10.1177/0149206305279602]
42. Ahmad R, Nawaz MR, Ishaq MI, Khan MM, Ashraf HA. Social exchange theory: systematic review and future directions. Front Psychol 2023;13:1015921. [doi: 10.3389/fpsyg.2022.1015921] [Medline: 36710813]
43. Brabers AE, de Jong JD, Groenewegen PP, van Dijk L. Social support plays a role in the attitude that people have towards taking an active role in medical decision-making. BMC Health Serv Res 2016 Sep 21;16(1):502. [doi: 10.1186/s12913-016-1767-x] [Medline: 27655113]
44. Umberson D, Montez JK. Social relationships and health: a flashpoint for health policy. J Health Soc Behav 2010;51 Suppl(Suppl):S54-S66. [doi: 10.1177/0022146510383501] [Medline: 20943583]
45. Mognard E, Naidoo K, Laporte C, et al. “Eating out”, spatiality, temporality and sociality. A database for China, Indonesia, Japan, Malaysia, Singapore and France. Front Nutr 2023;10:1066737. [doi: 10.3389/fnut.2023.1066737] [Medline: 36819687]

Abbreviations

AOR: adjusted odds ratio

Edited by A Mavragani, T Sanchez; submitted 27.Oct.2025; peer-reviewed by P Dootson, S Eshun; revised version received 13.Apr.2026; accepted 15.Apr.2026; published 14.May.2026.

Please cite as:
Aithal S, Guo H, Hildon ZLJ, Lwin MO, Chow A
Determinants of Willingness to Receive Health Information From Neighborhood Food and Beauty Establishments: Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e86435
URL: https://publichealth.jmir.org/2026/1/e86435
doi:10.2196/86435

© Seema Aithal, Huiling Guo, Zoe Lara-Jane Hildon, May O Lwin, Angela Chow. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 14.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution,

XSL•FO
RenderX

and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Willingness to Undergo Human Papillomavirus Testing Among Men Who Have Sex With Men in China Based on the Information-Motivation-Behavioral Skills Model: Online Cross-Sectional Study

Xin Liu[1,2*], MPH; Xuelin Yang[1*], BSc; Mohan Lyu[3], MPH; Zhenwei Dai[4], MPH; Shu Jing[1], BSc; You Xin[1], BA; Fei Yu[5], MSW; Shenglan Tang[3], PhD; Xiaoyou Su[1,6], PhD

[1]School of Population Medicine and Public Health, Chinese Academy of Medical Sciences & Peking Union Medical College, No. 9, Dongdan 3rd Street, Dongcheng District, Beijing, Beijing, China

[2]MRC Centre for Global Infectious Disease Analysis, Jameel Institute, School of Public Health, Imperial College London, London, United Kingdom

[3]Duke Global Health Institute, Duke University, Durham, NC, United States

[4]Peking University Sixth Hospital, Peking University Institute of Mental Health, NHC Key Laboratory of Mental Health (Peking University), National Clinical Research Center for Mental Disorders (Peking University Sixth Hospital), Beijing, China

[5]Danlan Public Welfare, Beijing, China

[6]State Key Laboratory of Respiratory Health and Multimorbidity, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Xiaoyou Su, PhD
School of Population Medicine and Public Health, Chinese Academy of Medical Sciences & Peking Union Medical College, No. 9, Dongdan 3rd Street, Dongcheng District, Beijing, Beijing, China

Abstract

Background: Men who have sex with men (MSM) are at high risk of human papillomavirus (HPV) infection, and HPV testing can facilitate early detection and timely intervention. However, evidence on the willingness to undergo different HPV testing modalities among MSM remains limited. The information-motivation-behavioral skills model provides a theoretical framework for understanding factors associated with the willingness to undergo HPV testing.

Objective: This study aimed to investigate the current status of HPV testing and screening among MSM in China; assess their willingness to undergo professional-collected HPV testing, self-collected HPV testing, and regular HPV testing; and identify associated factors.

Methods: A cross-sectional study was conducted through the Blued application (Danlan Public Welfare) from July 15, 2024, to August 2, 2024. The questionnaire collected information on participants' sociodemographic characteristics, sexual behaviors, willingness to undergo HPV testing, HPV testing and infection history, HPV knowledge level, motivation, and behavioral skills related to HPV testing. Chi-square tests and 2-tailed *t* tests were used to compare the distribution of variables between groups. Multivariable logistic regression analyses were used to identify factors associated with willingness to undergo each HPV testing modality among MSM.

Results: Among the 1080 participants recruited, 88.3% (n=954) were willing to receive professional-collected HPV testing, 83.6% (n=903) were willing to receive self-collected HPV testing, and 78.7% (n=850) were willing to engage in regular HPV testing. In multivariable analyses, willingness to undergo professional-collected HPV testing was positively associated with higher HPV knowledge (odds ratio [OR] 1.04, 95% CI 1.00‐1.08), greater perceived susceptibility (OR 1.24, 95% CI 1.01‐1.53), and higher self-efficacy (OR 1.69, 95% CI 1.14‐2.50), while it was negatively associated with perceived barriers (OR 0.61, 95% CI 0.47‐0.78). Participants who were retired or unemployed (OR 0.48, 95% CI 0.27‐0.83) and those with prior testing experience (OR 0.49, 95% CI 0.33‐0.73) were less likely to undergo self-collected HPV testing. Moreover, willingness to undergo regular HPV testing was positively associated with perceived susceptibility (OR 1.20, 95% CI 1.02‐1.40), whereas living with a partner (OR 0.54, 95% CI 0.33‐0.88) and living with family members (OR 0.69, 95% CI 0.48‐0.99) were negatively associated.

Conclusions: This study suggests that willingness to undergo HPV testing among MSM in China is high, whereas prior HPV testing uptake remains low, indicating a gap between testing intention and behavior. By applying the information-motivation-behavioral skills framework to different HPV testing modalities, this study extends the previous evidence from testing willingness to understanding of testing initiation and sustained HPV screening behavior. These findings suggest that

XSL•FO
RenderX

personalized health education initiatives and differentiated, privacy-sensitive, and accessible HPV testing strategies may help improve HPV testing uptake and support sustained HPV testing among MSM in China.

(JMIR Public Health Surveill 2026;12:e85543) doi:10.2196/85543

KEYWORDS

men who have sex with men; human papillomavirus; testing willingness; information-motivation-behavioral skills; associated factors

Introduction

Human papillomavirus (HPV) is one of the most common sexually transmitted infections (STIs), with over 200 different types [1,2]. According to a global meta-analysis, the prevalence of HPV infection among males was estimated to be 31%, peaking in the 25 - 29 years age group and remaining at a high level before the age of 50 [3]. High-risk HPV types, which are responsible for malignancies such as cervical cancer, penile cancer, and anal cancer, accounted for 21% of cases. HPV-16 was identified as the most prevalent genotype (5%), followed by HPV-6 (4%) [3]. Although HPV-6 is classified as a low-risk HPV type, it can still cause benign lesions such as genital warts. Men who have sex with men (MSM) bear a disproportionately high burden of HPV infection due to their sexual networks and behavioral characteristics. Previous research showed that the prevalence of anal high-risk HPV was 41.2% among MSM, compared to only 6.9% in other male populations [4]. Another meta-analysis of 107 international studies, including 36,773 MSM, reported pooled prevalences for anal, oral, and urethral HPV infections at 78.4%, 36.2%, and 17.3%, respectively [5].

Screening and vaccination are central strategies for preventing HPV infection and related diseases among MSM. Previous research has shown that HPV vaccines provide long-term protection, and vaccination among MSM has been reported to be cost-effective in multiple settings [6,7]. Accordingly, several countries have implemented targeted HPV vaccination programs for MSM [8,9]. In January 2025, the quadrivalent HPV vaccine (Gardasil) developed by Merck was approved in Chinese mainland for males aged 9 to 26 years, making it the first HPV vaccine indication for males in Chinese mainland [10]. This provided new opportunities for HPV prevention and control among MSM in China. However, in the short term, achieving high vaccine coverage among MSM may remain challenging due to limited vaccine supply, vaccine hesitancy, and cost-related factors [11]. Additionally, due to the sexual behavioral patterns of MSM, the herd immunity established by female HPV vaccination may provide limited protection for MSM. Moreover, some MSM also engage in sexual activity with women, which may contribute to HPV transmission to their female partners and thus have implications for cervical cancer elimination efforts in China.

HPV testing can help MSM understand their infection status, enabling them to take timely actions to reduce HPV transmission, facilitate early detection of persistent infections or lesions, and thereby reduce the risk of HPV-related diseases [12-14]. Several approaches are available for detecting HPV infections and associated lesions among MSM, including visual inspection, HPV testing, anal cytology, digital anorectal examination, and high-resolution anoscopy [15,16]. In some developed countries, regular anal screening is recommended for HIV-positive MSM to detect precancerous lesions at an early stage [17,18]. Therefore, improving HPV awareness and promoting proactive HPV testing remain critical for controlling HPV transmission among MSM. Even if HPV vaccination coverage expands in the future, regular testing and screening among MSM will still be necessary to facilitate early detection and treatment of HPV-related lesions and to reduce the disease burden.

However, implementing HPV testing and screening among MSM still faces multiple challenges. A systematic review reported that the uptake of anal cancer screening among MSM was relatively low, and limited knowledge of HPV-related risks was identified as a significant barrier, whereas greater perceived understanding was the primary facilitator [19]. Additionally, stigma and embarrassment, fear of abnormal results, and concerns about disclosure of sexual orientation in health care settings have been identified as barriers to undergoing HPV screening [20]. In China, although some institutions and community-based organizations have initiated HPV testing and screening programs for MSM, supportive programs and research evidence on HPV testing and anal screening for MSM remain limited [21]. Meanwhile, advances in HPV testing technologies have facilitated the development of self-collected HPV testing, which is considered a potential strategy for improving testing coverage among MSM due to its greater privacy, convenience, and high acceptability [22]. Therefore, it is necessary to investigate the willingness to undergo HPV testing and regular testing among MSM, to better understand their behavioral patterns and identify modifiable determinants, providing evidence for developing more targeted health education and service strategies.

This study uses the information-motivation-behavioral skills (IMB) model as its theoretical framework [23]. The IMB model categorizes behavior-influencing factors into 3 components: information (knowledge and awareness related to diseases and health behaviors), motivation (the willingness to adopt health behaviors based on personal factors and social support), and behavioral skills (the ability and skills necessary to change health behaviors). According to the IMB model, individuals with sufficient health information and motivation can initiate and maintain health behaviors through behavioral skills. The IMB model has been widely applied in research and interventions related to STIs [24-26]. A previous study used a modified IMB model to identify predictors of condom use among MSM in China, providing evidence for interventions to promote condom use [27]. Some researchers have combined mobile reminder tools with the IMB model, such as medication

XSL•FO
RenderX

reminders and risk feedback, to promote pre-exposure prophylaxis adherence among MSM [28]. Another study applied an online intervention based on the IMB model to improve awareness of and willingness to receive HPV vaccination [29]. Guided by the IMB model, this study aimed to investigate the current status of HPV testing and screening among MSM in China, assess willingness to undergo professional-collected HPV testing, self-collected HPV testing, and regular HPV testing, and identify associated factors.

Methods

Study Design and Participants

This study was an internet-based cross-sectional survey conducted from July 15, 2024, to August 2, 2024, using the Blued platform (Supplier: Danlan public welfare), one of the most popular social networking apps for MSM in China. The technical team randomly selected 300,000 users nationwide who had logged into Blued within the past month and sent them a recruitment poster. Interested users could click the link or scan the QR code on the poster to access the questionnaire. All participants were required to provide informed consent before completing the questionnaire. The inclusion criteria were as follows: (1) aged 18 years and older; (2) male at birth; (3) had anal intercourse with a male partner in the past 6 months; and (4) consented to participate in this study. Each IP address was allowed to submit the questionnaire only once, and all items were required to be completed before submission.

Study Size

The sample size was estimated using the formula for cross-section studies (Eq. 1). Based on our pilot survey, the willingness to undergo HPV testing among MSM was 95% (57/60). With a 95% CI and an allowable absolute error set at 1.5%, the required sample size was 811. Allowing for an anticipated 20% invalid response rate, the minimum required sample size was 1014.

$$(1)n=Z1-\alpha/22p(1-p)d2$$

Here n is the sample size, Z is the z-score (the critical value from the standard normal distribution), p is the anticipated or estimated population proportion, and d is precision or margin of error.

Measurements

The survey questionnaire used in this study was designed by a panel consisting of epidemiologists, physicians, statisticians, nongovernmental organization workers, and linguists, based on the findings of prior research and preliminary studies [24,30]. The questionnaire development followed a 4-step process. First, a targeted review of previous studies on HPV testing, HPV vaccination, and the IMB framework was conducted to identify candidate items. These items were adapted and translated to form an initial item pool. Second, we recruited 28 MSM and conducted semistructured interviews guided by the IMB model. Themes derived from the qualitative analysis were used to refine the constructs and revise item wording. Third, the draft questionnaire was reviewed by the panel to assess content relevance, clarity, and scientific accuracy. Fourth, a pilot survey was conducted among MSM to evaluate feasibility, comprehensibility, and skip logic, and the questionnaire was finalized accordingly.

The questionnaire included sections on sociodemographic information, sexual behavioral characteristics, health status, the constructs of the IMB model, including scales and items assessing the HPV-related knowledge, motivation, and behavioral skills regarding HPV testing, and MSM's willingness to undergo HPV testing. In total, 1094 questionnaires were completed. After quality control, 14 questionnaires were excluded, leaving 1080 valid questionnaires for analysis. Exclusions were made for the following reasons: response time of less than 200 seconds (n=10), residing outside of China in the past 6 months (n=3), and inconsistent or nonsensical responses (n=1).

Sociodemographic Characteristics

We collected sociodemographic information on participants, including age, gender at birth, gender identity, sexual orientation, city of residence, educational level, current employment status, monthly personal income, and cohabitation status. Sexual behavioral characteristics were also collected, including sexual role, number of sexual partners, whether they had engaged in anal sex, oral sex, commercial sex, or group sex in the past 6 months, and whether they had sex with regular partners, temporary partners, or female partners in the past 6 months. Health-related variables included circumcision status, history of HIV testing, HIV status, history of STIs, and history of HPV testing.

Willingness to Undergo HPV Testing

Participants were first introduced to professional-collected HPV testing. "*Professional-collected HPV testing requires visiting a testing site (such as a hospital), where a professional uses a swab to collect samples from the anus, coronal sulcus (glans penis), and/or oral cavity.*" Participants were then asked whether they would be willing to undergo professional-collected HPV testing if the test were free. Subsequently, participants were introduced to self-collected HPV testing. "*Self-collected HPV testing requires users to obtain a testing kit online, follow the instructions to collect samples from the anus, coronal sulcus (glans penis), and/or oral cavity by themselves, and then send the samples in a preservation tube to a designated lab for testing.*" Participants were then asked whether they would be willing to undergo self-collected HPV testing if it were free. Among MSM who were willing to undergo at least 1 type of HPV testing, further inquiries were made regarding their reasons for acceptance. Moreover, all participants were asked about their preferred testing frequency if testing were needed (response options: 1=every 6 months, 2=annually, 3=only when necessary, and 4=other). Participants who chose every 6 months, annually, or any other specific frequency were classified as having an intention for regular HPV testing. MSM who were unwilling to undergo either type of HPV testing were asked about their reasons for refusal.

Information of HPV Testing

HPV knowledge among MSM was assessed using a scale developed by Waller [31] in 2013. This scale includes general

XSL•FO
RenderX

HPV knowledge subscale (GK), HPV testing knowledge subscale (TK), and HPV vaccine knowledge subscale (VK), which assess individuals' knowledge of HPV. It has been applied in Chinese populations and has demonstrated good reliability among college students [32]. In this study, we used the GK including 16 items to measure participants' general understanding of HPV and TK consisting of 6 items to assess knowledge related to HPV testing. Each item offers 3 response options: "True," "False," and "Don't know". Correct answers were scored as 1, whereas incorrect or "Don't know" responses were scored as 0. Higher scores indicated greater knowledge level of HPV. In this study, the Cronbach α for GK and TK were 0.927 and 0.839, respectively.

Motivation for HPV Testing

Motivation for HPV testing was assessed using 17 items across 5 domains, adapted from previous studies and our unpublished qualitative study (Table S1 in Multimedia Appendix 1) [24,33]. Perceived susceptibility was assessed by 3 items evaluating the perceived risk of HPV infection and related diseases among participants. Perceived severity was measured using 3 items that assessed participants' perceived seriousness of HPV-related diseases. Perceived benefits were evaluated through 3 items that gauged participants' perceived health benefits of HPV testing. Perceived barriers included 5 items assessing perceived difficulties that might be encountered during decision-making or action-taking processes. Subjective norms were assessed by 3 items measuring perceived support from family and peers for the participants' decision to undergo HPV testing. All items were rated on a 5-point Likert scale (1=Strongly disagree, 2=Disagree, 3=Neutral, 4=Agree, and 5=Strongly agree). Higher scores indicated stronger motivation for HPV testing, except for perceived barriers, for which higher scores indicated greater perceived barriers. In this study, the Cronbach α of motivation for HPV testing was 0.850.

Behavioral Skills for HPV Testing

Behavioral skills for HPV testing were assessed through 5 items across 3 domains, based on previous studies and our unpublished qualitative study (Table S1 in Multimedia Appendix 1) [24,34]. Decision-making was assessed by 1 item evaluating participants' agency in deciding whether to undergo HPV testing. Self-efficacy contained 2 items assessing participants' confidence in overcoming potential difficulties to take action. Objective skills consisted of 2 items evaluating participants' skills and abilities necessary for undergoing HPV testing. All items were rated on a 5-point Likert scale (1=Strongly disagree, 2=Disagree, 3=Neutral, 4=Agree, and 5=Strongly agree). Higher scores indicated greater behavioral skills for HPV testing. In this study, the Cronbach α of behavioral skills for HPV testing was 0.884.

Bias

Several measures were taken to address potential sources of bias. The survey was administered anonymously, which may have reduced social desirability bias when participants reported sensitive sexual behaviors and STI history. For questions on sexual behaviors, a 6-month recall period was used to improve recall accuracy. Explanations of professional-collected and self-collected HPV testing were provided to reduce differences in participants' understanding of the testing procedures. All questionnaire items were mandatory to ensure completeness of responses and avoid missing data. No financial compensation was provided to participants, which may have reduced participation driven by monetary incentives. However, because this was a cross-sectional survey based on self-reported data, bias could not be fully eliminated.

Statistical Analysis

In this study, the distribution of variables between participants who were willing and unwilling to undergo the HPV testing was reported using means and SDs for continuous variables and frequencies and percentages (%) for categorical variables. Chi-squared tests and 2-tailed *t* tests were used to compare differences in variable distributions between groups. To reduce the risk of excluding potentially important predictors, variables with a *P* value of <.10 in the univariate analysis, as well as variables with professional relevance, were included in the multivariable logistic regression analysis. Multivariable logistic regressions using the Enter method, in which all selected covariates were entered into the model simultaneously, were used to identify factors associated with willingness to undergo HPV testing among MSM. Adjusted odds ratios [OR] and their corresponding 95% CI were calculated. Multicollinearity was assessed using the variance inflation factor, and all variance inflation factor values were <5, indicating no significant multicollinearity in the model. Statistical significance was assessed using 2-sided tests with a *P* value <.05 for all analyses.

Ethical Considerations

This study received ethics approval from the institutional ethics committee of the Chinese Academy of Medical Sciences and Peking Union Medical College (CAMS&PUMC-IEC-2022-078). The study followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines [35]. Digital informed consent was obtained from all participants. All data used in this study were deidentified, and no participants were identifiable in any images included in the manuscript or supplementary material. Participants received no compensation.

Results

Sociodemographic Characteristics of Participants

This study recruited 1080 participants, with a mean age of 31.14 (SD 9.45) years. Most of them were identified as male participants (n=1002, 92.8%), homosexual (n=825, 76.4%), were employed full-time (n=840, 77.8%), and lived in the northern (n=258, 23.9%) and eastern (n=258, 23.9%) regions of China. Among the participants, 34.1% (n=368) of MSM reported monthly incomes ranging from CN ¥3000 to ¥6999 (US $435-$1015), and 42.6% (n=460) lived alone (Table 1).

XSL•FO
RenderX

Table . Baseline characteristics of men who have sex with men (MSM) recruited online in China, 2024.

Variables	Overall (N=1080)
Age (y), mean (SD)	31.14 (9.45)
Gender identity, n (%)	
Male	1002 (92.8)
Transgender woman	78 (7.2)
Region, n (%)	
Northeast	73 (6.7)
North	258 (23.9)
Central	135 (12.5)
East	258 (23.9)
South	177 (16.4)
Southwest	125 (11.6)
Northwest	54 (5)
Education level, n (%)	
High school and below	210 (19.4)
High school above	870 (80.6)
Current employment status, n (%)	
Full-time	840 (77.8)
Student	163 (15.1)
Unemployed or retired	77 (7.1)
Personal monthly income (CN ¥; CN ¥1=US $0.14)	
<3000	232 (21.5)
3000 - 6999	368 (34.1)
7000 - 9999	191 (17.7)
≥10,000	289 (26.7)
Cohabitation status, n (%)	
Alone	460 (42.6)
Partner	118 (10.9)
Family	318 (29.5)
Friends or roommates	172 (15.9)
Others	12 (1.1)
Sexual orientation, n (%)	
Homosexuality	825 (76.4)
Bisexuality	190 (17.6)
Heterosexual and others	65 (6)
Sex role, n (%)	
Receptive	366 (33.9)
Versatile	393 (36.4)
Insertive	321 (29.7)
Number of sexual partners, n (%)	
1	354 (32.8)
≥2	726 (67.2)
Condom usage frequency during anal sex, n (%)	

XSL•FO
RenderX

Variables	Overall (N=1080)
Every time	531 (49.2)
Often use	210 (19.4)
Occasionally use	252 (23.3)
Never use	87 (8.1)
Performed oral sex in the past 6 mo, n (%)	
No	250 (23.1)
Yes	830 (76.9)
Received oral sex in the past 6 mo, n (%)	
No	423 (39.2)
Yes	657 (60.8)
Condom use frequency during oral sex, n (%)	
Every time	33 (3.1)
Often use	12 (1.1)
Occasionally use	65 (6)
Never use	926 (85.7)
No oral sex before	44 (4.1)
Had sex with regular partners in the past 6 mo, n (%)	
No	464 (43)
Yes	616 (57)
Had sex with temporary partners in the past 6 mo, n (%)	
No	465 (43.1)
Yes	615 (56.9)
Had commercial sex in the past 6 mo, n (%)	
No	1013 (93.8)
Yes	67 (6.2)
Had group sex in the past 6 mo, n (%)	
No	897 (83.1)
Yes	183 (16.9)
Had sex with female in the past 6 mo, n (%)	
No	962 (89.1)
Yes	118 (10.9)
Had a circumcision, n (%)	
No	861 (79.7)
Yes	219 (20.3)
Self-reported HIV status, n (%)	
Positive	105 (9.7)
Negative	790 (73.2)
Unknown	185 (17.1)
History of STIs[a], n (%)	
No	702 (65)
Yes	378 (35)
Information (score), mean (SD)	

XSL•FO
RenderX

Variables	Overall (N=1080)
Knowledge of HPV[b]	13.45 (5.40)
Motivation (score), mean (SD)	
Perceived susceptibility	3.23 (1.02)
Perceived severity	3.79 (0.99)
Perceived benefits	4.24 (0.82)
Perceived barriers	3.33 (0.94)
Subjective norms	3.44 (0.78)
Behavioral skills (score), mean (SD)	
Decision-making	4.11 (0.88)
Self-efficacy	3.98 (0.82)
Objective skills	3.83 (0.89)

[a]STI: sexually transmitted infection.

[b]HPV: human papillomavirus.

Characteristics of Sexual Behaviors and Health Status

In the past 6 months, all 1080 participants had engaged in anal sex with men: 36.4% (n=393) of them were versatile, 49.2% (n=531) consistently used condoms, and 95.9% (n=1036) had engaged in oral sex. Additionally, 76.9% (n=830) reported that they had performed oral sex and 60.8% (n=657) had received oral sex in the past 6 months. However, only 3.1% (n=33) reported using condoms during oral sex. Furthermore, 57% (n=616) of the MSM had sex with regular partners, 56.9% (n=615) with temporary partners, 6.2% (67/1080) had commercial sex, 10.9% (n=118) had sex with females, 16.9% (n=183) had group sex, and 67.2% (n=726) had 2 or more sexual partners. Additionally, 20.3% (n=219) had undergone circumcision, 9.7% (n=105) reported being HIV-positive, and 35% (n=378) had a history of STIs (Table 1).

Dimensions of the IMB Model

In this study, participants had a mean HPV knowledge score of 13.45 (SD 5.40). Regarding motivation for HPV testing, the mean scores were 3.23 (SD 1.02) for perceived susceptibility, 3.79 (SD 0.99) for perceived severity, 4.24 (SD 0.82) for perceived benefits, 3.33 (SD 0.94) for perceived barriers, and 3.44 (SD 0.78) for subjective norms. In terms of behavioral skills for HPV testing, the mean scores were 4.11 (SD 0.88) for decision-making, 3.98 (SD 0.82) for self-efficacy, and 3.83 (SD 0.89) for objective skills (Table 1).

Characteristics and Reasons for HPV Screening Among MSM

Among participants, 35.2% (380/1080) had previously undergone HPV screening and 23.3% (252/1080) had undergone HPV testing, with 43.3% of them (109/252) testing positive. Additionally, 17.8% (192/1080) had undergone an acetic acid test, 16.9% (182/1080) had undergone anoscopy, 5.5% (59/1080) had undergone anal cytology, and 22.6% (244/1080) reported a history of genital warts. This study investigated the reasons for HPV screening among participants with an HPV screening history. The most common reasons were "*appearance or suspected appearance of HPV-related symptoms*" (224/380, 58.9%), "*voluntary proactive testing*" (211/380, 55.5%), "*engagement in high-risk sexual behavior*" (96/380, 25.3%), "*recommendation by healthcare professionals*" (57/380, 15%), "*had sex with an HPV-positive individual*" (33/380, 8.7%), and "*participation in an HPV vaccine clinical trial*" (13/380, 3.4%) (Table 2 and Table S2 in Multimedia Appendix 1).

XSL•FO
RenderX

Table . HPV[a] testing history and HPV infection status among men who have sex with men (MSM) recruited online in China, 2024.

Variables	Total (N=1080)	HIV status	
		Negative or unknown (n=975)	Positive (n=105)
Ever received any HPV-related examination, n (%)			
No	700 (64.8)	639 (65.5)	61 (58.1)
Yes	380 (35.2)	336 (34.5)	44 (41.9)
Ever received HPV DNA testing, n (%)			
No	828 (76.7)	751 (77)	77 (73.3)
Yes	252 (23.3)	224 (23)	28 (26.7)
Ever received acetic acid test, n (%)			
No	888 (82.2)	806 (82.7)	82 (78.1)
Yes	192 (17.8)	169 (17.3)	23 (21.9)
Ever received anoscopy, n (%)			
No	898 (83.1)	816 (83.7)	82 (78.1)
Yes	182 (16.9)	159 (16.3)	23 (21.9)
Ever received anal cytology, n (%)			
No	1021 (94.5)	921 (94.5)	100 (95.2)
Yes	59 (5.5)	54 (5.5)	5 (4.8)
HPV test result (n=252), n (%)			
Positive	109 (43.3)	93 (41.5)	16 (57.1)
Negative	143 (56.7)	131 (58.5)	12 (42.9)
History of genital warts, n (%)			
No	836 (77.4)	779 (79.9)	57 (54.3)
Yes	244 (22.6)	196 (20.1)	48 (45.7)

[a]HPV: human papillomavirus.

Willingness to Undergo HPV Testing and Regular HPV Testing

The overall willingness to undergo HPV testing in this study was 97% (1048/1080). Of the participants, 74.9% (n=809) were willing to undergo both types of HPV testing, 13.4% (n=145) were willing to undergo only the professional-collected HPV testing, and 8.7% (n=94) were willing to undergo only the self-collected testing. The willingness to undergo the professional-collected HPV testing was 88.3% (n=954), the willingness to undergo self-collected HPV testing was 83.6% (n=903), and the willingness to engage in regular HPV testing was 78.7% (n=850).

Reasons for Unwillingness to Undergo HPV Testing

In this study, among 1080 participants, 32 (3%) expressed their unwillingness to undergo either type of HPV testing. The 3 most common reasons were as follows: "*I am reluctant to expose my genital area to others*" (22/32, 68.8%); "*I am concerned that sampling myself may affect the accuracy of the results*" (20/32, 62.5%); and "*I am concerned about the leakage of my private information*" (18/32, 56.3%). Additionally, 94 (8.7%) participants were unwilling to undergo the professional-collected HPV testing. The 3 most common reasons were as follows: "*I am reluctant to expose my genital area to others*" (65/94, 69.1%); "*I am concerned about the leakage of my private information*" (61/94, 64.9%); and "*Undergoing HPV testing may bring negative comments to me*" (45/94, 47.9%). Furthermore, 145 (13.4%) participants expressed unwillingness to undergo self-collected HPV testing, with the 3 most common reasons being as follows: "*I am concerned that sampling myself may affect the accuracy of the results*" (119/145, 82.1%), "*I suspect the accuracy of self-collected HPV testing*" (97/145, 66.9%), and "*I think the testing process is complicated*" (62/145, 42.8%) (Table S3 in Multimedia Appendix 1).

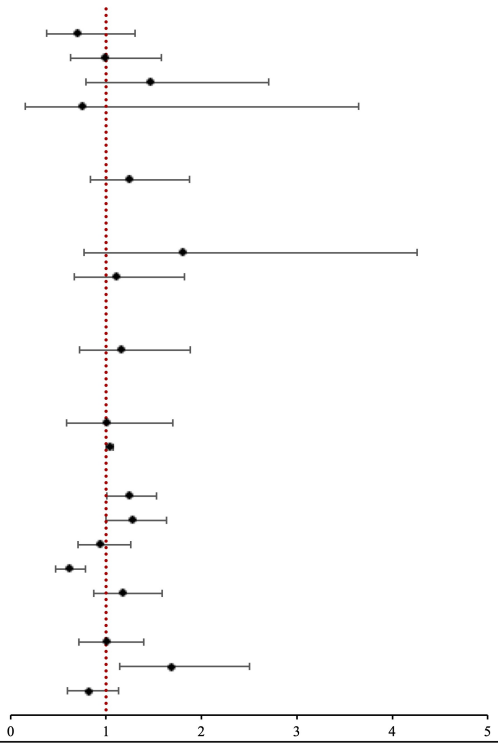
Factors Associated With Willingness to Undergo Professional-Collected HPV Testing

The factors associated with willingness to undergo professional-collected HPV testing are presented in Table S4 of Multimedia Appendix 1. The multivariable logistic regression analysis showed that participants with greater HPV knowledge (OR 1.04, 95% CI 1.00 - 1.08), perceived susceptibility (OR 1.24, 95% CI 1.01 - 1.53), and self-efficacy (OR 1.69, 95% CI 1.14 - 2.50) had a higher intention to undergo the professional-collected HPV test. Conversely, those with higher perceived barriers (OR 0.61, 95% CI 0.47 - 0.78) were less likely to be willing to undergo the professional-collected HPV test (Figure 1).

XSL•FO
RenderX

Figure 1. Factors influencing willingness for professional-collected human papillomavirus (HPV) testing among men who have sex with men (MSM) in China (N=1080). OR: odds ratio; STI: sexually transmitted infection; VIF: variance inflation factor.

Variables	OR (95% CI)	*P* value	VIF
Cohabitation status			1.02
Alone	1	.46	
Partner	0.70 (0.38-1.30)	.26	
Family	0.99 (0.63-1.58)	.99	
Friends or roommates	1.46 (0.79-2.71)	.23	
Others	0.75 (0.15-3.65)	.72	
Number of sexual partners			1.03
1	1		
≥2	1.25 (0.84-1.88)	.28	
Self-reported HIV status			1.06
Positive	1	.39	
Negative	1.81 (0.77-4.26)	.17	
Unknown	1.10 (0.67-1.82)	.70	
History of STI			1.18
No	1		
Yes	1.17 (0.72-1.88)	.53	
History of HPV testing			1.18
No	1		
Yes	0.99 (0.59-1.70)	.99	
Information (knowledge of HPV)	1.04 (1.00-1.08)	*.04*	1.20
Motivation			
Perceived susceptibility	1.24 (1.01-1.53)	*.04*	1.28
Perceived severity	1.28 (1.00-1.63)	.05	1.60
Perceived benefits	0.94 (0.71-1.26)	.69	1.79
Perceived barriers	0.61 (0.47-0.78)	*<.001*	1.27
Subjective norms	1.18 (0.87-1.59)	.29	1.43
Behavioral skills			
Decision-making	1.00 (0.72-1.40)	.99	2.78
Self-efficacy	1.69 (1.14-2.50)	*.009*	3.06
Objective skills	0.82 (0.60-1.13)	.23	2.09



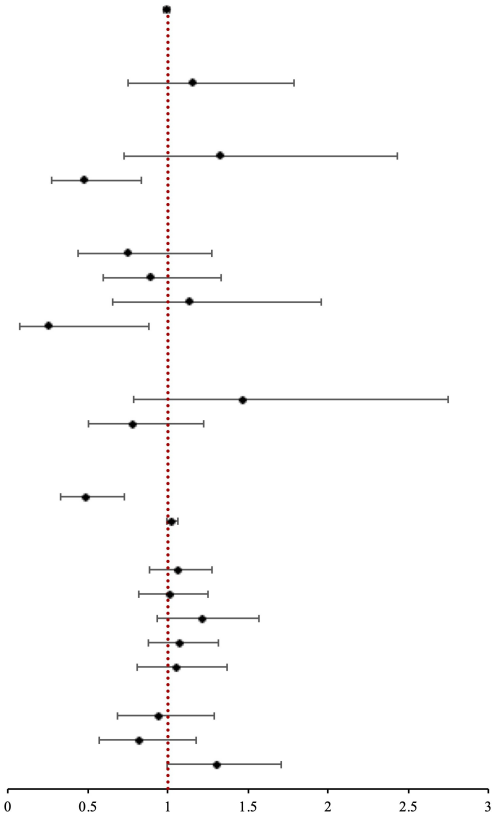
Factors Associated With Willingness to Undergo Self-Collected HPV Testing

The factors associated with the willingness to undergo self-collected HPV testing are presented in Table S5 of Multimedia Appendix 1. The multivariable logistic regression analysis showed that retired and unemployed participants (OR 0.48, 95% CI 0.27 - 0.83) were less likely to be willing to undergo the self-collected HPV test compared with those with full-time jobs. Additionally, compared with MSM who had never undergone HPV testing, those who had previously undergone testing were less willing to undergo self-collected testing (OR 0.49, 95% CI 0.33 - 0.73) (Figure 2).

XSL•FO
RenderX

Figure 2. Factors influencing willingness for self-collected human papillomavirus (HPV) testing among men who have sex with men in China (N=1080). OR: odds ratio; VIF: variance inflation factor.

Variables	OR (95% CI)	*P* value	VIF
Age (y)	0.99 (0.97-1.01)	.40	1.17
Education level			1.18
High school and below	1		
High school above	1.16 (0.75-1.79)	.51	
Current employment status			1.10
Full-time	1		
Student	1.33 (0.73-2.43)	.36	
Unemployed or retired	0.48 (0.27-0.83)	*.009*	
Cohabitation status			1.04
Alone	1		
Partner	0.75 (0.44-1.27)	.29	
Family	0.89 (0.60-1.33)	.58	
Friends or roommates	1.13 (0.66-1.96)	.66	
Others	0.26 (0.08-0.88)	*.03*	
Self-reported HIV status			1.09
Positive	1		
Negative	1.47 (0.79-2.75)	.23	
Unknown	0.79 (0.50-1.22)	.28	
History of HPV testing			1.13
No	1		
Yes	0.49 (0.33-0.73)	*<.001*	
Information (knowledge of HPV)	1.03 (0.99-1.06)	.12	1.30
Motivation			
Perceived susceptibility	1.06 (0.89-1.28)	.52	1.21
Perceived severity	1.01 (0.82-1.25)	.92	1.63
Perceived benefits	1.21 (0.93-1.57)	.15	1.80
Perceived barriers	1.08 (0.88-1.32)	.49	1.27
Subjective norms	1.05 (0.81-1.37)	.71	1.44
Behavioral skills			
Decision-making	0.94 (0.69-1.29)	.70	2.80
Self-efficacy	0.82 (0.57-1.18)	.28	3.07
Objective skills	1.303 (0.99-1.71)	.06	2.11

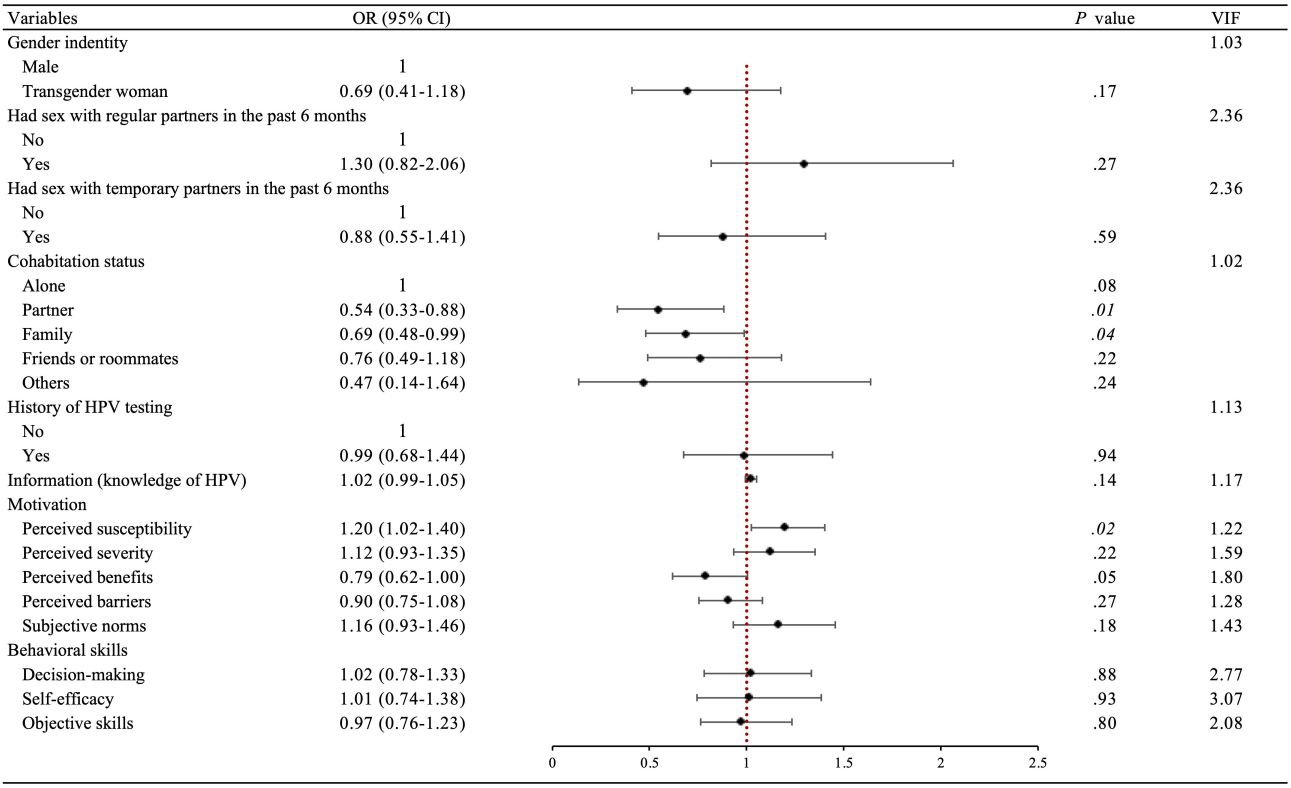


Factors Associated With Willingness to Undergo Regular HPV Testing

In this study, we explored factors associated with the willingness to undergo regular HPV testing (Table S6 in Multimedia Appendix 1). The results of the multivariable logistic regression analysis indicated that participants with higher perceived susceptibility (OR 1.20, 95% CI 1.02 - 1.40) were more likely to accept regular HPV testing. However, MSM living with a partner (OR 0.54, 95% CI 0.33 - 0.88) and those living with family members (OR 0.69, 95% CI 0.48 - 0.99) were less likely to undergo regular HPV testing (Figure 3).

XSL•FO
RenderX

Figure 3. Factors influencing willingness for regular human papillomavirus (HPV) testing among men who have sex with men in China (N=1080). OR: odds ratio; VIF: variance inflation factor.



Discussion

Principal Findings

This study investigated willingness to undergo HPV testing and its associated factors among MSM in China. Although most participants were willing to undergo HPV testing, the prior HPV testing uptake was only 23.3% (252/1080), which was similar to that reported in a study conducted in Hong Kong from 2019 to 2020 (67/350, 19.1%) [30]. However, the main reason for undergoing HPV testing was the presence or suspected presence of HPV-related symptoms, indicating a significant gap in HPV prevention among MSM. The analysis guided by the IMB model indicated that the gap between high willingness and low uptake of HPV testing among MSM reflects insufficient knowledge of HPV prevention, as well as multiple barriers that hinder the translation of willingness into actual testing behavior. These findings provide evidence for the development of more targeted public health intervention strategies.

In line with other studies [36,37], this study found that the level of HPV-related knowledge among MSM in China was relatively low. According to previous research, higher levels of knowledge are associated with a stronger intention to undergo HPV screening [38]. Individuals with greater HPV knowledge are generally more health-conscious and better able to understand the benefits of HPV testing, which is associated with a higher willingness to undergo HPV testing. Furthermore, due to sociocultural factors that contribute to unequal access to health information, many MSM lack sufficient awareness of the harms of HPV infection [39]. This was reflected in an underestimation of the threat posed by HPV to male health and inadequate recognition of the role of high-risk sexual behaviors in facilitating HPV transmission [40]. An intervention study in the United States showed that, after receiving HPV-related health education, participants' awareness that "HPV infection may cause anal cancer" increased, and their willingness to seek HPV testing improved [41]. These findings suggest that health authorities should work together with community-based organizations to deliver targeted health education for MSM and make greater use of internet-based platforms and new media technologies to expand outreach and improve HPV-related knowledge among MSM.

Perceived susceptibility was significantly associated with the willingness to undergo HPV testing and regular testing among MSM [42]. Previous research has shown that perceived risk of HIV infection significantly increases the intention to undergo regular testing among MSM [43]. Similarly, when MSM recognize that their sexual behaviors place them at higher risk of HPV infection, they are more likely to actively seek testing to identify their infection status [44]. Additionally, regular screening for HPV and anal cancer has been shown to increase life expectancy among MSM and is considered a cost-effective intervention [45]. Previous research also suggests that individuals with a better understanding of their potential health risks are more likely to engage in self-protective behaviors [46]. Therefore, MSM with higher perceived susceptibility to HPV tend to be more aware of their infection risk, show greater concern about their health status, and are more likely to initiate and maintain HPV testing behaviors [47]. This finding suggests that HPV-related services could be integrated with HIV prevention and other sexual health services within a broader STI prevention framework, which may provide a more effective approach to improving testing uptake and reducing STI transmission in this population.

XSL•FO
RenderX

In addition, perceived barriers also affect the willingness to undergo HPV testing among MSM. According to structural barriers theory, limited access to health care services and sociocultural pressures can significantly inhibit health behaviors [48]. A study on HIV testing willingness among MSM found that concerns about privacy leakage led 30% of respondents to avoid testing [49]. Moreover, STIs testing can cause stigma, which may increase their perceived barriers and reduce their willingness to undergo HPV testing [50]. When MSM are faced with issues such as shame, inadequate privacy protection, and stigma during the testing process, they may become more reluctant to undergo professional-collected HPV testing. Therefore, exploring the feasibility of delivering confidential and MSM-friendly professional-collected HPV testing, or professionally supported self-collected HPV testing, through community-based organizations may be a valuable strategy. Such service models may better meet the diverse needs of MSM and help improve testing coverage and adherence.

Improving self-efficacy among MSM is beneficial for promoting testing [51]. Previous research has shown that women with higher self-efficacy are more likely to undergo HPV testing [52]. MSM with greater awareness of self-protection and stronger confidence in HPV testing are more likely to overcome barriers associated with professional-collected HPV testing. Moreover, self-collected HPV testing provides a rapid, confidential, and cost-effective option [53]. Therefore, comprehensive efforts are needed to strengthen social advocacy that raises public awareness of HPV infection and reduces discrimination and stigma toward MSM. A meta-analysis showed that offering home-based self-collection kits contributes to higher participation rates compared with clinic-based professional sampling [54]. Self-sampling HPV testing has also been found to be acceptable and feasible among MSM [30]. Optimizing professional-collected HPV testing procedures to ensure greater privacy protection and promoting self-sampling HPV testing may help reduce psychological resistance and improve testing coverage among MSM [55].

In this study, retired or unemployed participants were less likely to be willing to undergo self-collected HPV testing compared with those who were employed. This may be related to lower occupational stability, which is often associated with reduced access to health care services [56]. Retired or unemployed individuals may face unstable income or insufficient health insurance coverage, making them more reliant on free or low-cost medical services [57]. Although self-collected HPV testing offers greater privacy, it typically requires purchasing test kits, which may impose an additional financial burden on retired and unemployed MSM [58]. Additionally, these groups are less likely to obtain testing information through online platforms or social media, leading to lower awareness and acceptance of self-collected HPV testing [59]. Therefore, governments and communities should consider providing subsidized or free HPV testing services or kits for retired and unemployed MSM. Targeted health education initiatives delivered through community for older adults may help improve awareness and acceptance of HPV testing [60].

This study extends existing evidence on HPV prevention among MSM by providing a theory-informed assessment of HPV testing willingness. First, this study distinguished between professional-collected and self-collected HPV testing, which helped identify modality-specific factors associated with different testing approaches. Second, by including regular HPV testing, this study extended the focus from one-time testing acceptance to sustained HPV screening behavior among MSM. Third, by applying the IMB model, this study interpreted testing initiation and sustained HPV testing behavior through the dimensions of information, motivation, and behavioral skills, providing a theoretical basis for designing differentiated interventions.

Limitations

This study has several limitations. First, participants were recruited using online convenience sampling, which may have introduced selection bias and limited the representativeness of the sample, as MSM with lower digital access and health literacy may have been underrepresented. Second, the data were collected through self-administered questionnaires, which may have resulted in recall bias and social desirability bias for sensitive information such as sexual behaviors and history of STIs. Third, motivation and behavioral skills related to HPV testing were assessed using self-developed, nonvalidated items, which may not have fully captured all relevant domains. Fourth, because this was a cross-sectional study, causal relationships between outcomes and associated factors could not be established. Accordingly, the findings should be interpreted with caution, and further longitudinal studies are needed to validate and extend these results.

Conclusion

This study investigated the willingness of MSM in China to undergo different HPV testing modalities and, based on the IMB model, explored the gap between willingness to undergo HPV testing and actual testing uptake in this population. The findings support the applicability of the IMB model in explaining willingness to undergo HPV testing among MSM and provide a basis for future research and strategy development. Our findings suggest that personalized health education should be strengthened through internet platforms and new media technologies to improve HPV knowledge among MSM. Meanwhile, diversified and privacy-sensitive testing service models should be explored to enhance the convenience and accessibility of HPV testing. These strategies may help translate testing willingness into actual testing behavior and promote the sustainable implementation of HPV testing services among MSM in China.

Acknowledgments

XL and XY are joint first authors of this work. The authors are grateful to Danlan Public Welfare and WenJuanXing. The authors confirm that no generative artificial intelligence tools were used to write, analyze, or interpret data in this manuscript.

XSL•FO
RenderX

Funding

This research was supported by the Non-profit Central Research Institute Fund of Chinese Academy of Medical Sciences (grant 2022-ZHCH330-01).

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: XL, XS
Data curation: XL, XY, ML, SJ, YX
Investigation: XL, XY, ML, ZD, YX, FY
Formal analysis: XL
Funding acquisition: XS
Software: FY
Supervision: ST, XS
Writing – original draft: XL, XY
Writing – review & editing: ZD, ST, XS
All authors approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Items assessing motivation and behavioral skills related to human papillomavirus (HPV) testing, reasons for HPV screening and unwillingness to undergo HPV testing, and univariate analyses of factors associated with willingness to undergo professional-collected HPV testing, self-collected HPV testing, and regular HPV testing among men who have sex with men (MSM) in China.
[DOCX File, 68 KB - publichealth_v12i1e85543_app1.docx]

References

1. de Sanjosé S, Brotons M, Pavón MA. The natural history of human papillomavirus infection. Best Pract Res Clin Obstet Gynaecol 2018 Feb;47:2-13. [doi: 10.1016/j.bpobgyn.2017.08.015] [Medline: 28964706]
2. Muñoz N, Bosch FX, de Sanjosé S, et al. Epidemiologic classification of human papillomavirus types associated with cervical cancer. N Engl J Med 2003 Feb 6;348(6):518-527. [doi: 10.1056/NEJMoa021641] [Medline: 12571259]
3. Bruni L, Albero G, Rowley J, et al. Global and regional estimates of genital human papillomavirus prevalence among men: a systematic review and meta-analysis. Lancet Glob Health 2023 Sep;11(9):e1345-e1362. [doi: 10.1016/S2214-109X(23)00305-4] [Medline: 37591583]
4. Wei F, Gaisa MM, D'Souza G, et al. Epidemiology of anal human papillomavirus infection and high-grade squamous intraepithelial lesions in 29 900 men according to HIV status, sexuality, and age: a collaborative pooled analysis of 64 studies. Lancet HIV 2021 Sep;8(9):e531-e543. [doi: 10.1016/S2352-3018(21)00108-9] [Medline: 34339628]
5. Farahmand M, Monavari SH, Tavakoli A. Prevalence and genotype distribution of human papillomavirus infection in different anatomical sites among men who have sex with men: a systematic review and meta-analysis. Rev Med Virol 2021 Nov;31(6):e2219. [doi: 10.1002/rmv.2219] [Medline: 33527636]
6. Goldstone SE, Giuliano AR, Palefsky JM, et al. Efficacy, immunogenicity, and safety of a quadrivalent HPV vaccine in men: results of an open-label, long-term extension of a randomised, placebo-controlled, phase 3 trial. Lancet Infect Dis 2022 Mar;22(3):413-425. [doi: 10.1016/S1473-3099(21)00327-3] [Medline: 34780705]
7. Soe NN, Ong JJ, Ma X, et al. Should human papillomavirus vaccination target women over age 26, heterosexual men and men who have sex with men? A targeted literature review of cost-effectiveness. Hum Vaccin Immunother 2018;14(12):3010-3018. [doi: 10.1080/21645515.2018.1496878] [Medline: 30024823]
8. Markowitz LE, Dunne EF, Saraiya M, et al. Human papillomavirus vaccination: recommendations of the Advisory Committee on Immunization Practices (ACIP). MMWR Recomm Rep 2014 Aug 29;63(RR-05):1-30. [Medline: 25167164]
9. Meites E, Szilagyi PG, Chesson HW, Unger ER, Romero JR, Markowitz LE. Human papillomavirus vaccination for adults: updated recommendations of the Advisory Committee on Immunization Practices. MMWR Morb Mortal Wkly Rep 2019 Aug 16;68(32):698-702. [doi: 10.15585/mmwr.mm6832a3] [Medline: 31415491]
10. Merck's GARDASIL receives expanded approval for males in China. : Merck; 2025 URL: https://www.merck.com/news/mercks-gardasil-receives-expanded-approval-for-males-in-china [accessed 2025-03-31]

XSL•FO
RenderX

11. Shen F, Du Y, Cao K, et al. Acceptance of the human papillomavirus vaccine among general men and men with a same-sex orientation and its influencing factors: a systematic review and meta-analysis. Vaccines (Basel) 2023 Dec 22;12(1):16. [doi: 10.3390/vaccines12010016] [Medline: 38250829]
12. Clarke MA, Deshmukh AA, Suk R, et al. A systematic review and meta-analysis of cytology and HPV-related biomarkers for anal cancer screening among different risk groups. Int J Cancer 2022 Dec 1;151(11):1889-1901. [doi: 10.1002/ijc.34199] [Medline: 35793241]
13. Viciana P, Milanés-Guisado Y, Fontillón M, et al. High-risk human papilloma virus testing improves diagnostic performance to predict moderate- to high-grade anal intraepithelial neoplasia in human immunodeficiency virus-infected men who have sex with men in low-to-absent cytological abnormalities. Clin Infect Dis 2019 Nov 27;69(12):2185-2192. [doi: 10.1093/cid/ciz144] [Medline: 30770528]
14. Giuliano AR, Anic G, Nyitray AG. Epidemiology and pathology of HPV disease in males. Gynecol Oncol 2010 May;117(2 Suppl):S15-S19. [doi: 10.1016/j.ygyno.2010.01.026] [Medline: 20138345]
15. McGovern J, Fuller C, Burris K. Anal cancer screening and prevention: a review for dermatologists. J Eur Acad Dermatol Venereol 2021 Aug;35(8):1622-1627. [doi: 10.1111/jdv.17263] [Medline: 33797819]
16. Leeds IL, Fang SH. Anal cancer and intraepithelial neoplasia screening: a review. World J Gastrointest Surg 2016 Jan 27;8(1):41-51. [doi: 10.4240/wjgs.v8.i1.41] [Medline: 26843912]
17. Stier EA, Clarke MA, Deshmukh AA, et al. International Anal Neoplasia Society's consensus guidelines for anal cancer screening. Int J Cancer 2024 May 15;154(10):1694-1702. [doi: 10.1002/ijc.34850] [Medline: 38297406]
18. Rao S, Guren MG, Khan K, et al. Anal cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol 2021 Sep;32(9):1087-1100. [doi: 10.1016/j.annonc.2021.06.015] [Medline: 34175386]
19. Sam I, Dang W, Iu N, Luo Z, Xiang YT, Smith RD. Barriers and facilitators to anal cancer screening among men who have sex with men: a systematic review with narrative synthesis. BMC Cancer 2025 Apr 1;25(1):586. [doi: 10.1186/s12885-025-13980-w] [Medline: 40169937]
20. Truesdale MD, Goldstone SE. The fear factor: drivers and barriers to follow-up screening for human papillomavirus-related anal cancer in men who have sex with men. Int J STD AIDS 2010 Jul;21(7):482-488. [doi: 10.1258/ijsa.2010.010070] [Medline: 20852198]
21. Ni Y, Lu Y, He X, et al. Self-sampled specimens demonstrate comparable accuracy and consistency to clinician-sampled specimens for HPV detection among men who have sex with men in China. Sex Transm Infect 2023 Mar;99(2):104-109. [doi: 10.1136/sextrans-2022-055408] [Medline: 35534231]
22. Wang Z, Chen S, Wong NS, et al. Differences in intention to receive clinician-collected and self-collected samples for HPV DNA testing and its determinants between heterosexual males and females in Hong Kong, China: findings of a territory-wide household survey. BMC Med 2024 Dec 2;22(1):569. [doi: 10.1186/s12916-024-03788-z] [Medline: 39623384]
23. Fisher JD, Fisher WA. Changing AIDS-risk behavior. Psychol Bull 1992 May;111(3):455-474. [doi: 10.1037/0033-2909.111.3.455] [Medline: 1594721]
24. Si M, Jiang Y, Su X, et al. Willingness to accept human papillomavirus vaccination and its influencing factors using information-motivation-behavior skills model: a cross-sectional study of female college freshmen in Mainland China. Cancer Control 2021;28:10732748211032899. [doi: 10.1177/10732748211032899] [Medline: 34634207]
25. Sheinfil AZ, Giguere R, Dolezal C, et al. Information and motivation predict HIV-serostatus among a population of high-risk men and transgender women who have sex with men. AIDS Behav 2020 Oct;24(10):2863-2871. [doi: 10.1007/s10461-020-02835-z] [Medline: 32200447]
26. Kinuthia J, Ronen K, Unger JA, et al. SMS messaging to improve retention and viral suppression in prevention of mother-to-child HIV transmission (PMTCT) programs in Kenya: a 3-arm randomized clinical trial. PLoS Med 2021 May;18(5):e1003650. [doi: 10.1371/journal.pmed.1003650] [Medline: 34029338]
27. Jiang H, Chen X, Li J, Tan Z, Cheng W, Yang Y. Predictors of condom use behavior among men who have sex with men in China using a modified information-motivation-behavioral skills (IMB) model. BMC Public Health 2019 Mar 4;19(1):261. [doi: 10.1186/s12889-019-6593-8] [Medline: 30832640]
28. Kawichai S, Songtaweesin WN, Wongharn P, et al. A mobile phone app to support adherence to daily HIV pre-exposure prophylaxis engagement among young men who have sex with men and transgender women aged 15 to 19 years in Thailand: pilot randomized controlled trial. JMIR Mhealth Uhealth 2022 Apr 21;10(4):e25561. [doi: 10.2196/25561] [Medline: 35451976]
29. Si M, Su X, Jiang Y, et al. An internet-based education program for human papillomavirus vaccination among female college students in Mainland China: application of the information-motivation-behavioral skills model in a cluster randomized trial. J Med Internet Res 2022 Sep 30;24(9):e37848. [doi: 10.2196/37848] [Medline: 36178723]
30. Wang Z, Fang Y, Wong NS, Ip M, Guo X, Wong SYS. Facilitators and barriers to take up clinician-collected and self-collected HPV tests among Chinese men who have sex with men. Viruses 2021 Apr 19;13(4):705. [doi: 10.3390/v13040705] [Medline: 33921752]
31. Waller J, Ostini R, Marlow LAV, McCaffery K, Zimet G. Validation of a measure of knowledge about human papillomavirus (HPV) using item response theory and classical test theory. Prev Med 2013 Jan;56(1):35-40. [doi: 10.1016/j.ypmed.2012.10.028] [Medline: 23142106]

XSL•FO
RenderX

32. Hu Q, Rui Y, Jiang J, et al. Awareness regarding human papillomavirus and willingness for vaccination among college students with or without medical background in Guizhou Province. Hum Vaccin Immunother 2024 Dec 31;20(1):2295992. [doi: 10.1080/21645515.2023.2295992] [Medline: 39693188]
33. Gerend MA, Madkins K, Crosby S, et al. A qualitative analysis of young sexual minority men's perspectives on human papillomavirus vaccination. LGBT Health 2019 Oct;6(7):350-356. [doi: 10.1089/lgbt.2019.0086] [Medline: 31556791]
34. Forster AS, McBride KA, Davies C, et al. Development and validation of measures to evaluate adolescents' knowledge about human papillomavirus (HPV), involvement in HPV vaccine decision-making, self-efficacy to receive the vaccine and fear and anxiety. Public Health (Fairfax) 2017 Jun;147:77-83. [doi: 10.1016/j.puhe.2017.02.006] [Medline: 28404501]
35. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol 2008 Apr;61(4):344-349. [doi: 10.1016/j.jclinepi.2007.11.008] [Medline: 18313558]
36. Lau JTF, Wang Z, Lau M, Lai CHY. Perceptions of HPV, genital warts, and penile/anal cancer and high-risk sexual behaviors among men who have sex with men in Hong Kong. Arch Sex Behav 2014 May;43(4):789-800. [doi: 10.1007/s10508-013-0172-3] [Medline: 23982568]
37. Li X, Cao X, Li Z, et al. Human papillomavirus awareness and vaccine acceptability among men who have sex with men from Mainland China. Sci Rep 2019 Jun 19;9(1):8763. [doi: 10.1038/s41598-019-45258-0] [Medline: 31217451]
38. Shrestha AD, Andersen JG, Gyawali B, et al. Cervical cancer screening utilization, and associated factors, in Nepal: a systematic review and meta-analysis. Public Health 2022 Sep;210:16-25. [doi: 10.1016/j.puhe.2022.06.007] [Medline: 35863158]
39. Fenkl EA, Jones SG, Schochet E, Johnson P. HPV and anal cancer knowledge among HIV-infected and non-infected men who have sex with men. LGBT Health 2016 Feb;3(1):42-48. [doi: 10.1089/lgbt.2015.0086] [Medline: 26654570]
40. Natipagon-Shah B, Lee E, Lee SY. Knowledge, beliefs, and practices among U.S. college students concerning papillomavirus vaccination. J Community Health 2021 Apr;46(2):380-388. [doi: 10.1007/s10900-020-00922-9] [Medline: 32939677]
41. Ong C, Winer J, Margolies L, Dai F, Patridge E, Chagpar AB. Awareness of HPV and its vaccine among lesbian, gay, bisexual, transgender, and queer (LGBTQ) populations. J Clin Oncol 2013 May 20;31(15_suppl):e17597. [doi: 10.1200/jco.2013.31.15_suppl.e17597]
42. Chan PSF, Fang Y, Chidgey A, Fong F, Ip M, Wang Z. Would Chinese men who have sex with men take up human papillomavirus (HPV) screening as an alternative prevention strategy to HPV vaccination? Front Med (Lausanne) 2022;9:904873. [doi: 10.3389/fmed.2022.904873] [Medline: 35721088]
43. Lin Y, Li C, Wang L, Jiao K, Ma W. The mediated effect of HIV risk perception in the relationship between peer education and HIV testing uptake among three key populations in China. AIDS Res Ther 2021 Mar 25;18(1):8. [doi: 10.1186/s12981-021-00334-2] [Medline: 33766062]
44. Reiter PL, McRee AL, Katz ML, Paskett ED. Human papillomavirus vaccination among young adult gay and bisexual men in the United States. Am J Public Health 2015 Jan;105(1):96-102. [doi: 10.2105/AJPH.2014.302095] [Medline: 25393178]
45. Goldie SJ, Kuntz KM, Weinstein MC, Freedberg KA, Palefsky JM. Cost-effectiveness of screening for anal squamous intraepithelial lesions and anal cancer in human immunodeficiency virus-negative homosexual and bisexual men. Am J Med 2000 Jun 1;108(8):634-641. [doi: 10.1016/s0002-9343(00)00349-1] [Medline: 10856411]
46. Brewer NT, Chapman GB, Gibbons FX, Gerrard M, McCaul KD, Weinstein ND. Meta-analysis of the relationship between risk perception and health behavior: the example of vaccination. Health Psychol 2007 Mar;26(2):136-145. [doi: 10.1037/0278-6133.26.2.136] [Medline: 17385964]
47. Alyafei A, Easton-Carr R. The health belief model of behavior change. In: StatPearls: StatPearls Publishing; 2026. URL: https://www.ncbi.nlm.nih.gov/books/NBK606120/ [accessed 2026-05-21]
48. McConnell EA, Janulis P, Phillips G 2nd, Truong R, Birkett M. Multiple minority stress and LGBT community resilience among sexual minority men. Psychol Sex Orientat Gend Divers 2018 Mar;5(1):1-12. [doi: 10.1037/sgd0000265] [Medline: 29546228]
49. Zhang C, Li X, Heilemann MV, Chen X, Wang H, Koniak-Griffin D. Facilitators and barriers of HIV self-testing among Chinese men who have sex with men: a qualitative study. J Assoc Nurses AIDS Care 2021;32(5):599-609. [doi: 10.1097/JNC.0000000000000218] [Medline: 33165181]
50. Philbin MM, Hirsch JS, Wilson PA, Ly AT, Giang LM, Parker RG. Structural barriers to HIV prevention among men who have sex with men (MSM) in Vietnam: diversity, stigma, and healthcare access. PLoS ONE 2018;13(4):e0195000. [doi: 10.1371/journal.pone.0195000] [Medline: 29614104]
51. Tan RKJ, Kaur N, Kumar PA, et al. Clinics as spaces of costly disclosure: HIV/STI testing and anticipated stigma among gay, bisexual and queer men. Cult Health Sex 2020 Mar;22(3):307-320. [doi: 10.1080/13691058.2019.1596313] [Medline: 30975036]
52. Tiraki Z, Yılmaz M. Cervical cancer knowledge, self-efficacy, and health literacy levels of married women. J Cancer Educ 2018 Dec;33(6):1270-1278. [doi: 10.1007/s13187-017-1242-3] [Medline: 28668992]
53. Yared NF, Horvath KJ, Baker JV, Thyagarajan B, Waterboer T, Kulasingam S. Concordance of self- and clinician-collected anal swabs to detect human papillomavirus in a sample of HIV-negative men. J Low Genit Tract Dis 2019 Jul;23(3):200-204. [doi: 10.1097/LGT.0000000000000475] [Medline: 30946259]

XSL•FO
RenderX

54. Verdoodt F, Jentschke M, Hillemanns P, Racey CS, Snijders PJF, Arbyn M. Reaching women who do not participate in the regular cervical cancer screening programme by offering self-sampling kits: a systematic review and meta-analysis of randomised trials. Eur J Cancer 2015 Nov;51(16):2375-2385. [doi: 10.1016/j.ejca.2015.07.006] [Medline: 26296294]
55. Heid-Picard B, Cochand-Priollet B, Rozenberg F, et al. Ambulatory anal self-sampling in MSM living with HIV, an acceptable and reliable screening method. PLoS One 2021;16(2):e0246338. [doi: 10.1371/journal.pone.0246338] [Medline: 33561135]
56. Driscoll AK, Bernstein AB. Health and access to care among employed and unemployed adults: United States, 2009-2010. NCHS Data Brief 2012 Jan(83):1-8. [Medline: 22617552]
57. Figueroa JF, Tosin-Oni M, Phelan J, Orav EJ, Epstein AM. Changes in employment status and access to care during COVID-19 pandemic among low-income adults in 4 southern states. J Gen Intern Med 2022 Aug;37(11):2795-2802. [doi: 10.1007/s11606-022-07547-9] [Medline: 35428901]
58. Melon M, Kombo B, Mugambi M, et al. Expanding options for HIV testing: a process evaluation of a community-led HIV self-testing intervention among men who have sex with men in Kenya. Gates Open Res 2023;7:127. [doi: 10.12688/gatesopenres.14819.2] [Medline: 39035469]
59. Vera JH, Soni S, Pollard A, et al. Acceptability and feasibility of using digital vending machines to deliver HIV self-tests to men who have sex with men. Sex Transm Infect 2019 Dec;95(8):557-561. [doi: 10.1136/sextrans-2018-053857] [Medline: 31101723]
60. Zhao YC, Zhao M, Song S. Online health information seeking behaviors among older adults: systematic scoping review. J Med Internet Res 2022 Feb 16;24(2):e34790. [doi: 10.2196/34790] [Medline: 35171099]

Abbreviations

GK: general HPV knowledge subscale
HPV: human papillomavirus
IMB: information-motivation-behavioral skills
MSM: men who have sex with men
OR: odds ratio
STI: sexually transmitted infection
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
TK: HPV testing knowledge subscale
VK: HPV vaccine knowledge subscale

Edited by S Brini; submitted 09.Oct.2025; peer-reviewed by J Jones, S Chen; revised version received 27.Apr.2026; accepted 02.May.2026; published 08.Jun.2026.

Please cite as:
Liu X, Yang X, Lyu M, Dai Z, Jing S, Xin Y, Yu F, Tang S, Su X
Willingness to Undergo Human Papillomavirus Testing Among Men Who Have Sex With Men in China Based on the Information-Motivation-Behavioral Skills Model: Online Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e85543
URL: https://publichealth.jmir.org/2026/1/e85543
doi:10.2196/85543

© Xin Liu, Xuelin Yang, Mohan Lyu, Zhenwei Dai, Shu Jing, You Xin, Fei Yu, Shenglan Tang, Xiaoyou Su. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 8.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Association Between Family Health and Proactive Health Risk Management With the Mediating Role of Health Literacy: Nationwide Cross-Sectional Study

Shangfeng Tang[1], PhD; Yilin Wei[1], PhD; Heng Dong[1], MD; Kangkang Zhang[1], PhD; Chunying Wang[1], MD; Jia Song[1], PhD; Hua Qing[1], MD; Yibo Wu[2], PhD

[1]School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology, Hangkong Road No.13, Wuhan, Hubei, China

[2]School of Public Health, Peking University, Beijing, Beijing, China

Corresponding Author:
Yilin Wei, PhD
School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology, Hangkong Road No.13, Wuhan, Hubei, China

Abstract

Background: Modifiable unhealthy behaviors account for over two-thirds of new cases of noncommunicable diseases. Behavioral risk factor reduction is a potentially cost-effective means to improve long-term health outcomes. Although family serves as a pivotal cornerstone for fostering and maintaining individuals' health, the associations between family health (FH) and the proactive health risk management index (PHRMI) remain unclear.

Objective: This study aimed to construct a comprehensive index to measure the PHRMI and examine the mediating effect of health literacy on the association between FH and the PHRMI, as well as the moderating effect of family communication on the associations among the PHRMI, health literacy, and FH.

Methods: A cross-sectional questionnaire survey was conducted with 30,044 participants from 34 provinces or regions in China who were recruited using a multistage stratified sampling strategy from June 20, 2023, to August 31, 2023. This study constructed the PHRMI for the general population by encompassing BMI, physical activity (International Physical Activity Questionnaire–Short Form), depression (Patient Health Questionnaire-9), sleep quality (Brief version of the Pittsburgh Sleep Quality Index), smoking behavior, and drinking behavior. Further, we assessed FH (Short Form of the Family Health Scale), health literacy (Short-Form Health Literacy Questionnaire-4), and family communication (Family Communication Scale–Short Form). In addition, we collected the sociodemographic characteristics of the participants. We used model 4 of the IBM SPSS macro PROCESS to verify the mediating effect of health literacy between FH and the PHRMI, while model 7 was adopted to test the moderated mediation of family communication among the PHRMI, health literacy, and FH.

Results: Higher levels of FH were significantly associated with higher PHRMI levels (β=.710, 95% CI 0.669-0.752). Health literacy significantly mediated the association between FH and the PHRMI (β=.207, 95% CI 0.168-0.245), playing a partial mediating role. Family communication significantly moderated the association between FH and health literacy (β=.117, 95% CI 0.105-0.130). The simple slope analysis showed that higher levels of family communication exacerbated the effects of FH on health literacy. Subsequently, we performed a sensitivity analysis, and the main results aligned with the findings of prior studies. Nevertheless, the subgroup analysis revealed that the mediating effect of health literacy was not significant in the group aged >60 years (β=.066, 95% CI −0.024 to 0.157).

Conclusions: FH can be an important target that appears to be positively linked to proactive health risk management and health literacy. FH promotion for older adults should pay more attention to family or intergenerational communication.

(JMIR Public Health Surveill 2026;12:e73659) doi:10.2196/73659

KEYWORDS

family health; health literacy; proactive health risk management; middle-aged and older adults; health behavior

Introduction

The presence of multiple health risk factors (environmental and occupational hazards, behavioral risks, and metabolic risks) exerts a cumulative negative influence on health [1]. From the perspective of the full life cycle, in addition to genetic factors, the accumulation of risk factors during the early stages of life exacerbates health hazards in the later stages of life [2,3]. Modifiable unhealthy behaviors, such as tobacco use,

XSL•FO
RenderX

consumption of foods high in saturated and trans fats, salt, and sugar, physical inactivity, and the harmful consumption of alcohol, account for more than two-thirds of all new cases of noncommunicable diseases and increase the risk of complications in individuals with noncommunicable diseases [4]. The rapid expansion and high incidence of chronic diseases worldwide, combined with population aging, have resulted in an increasing disease burden on individuals [5], families, and society [6]. Behavioral risk factor reduction is a potentially cost-effective means to improve long-term health outcomes and reduce the disease burden. Therefore, numerous countries have implemented policies to encourage individuals to actively practice a healthy lifestyle, aligning with the "proactive health" initiative actively promoted by China in recent years.

Despite being a novel concept proposed in 2015, proactive health has a lengthy research history concerning its concept and practice [7-11]. Most relevant research focuses on the connotation, subject matter, objectives, essential components, and implementation strategies of proactive health. Current academic discourse presents two main perspectives on the concept of proactive health [12]. First, based on the micro perspective of residents, proactive health is defined as guiding residents to identify and control risk factors [8], advocating self-active health management, and promoting the formation of healthy living habits [13]. Second, proactive health inherently leverages information technology to achieve its "proactive and full-lifecycle" characteristics in health management services [14]. From this perspective, proactive health is defined as health care professionals regularly monitoring residents' health risk indicators through means such as model design and platform construction [15], aiming to detect service needs early and intervene in the medical and health service process in advance. Despite the variations in conceptualizations of proactive health, scholars in China have reached a consensus on the identification and control of health risk factors within the core essence of proactive health. Some scholars have focused on special groups, such as adolescents [16], disabled older adults [17], patients with hypertension, patients with coronary heart disease [18], and others, to develop appropriate proactive health behavior measurement scales. The development approach mainly revolves around the specific health needs or unique health risk factors of the participants being assessed. For instance, proactive health management for adolescents [16] is composed of 5 dimensions: health responsibility, physical activity, nutritional diet, mental health, and self-discipline. In line with the above studies, we developed the proactive health risk management index (PHRMI), integrating the most common health risks for the general population. Proactive health risk management is defined as a resident's active engagement in health management activities encompassing psychological, dietary, exercise, sleep, and lifestyle aspects. It involves identifying and modifying detrimental habits to prevent health risks from the source, encouraging individuals to assume health management responsibilities, thereby comprehensively safeguarding individual and collective well-being. Risk factors were extracted based on the "four cornerstones of health" outlined in the "Victoria Declaration," namely "balanced diet," "moderate exercise," "smoking cessation and alcohol restriction," and "psychological balance" [19]. In addition, sleep of sufficient quality and duration is essential for good health [20] and reduces the risk of total mortality, obesity [21], cardiovascular disease [22,23], cerebrovascular diseases [24,25], and dementia. Therefore, sleep quality is incorporated as one of the components of the proactive health index.

Subsequently, this study aims to investigate family-level and individual-level determinants influencing the PHRMI. Entering the era of "Public Health 3.0," the family's role as a fundamental contributor to individual and community health remains unquestioned. The family, owing to its fundamental capacity to nurture, care for, protect, educate, and influence across the lifespan, serves as a pivotal cornerstone for fostering and maintaining individuals' health. Engaging families as proactive participants in health promotion initiatives can enhance family members' acceptance and use [19] and may facilitate longer-lasting and larger-scale behavioral changes [26]. Ho et al [27] have developed a more comprehensive conceptual framework for "family health" (FH) by using the Delphi method. The framework posits FH as a multifaceted resource arising at the family-unit level through the interplay of the members' health, their interactions, and the family's array of physical, social, emotional, economic, and medical resources. FH is a comprehensive construct that encompasses intrafamily dynamics, such as associations, communication, and the domestic environment, as well as the accessibility of social support and health resources. The influence of FH on proactive health, both in terms of its presence and magnitude, remains uncertain. Clarifying these ambiguities may have important implications for individual, clinical, and public health, especially amid the escalating incidence of chronic diseases [28].

Health literacy serves as a key measure reflecting an individual's accumulation of health knowledge and capabilities [29]. It is hypothesized that FH would result in a positive influence on health literacy, given that the family constitutes a primary environment through which health-related attitudes, behaviors, and competencies are developed. Existing studies have demonstrated that individuals with low health literacy are associated with inadequate health-related behaviors [30], leading to adverse health-related outcomes [31], including poorer general health [32], more hospitalizations [33], compromised self-management ability, less effective use of health services [34], and higher health care costs [35]. Although lower health literacy is often linked to poorer health outcomes, some patients with chronic diseases can successfully manage their conditions despite inadequate health literacy. The buffering effect of valuable resources and support from social networks appears to mitigate the negative impacts of low health literacy [36]. Humans, inherently endowed with social attributes, consistently inhabit a collective milieu characterized by diverse levels of social support and resource availability. Family-level resources and support are pivotal in an individual's social network. Research across Brazil, the United States, and China has established that family structure, income levels, and cohesion, along with other factors, significantly influence the development and maintenance of health literacy [37-39].

The enhancement of health literacy through FH promotion does not occur spontaneously but rather through dynamic interactions within familial systems. The circumplex model emphasizes that

XSL•FO
RenderX

family communication is considered a facilitating dimension and is essential for family cohesion and flexibility [40]. Conversation-oriented communication patterns significantly facilitate intermember health information exchange [41]. According to the "stimulus-awareness-internalization" pathway, family communication serves as a "stimulus" by facilitating information exchange and sharing [41-43]. This process enables individuals to receive health information and resources transmitted at the family level, subsequently internalizing them into cognitive components such as health literacy. Research has shown that better family communication is associated with better health status and higher cervical cancer literacy scores [37]. Another anorexia family therapy trial showed that family communication shifting from covert aggression to overt aggression during systemic family therapy significantly improved patients' eating attitudes [44]. The positive transformation of patients' eating attitudes and self-health cognition was promoted by coordinating the conflicts of health concepts within the family. Therefore, this study attempts to examine whether family communication plays a moderating role in the association between FH and health literacy.

Well-established evidence indicates that over two-thirds of the risk factors for chronic diseases stem from modifiable unhealthy behaviors [2,4,45]. Therefore, encouraging individuals to adopt a healthy lifestyle is crucial and urgent, with FH potentially serving as an effective intervention target. This study explores whether and how FH can enhance the PHRMI in China. We propose the following three hypotheses: (1) FH is positively associated with proactive health risk management, (2) the association between FH and the PHRMI is mediated by health literacy, and (3) family communication moderates the mediating effect of health literacy between FH and the PHRMI.

Methods

Sampling and Participants

The data originated from a nationwide, representative, population-based survey conducted from June 20, 2023, to August 31, 2023, encompassing 150 cities from 23 provinces, 5 autonomous regions, 4 municipalities directly under the Central Government, and 2 special administrative regions in China [46,47]. To ensure the overall representativeness of the study population, the survey used a multistage sampling method, incorporating stratified sampling at various levels, including province, city, and community. During the first stage, the city selection included capital cities across China, and 2 to 12 cities were selected from each noncapital prefecture within every province and autonomous region, totaling 150 cities. The number of cities selected from each province was determined by the population size weight of that province. Moving on to the second stage, of the 150 municipalities sampled, the number of communities sampled within each municipality was proportional to the population base of the first-level administrative district. A total of 10 to 60 communities were sampled, with an urban-to-rural ratio of 3:2 in the municipalities sampled in each province, resulting in 800 communities. In the last stage, quota sampling was used at the community and village levels as well as at the individual level, using quotas that were determined based on sex and age attributes from the data of "The 7th National Census in 2021." The sample size was estimated to be 45,830. After implementing quota sampling, the final sample included 30,054 participants. This sample specifically accounted for attributes such as sex and age.

The survey protocol has been published [48]. Inclusion criteria for our survey include: (1) being aged 18 years or older, (2) holding the nationality of the People's Republic of China, (3) being permanent residents with annual travel of 1 month or less, (4) being able to complete the online questionnaire independently or with assistance, and (5) being able to understand each questionnaire item. Individuals who were delirious, suffering from cognitive dysfunction, participating in other similar research, or unwilling to participate were excluded. In this study, we excluded samples with missing data on critical study variables. Specifically, 10 participants were removed due to incomplete bedtime and wake-up time records, which precluded reliable assessment of their sleep quality.

Ethical Considerations

This study was conducted in compliance with the ethical guidelines of the Declaration of Helsinki and received approval from the ethics research committee of the Shandong Provincial Hospital (SWYX:NO 2023 - 198). All data in this research did not involve participant compensation. All participants were required to sign informed consent prior to participating in the study. Meanwhile, we ensured the confidentiality of the collected data through anonymity and stringent security measures.

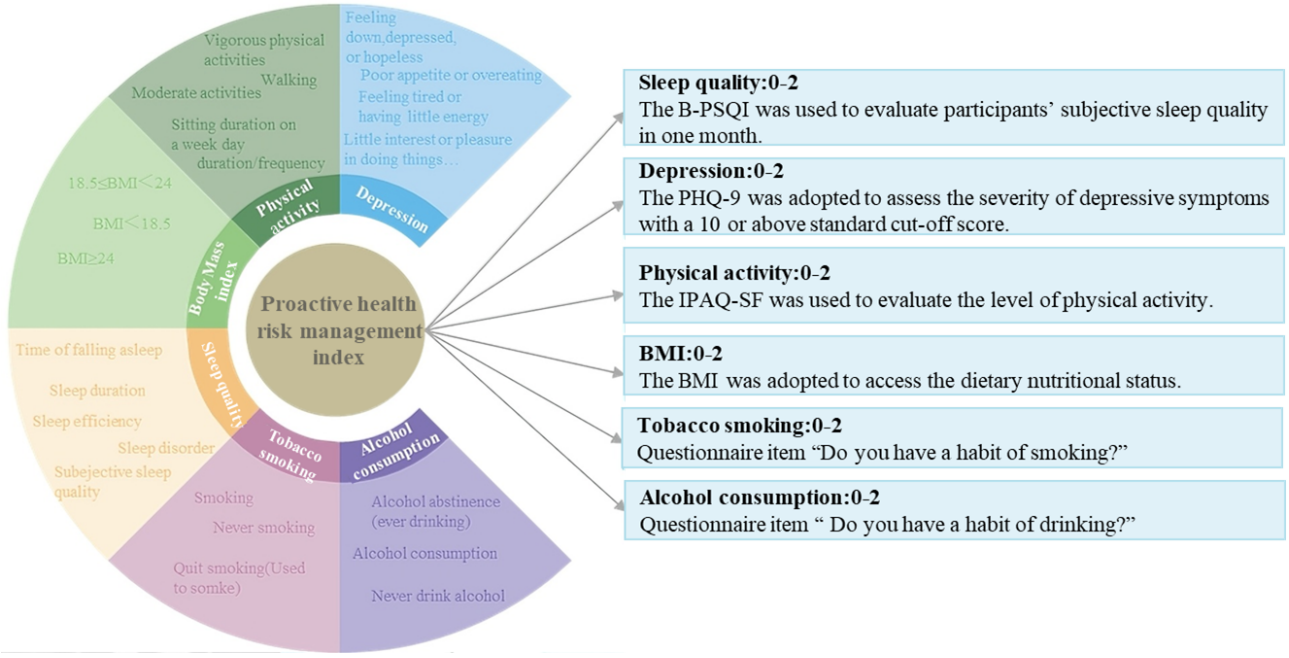
Proactive Health Risk Management Index

Referring to previous relevant studies [49-51], PHRMIs were measured through a self-created comprehensive instrument, including BMI, physical activity, depression, sleep quality, smoking behavior, and drinking behavior. We defined the PHRMI as comprising 6 aspects, with its measurement based on the 4 cornerstones of health proposed by the World Health Organization. Given the detrimental health effects of tobacco smoking and alcohol consumption, we incorporated questions about smoking and drinking behavior, offering response options such as 0 ("yes, smoking or drinking"), 1 ("ever smoked or drank [quit]"), and 2 ("never smoked or drank"), alongside "Do you have a habit of smoking?" and "Do you have a habit of drinking?" The International Physical Activity Questionnaire–Short Form (IPAQ-SF) was used to evaluate the level of physical activity, assessing participants' vigorous activity, moderate activity, and walking activities in terms of frequency and duration from the previous week [52,53]. Previous studies have indicated a moderate-to-high reliability of IPAQ-SF across 12 countries, as well as the Chinese version. According to each participant's total metabolic equivalent–minute/week in walking, moderate, and vigorous activity for 7 days, we categorized them as high, moderate, and low physical activity and assigned the values "2," "1," or "0," respectively. The 9-item Patient Health Questionnaire was adopted to assess the severity of depressive symptoms [54], using a cutoff score of 10 or higher. We encoded Patient Health Questionnaire-9 scale scores of "15 or more" as "0," "5 to 14" as "1," and "less than 4" as "2." The Brief Version of the

XSL•FO
RenderX

Pittsburgh Sleep Quality Index (B-PSQI) was used to evaluate participants' subjective sleep quality over 1 month [55], and it has been proven to be a reliable and valid sleep-quality measure for the general population. This scale encompassed sleep efficiency, sleep latency, sleep duration, sleep disorders, and sleep quality, with a 5 or higher standard cutoff score. We encoded B-PSQI scale scores of "5 or more" as 0 and "less than 5" as 2. BMI was selected as an indicator for the examination of reasonable diet. Participants were classified as overweight (BMI≥24 kg/m^2), normal weight (18.5 kg/m^2≤BMI<24 kg/m^2), and underweight (BMI<18.5 kg/m^2), respectively. The indicator was coded as 0 (overweight or underweight) or 2 (normal weight). As shown in Figure 1, the total scores were summed to calculate a PHRMI score ranging from 0 to 12, where a higher score signifies better proactive health risk management.

Figure 1. Schematic diagram of the components of proactive health risk management, encompassing sleep quality (Brief version of the Pittsburgh Sleep Quality Index [B-PSQI]), depression (Patient Health Questionnaire-9 [PHQ-9]), physical activity (International Physical Activity Questionnaire–Short Form [IPAQ-SF]), BMI, tobacco smoking, and alcohol consumption.



Assessment of FH

FH was measured using the Short Form of the Family Health Scale [56], which was modified to align with China's national conditions and cultural characteristics. The scale has been validated and proven to be reliable and effective. It contains four dimensions: (1) family healthy lifestyle; (2) FH resources; (3) family, social, or emotional health processes; and (4) family external social support. The 3 dimensions (7 items) are scored on a Likert-5 scale ranging from 1 (strongly disagree) to 5 (strongly agree), except for "family health resources" (3 items). Scores ranged from 10 to 50 points, with a high score indicating higher levels of FH. The Cronbach α for the sample was 0.807, and the Kaiser-Meyer-Olkin value was 0.879, both indicating acceptable reliability and validity.

Assessment of Health Literacy

The mediator variable in this study was health literacy, assessed by a 4-item version of the Short-Form Health Literacy Questionnaire-4. The simplified health literacy scales, derived from the Short-Form Health Literacy Questionnaire-12 [57], demonstrated good reliability and validity and were reliable and effective tools for assessing the health literacy of the Chinese population. The total score ranges from 4 to 16, with a higher score indicating better levels of health literacy.

Assessment of Family Communication

Family communication was the moderator in this study, assessed using the Family Communication Scale-Short Form (FCS-SF). For example, participants were asked, "Family members can discuss things calmly with each other." Items were scored on a 5-point scale from 1 to 5, from 1 ("strongly disagree") to 5 ("strongly agree"). The sum of scores on the FCS-SF, ranging from 4 to 20 points, indicates better communication between family members with higher scores. The Cronbach α for the FCS-SF was 0.926 in this study.

Covariates

Covariates were identified initially based on relevant research and general knowledge, including age, sex, nationality, education level, family type, and employment status. Enrolled covariates were analyzed subsequently for intergroup differences using the bilateral 2-tailed Student *t* test or one-way ANOVA. Covariates were considered for inclusion as potential confounders in the final models if significant differences between the groups were observed. The final demographic characteristics included age, sex (1="male" and 2="female"), nationality (1="Han nationality" and 2="nationality minority"), educational level (1="illiterate or semi-illiterate"; 2="junior high school and below"; 3="secondary technical school, senior high school, and college for professional training"; and 4="Bachelor's degree and higher"), residence (1="urban" and

XSL•FO
RenderX

2="rural"), family type (1="core family," 2="backbone family," 3="joint family," 4="conjugal family," 5="single-parent family," and 6="other family"), household income monthly per capita (1="≤￥3000" [US $421], 2="￥3001-￥6000" [US $421-US $841], and 3="≥￥6001" [US $841]), children (0="no" and 1="yes"), and insurance (0="no" and 1="yes"). Age and self-assessment of social status (scoring 0-7) were set as continuous variables.

Statistical Analysis

Demographic characteristics, FH, PHRMIs, health literacy, and family communication of participants were summarized using frequencies (percentages) or means and SDs. The disparity in PHRMIs across different groups was analyzed using the 2-tailed Student *t* test and one-way ANOVA. We used SPSS 26.0 (IBM Corp) and the Hayes [58] SPSS macro program PROCESS to test the mediating effect using model 4. If the coefficient *a* (representing the association between X and M), *b* (the mediator's effect on Y), and *c* (the relationship between X and Y) in the regression model including X, M, and Y were statistically significant, this indicated that the mediating effect was significant. Next, we constructed a new model containing the aforementioned variables and the product term of the moderator and independent variable to verify the moderating effect of family communication. Finally, the PROCESS macro for SPSS was used to examine whether family communication moderated the direct and indirect effects of FH on the PHRMI. All statistical analyses were conducted using SPSS 26.0. In addition, all models controlled for covariates, and the significance level was set at a *P* value of <.05 for all the hypothesis tests [59].

Sensitivity and Subgroup Analysis

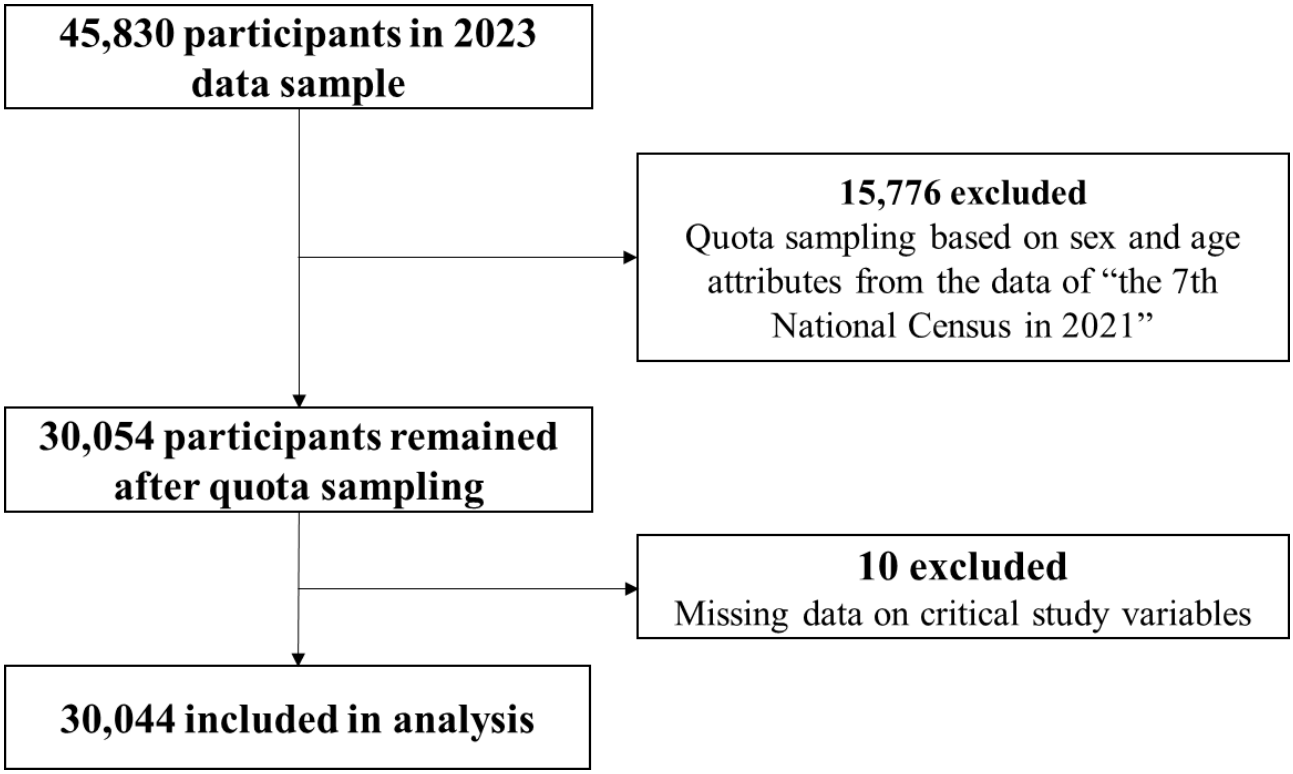
To validate the results, we conducted a sensitivity analysis. Specifically, we replaced depression with generalized anxiety disorder as a measure of the mood component and recalculated the PHRMI. Following this adjustment, we repeated the aforementioned statistical analysis steps to ascertain whether the direct and indirect effects remained statistically significant. In addition, subgroup analysis was used to explore the heterogeneity of PHRMIs. Considering the wide age spectrum of participants (aged 18-106 y), we further stratified the age cohorts to compare the association between familial health and proactive health levels across different age groups, as well as to examine the mediating role of health literacy.

Results

Sample Characteristics

The process for screening participants is shown in Figure 2. A total of 30,044 participants were included in the final analysis. All participants had a mean age of 43.99 (SD 16.55) years, ranging from 18 to 106 years, with 15,004 (49.90%) males and 15,040 (50.10%) females. Most participants were married (n=19,391, 64.50%), lived in urban areas (n=20,728, 69.00%), and reported having a Bachelor's degree or higher (n=10,609, 35.30%). The mean scores for FH and the PHRMI were 39.86 (SD 6.15) and 5.88 (SD 1.78), respectively. Subsequently, we performed an analysis of intergroup differences to assess the disparity in PHRMIs across demographic variables, using a 2-tailed *t* test and one-way ANOVA. Significant differences were found among various demographic groups, including sex, nationality, education level, and residence (Table 1).

Figure 2. Flowchart of participant selection in the 2023 Psychology and Behavior Investigation of Chinese Residents for this study, conducted from June 20, 2023, to August 31, 2023.



XSL•FO
RenderX

Table . Characteristics of participants and the scores of family health (FH) and the proactive health risk management index (PHRMI) based on demographic variables in this cross-sectional study (2023, China; N=30,044).

Characteristics	Total, n (%)	FH, mean (SD)	PHRMI, mean (SD)	*P* value
Age (y)	42.92 (16.56)	3.99 (0.61)	7.57 (2.29)	<.001
Sex				<.001
Male	15,004 (49.9)	3.95 (0.63)	7.08 (2.43)	
Female	15,040 (50.1)	4.02 (0.60)	8.05 (2.03)	
Nationality				<.001
Han	27,290 (90.8)	4.00 (0.61)	7.62 (2.26)	
National minority	2754 (9.2)	3.85 (0.65)	7.08 (2.50)	
Education level				<.001
Illiterate or semi-illiterate	1458 (4.9)	3.83 (0.61)	7.54 (2.09)	
Junior high school and below	7783 (25.9)	3.92 (0.59)	7.57 (2.25)	
Secondary technical school, senior high school, and college for professional training	10,194 (33.9)	3.98 (0.62)	7.47 (2.33)	
Bachelor degree and higher	10,609 (35.3)	4.06 (0.62)	7.66 (2.31)	
Family type				<.001
Core family	15,141 (50.5)	4.06 (0.59)	7.77 (2.27)	
Backbone family	6066 (20.2)	3.97 (0.60)	7.49 (2.22)	
Joint family	1220 (4.1)	3.92 (0.62)	7.19 (2.40)	
Conjugal family	4963 (16.5)	3.90 (0.62)	7.42 (2.33)	
Single-parent family	1047 (3.5)	3.87 (0.63)	7.04 (2.28)	
Other	1567 (5.2)	3.75 (0.68)	7.04 (2.37)	
Employment status				<.001
Incumbent	13,295 (44.3)	4.03 (0.61)	7.60 (2.35)	
Rolling stone	5042 (16.8)	3.92 (0.58)	7.44 (2.31)	
Retirement	4226 (14.1)	3.95 (0.64)	7.46 (2.18)	
Unemployment	7481 (24.9)	3.98 (0.62)	7.65 (2.24)	
Per capita monthly household income				<.001
≤￥3000 (≤US $421)	8779 (29.2)	3.84 (0.61)	7.37 (2.29)	
￥3001-￥6000 (US $421-US $841)	13,425 (44.7)	4.00 (0.60)	7.64 (2.27)	
≥￥6001(≥ US $841)	7840 (26.1)	4.13 (0.61)	7.65 (2.32)	
Children				<.001
No	10,483 (34.9)	3.96 (0.65)	7.57 (2.33)	
Yes	19,561 (65.1)	4.00 (0.60)	7.56 (2.27)	
Marital status				.105
Married	19,391 (64.5)	4.00 (0.60)	3.95 (0.64)	
Divorced, widowed, or unmarried	10,653 (35.5)	3.95 (0.64)	5.86 (1.81)	
Insurance				<.001
No	1441 (4.8)	3.74 (0.67)	5.67 (1.76)	
Yes	28,603 (95.2)	4.00 (0.61)	5.89 (1.78)	
Residence				<.001

XSL•FO
RenderX

Characteristics	Total, n (%)	FH, mean (SD)	PHRMI, mean (SD)	*P* value
Urban	20,728 (69)	4.02 (0.61)	7.59 (2.30)	
Rural	9316 (31)	3.91 (0.61)	7.52 (2.28)	
Self-assessment of social status	3.98 (1.36)	3.99 (0.61)	7.57 (2.29)	<.001

Mediation Analyses

Model 4 is a simple mediating model in SPSS macro PROCESS compiled by Hayes. We adopted to test the mediating effect of health literacy on the association between FH and the PHRMI, while controlling for age, gender, nationality, education level, family type, employment status, income, children, marital status, insurance, residence, self-assessment of social status, and registered permanent residence. Model 1 of Table 2 shows that the positive predictive effect of FH on the PHRMI was significant (β=.710, 95% CI 0.669-0.752). Model 2 of Table 2 shows that FH had a significant positive predictive effect on health literacy (β=.136, 95% CI 0.124-0.148), while health literacy also had a significant positive predictive effect on the PHRMI (β=.207, 95% CI 0.168-0.245). Moreover, when the mediating variable was added, the direct predictive effect of FH on the PHRMI was still significant (β=.682, 95% CI 0.641-0.724), as shown in model 3 in Table 2. In addition, the upper and lower bounds of the bootstrapped 95% CI for the direct effect of FH on proactive health risk management and the mediating effect of health literacy did not include zero in Figure 3, indicating that the mediating effect was significant (indirect effect=0.028; *P*<.001). The results implied that the association between FH and the PHRMI was partially mediated by health literacy.

Table . Results of the mediating effect of health literacy in the association between family health (FH) and the proactive health risk management index (PHRMI) in the cross-sectional study based on the 2023 Psychology and Behavior Investigation of Chinese Residents (N=30,044). These models incorporated covariates, including age, gender, nationality, education level, family type, employment status, income, children, marital status, insurance, residence, self-assessment of social status, and registered permanent residence.

Predictors	Model 1[a] (DV[b]=PHRMI)	Model 2[c] (DV=HL[d])	Model 3[e] (DV=PHRMI)
FH, β (95% CI)	.710 (0.669-0.752)	.136 (0.124-0.148)	.682 (0.641-0.724)
HL, β (95% CI)	__[f]	—	.207 (0.168-0.245)

[a]$F_{13, 30,030}$=166.605, *P*<0.001, R^2=0.109.

[b]DV: dependent variable.

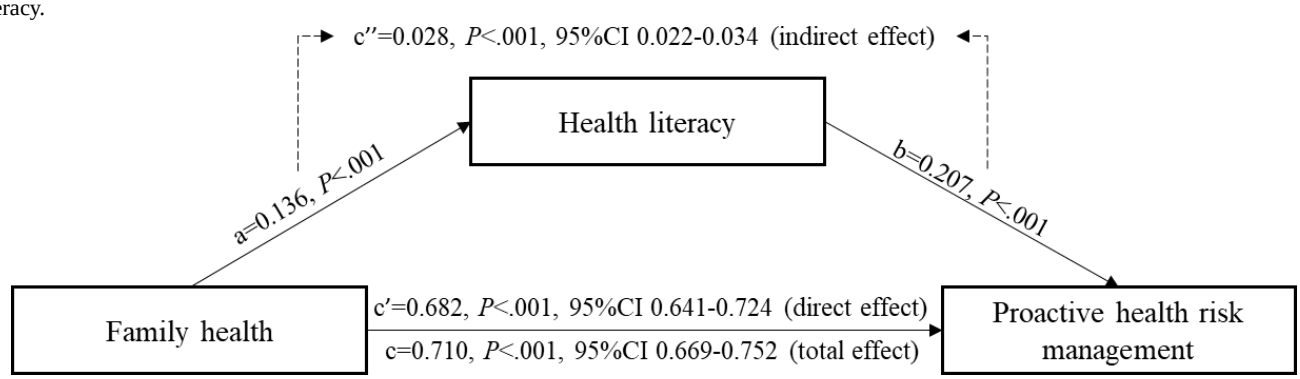
[c]$F_{13, 30,030}$=226.421, *P*<0.001, R^2=0.142.

[d]HL: health literacy.

[e]$F_{14, 30,029}$=164.701, *P*<0.001, R^2=0.112.

[f]Not applicable.

Figure 3. Mediating model of the effect of family health on proactive health risk management via health literacy in this cross-sectional study based on the Psychology and Behavior Investigation of Chinese Residents (2023; N=30,044). "a" represents the coefficient between family health and health literacy; "b" represents the coefficient between health literacy and proactive health risk management; and "c" represents the total coefficient when combined direct coefficient (c') and from family health and proactive health risk management and indirect coefficient (c") between them via health literacy.



Moderated Mediation Analyses

To verify the moderated mediation model while controlling for all covariates, we used model 7 of the SPSS macro PROCESS compiled by Hayes. The results of the family communication moderation test are shown in Table 3 and Figure 4. After putting family communication into the model, the product (interaction term) of FH and family communication had a significant predictive effect on health literacy (β=.117, 95% CI 0.105-0.130), as shown in model 4 of Table 3. The result showed that family communication played a moderating role in the association between FH and health literacy. Higher levels of

XSL•FO
RenderX

family communication exacerbated the effects of FH on health literacy, as evidenced by the fact that the linear slope for higher levels of family communication is steeper than that for lower levels in Figure 4. Furthermore, the effect size estimate for the moderated mediating effect was 0.024, and the 95% CI was (0.019-0.031), excluding 0, which also supported the existence of a moderated mediation effect in this model.

Table . Results of the moderated mediation model of family communication and health literacy in the association between family health (FH) and the proactive health risk management index (PHRMI) in this cross-sectional study (2023, China; N=30,044). All models incorporated covariates.

Predictors	Model 4[a] (DV[b]=HL[c]), β (95% CI)	Model 5[d] (DV=PHRMI[e]), β (95% CI)
FH	.140 (0.126-0.153)	.682 (9.641-0.724)
HL	—[f]	.207 (0.168-0.245)
FC[g]	.029 (0.019-0.153)	—
FH×FC[h]	.117 (0.105-0.130)	—

[a]$F_{15, 30,028}$=224.369; R^2=0.152.

[b]DV: dependent variable.

[c]HL: health literacy.

[d]$F_{14, 30,029}$=164.701; R^2=0.112.

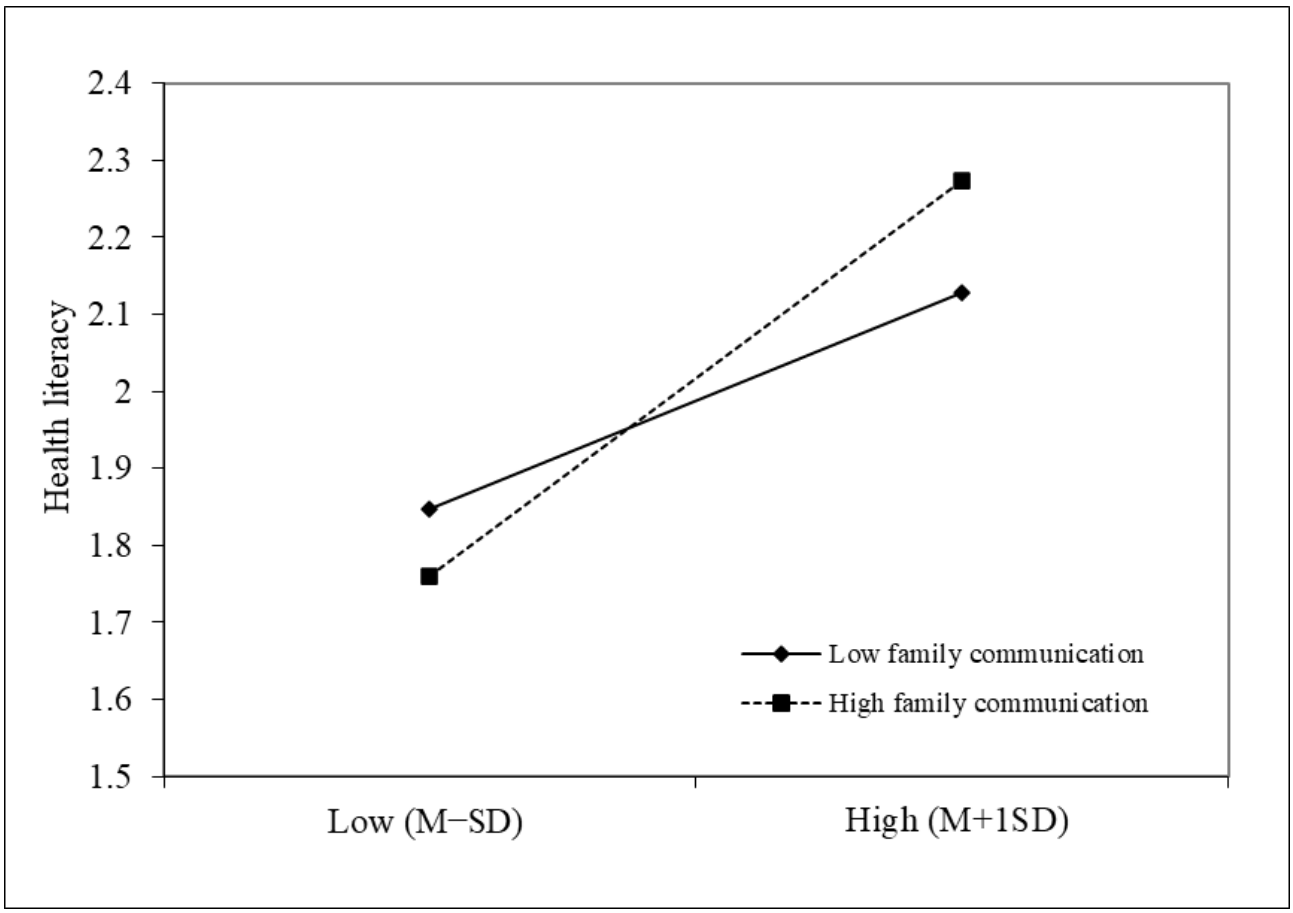
[e]PHRMI:proactive health risk management index

[f]Not applicable.

[g]FC: family communication.

[h]FH×FC: interaction of family health and family communication.

Figure 4. Simple slope analysis shows that family communication moderated the relationship between family health and health literacy in this cross-sectional study (2023, China; N=30,044). The function was graphed for 2 levels of the independent variable and moderator: 1 SD above the mean and 1 SD below the mean.



XSL•FO
RenderX

Sensitivity Analysis

In the sensitivity analysis (Table S2 in Multimedia Appendix 1), we assessed proactive health risk management using generalized anxiety disorder instead of depression. The result was consistent with that of the aforementioned analysis. Upon replacing depression with generalized anxiety disorder, the results revealed that better FH was associated with better health literacy (β=.198, 95% CI 0.160-0.236) and further increased proactive health risk management (β=.630, 95% CI 0.589-0.671), with a significant mediating effect. The moderating influence of family communication on the association between FH, health literacy, and the PHRMI was not assessed in this study, as variations in proactive health risk management did not affect the moderating effect of family communication.

Subgroup Analysis

To clarify further the association between FH and the PHRMI, we performed subgroup analyses in different age groups, including "aged 18 - 44 years," "aged 45 - 59 years," and "aged 60 years and older." First, the positive associations between FH and the PHRMI were significant in all subgroups (for those aged 18 - 44 y: β=.708, 95% CI 0.650-0.767; for those aged 45 - 59 y: β=.653, 95% CI 0.574-0.731; for those aged ≥60 y: β=.732, 95% CI 0.637-0.827). Notably, the mediating effect of health literacy remained significant in individuals aged 18 to 44 and 45 to 59 years. However, for those aged 60 years or older, the association between health literacy and the PHRMI (β=.073, 95% CI −0.016 to 0.161) and the mediating effect of health literacy between FH and the PHRMI (β=.066, 95% CI −0.024 to 0.157) was not significant. The details are presented in Table S2 in Multimedia Appendix 1.

Discussion

Principal Findings

This study revealed that increased FH was linked to an increased PHRMI (β=.710; *P*<.001). Health literacy acted as a mediator in this relationship (β=.207; *P*<.001). Notably, while both the direct and indirect paths reached statistical significance, their magnitudes differed substantially. The strong direct association of FH with the PHRMI dominated the mediation pathway. Furthermore, higher-level family communication shows a small but statistically significant positive association with health literacy (β=.029; *P*<.001) while simultaneously intensifying the promoting effect of FH on it (β=.117; *P*<.001).

Comparison With Prior Work

The study confirms a significant effect of family elements on an individual's health behaviors or outcomes. This aligns with the findings of most existing studies. Previous studies have already shown that various elements of family, such as family structure [60], sibling relationships [61], family social status [62], FH climate [63], family social capital [64], and so on, influence an individual's health. A similar effect also reflects on the influence on health behavior [65-68]. However, few studies have integrated family-level socialization dynamics, as described in family system theories [69], and explored the association between them and health or health behaviors. This critical deviation may lead to misjudging the role of family in members' health, limiting our ability to develop health promotion programs, policies, and goals and even blurring intervention priorities [70]. In this study, we used FH, encapsulating structure, function, and social network, which moves beyond examining isolated family factors to capture the holistic, systemic dynamics emphasized in family system theories. We discovered the positive association between FH and the PHRMI, aligning with a prior study [38]. In contrast, this study aims to examine the association between FH and clustered health behaviors, while the other focuses on exploring the mediating role of four independent health behaviors between FH and frailty. Currently, little research has examined how family-level integrated elements affect clustered health behaviors. A key innovation of this study is the construction of the PHRMI, encompassing important health risk factors. To some extent, it decreases the interdependencies among health behaviors that are often overlooked in single health behavior research. Consequently, this strengthens the argument that fostering overall FH may be a potent strategy for promoting proactive health behavior changes [71,72].

Higher health literacy is associated with increased participation in healthy behavior [30]. This research extends the outcome variables of health literacy by demonstrating that health literacy is positively associated with the PHRMI. In addition, individuals with lower FH tend to exhibit lower health literacy. Another cross-sectional survey in China also confirmed the positive association between FH and health literacy [73]. One possible explanation is that individuals with fewer family resources or family support may lack confidence and thus be unable to independently address health difficulties and manage health risk factors. In addition, health literacy mediates the positive association between FH and the PHRMI. Notably, the mediation analysis showed that the mediator only accounted for 4% of the total effect of FH on the PHRMI, suggesting a relatively modest mediated effect. FH is a complex and comprehensive concept influenced by multiple factors across family and individual domains. Complex familial interactions affect and shape individual health behavior through a variety of underlying mechanisms [74]. While our findings support the association between health literacy and proactive health risk management, exploring additional clinical, psychological, or physiological pathways remains an important direction for future research. On the other hand, family communication moderated the mediating effect of health literacy between FH and the PHRMI. This means the positive association between FH and health literacy was enhanced as family communication levels increased. The circumplex model proposes that families with a balance of cohesion and communication are more likely to set and achieve goals [75] by effectively integrating family members' individuality and independence. Therefore, a focus on both the family and the individual for health promotion could be more effective than efforts targeted solely at the individual.

Regarding the effects of variation among different age groups, subgroup analysis results showed that FH had a stronger positive effect on the PHRMI among older adult individuals compared to younger ones. Presumably, this result is due to the fact that older adults receive health information from a relatively singular

XSL•FO
RenderX

source [76,77]. As age increases, due to both subjective and objective factors such as the lack of credible internet-based health information sources [78], low-level electronic health literacy [79,80], and declines in cognitive [81,82] and physical function [83], the ability of older adults to discriminate health information gradually declines [84]. Therefore, they prefer to trust health interventions recommended by trusted family members [85]. In addition, consistent with previous studies, males tend to exhibit poorer health behaviors. Males' PHRMI scores are lower than females' at any age, indicating the need to strengthen males' health interventions in particular. We find that low self-assessed social status and resource-limited economic status are associated with lower FH and PHRMIs. Self-assessed social status was statistically associated with health literacy and the PHRMI, which is consistent with previous studies**.** Our study provides further empirical evidence that health interventions are not only about health but also about improving one's economic and social status. For instance, evidence from China shows that the national program to stabilize poverty alleviation has a significant health benefit, particularly for marginal poverty-prone households and people in resource-limited areas.

Limitations

There are some limitations to this study that need further investigation in future research. First, although this study designed the proactive health risk management score using validated scales and indicators, the weighting coefficients between indicators were not addressed in the calculation process. Second, the cross-sectional research design means that causality cannot be deduced from the results. Future research could adopt a longitudinal design to explore the causal association between FH and proactive health risk management. Ultimately, some of the scales used to construct the proactive health risk management score in this paper, such as the B-PSQI and the IPAQ-SF, were completed by patients relying on their recall; this may introduce bias compared to the real situation. Future research could consider collecting objective indicators for the exercise and sleep dimensions.

Conclusions

In conclusion, FH can be an important intervention target that appears to be positively linked to proactive health risk management. The mediating role of health literacy suggests its potential as a pathway through which improved FH may correspond to enhanced proactive health risk management. Our study provides evidence that maintaining FH is advantageous for promoting healthy behavior in Chinese individuals, particularly within the older adult population. Meanwhile, in the context of the family, health interventions should address a wider range of familial factors, such as family communication, family healthy lifestyle, FH resource accessibility, and family external social support, rather than focusing solely on health-related interventions.

Acknowledgments

The authors are extremely grateful to all participants who took part in the study for their time and involvement. The authors would also like to thank the survey team for conducting the survey and cleaning the data.

Funding

This work was supported by the National Key Research and Development Program of China (grant: 2023YFC3606400 and 2022YFE0133000), National Nature Foundation of China (grant 72004073 and 72374079), and Research Project of Humanities and Social Sciences of the Ministry of Education (grant 20YJC630134). The funder had no involvement in the study design; the data collection, analysis, and interpretation; or the writing of the manuscript.

Data Availability

The data used in this study are publicly accessible at Psychology and Behavior Investigation of Chinese Residents (PBICR) [86]. The PBICR data are managed by the Psychology and Behavior Investigation of Chinese Residents Project Team and are not publicly available. Researchers interested in accessing PBICR data for replication or further analysis need to submit a data request to the project team (email: pbicr_zgjmxlyxwdc@163.com).

Authors' Contributions

All authors reviewed and edited the manuscript. Y Wei and ST conceptualized the study. YW performed data analysis and wrote the original draft. Y Wu, HD, KZ, JS, HQ, and CW contributed to the writing of the manuscript. ST was responsible for funding acquisition. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Co-corresponding author: Yibo Wu, School of Public Health, Peking University, No.38, Xueyuan Road, Beijing 100191, China, bjmuwuyibo@outlook.com

Conflicts of Interest

None declared.

Multimedia Appendix 1

XSL•FO
RenderX

Additional data tables.
[DOCX File, 21 KB - publichealth_v12i1e73659_app1.docx]

References

1. James E, Freund M, Booth A, et al. Comparative efficacy of simultaneous versus sequential multiple health behavior change interventions among adults: a systematic review of randomised trials. Prev Med 2016 Aug;89:211-223. [doi: 10.1016/j.ypmed.2016.06.012] [Medline: 27311332]
2. Ng R, Sutradhar R, Yao Z, Wodchis WP, Rosella LC. Smoking, drinking, diet and physical activity-modifiable lifestyle risk factors and their associations with age to first chronic disease. Int J Epidemiol 2020 Feb 1;49(1):113-130. [doi: 10.1093/ije/dyz078] [Medline: 31329872]
3. Lacombe J, Armstrong MEG, Wright FL, Foster C. The impact of physical activity and an additional behavioural risk factor on cardiovascular disease, cancer and all-cause mortality: a systematic review. BMC Public Health 2019 Jul 8;19(1):900. [doi: 10.1186/s12889-019-7030-8] [Medline: 31286911]
4. Hanson M, Gluckman P, Nutbeam D, Hearn J. Priority actions for the non-communicable disease crisis. Lancet 2011 Aug 13;378(9791):566-567. [doi: 10.1016/S0140-6736(11)61286-5] [Medline: 21840453]
5. Li Y, Pan A, Wang DD, et al. Impact of healthy lifestyle factors on life expectancies in the US population. Circulation 2018 Jul 24;138(4):345-355. [doi: 10.1161/CIRCULATIONAHA.117.032047] [Medline: 29712712]
6. Bauer UE, Briss PA, Goodman RA, Bowman BA. Prevention of chronic disease in the 21st century: elimination of the leading preventable causes of premature death and disability in the USA. Lancet 2014 Jul 5;384(9937):45-52. [doi: 10.1016/S0140-6736(14)60648-6] [Medline: 24996589]
7. Yilin W, Li W, Fangfei C. An active health behavior scale for hypertensive patients: development and reliability and validity evaluation [Article in Chinese]. Chin J Public Health 2023;39(3):370-374. [doi: 10.11847/zgggws1139802]
8. Liu J, Li W, Yao H, Liu J. Proactive health: an imperative to achieve the goal of healthy China. China CDC Wkly 2022 Sep 9;4(36):799-801. [doi: 10.46234/ccdcw2022.156] [Medline: 36285279]
9. Liujue L, Yaohongyan L. Progress and prospects of active health research [Article in Chinese]. Chin Prev Med 2024;24:750-752 [FREE Full text] [doi: 10.16506/j.1009-6639]
10. Lvlin XXZ. Value realization of basic-level health services from the perspective of active health frontier exploration and policy implication [Article in Chinese]. Chin Public Adm 2023;39:102-108. [doi: 10.19735/j.issn.1006-0863.2023.11.13]
11. Liao Y, Hongwu L, Xianliang W. The theoretical value, realistic dilemma, and relief strategy of implementing active health strategy in China [Article in Chinese]. Chin Health Econ 2023;42(2023):1-5 [FREE Full text]
12. Chaoyi C, Hao C, Zhanchun F, et al. Exploring the connotation and optimization pathways of proactive health management from the perspective of digital health [Article in Chinese]. Chin Prev Med 2024;25:1463-1468. [doi: 10.16506/j.1009-6639.2024.11.021]
13. Wangyu LI, Hua JIN, Dehua YU. Implementation strategies of community-based proactive care management [Article in Chinese]. Chin Prev Med 2022;25:3928-3932 [FREE Full text]
14. Jianping H, Ling X, Wen F, Xiaohui Q, Wensheng H, Qun M. Researches on the theoretical framework of initiative medical health service model [Article in Chinese]. Chin J Health Inform Manag 2016;13:227-231 [FREE Full text]
15. Mingliang SU, Shiquan W, Wei LI. Development of hospital-sanatorium intelligent service management platform based on active health access technology [Article in Chinese]. Chin Med Equip J 2023;40:31-35. [doi: 10.19745/j.1003-8868.2019137]
16. Fan K, Lu Z, Li S, Yu L, Hu T, Zhou Y. Development and reliability and validity testing of the proactive health behavior scale for adolescents [Article in Chinese]. Mod Prev Med 2024;51:3529-3534. [doi: 10.20043/j.cnki.MPM.202405382]
17. Liyu C, Xue X, Hong Z, et al. The level and factors associated with active health behavior capability among housebound disabled older adults [Article in Chinese]. J Nurs Sci 2025;40:24-28 [FREE Full text]
18. Cheng X. Development and reliability and validity evaluation of an active health behavior assessment seale for patients with coronary heart disease [Article in Chinese]. Chin Clin Nurs 2024;16:672-676 [FREE Full text]
19. Barnes MD, Hanson CL, Novilla LB, Magnusson BM, Crandall AC, Bradford G. Family-centered health promotion: Perspectives for engaging families and achieving better health outcomes. Inquiry 2020;57:46958020923537. [doi: 10.1177/0046958020923537] [Medline: 32500768]
20. Provini F. Sleep research in 2023: Beyond the bare necessities. Lancet Neurol 2024 Jan;23(1):25-27. [doi: 10.1016/S1474-4422(23)00466-0] [Medline: 38101892]
21. Arora T, Grey I, Östlundh L, et al. A systematic review and meta-analysis to assess the relationship between sleep duration/quality, mental toughness and resilience amongst healthy individuals. Sleep Med Rev 2022 Apr;62:101593. [doi: 10.1016/j.smrv.2022.101593] [Medline: 35462348]
22. Huang BH, Duncan MJ, Cistulli PA, Nassar N, Hamer M, Stamatakis E. Sleep and physical activity in relation to all-cause, cardiovascular disease and cancer mortality risk. Br J Sports Med 2022 Jul;56(13):718-724. [doi: 10.1136/bjsports-2021-104046] [Medline: 34187783]

XSL•FO
RenderX

23. Cappuccio FP, Cooper D, D'Elia L, Strazzullo P, Miller MA. Sleep duration predicts cardiovascular outcomes: a systematic review and meta-analysis of prospective studies. Eur Heart J 2011 Jun;32(12):1484-1492. [doi: 10.1093/eurheartj/ehr007] [Medline: 21300732]
24. Bassetti CLA. Sleep and stroke: A bidirectional relationship with clinical implications. Sleep Med Rev 2019 Jun;45:127-128. [doi: 10.1016/j.smrv.2019.04.005] [Medline: 31080163]
25. Zhou L, Yu K, Yang L, et al. Sleep duration, midday napping, and sleep quality and incident stroke: the Dongfeng-Tongji cohort. Neurology (ECronicon) 2020 Jan 28;94(4):e345-e356. [doi: 10.1212/WNL.0000000000008739] [Medline: 31827003]
26. Secker-Walker RH, Flynn BS, Solomon LJ, Skelly JM, Dorwaldt AL, Ashikaga T. Helping women quit smoking: results of a community intervention program. Am J Public Health 2000 Jun;90(6):940-946. [doi: 10.2105/ajph.90.6.940] [Medline: 10846513]
27. Ho YCL, Mahirah D, Ho CZH, Thumboo J. The role of the family in health promotion: a scoping review of models and mechanisms. Health Promot Int 2022 Dec 1;37(6):daac119. [doi: 10.1093/heapro/daac119] [Medline: 36398941]
28. Chudasama YV, Khunti K, Gillies CL, et al. Healthy lifestyle and life expectancy in people with multimorbidity in the UK Biobank: a longitudinal cohort study. PLoS Med 2020 Sep;17(9):e1003332. [doi: 10.1371/journal.pmed.1003332] [Medline: 32960883]
29. Health literacy. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/health-literacy [accessed 2025-07-20]
30. Svendsen MT, Bak CK, Sørensen K, et al. Associations of health literacy with socioeconomic position, health risk behavior, and health status: a large national population-based survey among Danish adults. BMC Public Health 2020 Apr 28;20(1):565. [doi: 10.1186/s12889-020-08498-8] [Medline: 32345275]
31. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. Ann Intern Med 2011 Jul 19;155(2):97-107. [doi: 10.7326/0003-4819-155-2-201107190-00005] [Medline: 21768583]
32. Heijmans M, Waverijn G, Rademakers J, van der Vaart R, Rijken M. Functional, communicative and critical health literacy of chronic disease patients and their importance for self-management. Patient Educ Couns 2015 Jan;98(1):41-48. [doi: 10.1016/j.pec.2014.10.006] [Medline: 25455794]
33. Wu JR, Holmes GM, DeWalt DA, et al. Low literacy is associated with increased risk of hospitalization and death among individuals with heart failure. J Gen Intern Med 2013 Sep;28(9):1174-1180. [doi: 10.1007/s11606-013-2394-4] [Medline: 23478997]
34. Cho YI, Lee SYD, Arozullah AM, Crittenden KS. Effects of health literacy on health status and health service utilization amongst the elderly. Soc Sci Med 2008 Apr;66(8):1809-1816. [doi: 10.1016/j.socscimed.2008.01.003] [Medline: 18295949]
35. Eichler K, Wieser S, Brügger U. The costs of limited health literacy: a systematic review. Int J Public Health 2009;54(5):313-324. [doi: 10.1007/s00038-009-0058-2] [Medline: 19644651]
36. Lee SYD, Arozullah AM, Cho YI. Health literacy, social support, and health: a research agenda. Soc Sci Med 2004 Apr;58(7):1309-1321. [doi: 10.1016/S0277-9536(03)00329-0]
37. Zambrana RE, Meghea C, Talley C, Hammad A, Lockett M, Williams KP. Association between family communication and health literacy among underserved racial/ethnic women. J Health Care Poor Underserved 2015 May;26(2):391-405. [doi: 10.1353/hpu.2015.0034] [Medline: 25913337]
38. Li H, Wu Y, Bai Z, et al. The association between family health and frailty with the mediation role of health literacy and health behavior among older adults in China: nationwide cross-sectional study. JMIR Public Health Surveill 2023 Jun 27;9:e44486. [doi: 10.2196/44486] [Medline: 37368463]
39. Lau RR, Quadrel MJ, Hartman KA. Development and change of young adults' preventive health beliefs and behavior: influence from parents and peers. J Health Soc Behav 1990 Sep;31(3):240-259. [Medline: 2133479]
40. Thomas V, Ozechowski TJ. A test of the Circumplex Model of Marital and Family Systems using the Clinical Rating Scale. J Marital Fam Ther 2000 Oct;26(4):523-534. [doi: 10.1111/j.1752-0606.2000.tb00321.x] [Medline: 11042844]
41. Gao F, Zhang J. Research on how information attention and family communication patterns influence people's sharing of health information with parents [Article in Chinese]. Chin J Journal Commun 2023;45:117-135. [doi: 10.13495/j.cnki.cjjc.2023.12.007]
42. Zhang N, Qi J, Liu Y, et al. Relationship between big five personality and health literacy in elderly patients with chronic diseases: the mediating roles of family communication and self-efficacy. Sci Rep 2024;14(1):24943. [doi: 10.1038/s41598-024-76623-3]
43. Sit SMM, Gong WJ, Ho SY, et al. A population study on COVID-19 information sharing: sociodemographic differences and associations with family communication quality and well-being in Hong Kong. Int J Environ Res Public Health 2022 Mar 17;19(6):3577. [doi: 10.3390/ijerph19063577] [Medline: 35329264]
44. Shugar G, Krueger S. Aggressive family communication, weight gain, and improved eating attitudes during systemic family therapy for anorexia nervosa. Int J Eat Disord 1995;17:23-31. [doi: 10.1002/1098-108x(199501)17:13.0.co;2-8] [Medline: 7894449]
45. Global status report on noncommunicable diseases 2014. : World Health Organization; 2014 URL: https://www.who.int/publications/i/item/9789241564854 [accessed 2024-09-20]

XSL•FO
RenderX

46. Yang Y, Fan S, Chen W, Wu Y. Broader open data needed in psychiatry: practice from the Psychology and Behavior Investigation of Chinese Residents. Alpha Psychiatry 2024 Aug;25(4):564-565. [doi: 10.5152/alphapsychiatry.2024.241804] [Medline: 39360297]
47. Wu Y, Fan S, Liu D, Sun X. Psychological and behavior investigation of Chinese residents: concepts, practices, and prospects. Chin Gen Pract J 2024 Sep;1(3):149-156. [doi: 10.1016/j.cgpj.2024.07.006]
48. Liu D, Fan S, Huang X, et al. Study protocol: A national cross-sectional study on psychology and behavior investigation of Chinese residents in 2023. Health Care Sci 2024 Dec;3(6):475-492. [doi: 10.1002/hcs2.125] [Medline: 39735279]
49. Jia J, Zhao T, Liu Z, et al. Association between healthy lifestyle and memory decline in older adults: 10 year, population based, prospective cohort study. BMJ 2023 Jan 25;380:e072691. [doi: 10.1136/bmj-2022-072691] [Medline: 36696990]
50. Wang K, Fang Y, Zheng R, et al. Associations of socioeconomic status and healthy lifestyle with incident dementia and cognitive decline: two prospective cohort studies. EClinicalMedicine 2024 Oct;76:102831. [doi: 10.1016/j.eclinm.2024.102831] [Medline: 39318786]
51. Zhang YB, Chen C, Pan XF, et al. Associations of healthy lifestyle and socioeconomic status with mortality and incident cardiovascular disease: two prospective cohort studies. BMJ 2021 Apr 14;373:n604. [doi: 10.1136/bmj.n604] [Medline: 33853828]
52. Craig CL, Marshall AL, Sjöström M, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
53. Lee PH, Macfarlane DJ, Lam TH, Stewart SM. Validity of the International Physical Activity Questionnaire Short Form (IPAQ-SF): a systematic review. Int J Behav Nutr Phys Act 2011 Oct 21;8(1):115. [doi: 10.1186/1479-5868-8-115] [Medline: 22018588]
54. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: Validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613. [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
55. Sancho-Domingo C, Carballo JL, Coloma-Carmona A, Buysse DJ. Brief version of the Pittsburgh Sleep Quality Index (B-PSQI) and measurement invariance across gender and age in a population-based sample. Psychol Assess 2021 Feb;33(2):111-121. [doi: 10.1037/pas0000959] [Medline: 33119375]
56. Wang F, Wu Y, Sun X, et al. Reliability and validity of the Chinese version of a short form of the family health scale. BMC Prim Care 2022 May 6;23(1):108. [doi: 10.1186/s12875-022-01702-1] [Medline: 35524178]
57. Sun X, Lv K, Wang F, et al. Validity and reliability of the Chinese version of the Health Literacy Scale Short-Form in the Chinese population. BMC Public Health 2023 Feb 23;23(1):385. [doi: 10.1186/s12889-023-15237-2] [Medline: 36823591]
58. Hayes AF. Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach: The Guilford Press; 2013:507.
59. Zhao X, Zhang D, Wu M, et al. Loneliness and depression symptoms among the elderly in nursing homes: a moderated mediation model of resilience and social support. Psychiatry Res 2018 Oct;268:143-151. [doi: 10.1016/j.psychres.2018.07.011] [Medline: 30025285]
60. Dong X, Ng N, Santosa A. Family structure and depressive symptoms among older adults in China: a marginal structural model analysis. J Affect Disord 2023 Mar 1;324:364-369. [doi: 10.1016/j.jad.2022.12.067] [Medline: 36566935]
61. Feinberg ME, Solmeyer AR, McHale SM. The third rail of family systems: Sibling relationships, mental and behavioral health, and preventive intervention in childhood and adolescence. Clin Child Fam Psychol Rev 2012 Mar;15(1):43-57. [doi: 10.1007/s10567-011-0104-5] [Medline: 22105663]
62. Rivenbark J, Arseneault L, Caspi A, et al. Adolescents' perceptions of family social status correlate with health and life chances: a twin difference longitudinal cohort study. Proc Natl Acad Sci U S A 2020 Sep 22;117(38):23323-23328. [doi: 10.1073/pnas.1820845116] [Medline: 31907315]
63. Niermann CYN, Kremers SPJ, Renner B, Woll A. Family health climate and adolescents' physical activity and healthy eating: a cross-sectional study with mother-father-adolescent triads. PLoS ONE 2015;10(11):e0143599. [doi: 10.1371/journal.pone.0143599] [Medline: 26606157]
64. Quick V, Delaney C, Eck K, Byrd-Bredbenner C. Family social capital: Links to weight-related and parenting behaviors of mothers with young children. Nutrients 2021 Apr 23;13(5):1428. [doi: 10.3390/nu13051428] [Medline: 33922656]
65. Ali S, Malik JA. Consistency of prediction across generation: explaining quality of life by family functioning and health-promoting behaviors. Qual Life Res 2015 Sep;24(9):2105-2112. [doi: 10.1007/s11136-015-0942-6] [Medline: 25724696]
66. Rossow I. Intrafamily influences on health behavior. A study of interdental cleaning behavior. J Clin Periodontol 1992 Nov;19(10):774-778. [doi: 10.1111/j.1600-051x.1992.tb02169.x] [Medline: 1452803]
67. Stahlmann K, Hebestreit A, DeHenauw S, et al. A cross-sectional study of obesogenic behaviours and family rules according to family structure in European children. Int J Behav Nutr Phys Act 2020 Mar 5;17(1):32. [doi: 10.1186/s12966-020-00939-2] [Medline: 32138743]
68. Poutiainen H, Levälahti E, Hakulinen-Viitanen T, Laatikainen T. Family characteristics and health behaviour as antecedents of school nurses' concerns about adolescents' health and development: a path model approach. Int J Nurs Stud 2015 May;52(5):920-929. [doi: 10.1016/j.ijnurstu.2015.01.001] [Medline: 25744429]

XSL•FO
RenderX

69. Malik N. Encyclopedia of behavioral medicine. In: Gellman MD, editor. Family Systems Theory: Springer; 2020. [doi: 10.1007/978-3-030-39903-0_746]
70. Weiss-Laxer NS, Crandall A, Okano L, Riley AW. Building a Foundation for Family Health Measurement in National Surveys: A Modified Delphi Expert Process. Matern Child Health J 2020 Mar;24(3):259-266. [doi: 10.1007/s10995-019-02870-w] [Medline: 31912378]
71. Cox MJ, Paley B. Families as systems. Annu Rev Psychol 1997;48:243-267. [doi: 10.1146/annurev.psych.48.1.243] [Medline: 9046561]
72. Rhodes RE, Guerrero MD, Vanderloo LM, et al. Development of a consensus statement on the role of the family in the physical activity, sedentary, and sleep behaviours of children and youth. Int J Behav Nutr Phys Act 2020 Jun 16;17(1):74. [doi: 10.1186/s12966-020-00973-0] [Medline: 32539730]
73. Wang D, Sun X, He F, Liu C, Wu Y. The mediating effect of family health on the relationship between health literacy and mental health: a national cross-sectional survey in China. Int J Soc Psychiatry 2023 Sep;69(6):1490-1500. [doi: 10.1177/00207640231166628] [Medline: 37095729]
74. Bowen M. The use of family theory in clinical practice. Compr Psychiatry 1966 Oct;7(5):345-374. [doi: 10.1016/s0010-440x(66)80065-2] [Medline: 5922263]
75. Olson DH, Russell CS, Sprenkle DH. Circumplex Model of Marital and Family Systems: VI. theoretical update. Fam Process 1983 Mar;22(1):69-83. [doi: 10.1111/j.1545-5300.1983.00069.x] [Medline: 6840263]
76. Xing Z, Yuan X, Vizer L. The age-related differences in web information search process. arXiv. Preprint posted online on Oct 26, 2020. [doi: 10.48550/arXiv.2010.13352]
77. Chen Y, Zhuang J, Tan D. Research on the differences of the elders' choice of health information sources in health situations [Article in Chinese]. Libr Inf Serv 2020;64:90-98. [doi: 10.13266/j.issn.0252-3116.2020.21.012]
78. Zhang Y, Sun Y, Xie B. Quality of health information for consumers on the web: a systematic review of indicators, criteria, tools, and evaluation results. J Assoc Inf Sci Technol 2015 Oct;66(10):2071-2084. [doi: 10.1002/asi.23311]
79. Choi NG, Dinitto DM. The digital divide among low-income homebound older adults: internet use patterns, eHealth literacy, and attitudes toward computer/internet use. J Med Internet Res 2013 May 2;15(5):e93. [doi: 10.2196/jmir.2645] [Medline: 23639979]
80. Xie L, Zhang S, Xin M, Zhu M, Lu W, Mo PKH. Electronic health literacy and health-related outcomes among older adults: a systematic review. Prev Med 2022 Apr;157:106997. [doi: 10.1016/j.ypmed.2022.106997] [Medline: 35189203]
81. Yang L, Xu Y, Zhao H, Wang K, Zheng C. Longitudinal patterns of cognitive function and depression: insights from the China Health and Retirement Longitudinal Study. J Glob Health 2025 Feb 28;15:04060. [doi: 10.7189/jogh.15.04060] [Medline: 40017457]
82. Clare L, Wu YT, Teale JC, et al. Potentially modifiable lifestyle factors, cognitive reserve, and cognitive function in later life: a cross-sectional study. PLoS Med 2017 Mar;14(3):e1002259. [doi: 10.1371/journal.pmed.1002259] [Medline: 28323829]
83. Gore PG, Kingston A, Johnson GR, Kirkwood TBL, Jagger C. New horizons in the compression of functional decline. Age Ageing 2018 Nov 1;47(6):764-768. [doi: 10.1093/ageing/afy145] [Medline: 30165372]
84. Grinberg N, Joseph K, Friedland L, Swire-Thompson B, Lazer D. Fake news on Twitter during the 2016 U.S. presidential election. Science 2019 Jan 25;363(6425):374-378. [doi: 10.1126/science.aau2706] [Medline: 30679368]
85. Jin W, Wang J, Hu X. Preference of urban and rural older people in Shandong Province for long-term care insurance: based on discrete choice experiment. Front Public Health 2024;12:1445273. [doi: 10.3389/fpubh.2024.1445273] [Medline: 39473595]
86. Psychology and Behavior Investigation of Chinese Residents (PBICR). URL: https://www.x-mol.com/groups/pbicr [accessed 2026-04-10]

Abbreviations

B-PSQI: Brief version of the Pittsburgh Sleep Quality Index
FCS-SF: Family Communication Scale-Short Form
FH: family health
IPAQ-SF: International Physical Activity Questionnaire—Short Form
PHRMI: proactive health risk management index

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 09.Mar.2025; peer-reviewed by C Onah, K Kaphingst; revised version received 22.Jul.2025; accepted 23.Jul.2025; published 23.Apr.2026.

Please cite as:
Tang S, Wei Y, Dong H, Zhang K, Wang C, Song J, Qing H, Wu Y
The Association Between Family Health and Proactive Health Risk Management With the Mediating Role of Health Literacy: Nationwide Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e73659
URL: https://publichealth.jmir.org/2026/1/e73659
doi:10.2196/73659

© Shangfeng Tang, Yilin Wei, Heng Dong, Kangkang Zhang, Chunying Wang, Jia Song, Hua Qing, Yibo Wu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 23.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Prevalence and Associated Factors of Renal Disease in Saudi Residents Attending Primary Health Care Centers in Riyadh, Saudi Arabia: Cross-Sectional Study

Saad Alshahrani[1], MD; Ashraf A El-Metwally[2,3,4], MD, PhD; Awad Alshahrani[3,5,6], MD; Badr F Al-Khateeb[3,5,6], MD, SBFM, ABFM; Aljohrah Ibrahim Aldubikhi[7], PhD; Khadijah Angawi[8], PhD; Amani Alharthy[9], MPH, DrPH; Lubna Alnaim[3,4,10], PhD; Amal Mousa Zaidan[2,3,4], PhD; Raed Aldahash[3,5,6], MD

[1]Department of Surgery, Division of Urology, College of Medicine, Prince Sattam bin Abdulaziz University, Alkharj, Saudi Arabia

[10]Department of Clinical Nutrition, College of Applied Medical Sciences, King Saud bin Abdulaziz University for Health Science, Al Ahsa, Saudi Arabia

[2]College of Public Health & Health Informatics, King Saud bin Abdulaziz University for Health Science, KSAU-HS, Riyadh, Saudi Arabia

[3]King Abdullah International Medical Research Center, Riyadh, Riyadh Region, Saudi Arabia

[4]Ministry of National Guard-Health Affairs, Riyadh, Saudi Arabia

[5]College of Medicine, King Saud bin Abdulaziz University for Health Science, Riyadh, Saudi Arabia

[6]King Abdulaziz Medical City, Ministry of the National Guard-Health Affairs, Riyadh, Saudi Arabia

[7]College of Health Sciences, Saudi Electronic University, Riyadh, Saudi Arabia

[8]Department of Health Services and Hospital Administration, Faculty of Economics and Administration, King Abdulaziz University, Jeddah, Saudi Arabia

[9]Department of Health Science, College of Health and Rehabilitation Sciences, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia

Corresponding Author:
Ashraf A El-Metwally, MD, PhD
College of Public Health & Health Informatics, King Saud bin Abdulaziz University for Health Science, KSAU-HS, Riyadh, Saudi Arabia

Abstract

Background: Renal disease represents a significant and growing public health concern globally and within Saudi Arabia. Despite the increasing burden of noncommunicable diseases, population-based data on the prevalence and determinants of renal disease in Saudi Arabia remain limited. Understanding epidemiology, including prevalence and associated risk factors of renal disease in the context of Saudi Arabia, is essential for designing preventive and early detection strategies.

Objective: This study aims to estimate the prevalence of renal disease and to identify sociodemographic, behavioral, and clinically associated factors among adults attending primary health care centers (PHCs) in Riyadh, Saudi Arabia.

Methods: A cross-sectional, community-based study was conducted between March 2023 and July 2023 across 48 PHCs within Riyadh, using a consecutive sampling strategy. Adults aged 18 years and older were recruited and completed a validated, interviewer-administered electronic questionnaire assessing sociodemographic characteristics, lifestyle behaviors, and medical history, including comorbid conditions and health care access. Multivariable logistic regression was used to determine factors associated with self-reported renal disease, with results expressed as adjusted odds ratios (AORs) and 95% CIs, ensuring model adequacy and precision.

Results: A total of 14,239 participants were surveyed (n=7519, 52.8% female; mean age 41.6, SD 13.2 y). The prevalence of self-reported renal disease (including chronic kidney disease and kidney stones) was 3.5% (n=504). Individuals aged younger than 50 years had higher odds of renal disease (AOR 1.22, 95% CI 1.02 - 1.47). Female individuals were more likely to report renal disease than males (AOR 1.51, 95% CI 1.24 - 1.84). Participants with health insurance had increased odds (AOR1.74; 95% CI 1.44 - 2.11). Smoking was strongly associated with renal disease (AOR 3.07, 95% CI 2.38 - 3.96). Other important associated factors included comorbidities such as diabetes (AOR 1.51, 95% CI 1.12 - 2.04), hypertension (AOR 2.27, 95% CI 1.67 - 3.08), obesity (AOR 12.75, 95% CI 9.97 - 16.30), hypercholesterolemia (AOR 1.93, 95% CI 1.43 - 2.60), and heart disease (AOR 6.84, 95% CI 5.14 - 9.10).

Conclusions: This large, community-based cross-sectional study identifies a significant burden of renal disease among adults in Riyadh and highlights several modifiable risk factors that could be targeted in preventive health programs. The findings emphasize the importance of integrating renal health screening within PHCs and strengthening public health strategies addressing obesity, hypertension, and smoking. While the study relied on self-reported data without biomarker or clinical confirmation,

XSL•FO
RenderX

potentially underestimating true prevalence and introducing misclassification bias, it provides a crucial population-level baseline that can guide resource allocation and inform the development of nationwide surveillance systems for early detection of renal disease.

(JMIR Public Health Surveill 2026;12:e81849) doi:10.2196/81849

KEYWORDS

renal disease; prevalence; risk factors; Saudi Arabia; primary health care; PHC; associated factors

Introduction

Renal disease represents a significant and growing challenge for health care systems worldwide [1-3]. In its early stages, kidney disease is often asymptomatic, progressing silently until substantial renal function is lost [4]. Without timely detection and management, it may advance to end-stage renal disease, necessitating dialysis or kidney transplantation [5,6]. Traditionally associated with older populations, renal dysfunction increasingly affects younger age groups due to rising rates of diabetes mellitus, hypertension, and obesity, conditions particularly prevalent in Middle Eastern populations [7-9]. Recent global estimates from the Global Burden of Disease Study highlight chronic kidney disease (CKD) among the top 10 causes of mortality worldwide, reflecting a steady increase in prevalence and disability-adjusted life years over the past decade [10,11]. Emerging studies also emphasize the link between metabolic syndrome, cardiovascular disease, and renal dysfunction, reinforcing shared risk factor pathways for CKD progression [12-14].

Saudi Arabia is undergoing rapid demographic, nutritional, and lifestyle transitions that have contributed to the rising burden of noncommunicable diseases, including diabetes and cardiovascular disorders, with downstream effects on renal health [15-17]. Despite expanding tertiary care services, early identification and risk stratification of individuals at increased risk of kidney disease remain limited [18-20]. National reports indicate rising prevalence of CKD and associated risk factors, with projections suggesting substantial increases in the demand for renal replacement therapies over the next decade [21,22]. Previous research has mainly focused on high-risk groups such as patients with diabetes, those on dialysis, or individuals already diagnosed with CKD [23-25]. Population-based data remains scarce, and existing studies show heterogeneity in methodologies, diagnostic criteria, and definitions of kidney dysfunction, limiting comparability across regions and subpopulations [26-28]. Furthermore, few studies have investigated lifestyle and behavioral determinants, such as physical inactivity, dietary patterns, and water consumption that may influence kidney health [29,30].

A critical gap remains in understanding the factors associated with renal disease across diverse demographic and clinical backgrounds in Saudi Arabia. Identifying these factors is essential for targeted interventions, national screening strategies, and efficient health care resource allocation [31]. Health transformation emphasizes preventive health care, precision public health, and early detection programs for chronic diseases, highlighting the importance of timely epidemiological insights into renal disease [31-33]. In this context, this study aims to examine both the prevalence and associated factors of renal disease among adults attending primary health care centers (PHCs) in Riyadh, Saudi Arabia. By identifying demographic, clinical, and lifestyle factors associated with renal dysfunction, this study seeks to provide evidence to inform public health policy, guide clinical practice, and support national prevention programs. Ultimately, our objective is to generate a clearer understanding of the epidemiological landscape of kidney disease in Saudi Arabia and to provide a foundation for future longitudinal and interventional research.

Methods

Study Design, Duration, Setting, and Sampling Strategy

A cross-sectional study was conducted between March 2023 and July 2023 across PHCs in Riyadh, Saudi Arabia. The survey included adult Saudi residents aged 18 years and older who attended PHCs during the study period. To minimize selection bias, consecutive sampling was used whereby all eligible individuals attending the PHCs during the data collection period were invited to participate until the required sample size was achieved. Data collection took place in the outpatient waiting areas of participating PHCs. Trained data collectors approached eligible individuals, explained the purpose of the study, obtained written informed consent, and administered a structured, interviewer-guided electronic questionnaire. The study period spanned from March 1, 2023, to July 31, 2023. This recruitment approach allowed inclusion of a broad and diverse population of PHC attendees across Riyadh, enhancing the representativeness and external validity of the findings. The reporting of this research study adheres to the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines for cross-sectional studies to ensure methodological transparency and completeness of reporting [34]. The completed STROBE checklist is provided as a supplementary file (Checklist 1).

Eligibility Criteria and Study Participants

The study included adult individuals aged 18 years and older who attended selected PHCs in Riyadh during the data collection period. Eligibility was not restricted by nationality or residency status; both Saudi and non-Saudi residents were invited to participate. Exclusion criteria included individuals aged younger than 18 years and PHC staff or employees. Eligible individuals were identified on-site and invited to participate by trained data collectors. Those who declined to provide written informed consent were excluded.

XSL•FO
RenderX

Sample Size

A total of 14,239 participants completed the survey and were included in the final analysis. The sample size was determined using the single population proportion formula, assuming a 50% estimated prevalence of renal disease (for maximum variability), a 95% confidence level, and a 2% margin of error. The minimum required sample was 2,401; however, a substantially larger sample was achieved to enhance precision and generalizability across subgroups.

Study Questionnaire Development, Adaptation, and Validation

Study Questionnaire Development

The survey instrument used in this study was developed to comprehensively assess health-related behaviors, medical history, and sociodemographic characteristics among adults attending PHCs in Riyadh. The use of previously validated survey items, expert review, pilot testing, and reliability assessment was intended to minimize measurement bias and enhance the accuracy and consistency of self-reported data.

Adaptation and Validation

The questionnaire was primarily adapted from existing validated instruments used in population health surveys internationally, ensuring relevance, comparability, and methodological rigor [35-37]. Additional items were included to capture specific national health care priorities and contextual factors relevant to Saudi Arabia. This approach ensured that the tool combined validated measures with region-specific questions to reflect the study objectives accurately.

Translation and Cultural Adaptation

The original English version of the questionnaire underwent a forward translation into Arabic, followed by back-translation into English by independent bilingual experts to confirm semantic equivalence and cultural appropriateness. This process ensured that both language versions conveyed the intended meaning consistently and were suitable for the target population.

Content and Face Validity

Content validity was evaluated by a panel of 15 subject matter experts, including clinicians and public health professionals from multiple regions of Saudi Arabia including nephrologists, urologists, primary care physicians, endocrinologists, cardiologists, and public health professionals. These experts evaluated the questionnaire for clarity, relevance, and suitability, ensuring that items adequately captured factors pertinent to renal health, lifestyle behaviors, and comorbid conditions. Based on their recommendations, several items were refined or removed to improve the instrument's precision and effectiveness. Experts also assessed each item for relevance, clarity, and appropriateness within the national context. Following content validation, the Arabic version of the questionnaire was pilot-tested for face validity among 200 participants to assess clarity, comprehensibility, and overall acceptability. Feedback from pilot participants informed minor revisions to wording and phrasing, enhancing readability and interpretability.

Reliability Assessment

To evaluate consistency, a test-retest reliability assessment was conducted with 100 pilot participants who completed the questionnaire on 2 separate occasions. The second administration was conducted over the phone after minor refinements were made based on participant feedback. The resulting test-retest reliability coefficient of 0.83 demonstrated strong consistency over time.

Pilot Testing and Justification for Site Selection

The pilot study was conducted in Hail City, which was selected by the Central Health Services Reform Management Team due to its demographic and health characteristics being broadly comparable to those of the Riyadh population. Hail's population structure, health care use patterns, and prevalence of major chronic diseases closely mirror those of urban regions such as Riyadh, making it an appropriate and representative site for pretesting the instrument. A total of 100 patients and 20 focus group participants provided feedback on the clarity, comprehensibility, and cultural relevance of the questionnaire. Based on their input, minor revisions were implemented to enhance wording and ease of interpretation. The finalized questionnaire, following validation and pilot testing, was deemed ready for implementation across 48 PHCs in Riyadh.

Data Collection Procedures at PHCs

Data collection was conducted through an interviewer-assisted electronic survey. To reduce information and interviewer bias, data collectors received standardized training, and a structured, interviewer-guided electronic questionnaire was used to ensure uniform administration across all sites. Trained data collectors stationed at PHCs in Riyadh, Saudi Arabia, administered the digital questionnaire using handheld devices such as iPads or Android tablets. Prior to initiating the survey, each data collector verified the eligibility of individuals present at the PHCs, confirming that participants were aged 18 years or older. Eligible individuals were then approached, provided with a clear explanation of the study's purpose and procedures, and asked to give written informed consent. Participation in the survey was entirely voluntary. Those who agreed to participate completed the questionnaire under the guidance of the data collectors, who were available to clarify questions and ensure accurate responses.

The survey instrument captured a wide range of variables. Sociodemographic information included age, sex, marital status, household size, educational attainment, employment status, and self-rated health. Additionally, the questionnaire explored lifestyle and behavioral factors such as smoking habits, fast food consumption, alcohol use, and levels of physical activity or exercise. Participants were also asked to report any known health conditions, including hypertension, diabetes, obesity, and hypercholesterolemia. In total, 14,239 individuals completed the survey and were included in the final dataset for analysis.

Assessments and Data Sources

Outcome Variable

The primary outcome was self-reported renal disease (yes or no), determined based on participants' responses to the question:

XSL•FO
RenderX

"Have you ever been told by a physician that you have any kidney-related condition (such as kidney stones, infection, or chronic kidney disease)?" Responses were coded dichotomously (yes=1; no=0). The outcome, therefore, reflects any type of renal disease diagnosed by a physician, encompassing both medical and surgical conditions. No laboratory confirmation (eg, serum creatinine, estimated glomerular filtration rate [eGFR], and urinalysis) was obtained as part of this study, consistent with the survey-based design. Recall bias was mitigated by limiting outcome assessment to physician-diagnosed conditions and by administering the questionnaire in a supervised setting with trained interviewers available for clarification.

Independent Variables

Independent variables included sociodemographic characteristics, such as age, sex, marital status, educational attainment, employment status, nationality, and health insurance coverage, as well as clinical factors, including the presence of participants' self-reported physician diagnoses such as diabetes, hypertension, obesity, cardiovascular disease, and hyperlipidemia. Age was categorized into 2 groups (<50 and ≥50 y) based on clinical and epidemiological considerations. Evidence from renal disease research indicates that the risk of kidney dysfunction increases substantially after middle age, largely due to cumulative exposure to chronic conditions such as hypertension, diabetes, and cardiovascular disease. This categorization allowed differentiation between younger adults with potential early-onset or lifestyle-related renal risk and older adults at higher risk of age-related and comorbidity-associated renal disease. The selected cutoff is consistent with those used in previous regional and international studies, facilitating comparability and public health interpretation. Lifestyle factors were also assessed, encompassing smoking, dietary habits (eg, fast-food consumption), and physical activity. Perceived health status was measured using a 5-point Likert scale, with the response options scored as follows: 0 ("poor"), 1 ("fair"), 2 ("good"), 3 ("very good"), and 4 ("excellent").

Data Sources

All data were obtained through a structured, interviewer-administered electronic questionnaire. The tool was developed by the Central Health Services Reform Management Team in collaboration with regional experts as part of a national health assessment initiative. The survey collected self-reported information only, and no clinical or laboratory assessments were performed. There were no missing data in this study, as the electronic data collection system was designed to require completion of all questionnaire items before submission.

Statistical Analysis Approach

The distribution of all variables was initially examined using histograms to assess normality. Descriptive statistics were then used to summarize the data. For continuous variables that followed normal distributions, such as age and results were expressed as means and SDs. To facilitate further analysis, age was also categorized into distinct groups. Categorical variables, including insurance status, perceived health, marital status, employment status, and educational level, were summarized using frequencies and percentages. Given that the main outcome variable, renal disease, was dichotomous (yes or no), binary logistic regression was used for analysis. Initially, univariate logistic regression was conducted to explore associations between each independent variable and the likelihood of having renal disease. Variables demonstrating a P value less than 0.25 in the univariate analysis were selected for inclusion in the multivariable logistic regression model. The multivariable logistic regression was then carried out to identify factors associated independently with renal disease, with statistical significance set at $P<.05$.

To enhance model interpretability and stability, a parsimonious modeling strategy was adopted to reduce the risk of overfitting and potential multicollinearity, given the relatively large number of predictors in relation to the number of outcome events. Accordingly, 2 separate multivariable logistic regression models were constructed. The first model included sociodemographic variables, while the second model examined behavioral and comorbidity-related factors. The behavioral or comorbidity model was adjusted for key demographic covariates, specifically age and sex, to account for major confounding effects while preserving model robustness.

Results from the multivariable analyses were reported as adjusted odds ratios (AORs) with corresponding 95% CIs, and statistical significance was set at $P<.05$. All statistical analyses were performed using IBM SPSS Statistics for Windows (version 26).

Ethical Considerations

The study was conducted in accordance with the principles outlined in the Declaration of Helsinki and received approval from the Institutional Review Board of King Fahad Medical City, Riyadh (protocol code: 22-397E; approval date: October 6, 2022). Written informed consent was obtained from all participants prior to data collection, and participants were informed of their right to withdraw from the study at any stage without any consequences. Data were collected anonymously, and no personally identifiable information was recorded to ensure privacy and confidentiality. All electronic data were stored on password-protected devices accessible only to authorized research personnel. Participants did not receive any financial or nonfinancial compensation for their involvement in the study. Furthermore, no images or identifiable participant data were collected or included in this paper or its supplementary materials.

Results

Sociodemographic Profile of Surveyed Saudi Residents

The sociodemographic profile of individuals who participated in the survey at PHCs in Riyadh is detailed in Table 1. The average age of the surveyed population was approximately 59.8 (SD 16.4) years. The largest age bracket comprised individuals aged 50 years or older, representing 66% (n=9391) of the participants, and most of the surveyed individuals were married (n=9300, 65.3%). Regarding educational attainment, 31.7% (n=4509) had attained less than a college education, whereas 51.4% (n=7317) were employed. A significant proportion of

XSL•FO
RenderX

participants perceived their health as very good (n=5076, 35.6%) or excellent (n=4798, 33.7%). The majority of the participants (n=10782, 75.7%) indicated a lack of insurance coverage. The prevalence of smoking was 27.7% (n=3942). In terms of physical activity, a smaller majority (n=8641, 60.7%) reported engaging in exercise. The prevalence of obesity was 5.2% (n=737), whereas the prevalence of diabetes, hypertension, and renal disease was 12.4% (n=1765), 11.1% (n=1580), and 3.5% (n=504), respectively.

XSL•FO
RenderX

Table . Sociodemographic characteristics of Saudi residents (N=14,239) attending primary health care centers in Riyadh, Saudi Arabia, from March 2023 to July 2023: findings from a cross-sectional study on renal disease prevalence and associated factors.

Variable	Value, n (%)
Age (y)	
< 50	4848 (34.0)
≥50	9391 (66.0)
Education attainment	
Less than a college education	4509 (31.7)
At least a college education	9730 (68.3)
Biological sex	
Female	8062 (56.6)
Male	6177 (43.4)
Marital status	
Not married	4939 (34.7)
Married	9300 (65.3)
Employment	
Employed	7317 (51.4)
Unemployed	6922 (48.6)
Perceived health status	
Excellent	4798 (33.7)
Very good	5076 (35.6)
Good	2815 (19.8)
Fair	1256 (8.8)
Poor	294 (2.1)
Insurance coverage	
Yes	3457 (24.3)
No	10,782 (75.7)
Smoking status	
No	10,297 (72.3)
Yes	3942 (27.7)
Exercise	
No	5598 (39.3)
Yes	8641 (60.7)
Obesity	
No	13,502 (94.8)
Yes	737 (5.2)
Diabetes	
No	12,474 (87.6)
Yes	1765 (12.4)
Hypertension	
No	12,659 (88.9)
Yes	1580 (11.1)
Hypercholesterolemia	
No	13,386 (89.4)

XSL•FO
RenderX

Variable	Value, n (%)
Yes	853 (10.6)
Heart disease	
No	13,845 (95.1)
Yes	394 (4.9)
Renal disease	
No	13,735 (96.5)
Yes	504 (3.5)

Sociodemographic Factors Associated With Renal Disease: Univariate and Multivariable Analysis

Univariate Analysis

Table 2 presents the univariable analysis exploring the relationship between various sociodemographic factors and the presence of renal disease among Saudi residents who participated in the survey. Age was significantly associated with renal disease (*P*=.03). Compared to individuals aged over 50 years, those aged less than 50 years had higher odds of renal disease (OR 1.26, 95% CI 1.05 - 1.52). The relatively narrow CI indicates a moderately precise estimate, suggesting that younger participants consistently had greater odds of renal disease across the sample. Sex showed a significant association (*P*<.001), with female individuals having 1.58 times higher odds of renal disease compared to male individuals (OR 1.58, 95% CI 1.31 - 1.91). This estimate is precise, as reflected by the narrow CI, indicating a robust relationship between sex and renal disease. Compared to those with less than a college education, individuals with at least a college education had slightly higher odds (OR 1.15, 95% CI 0.95 - 1.40) of renal disease. However, the CI includes 1.0, suggesting that the association lacks statistical precision and may not be clinically meaningful. Employment status showed a significant association, with unemployed individuals having 1.20 times higher odds of renal disease compared to employed individuals (OR 1.20, 95% CI 1.00 - 1.43). The CI borders unity, indicating limited precision and a possible weak association. Finally, individuals having insurance coverage exhibited 1.71 times higher odds of renal disease compared to those without insurance (OR 1.71, 95% CI 1.42 - 2.06). The narrow CI suggests a highly precise and consistent estimate.

Table . Sociodemographic factors associated with renal disease among Saudi residents attending primary health care centers in Riyadh, Saudi Arabia, from March 2023 to July 2023: results from univariate and multivariable logistic regression analyses in a cross-sectional study.

Associated factors	Univariable analysis		Multivariable analysis	
	OR[a] (95% CI)	*P* value	AOR[b] (95% CI)	*P* value
Age (y) (reference: ≥50)				
<50	1.26 (1.05-1.52)	.01	1.22 (1.02-1.47)	.03
Sex (reference: male)				
Female	1.58 (1.31-1.91)	<.001	1.51 (1.24-1.84)	<.001[c]
Education (reference: less than a college education)				
At least a college education	1.15 (0.95-1.4)	.16	1.16 (0.95-1.43)	.15
Marital status (reference: single)				
Married	0.98 (0.81-1.18)	.83	__[d]	—
Employment status (reference: employed)				
Unemployed	1.2 (1-1.43)	.04	1.16 (0.96-1.42)	.13
Insurance coverage (reference: no)				
Yes	1.71 (1.42-2.06)	<.001	1.74 (1.44-2.11)	<.001[c]

[a]OR: odds ratio.

[b]AOR: adjusted odds ratio, which was calculated after controlling for sociodemographic factors that were significant at *P*<0.25 in univariate analysis

[c]*P*<.05 was considered significant in multivariable analysis.

[d]not applicable.

XSL•FO
RenderX

Multivariable Analysis

After adjusting for the other factors, age was found to be a significant factor for renal disease. Compared to individuals aged over 50 years, those aged less than 50 years had higher odds of renal disease (AOR 1.22, 95% CI 1.02 - 1.47). Sex was also significantly associated with renal disease, with female individuals having 1.51 times higher AOR compared to male individuals (AOR 1.51, 95% CI 1.24 - 1.84). This narrow CI demonstrates high precision and strengthens confidence in the observed sex difference. Insurance coverage remained significantly associated with renal disease, with individuals having insurance coverage exhibiting 1.74 times higher AOR compared to those without insurance (AOR 1.74, 95% CI 1.44 - 2.11). The tight CI further supports the precision and consistency of this association. Education and employment status did not retain statistical significance in the multivariable model as shown in Table 2.

Behavioral Risk Factors and Comorbidities Associated With Renal Disease: Univariate and Multivariable Analysis

Univariate Analysis

Individuals who reported smoking had 5.22 times higher odds of having renal disease compared to non-smokers (OR 5.22, 95% CI 4.33 - 6.29). The narrow CI indicates a highly precise and consistent estimate. Participants with diabetes had 5.12 times higher odds of renal disease compared to those without diabetes (OR 5.12, 95% CI 4.25 - 6.16), demonstrating good precision and suggesting a strong, stable association. The presence of hypertension was associated with 10.11 times higher odds of renal disease (OR 10.11, 95% CI 8.41 - 12.15). The relatively narrow CI supports the reliability of this association. Obese individuals had a substantially higher odds of renal disease at 38.64 times compared to nonobese individuals (OR 38.64, 95% CI 31.65 - 47.18). Although the CI range is wider than others, it remains far from unity, reflecting a strong and highly significant relationship. Hypercholesterolemia was associated with 14.83 times higher odds of renal disease (OR 14.83, 95% CI 12.30 - 17.87), indicating a consistent and precise association. Finally, individuals with heart disease had a remarkably higher odds of renal disease at 40.69 times compared to those without heart disease (OR 40.69, 95% CI 33.12 - 49.99). Despite a wider CI, the estimate remains highly precise due to its magnitude and narrow relative interval.

Multivariable Analysis

After adjusting for the other variables in the model, smoking remained significantly associated with renal disease ($P<.001$), with smokers having 3.07 times higher AOR of renal disease compared to nonsmokers (AOR 3.07, 95% CI 2.38 - 3.96). The narrow CI reflects a precise estimate, confirming the robustness of the association. Diabetes remained significantly associated with renal disease, with individuals with diabetes having 1.51 times higher adjusted odds of renal disease (AOR 1.51, 95% CI 1.12 - 2.04) compared to those without diabetes. The CI does not cross 1.0, indicating a statistically meaningful and fairly precise association. Hypertension remained significantly associated with renal disease ($P<.001$), with individuals with hypertension having 2.27 times higher adjusted odds of renal disease (AOR 2.27, 95% CI 1.67 - 3.08) compared to those without hypertension. The narrow CI suggests precision and reinforces the strength of this relationship. Obesity remained strongly associated with renal disease ($P<.001$), with obese individuals having 12.75 times higher adjusted odds of renal disease (AOR 12.75, 95% CI 9.97 - 16.30) compared to nonobese individuals. Although the CI is somewhat wide, the lower bound is well above 1, indicating a very strong and precise association. Hypercholesterolemia remained significantly associated with renal disease ($P<.001$), with individuals having hypercholesterolemia exhibiting 1.93 times higher adjusted odds of renal disease (AOR 1.93, 95% CI 1.43 - 2.60) compared to those without hypercholesterolemia. The relatively narrow CI indicates good precision and reliability. Finally, heart disease remained significantly associated with renal disease ($P<.001$), with individuals reporting heart disease exhibiting 6.84 times higher adjusted odds of renal disease (AOR 6.84, 95% CI 5.14 - 9.10), as shown in Table 3. The narrow CI demonstrates strong precision and reinforces the robustness of this association.

XSL•FO
RenderX

Table . Behavioral risk factors and comorbidities associated with renal disease among Saudi residents attending primary health care centers in Riyadh, Saudi Arabia, from March 2023 to July 2023: univariate and multivariable logistic regression findings from a cross-sectional study.

Associated factors	Univariate analysis		Multivariable analysis	
	OR[a] (95% CI)	*P* value[b]	AOR[c] (95% CI)	*P* value[d]
Smoking (reference: no)				
Yes	5.22 (4.33-6.29)	<.001	3.07 (2.38-3.96)	<.001
Diabetes (reference: no)				
Yes	5.12 (4.25-6.16)	<.001	1.51 (1.12-2.04)	<.001
Hypertension (reference: no)				
Yes	10.11 (8.41-12.15)	<.001	2.27 (1.67-3.08)	<.001
Obesity (reference: no)				
Yes	38.64 (31.65-47.18)	<.001	12.75 (9.97-16.3)	<.001
Hypercholesterolemia (reference: no)				
Yes	14.83 (12.3-17.87)	<.001	1.93 (1.43-2.6)	<.001
Heart disease (reference: no)				
Yes	40.69 (33.12-49.99)	<.001	6.84 (5.14-9.1)	<.001

[a]OR: odds ratio.

[b]*P* value of <.25 was considered significant in univariable analysis.

[c]AOR: adjusted odds ratio, which was calculated in the multivariable analysis after controlling for sex, age, and mutual adjustment with other behavioral risk factors and comorbidities.

[d]*P* value of <.05 was considered significant in multivariable analysis.

Discussion

Principal Findings

This study aimed to investigate the prevalence and associated factors of self-reported, physician-diagnosed renal disease among adults attending PHCs in Riyadh, Saudi Arabia. Renal disease in this study encompassed both medical and surgical conditions, as reported by participants, without laboratory confirmation of diagnosis. The findings revealed a self-reported renal disease prevalence of 3.5% (n=504), which reflects disease awareness rather than actual prevalence, given that diagnoses were based solely on self-report. Multiple demographic, behavioral, and clinical factors, including age, sex, insurance coverage, smoking, diabetes, hypertension, obesity, hypercholesterolemia, and heart disease, were independently associated with higher odds of renal disease. These findings underscore the multifactorial nature of renal pathology and provide valuable insights into population-level risk determinants that can inform public health and clinical strategies in Saudi Arabia.

The observed prevalence of 3.5% (n=504) reflects the burden within a primary care-attending population, consistent with reports from other regional studies. While direct comparisons are complicated by differences in diagnostic criteria and population characteristics [26-28], this prevalence aligns with previous estimates indicating a notable burden of kidney dysfunction ranging between 4.76% to 10% [29,38,39]. Although comparable, this prevalence likely under-represents the true magnitude of renal disease, the outcome measure relied entirely on self-report, and it likely underestimates true disease prevalence and instead reflects participants' awareness of having a renal condition. The absence of objective biomarkers such as serum creatinine, eGFR, or urinalysis limits diagnostic precision and may have led to both under- and over-reporting. This limitation emphasizes the need for population-based screening incorporating objective measures to capture the full spectrum of renal impairment within the community.

An unexpected finding was that individuals aged younger than 50 years reported higher odds of renal disease compared to those over 50 years. This result diverges from conventional epidemiological evidence, where older age is typically associated with higher CKD risk due to cumulative exposure to risk factors and age-related decline in renal function [40,41]. A plausible explanation is that younger, more educated, and health-conscious individuals are more likely to recognize and report a previous renal diagnosis, leading to differential awareness rather than a true shift in disease burden. Additionally, this pattern may reflect the contribution of renal stones, conditions more common among younger and middle-aged adults in hot climates like Riyadh, which could explain much of this association. These findings highlight the importance of distinguishing between chronic and acute, as well as surgical renal conditions in future studies.

Female individuals were found to have a significantly higher likelihood of renal disease compared to male individuals. This aligns with prior studies suggesting that women may be at increased risk due to biological differences, including hormonal influences on kidney function, higher tendency to develop kidney stones, or sociocultural factors influencing health care−seeking behavior and disease reporting [39,42-44]. This finding also highlights the importance of gender-sensitive

XSL•FO
RenderX

approaches to renal health promotion and targeted screening for women at risk.

Participants with health insurance exhibited higher odds of renal disease. This is likely attributable to improved health care access and diagnostic opportunities among insured individuals, leading to higher detection and reporting rates rather than increased biological risk [45-47]. This finding supports the importance of health care access in promoting disease recognition and timely intervention.

Smoking was strongly associated with renal disease, with smokers having over 3 times the odds compared to nonsmokers. This is consistent with a robust body of evidence linking tobacco to kidney damage through mechanisms including increased oxidative stress, inflammation, and vascular injury [48-50]. Smoking accelerates the progression of both diabetic and hypertensive nephropathy and is a modifiable risk factor that should be a primary focus of prevention campaigns [48,49]. The high odds ratio observed underscores the critical need for integrating smoking cessation programs within renal disease prevention efforts in Saudi Arabia.

Several chronic health conditions were significantly associated with renal disease in this study, consistent with global literature. Individuals with diabetes had 1.5 times higher odds of renal disease, reaffirming diabetes as a leading cause of chronic kidney disease worldwide as found in other similar studies [43,51]. Poor glycemic control contributes to microvascular damage in the kidneys, leading to diabetic nephropathy [52]. Given the high prevalence of diabetes in Saudi Arabia, intensified efforts toward diabetes management and renal monitoring are essential. Consistent with other studies, hypertension was also associated with renal disease [39,42,43]. Hypertension contributes to kidney damage via increased glomerular pressure and vascular remodeling. The bidirectional relationship between hypertension and renal disease calls for integrated management approaches. However, as our outcome combined multiple renal conditions, these associations may reflect general comorbidity clustering or increased diagnostic attention among individuals with multiple chronic diseases.

As found in previous literature, obesity was strongly associated with renal disease indicating its substantial role in kidney disease risk [43]. Obesity induces renal hyperfiltration and promotes inflammation, hypertension, and diabetes, all of which contribute to renal impairment. This finding is particularly relevant given rising obesity rates in Saudi Arabia and suggests that weight management should be a cornerstone of kidney disease prevention. Individuals with high cholesterol levels had nearly twice the odds of renal disease as found in other studies [53,54]. Dyslipidemia can contribute to renal damage through atherosclerotic changes and glomerular injury, supporting the importance of lipid control in at-risk populations [55]. The presence of heart disease was strongly associated with renal disease, increasing the odds nearly 7-fold. This relationship reflects the cardiorenal syndrome, where heart and kidney dysfunction coexist and exacerbate each other [56]. Patients with cardiovascular disease require careful renal monitoring and vice versa. Moreover, the identified associated factors largely concur with international studies emphasizing the interplay between metabolic, cardiovascular, and behavioral risk factors in renal disease pathogenesis. Similar findings in Middle Eastern cohorts have reported obesity, hypertension, diabetes, and smoking as key drivers of kidney disease [57,58].

Overall, the findings should be interpreted as associations with self-reported awareness of kidney-related diagnoses rather than confirmed clinical disease. The study highlights the importance of health literacy, access to care, and lifestyle risk factors in shaping population awareness of renal health in Saudi Arabia.

Strengths and Limitations

One of the major strengths of this study lies in its robust sampling methodology with a large and representative sample of individuals accessing primary health care services across Riyadh's diverse urban and suburban populations. This sampling strategy enhances the generalizability of the findings within the region and provides a solid basis for informing regional health care planning and policy development. Additionally, the study benefits from comprehensive data collection encompassing a wide range of sociodemographic, behavioral, and clinical variables. The rigorous validation and reliability assessment of the survey instrument ensured high-quality data, strengthening the internal validity of the study. The application of multivariable logistic regression allowed for the identification of independently associated factors of renal disease, providing nuanced insights into the complex interplay of risk factors in this population.

Despite its strengths, several limitations should be noted. The cross-sectional design precludes causal inference and limits understanding of temporal relationships between exposures and renal disease. First, the definition of "renal disease" in this study was based on self-reported, physician-diagnosed conditions without specifying subtypes (eg, chronic kidney disease, renal stones, or infections), introducing potential ambiguity in interpretation. Second, the absence of laboratory or imaging confirmation (eg, serum creatinine, eGFR, and urinalysis) limits diagnostic accuracy and may have resulted in both underestimation and misclassification of true disease prevalence. Third, reliance on self-reported data introduces the possibility of recall and reporting bias, as participants' awareness and understanding of their renal condition may vary by demographic and educational background. Fourth, the sample's relatively high mean age may reflect a degree of sampling bias, as older individuals are more likely to attend PHCs, potentially influencing prevalence estimates. Although these limitations are acknowledged, reliance on self-report also reflects real-world challenges in laboratory-based surveillance within many PHC settings. Some observed ORs, particularly for obesity, are notably high and may reflect residual confounding from unmeasured factors such as diet, physical activity, socioeconomic status, or genetic predisposition. Nonetheless, the findings remain highly relevant for public health planning, providing the first large-scale population estimate of self-reported renal disease in Riyadh and serving as a benchmark for future screening and surveillance efforts that incorporate objective clinical measures.

XSL•FO
RenderX

Policy Implications

While this study provides valuable insight into self-reported renal disease among PHC attendees in Riyadh, the findings should be interpreted cautiously given the reliance on self-reported, unverified diagnoses. Rather than implying a true epidemiologic shift, the observed associations, particularly the higher prevalence among younger, insured, and female participants, may reflect greater health awareness, reporting behavior, and access to health care services rather than genuine differences in disease burden. Accordingly, these findings highlight important awareness and surveillance gaps rather than immediate policy directives. Strengthening national health surveillance systems that integrate laboratory-based diagnostic confirmation (eg, eGFR and urinalysis) would allow more accurate estimation of kidney disease prevalence and its risk factors. Nevertheless, the study underscores the continued importance of public health strategies targeting modifiable risk factors, such as obesity, diabetes, hypertension, and smoking. Integrating kidney health into broader non-communicable disease prevention programs, alongside improved health literacy and early detection at the PHC level, could substantially improve renal health outcomes. Enhanced data collection infrastructure and population-based screening programs would further enable policymakers to design evidence-informed interventions and track progress toward national health goals.

Conclusions

This study offers valuable insight into the self-reported prevalence and associated factors of renal disease among individuals attending PHCs in Riyadh, Saudi Arabia. While the findings indicate that a notable proportion of participants reported a prior physician diagnosis of kidney-related conditions, the reliance on self-reported data without laboratory confirmation suggests that these results more accurately reflect awareness and detection patterns rather than the true epidemiological burden of renal disease. Several modifiable factors, including smoking, obesity, hypertension, diabetes, and hypercholesterolemia, were found to be significantly associated with self-reported renal disease, consistent with established risk profiles in the literature. Demographic associations, such as those observed with age, sex, and insurance coverage, likely reflect differences in health care access and health awareness rather than causal relationships.

Overall, the study underscores the importance of strengthening surveillance systems and integrating laboratory-based screening within primary care to improve early identification and monitoring of kidney disease in Saudi Arabia. Future research using objective measures and longitudinal designs is needed to validate these associations and to better understand the true epidemiology of renal disorders in the population.

Funding

This research was funded by Prince Sattam bin Abdulaziz University (project PSAU/2025/03/32295).

Conflicts of Interest

None declared.

Checklist 1
STROBE checklist.
[PDF File, 157 KB - publichealth_v12i1e81849_app1.pdf]

References

1. Jha V, Al-Ghamdi SMG, Li G, et al. Global economic burden associated with chronic kidney disease: a pragmatic review of medical costs for the inside CKD research programme. Adv Ther 2023 Oct;40(10):4405-4420. [doi: 10.1007/s12325-023-02608-9] [Medline: 37493856]
2. Fletcher BR, Damery S, Aiyegbusi OL, et al. Symptom burden and health-related quality of life in chronic kidney disease: a global systematic review and meta-analysis. PLoS Med 2022 Apr;19(4):e1003954. [doi: 10.1371/journal.pmed.1003954] [Medline: 35385471]
3. Jadoul M, Aoun M, Masimango Imani M. The major global burden of chronic kidney disease. Lancet Glob Health 2024 Mar;12(3):e342-e343. [doi: 10.1016/S2214-109X(24)00050-0] [Medline: 38365398]
4. Rennke H, Denker BM. Renal Pathophysiology: The Essentials: Lippincott Williams & Wilkins; 2023.
5. Gupta R, Woo K, Yi JA. Epidemiology of end-stage kidney disease. Semin Vasc Surg 2021 Mar;34(1):71-78. [doi: 10.1053/j.semvascsurg.2021.02.010] [Medline: 33757639]
6. Imamah NF, Lin HR. Palliative care in patients with end-stage renal disease: a meta synthesis. Int J Environ Res Public Health 2021 Oct 11;18(20):10651. [doi: 10.3390/ijerph182010651] [Medline: 34682395]
7. Alkaabi AJ, Alkous A, Mahmoud K, et al. The prevalence and correlates of depression among patients with chronic diseases in the United Arab Emirates. PLoS ONE 2022;17(12):e0278818. [doi: 10.1371/journal.pone.0278818] [Medline: 36516141]
8. Bashir M, Alshamsi M, Almahrooqi S, et al. Prevalence of chronic diseases among United Arab Emirates University students: cross-sectional study. Eur J Public Health 2022 Oct 21;32(Supplement_3):ckac131-ckac142. [doi: 10.1093/eurpub/ckac131.142]

XSL•FO
RenderX

9. Al-Jawaldeh A, Abbass MMS. Unhealthy dietary habits and obesity: the major risk factors beyond non-communicable diseases in the eastern mediterranean region. Front Nutr 2022;9:817808. [doi: 10.3389/fnut.2022.817808] [Medline: 35369054]
10. Ke C, Liang J, Liu M, Liu S, Wang C. Burden of chronic kidney disease and its risk-attributable burden in 137 low-and middle-income countries, 1990-2019: results from the global burden of disease study 2019. BMC Nephrol 2022 Jan 5;23(1):17. [doi: 10.1186/s12882-021-02597-3] [Medline: 34986789]
11. Li Y, Ning Y, Shen B, et al. Temporal trends in prevalence and mortality for chronic kidney disease in China from 1990 to 2019: an analysis of the Global Burden of Disease Study 2019. Clin Kidney J 2023 Feb;16(2):312-321. [doi: 10.1093/ckj/sfac218] [Medline: 36755850]
12. Scurt FG, Ganz MJ, Herzog C, Bose K, Mertens PR, Chatzikyrkou C. Association of metabolic syndrome and chronic kidney disease. Obes Rev 2024 Jan;25(1):e13649. [doi: 10.1111/obr.13649] [Medline: 37783465]
13. Massy ZA, Drueke TB. Combination of cardiovascular, kidney, and metabolic diseases in a syndrome named cardiovascular-kidney-metabolic, with new risk prediction equations. Kidney Int Rep 2024 Sep;9(9):2608-2618. [doi: 10.1016/j.ekir.2024.05.033] [Medline: 39291205]
14. Rafferty Q, Stafford LK, Vos T, et al. Global, regional, and national prevalence of kidney failure with replacement therapy and associated aetiologies, 1990–2023: a systematic analysis for the global burden of disease study 2023. Lancet Glob Health 2025 Aug;13:e1378-e1395. [doi: 10.1016/S2214-109X(25)00198-6]
15. Alageel S, Alhujaili M, Altwaijri Y, Bilal L, Alsukait R. Barriers and facilitators to adopting healthier lifestyle among low-income women in Saudi Arabia: a qualitative study. Health Expect 2023 Jun;26(3):1202-1212. [doi: 10.1111/hex.13735] [Medline: 36806821]
16. Al-Hazzaa HM, Albawardi NM. Obesity, lifestyle behaviors, and dietary habits of Saudi adolescents living in Riyadh (ATLS-2 project): revisited after a ten-year period. Life (Basel) 2021 Oct 13;11(10):1078. [doi: 10.3390/life11101078] [Medline: 34685449]
17. Alhejely MMM, Shibli KY, Almalki WAH, et al. Influence of lifestyle changes on cardiovascular diseases in Saudi Arabia: a systematic literature review. Cureus 2023;15(6):e40075. [doi: 10.7759/cureus.40075]
18. Yousef L, AlAngari D, AlShehri R, AlSharif B, Bayameen O, Alnemer Z. Healthcare transformation journey in the eastern egion of Saudi Arabia: an overview, challenges and lessons learned. J Med Life 2023 Apr;16(4):583-592. [doi: 10.25122/jml-2023-0010] [Medline: 37305832]
19. Al-Ghamdi S, Abu-Alfa A, Alotaibi T, et al. Chronic kidney disease management in the Middle East and Africa: concerns, challenges, and novel approaches. Int J Nephrol Renovasc Dis 2023;16:103-112. [doi: 10.2147/IJNRD.S363133] [Medline: 37051319]
20. Vachharajani T, Jasuja S, AlSahow A, et al. Current status and future of end-stage kidney disease in Gulf Cooperation Council countries: challenges and opportunities. Saudi J Kidney Dis Transpl 2021;32(4):1073-1088. [doi: 10.4103/1319-2442.338281] [Medline: 35229807]
21. Alkubati SA, Alfaraaj AA, Almoliky MA, Alsaqri SH, Saleh KA, Siam BG. Exploring the readiness of critical care in implementing continuous renal replacement therapy in Hail hospitals, Saudi Arabia: findings for acute kidney injury patient care improvement. Healthcare (Basel) 2025;13(18):2320. [doi: 10.3390/healthcare13182320]
22. Al-Za'abi RK, Tolmie EP, Rice AM, Mohsin N, Al-Busadi AS. Epidemiology of end-stage kidney disease on renal replacement therapy: 35 years analysis study nephrologists. Saudi J Kidney Dis Transpl 2021;32(1):183-190. [doi: 10.4103/1319-2442.318521] [Medline: 34145129]
23. Mousa D, Alharbi A, Helal I, et al. Prevalence and associated factors of chronic kidney disease among relatives of hemodialysis patients in Saudi Arabia. Kidney Int Rep 2021 Mar;6(3):817-820. [doi: 10.1016/j.ekir.2020.12.029] [Medline: 33732996]
24. Al-Zahrani N, AlSwat HK, AlQarni AM, et al. Prevalence and risk factors of diabetic nephropathy among Saudi type-1 diabetic patients in Taif City, Saudi Arabia. Diabetes Metab Syndr Obes 2023;16:3609-3616. [doi: 10.2147/DMSO.S432700] [Medline: 37964940]
25. Naser AY, Alwafi H, Alotaibi B, et al. Prevalence of chronic kidney diseases in patients with diabetes mellitus in the Middle East: a systematic review and meta-analysis. Int J Endocrinol 2021;2021(1):4572743. [doi: 10.1155/2021/4572743] [Medline: 34497644]
26. Safdar OY, Alblowi SS, Aboulola NA, Alharazy DT. Renal stones and risk factors in Jeddah and Riyadh. Saudi J Kidney Dis Transpl 2021;32(1):191-198. [doi: 10.4103/1319-2442.318523] [Medline: 34145130]
27. Alzahrani NS, Almarwani AM, Alhowaymel F, Alharbi HF. The lived experience of middle-aged Saudi patients with end-stage chronic renal disease on hemodialysis treatment. J Transcult Nurs 2023 Jul;34(4):263-269. [doi: 10.1177/10436596231166041] [Medline: 37086181]
28. Aljohani R, Alenzi F, Aboabat A, et al. Disease characteristics and outcomes of lupus nephritis in Saudi Arabia: a systematic review. Lupus (Los Angel) 2025 Oct;34(11):1184-1198. [doi: 10.1177/09612033251360628] [Medline: 40682252]
29. Alshehri MA, Alkhlady HY, Awan ZA, et al. Prevalence of chronic kidney disease in Saudi Arabia: an epidemiological population-based study. BMC Nephrol 2025 Jan 24;26(1):37. [doi: 10.1186/s12882-025-03954-2] [Medline: 39849386]

XSL•FO
RenderX

30. Alshahrani M, Almalki LZ, Hasoosah NA, et al. Prevalence and assessment of risk factors of chronic kidney disease in the ASIR region of Saudi Arabia. Ann Med Surg (Lond) 2024 Jul;86(7):3909-3916. [doi: 10.1097/MS9.0000000000001755] [Medline: 38989199]
31. Suleiman AK, Ming LC. Transforming healthcare: Saudi Arabia's vision 2030 healthcare model. J Pharm Policy Pract 2025;18(1):2449051. [doi: 10.1080/20523211.2024.2449051] [Medline: 39845746]
32. Alasiri AA, Mohammed V. Healthcare transformation in Saudi Arabia: an overview since the launch of vision 2030. Health Serv Insights 2022;15:11786329221121214. [doi: 10.1177/11786329221121214] [Medline: 36081830]
33. Alzahrani AS, Alghamdi MS, Alzahrani AYA, et al. Advancements in chronic disease management: evaluating the role of family medicine physicians to achieving the goals of Saudi Vision 2030 through systematic review. J Int Crisis Risk Commun Res 2024;7(S11):1870-1889. [doi: 10.63278/jicrcr.vi.1538]
34. von Elm E, Altman DG, Egger M, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ 2007 Oct 20;335(7624):806-808. [doi: 10.1136/bmj.39335.541782.AD] [Medline: 17947786]
35. Noncommunicable disease surveillance, monitoring and reporting. World Health Organization. 2022. URL: https://www.who.int/teams/noncommunicable-diseases/surveillance/systems-tools/steps/instrument [accessed 2026-02-21]
36. Publications catalog. The Demographic and Health Surveys Program. 2023. URL: https://dhsprogram.com/publications/publication-search.cfm?type=35 [accessed 2026-02-21]
37. Needs Assessment Handbook. : United Nations High Commissioner for Refugees; 2017 URL: https://www.unhcr.org/handbooks/assessment/sites/assessment/files/2023-10/UNHCR%20Needs%20Assessment%20Handbook.pdf [accessed 2026-02-21]
38. AlKetbi LB, Hemaid A, AlMansoori M, et al. Risk factors of chronic kidney disease incidence: a retrospective cohort study in Abu Dhabi. Research Square. Preprint posted online on Dec 18, 2024. [doi: 10.21203/rs.3.rs-5423040/v1]
39. Chukwuonye II, Ohagwu KA, Adelowo OO, et al. Prevalence and predictors of chronic kidney disease in a semiurban community in Lagos. Int J Nephrol 2019;2019(1):1625837. [doi: 10.1155/2019/1625837] [Medline: 31186960]
40. Fox CS, Larson MG, Leip EP, Culleton B, Wilson PWF, Levy D. Predictors of new-onset kidney disease in a community-based population. JAMA 2004 Feb 18;291(7):844-850. [doi: 10.1001/jama.291.7.844] [Medline: 14970063]
41. Kashgary A, Hassanein DE. Epidemiology and patient characteristics of chronic kidney disease in Saudi Arabia. Research Square. Preprint posted online on Jan 18, 2024. [doi: 10.21203/rs.3.rs-3830286/v1]
42. Moftakhar L, Jafari F, Ghoddusi Johari M, Rezaeianzadeh R, Hosseini SV, Rezaianzadeh A. Prevalence and risk factors of kidney stone disease in population aged 40-70 years old in Kharameh cohort study: a cross-sectional population-based study in southern Iran. BMC Urol 2022 Dec 19;22(1):205. [doi: 10.1186/s12894-022-01161-x] [Medline: 36536352]
43. Chukwuonye II, Ogah OS, Anyabolu EN, et al. Prevalence of chronic kidney disease in Nigeria: systematic review of population-based studies. Int J Nephrol Renovasc Dis 2018;11:165-172. [doi: 10.2147/IJNRD.S162230] [Medline: 29861638]
44. Hecking M, Bieber BA, Ethier J, et al. Sex-specific differences in hemodialysis prevalence and practices and the male-to-female mortality rate: the Dialysis Outcomes and Practice Patterns Study (DOPPS). PLoS Med 2014 Oct;11(10):e1001750. [doi: 10.1371/journal.pmed.1001750] [Medline: 25350533]
45. Giacovelli JK, Egorova N, Nowygrod R, Gelijns A, Kent KC, Morrissey NJ. Insurance status predicts access to care and outcomes of vascular disease. J Vasc Surg 2008 Oct;48(4):905-911. [doi: 10.1016/j.jvs.2008.05.010] [Medline: 18586449]
46. Hall YN, Rodriguez RA, Boyko EJ, Chertow GM, O'Hare AM. Characteristics of uninsured Americans with chronic kidney disease. J Gen Intern Med 2009 Aug;24(8):917-922. [doi: 10.1007/s11606-009-1028-3] [Medline: 19506974]
47. Ward MM. Medical insurance, socioeconomic status, and age of onset of endstage renal disease in patients with lupus nephritis. J Rheumatol 2007 Oct;34(10):2024-2027. [Medline: 17696272]
48. Oliveira Coelho F, Andrade L. Smoking and kidney disease: risk factors, challenges, and preventive strategies. Contrib Nephrol 2021;199:179-187. [doi: 10.1159/000517749] [Medline: 34344004]
49. Wang S, Qin A, Pei G, et al. Cigarette smoking may accelerate the progression of IgA nephropathy. BMC Nephrol 2021 Jun 29;22(1):239. [doi: 10.1186/s12882-021-02453-4] [Medline: 34187402]
50. Nagasawa Y, Yamamoto R, Rakugi H, Isaka Y. Cigarette smoking and chronic kidney diseases. Hypertens Res 2012 Mar;35(3):261-265. [doi: 10.1038/hr.2011.205] [Medline: 22158113]
51. Ghelichi-Ghojogh M, Fararouei M, Seif M, Pakfetrat M. Chronic kidney disease and its health-related factors: a case-control study. BMC Nephrol 2022 Jan 10;23(1):24. [doi: 10.1186/s12882-021-02655-w] [Medline: 35012483]
52. Kumar M, Dev S, Khalid MU, et al. The bidirectional link between diabetes and kidney disease: mechanisms and management. Cureus 2023;15(9):e45615. [doi: 10.7759/cureus.45615]
53. Mirghani HO, Abdul Khalig Alyoussef A, Mohammed OS. The relationship between dyslipidemia and lupus nephritis in systemic lupus erythematosus patients attending a Saudi Rheumatic Center, Tabuk. Sudan J Med Sci 2020;15(1):10-19. [doi: 10.18502/sjms.v15i1.6697]
54. Mohammedsaeed W, Al Malki A, Alsayed S. The risk factors for cardiovascular disease and chronic kidney disease in patients with nonalcoholic fatty liver disease in Saudi Arabia. Saudi Med J 2025 May;46(5):478-490. [doi: 10.15537/smj.2025.46.5.20240901] [Medline: 40335100]

XSL•FO
RenderX

55. AlGhonaim MA, Fathalla AA. Vascular calcification in patients with chronic kidney disease on dialysis in the Kingdom of Saudi Arabia: a cross-sectional study. Saudi J Kidney Dis Transpl 2019;30(3):571-580. [doi: 10.4103/1319-2442.261329] [Medline: 31249220]
56. Al-Shamsi S, Regmi D, Govender RD. Chronic kidney disease in patients at high risk of cardiovascular disease in the United Arab Emirates: a population-based study. PLoS ONE 2018;13(6):e0199920. [doi: 10.1371/journal.pone.0199920] [Medline: 29949629]
57. Ginawi IA, Ahmed HG, Al-Hazimi AM. Assessment of risk factors for chronic kidney disease in Saudi Arabia. Int J Sci Res 2014;3(7):446-450 [FREE Full text]
58. Tabatabaei-Malazy O, Saeedi Moghaddam S, Khashayar P, et al. Regional burden of chronic kidney disease in North Africa and Middle East during 1990-2019; results from Global Burden of Disease study 2019. Front Public Health 2022;10:1015902. [doi: 10.3389/fpubh.2022.1015902] [Medline: 36304241]

Abbreviations

AOR: adjusted odds ratio
CKD: chronic kidney disease
eGFR: estimated glomerular filtration rate
OR: odds ratio
PHC: primary health care center
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Mavragani, S Brini; submitted 04.Aug.2025; peer-reviewed by J Jamaluddin, M Alshehri; revised version received 07.Jan.2026; accepted 14.Jan.2026; published 26.Mar.2026.

Please cite as:
Alshahrani S, El-Metwally AA, Alshahrani A, Al-Khateeb BF, Aldubikhi AI, Angawi K, Alharthy A, Alnaim L, Zaidan AM, Aldahash R
Prevalence and Associated Factors of Renal Disease in Saudi Residents Attending Primary Health Care Centers in Riyadh, Saudi Arabia: Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e81849
URL: https://publichealth.jmir.org/2026/1/e81849
doi:10.2196/81849

© Saad Alshahrani, Ashraf A El-Metwally, Awad Alshahrani, Badr F Al-Khateeb, Aljohrah Ibrahim Aldubikhi, Khadijah Angawi, Amani Alharthy, Lubna Alnaim, Amal Mousa Zaidan, Raed Aldahash. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 26.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Trends and Core Competence Shifts in Nurses' Infectious Disease Emergency Response Competence Across COVID-19 Pandemic Phases: Repeated Cross-Sectional Survey and Network Analysis

Jiagen Xiang[1,2*], MSN; Qiuxia Liang[1*], BSN; Yufei Lu[1*], BSN; Meizhen Lin[3], MSN; Juan Liu[4], BSN; Hongman Li[1], BSN; Jingjing Cai[1], BSN; Yan Wu[1], PhD; Bingbing Li[1], PhD; Wenting Liu[1], PhD; Ming Yang[1], PhD; Xiaoning Sun[1], MSN; Yan Luo[5], BSN; Yibing Tan[1*], PhD; Hairun Liu[6*], MMed; Jiaying Li[7*], PhD; Liming Lu[2*], PhD; Zengjie Ye[1,8*], PhD

[1]School of Nursing, Guangzhou University of Chinese Medicine, 232 Outer Ring East Road, Guangzhou University City, Panyu District, Guangzhou, Guangdong, China

[2]Medical College of Acu-Moxi and Rehabilitation, Guangzhou University of Chinese Medicine, Guangzhou, Guangdong, China

[3]Infection Control Office, Guangdong Provincial Hospital of Traditional Chinese Medicine, Guangzhou, Guangdong, China

[4]Department of General Medicine, Guangdong Provincial Hospital of Traditional Chinese Medicine, Guangzhou, Guangdong, China

[5]Cancer Hospital Chinese Academy of Medical Sciences, Beijing, China

[6]Center of Cognitive and Sleeping, The People's Hospital of Guangxi Zhuang Autonomous Region, Nanning, China

[7]The Nethersole School of Nursing, Faculty of Medicine, Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

[8]School of Nursing, Guangzhou Medical University, Guangzhou, Guangdong, China

[*]these authors contributed equally

Corresponding Author:
Zengjie Ye, PhD
School of Nursing, Guangzhou University of Chinese Medicine, 232 Outer Ring East Road, Guangzhou University City, Panyu District, Guangzhou, Guangdong, China

Abstract

Background: The COVID-19 pandemic exposed structural vulnerabilities in the global health emergency workforce. Sustained monitoring of response competence dynamics is critical for maintaining health system resilience during protracted outbreaks.

Objective: This study aimed to investigate trends and structural transitions in nurses' infectious disease emergency response competence during different phases of the COVID-19 pandemic.

Methods: Using the Healthcare Workers' Infectious Disease Emergency Response Competence Questionnaire, 3 repeated cross-sectional surveys were conducted in February 2020, August 2021, and February 2023. The participants were matched in a 1:1:1 ratio by their demographic characteristics via propensity scores. Temporal trends were analyzed using ANOVA, and competence structure transitions were examined via network analysis.

Results: Three-wave analyses (n=2525 per wave) demonstrated a sustained decline in competence, from 4.05 (SD 0.63) in February 2020 to 3.81 (SD 0.66) in August 2021 and further to 3.70 (SD 0.66) in February 2023. All pairwise comparisons were statistically significant (all $P<.001$). Network analysis identified critical structural shifts in competence architecture: in February 2020, the core network node was T24 (emergency management of body fluid exposure; strength=1.764), whereas in August 2021 and February 2023, the core network node was T19 (acquisition of key information on emerging infectious diseases; strength=1.759 and 1.852, respectively). Network structure comparisons revealed significant differences between February 2020 and August 2021 ($P=.01$) and between February 2020 and February 2023 ($P=.01$), whereas the difference between August 2021 and February 2023 was not significant ($P=.07$).

Conclusions: Despite accumulated pandemic experience, nurses' infectious disease response competence declined substantially, indicating systemic fragility during prolonged crises. However, this experience reshaped competence architecture, evolving from technical infection prevention toward higher-order competence in information integration and clinical decision-making under uncertainty. To rebuild resilience, phase-specific training programs are essential: early-phase training emphasizes infection prevention, whereas prolonged outbreaks focus on information identification and decision support. Additionally, standardized information platforms and psychological support are essential to manage ongoing pandemic pressures.

(JMIR Public Health Surveill 2026;12:e83281) doi:10.2196/83281

XSL•FO
RenderX

KEYWORDS

infectious disease emergencies; response competences; emergency preparedness; health system resilience; propensity score matching; network analysis; COVID-19; nurses

Introduction

The threat of emerging infectious diseases has remained a persistent challenge to global public health, highlighting the critical need for robust health care system preparedness and response capabilities [1,2]. Nurses, who constitute the largest segment of the health care workforce, are indispensable at every stage of an outbreak. Their performance not only influences individual patient outcomes but also affects the resilience of the entire health care system [3,4]. Central to this role is the concept of infectious disease emergency response competence, defined as the integrated knowledge, skills, and experience required to deliver rapid and comprehensive responses during infectious disease crises [5,6].

The profound and prolonged experience of the COVID-19 pandemic compels us to learn how pandemics evolve through distinct phases. The World Health Organization has consequently advocated for phase-tailored strategies to underpin health system resilience [7]. Consequently, nursing emergency preparedness programs should be agile—continuously refined to match changes in epidemiology, resource availability, and clinical priorities [8]. Therefore, understanding how nurses' infectious disease emergency response competence shifted across these pandemic phases is essential not only for interpreting the past but also, crucially, for preparing effective responses when the next inevitable threat emerges.

Existing studies have reported inconsistent levels of infectious disease emergency response competence among nurses as the COVID-19 pandemic has evolved [9-11]. These inconsistencies may stem from unexamined differences between pandemic phases. Although individual-level predictors such as educational attainment, prior training, and clinical experience have previously been identified [12-14], the contextual influence of distinct pandemic phases—each with its own operational challenges—has not been fully explored, particularly regarding sustained impacts over an extended duration such as during the COVID-19 pandemic.

Conceptually, infectious disease emergency response competence trajectories may reflect the interplay of two opposing forces: (1) potential competence enhancement through cumulative experiential learning as nurses gain practical exposure during prolonged outbreaks [5,14] and (2) potential competence erosion resulting from sustained psychological stress, physical exhaustion [15], and depleted cognitive resources [16]. Critically, beyond overall levels, the structure of required competence is also likely to undergo significant shifts during a prolonged crisis such as the COVID-19 pandemic.

However, empirical evidence mapping both the temporal trends and structural shifts in competence across distinct pandemic phases is lacking, and existing data do not capture these complex, longitudinal interactions. This critical gap impedes our understanding of how experiential learning and prolonged stress interact to shape both the overall level and the structural configuration of infectious disease emergency response competence over time. Therefore, this study aimed to examine temporal trends and structural transitions in nurses' infectious disease emergency response competence across 3 COVID-19 pandemic phases. We hypothesized the following:

- Hypothesis 1: cumulative experiential learning will drive a progressive increase in overall infectious disease emergency response competence.
- Hypothesis 2: the configuration of core infectious disease emergency response competencies will shift to meet phase-specific demands.

To test these hypotheses, we conducted 3 repeated cross-sectional surveys of registered nurses in Guangdong Province across 3 pandemic phases (early surge, Delta containment, and post-Omicron transition). We used 1:1:1 propensity score matching to balance staffing characteristics and applied network analysis to decode structural transitions among core competencies. The results may help identify actionable intervention targets to support phase-tailored preparedness strategies.

Methods

Study Design

This study adopted a repeated cross-sectional survey design to investigate trends and structural transitions in nurses' infectious disease emergency response competence across 3 key COVID-19 pandemic phases in Guangdong Province: February 2020 (early surge), August 2021 (Delta containment), and February 2023 (post-Omicron transition).

In February 2020, Guangdong reported a rapid surge in cases, overwhelmed health care capacity, and a strong resolve among medical staff to fight the pandemic despite limited resources [17]. By August 2021, Guangdong had implemented a "dynamic zero-COVID" policy that included mass testing and strict mobility controls, while the Delta outbreak tested these strategies and drove protocol standardization [18]. In February 2023, following the relaxation of restrictions, Guangdong faced another surge in cases at a time when many health care workers had already been infected or were themselves becoming infected, further intensifying the strain on the health care system. The focus shifted toward long-term mitigation and resilience strategies as the health care system adjusted to endemicity [19]. Together, these 3 waves capture contrasting, policy-relevant phases of the pandemic, providing a clear basis for examining phase-dependent competence dynamics. For a timeline of key epidemiological milestones, policy interventions, and their implications for nursing competence, please refer to Table S1 in Multimedia Appendix 1.

Participant recruitment used convenience sampling of registered nurses from Guangdong Province at each of the 3 time points. Consistent with the repeated cross-sectional design, no active tracking of individuals occurred across survey waves; each wave

XSL•FO
RenderX

was treated as an independent sample despite the possibility of incidental reparticipation by some nurses.

Participants

Eligible participants were registered nurses with a valid license, aged ≥18 years, with more than 1 year of frontline clinical experience, and who provided informed consent. We excluded nurses who were in training programs, completing internships, or on leave during the survey period.

The sample size was estimated using a precision-based approach for the total competence score: $n = (Z_{1-\alpha/2} \times \sigma/\delta)^2$, where $Z_{1-\alpha/2}$ =1.96 for a 95% confidence level. On the basis of a previous study [20] reporting a total score of 116.13 (SD 22.84), we set σ=22.84 and an absolute precision of δ=1 point on the total score scale. This yielded a required sample size of 2004. Assuming an 80% response rate, the target sample size was 2505 per wave. Although the sample size calculation was based on the total score, the final analysis and results in this study were reported using the mean score for clarity and ease of interpretation.

To address between-wave comparisons across 3 survey waves, we further applied Bonferroni adjustment for the 3 pairwise comparisons (α=.0167) and conducted an achieved power (sensitivity) analysis using the final propensity score–matched sample size (n=2525 per wave).

Questionnaires

Demographic and Professional Characteristics

The participants' demographic and professional characteristics, including age, gender, clinical experience, professional title, education level, and department, were selected on the basis of prior studies [20,21].

Healthcare Workers' Infectious Disease Emergency Response Competence Questionnaire

The infectious disease emergency response competence questionnaire, developed by Kan et al [21] through a comprehensive literature review and expert consultation, is the most widely used assessment tool in China and has robust reliability and validity [10,12,22]. It comprises 36 items across 3 dimensions: emergency prevention competence (3 items), emergency preparedness competence (4 items), and emergency response competence (29 items). Each item is rated on a 5-point Likert scale (1="totally unknown"; 5 ="very familiar"), yielding a total score ranging from 36 to 180. Higher scores indicate stronger competence. For descriptive interpretation, mean scores <3, 3 to 4, and >4 were considered low, moderate, and high perceived competence, respectively; these are interpretive anchors for the 5-point Likert scale rather than empirically validated clinical cutoffs. In this study, internal consistency was excellent, with Cronbach α coefficients ranging from 0.85 to 0.98 for the total scale and each dimension. The full 36-item list is provided in Table S2 in Multimedia Appendix 1.

Data Collection

Surveys were administered via Wenjuanxing, a Chinese online survey platform. At each survey wave, the Wenjuanxing survey link and the recruitment invitation letter were first disseminated in a WeChat (Tencent Holdings Limited) working group of nursing department directors in Guangdong Province. Directors who agreed to participate forwarded the survey link, along with the standardized guidelines, to head nurses, who subsequently shared it with staff nurses. To ensure data integrity, each mobile number could submit the survey only once, and all items had to be completed before submission.

The link remained active for 7 days. If needed, brief reminders were sent via the same internal channels during the open period. After closure, the responses were screened for quality: those that failed the inclusion or exclusion criteria, were incomplete, or contained inconsistent answers, or were completed in under 120 seconds were excluded. Inconsistent answers referred to implausible or contradictory information in key demographic or professional fields (eg, clinical experience exceeding age-derived limits). The 120-second threshold was determined through pilot testing by the investigators as the conservative minimum time to complete the demographic items plus the 36 Likert scale items; responses below this threshold were considered insufficient engagement, raising concerns about data quality.

Data Analysis

Propensity Score Matching

We used DecisionLinnc software (version 1.0; DecisionLinnc Core Team) [23,24] to perform propensity score matching (PSM) across 3 survey waves (February 2020, August 2021, and February 2023). Propensity scores were estimated using baseline covariates, including gender, age, professional title, education level, department, and clinical experience. A 2-stage matching approach was applied: first, matching wave 2020 vs wave 2021 using nearest-neighbor matching with a caliper of 0.25 SD of the logit-transformed propensity score, and then matching the pooled sample from stage 1 with wave 2023. A 1:1:1 matching ratio was achieved. Adequate balance was defined as a standardized mean difference (SMD) <0.10 for each covariate. We also conducted 3-wave comparisons and covariate-adjusted linear regression using the pre-PSM data to verify the robustness of the postmatching results.

Descriptive and Comparative Statistics

All analyses were conducted in R (version 4.4.3; R Foundation for Statistical Computing). Continuous variables with a normal distribution are presented as the mean (SD); nonnormally distributed continuous variables are reported as the median (IQR). Categorical variables are expressed as frequencies and percentages. Between-group comparisons for continuous variables were performed via Student 2-tailed *t* test or 1-way ANOVA, as appropriate; χ^2 tests were used to assess differences in categorical variables. A 2-tailed $P<.05$ was considered statistically significant.

Psychometric Network Analysis

We estimated networks of the 36 items via the EBICglasso algorithm (bootnet package) and visualized them via the qgraph package. Each node's predictability is indicated by a blue outer ring, and edge width represents association strength (green=positive correlation; red=negative correlation). Strength

XSL•FO
RenderX

centrality was used to identify key nodes, as it is the most stable centrality measure [25]. In this context, strength centrality reflects the overall connectedness of an item to the rest of the network; items with higher strength can be interpreted as "hub" competencies that may exert broader influence on the competence architecture. Network stability was assessed via the correlation stability (CS) coefficient, with CS >0.50 indicating adequate stability [26]. We compared network structures across the 3 time points via the network comparison test to evaluate differences in both network structure and global strength.

Ethical Considerations

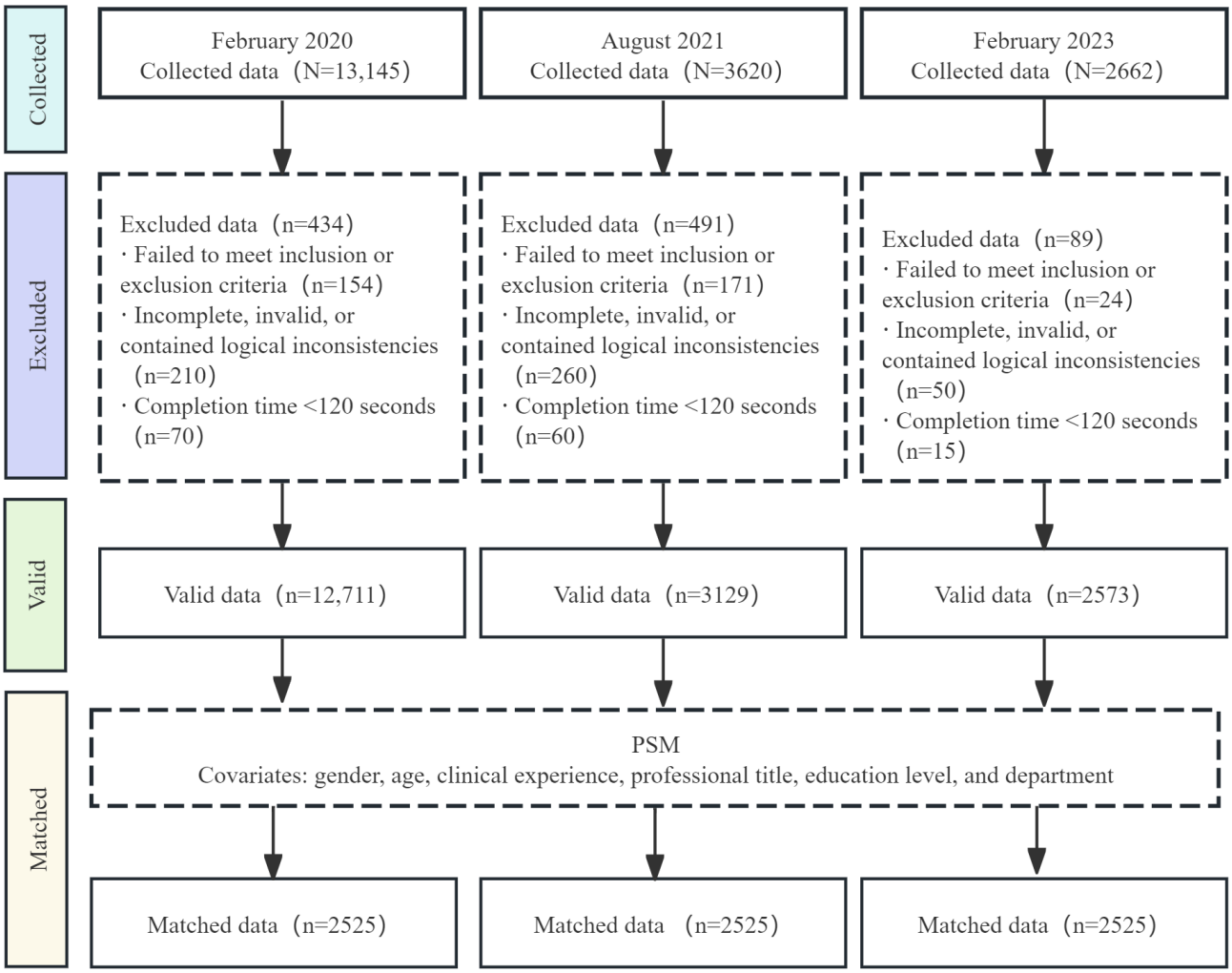
This study was approved by the Ethics Board of the First Affiliated Hospital of Guangzhou University of Chinese Medicine (ZYYEC-ERK[2020]132). Informed consent was obtained electronically from all participants. Specifically, participants provided their consent by clicking the survey link, reviewing the electronic informed consent form (which detailed the study procedures, objectives, and their rights), and subsequently completing and submitting the questionnaire. This process strictly adhered to the principles of the Declaration of Helsinki. Participants were assured of data confidentiality and anonymity throughout the study. Participants did not receive any financial incentives or compensation for completing the survey.

Results

Participant Flow and Sample Size

Overall, the surveys yielded 12,711 valid responses in February 2020, 3129 in August 2021, and 2573 in February 2023. After PSM, each wave retained 2525 (19.86% for February 2020, 80.7% for August 2021, and 98.13% for February 2023) participants. The participant flow, screening, and matching procedures are detailed in Figure 1.

Figure 1. Flow diagram of participant enrollment, screening, and propensity score matching (PSM) for 3 repeated cross-sectional online surveys of registered nurses in Guangdong Province, China, conducted during 3 COVID-19 pandemic phases (February 2020, August 2021, and February 2023). The flowchart shows exclusions based on eligibility and data quality criteria and the 1:1:1 PSM procedure yielding 2525 nurses per wave.



Demographic and Professional Characteristics of Nurses After PSM

Table 1 presents the demographic and professional characteristics of the participants in each wave after PSM, which achieved adequate balance across covariates (all |SMD|<0.10; all *P*>.20). For example, the February 2023 group exhibited a mean age of 32.95 (SD 7.70) years; 96.63% (2440/2525) were female, 62.97% (1590/2525) held junior professional titles, and 62.22% (1571/2525) held a bachelor's degree or higher, with

XSL•FO
RenderX

90.97% (2297/2525) working in nonemergency departments. Full prematching and postmatching characteristics and balance diagnostics (including SMDs) are provided in Table S3 in Multimedia Appendix 1.

Table . Demographic and professional characteristics of registered nurses in Guangdong Province, China, across 3 COVID-19 pandemic phases (February 2020, August 2021, and February 2023) after 1:1:1 propensity score matching (n=2525 per wave). Variables include gender, age, clinical experience, professional title, education level, and department. Between-wave balance is quantified using standardized mean differences (SMDs).

	February 2020	August 2021	February 2023	Chi-square (*df*)[a]	*P* value	SMD
Gender, n (%)				0.7 (2)	.70	0.015
Men	83 (3.29)	75 (2.97)	85 (3.37)			
Women	2442 (96.71)	2450 (97.03)	2440 (96.63)			
Age (years), mean (SD)[b]	32.93 (7.61)	32.64 (7.62)	32.95 (7.70)	__[c]	.27	0.027
Clinical experience (years), n (%)				0.5 (2)	.77	0.013
0 - 5	622 (24.63)	643 (25.47)	626 (24.79)			
≥6	1903 (75.37)	1882 (74.53)	1899 (75.21)			
Professional title, n (%)				0.1 (4)	.99	0.005
Junior	1583 (62.69)	1591 (63.01)	1590 (62.97)			
Intermediate	768 (30.42)	761 (30.14)	763 (30.22)			
Senior	174 (6.89)	173 (6.85)	172 (6.81)			
Education level, n (%)				0.2 (2)	.92	0.008
Associate degree or below	946 (37.47)	940 (37.23)	954 (37.78)			
Bachelor's degree or above	1579 (62.53)	1585 (62.77)	1571 (62.22)			
Department, n (%)				1.8 (2)	.42	0.024
Nonemergency	2302 (91.17)	2322 (91.96)	2297 (90.97)			
Emergency	223 (8.83)	203 (8.04)	228 (9.03)			

[a]Categorical variables were compared using chi-square tests. Chi-square values are reported to 1 decimal place, with df shown in parentheses.

[b]Age was compared using 1-way ANOVA ($F_{2,7572}$=1.317).

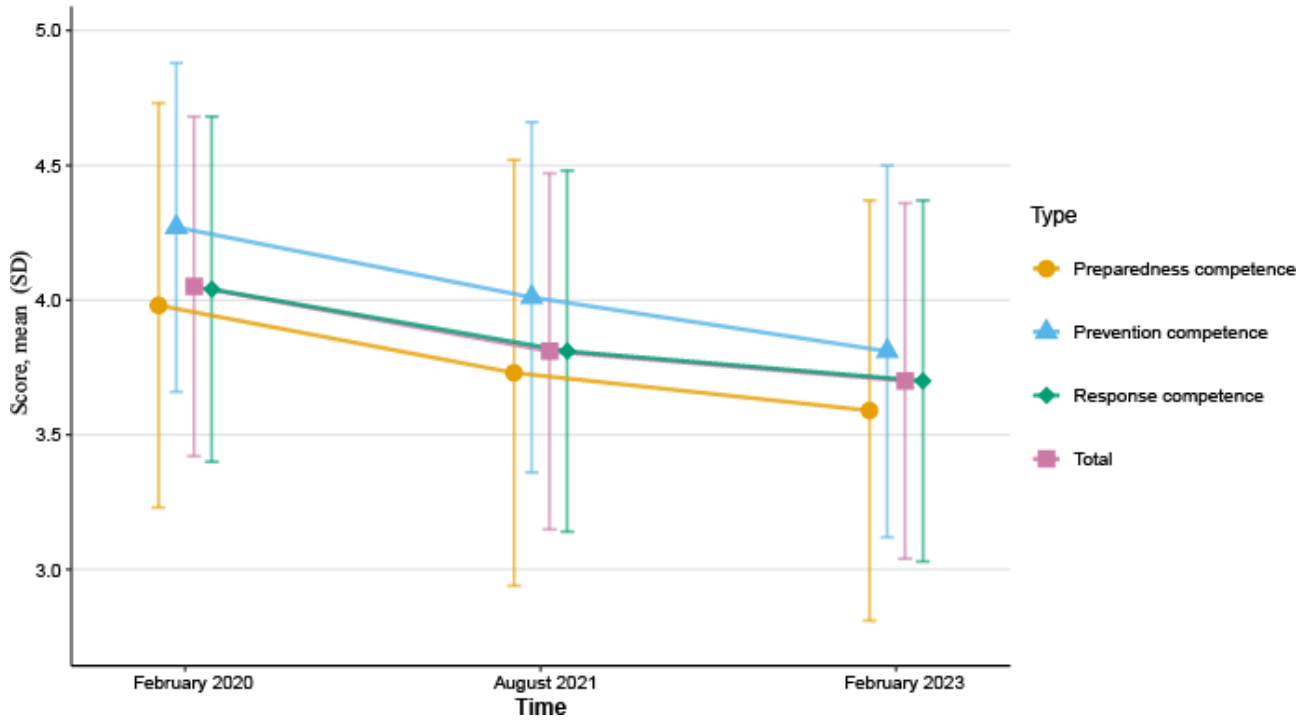
[c]Not applicable.

Nurses' Infectious Disease Emergency Response Competence

The mean total score decreased from 4.05 (SD 0.63) in February 2020 to 3.81 (SD 0.66) in August 2021 and further to 3.70 (SD 0.66) in February 2023. The 3 competence domains (prevention, preparedness, and response) followed a similar downward trajectory. One-way ANOVA with Bonferroni correction demonstrated significant differences between every pair of time points for both total and domain scores (all P<.001). The detailed results are presented in Figure 2 and Tables S4 in Multimedia Appendix 1. Table S4A in Multimedia Appendix 1 summarizes the mean (SD) scores and overall comparisons across waves, and Table S4B in Multimedia Appendix 1 provides pairwise mean differences with 95% CIs, effect sizes (Cohen *d*), and percentage changes for each comparison. Sensitivity analyses using the unmatched sample (including 3-wave comparisons and covariate-adjusted regression) yielded consistent conclusions (Table S5 in Multimedia Appendix 1).

XSL•FO
RenderX

Figure 2. Mean (SD) total and domain scores of nurses' infectious disease emergency response competence (36-item Healthcare Workers' Infectious Disease Emergency Response Competence Questionnaire; 5-point Likert scale) among propensity score–matched registered nurses in Guangdong Province, China, across 3 COVID-19 pandemic phases (February 2020, August 2021, and February 2023; n=2525 per wave). Between-wave differences were tested using 1-way ANOVA with Bonferroni-adjusted pairwise comparisons.



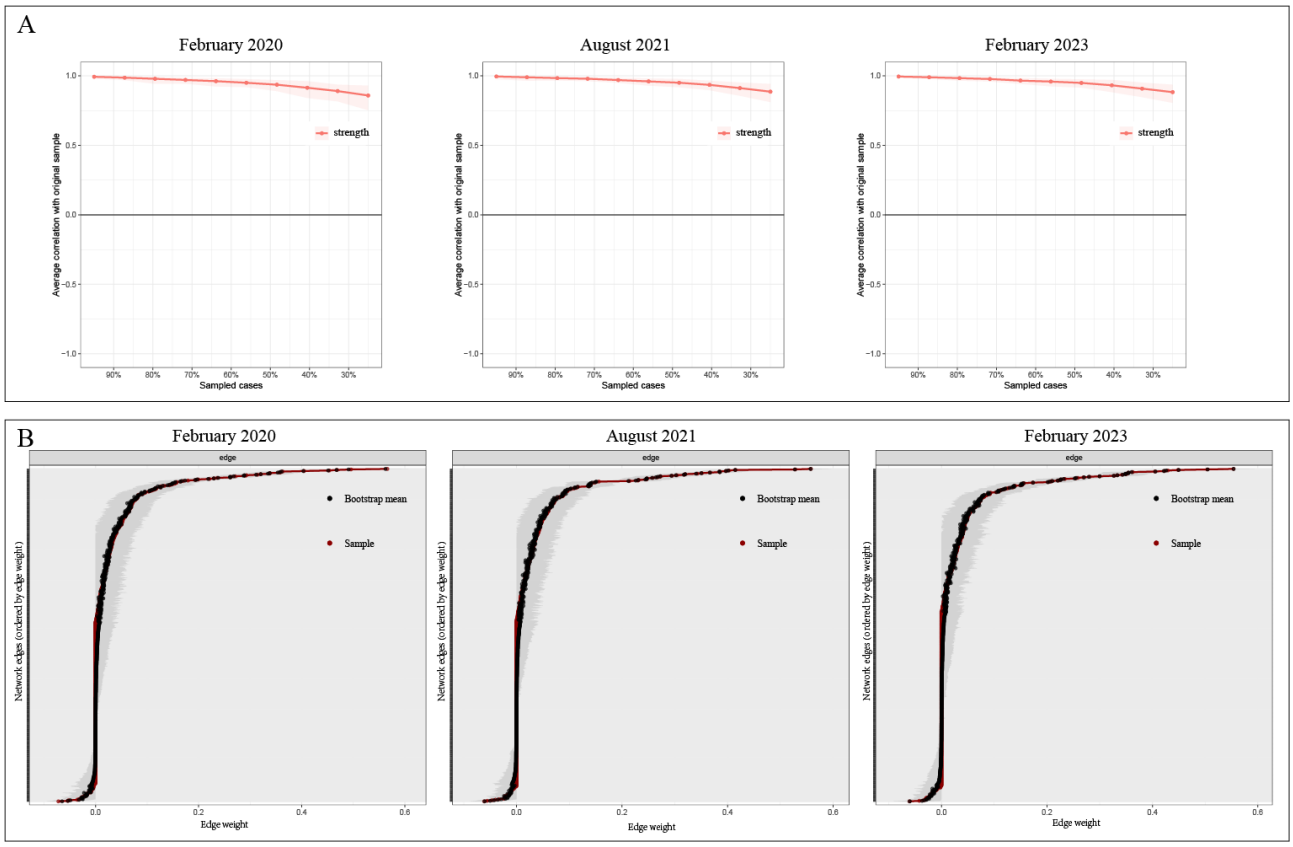
Network Analysis of Nurses' Infectious Disease Emergency Response Competence

Network Accuracy and Stability

We first evaluated the accuracy and stability of the estimated networks before interpreting network structure and centrality. These assessments supported the reliability of all 3 networks. First, the CS coefficient for strength centrality was 0.75 for each network, exceeding the recommended threshold of 0.5 (Figure 3A). Second, nonparametric bootstrapping indicated relatively narrow 95% CIs for edge weights, suggesting adequate precision in edge estimation (Figure 3B; gray bands indicate 95% CIs).

XSL•FO
RenderX

Figure 3. Accuracy and stability diagnostics for psychometric networks of nurses' infectious disease emergency response competence in 3 COVID-19 pandemic phases among propensity score–matched registered nurses in Guangdong Province, China (February 2020, August 2021, and February 2023; n=2525 per wave). (A) Correlation stability (CS) coefficients for strength centrality (CS >0.50 indicates adequate stability for interpreting centrality). (B) Bootstrap 95% CIs (gray shading) for edge weights (narrower CIs indicate more precise edge-weight estimation).

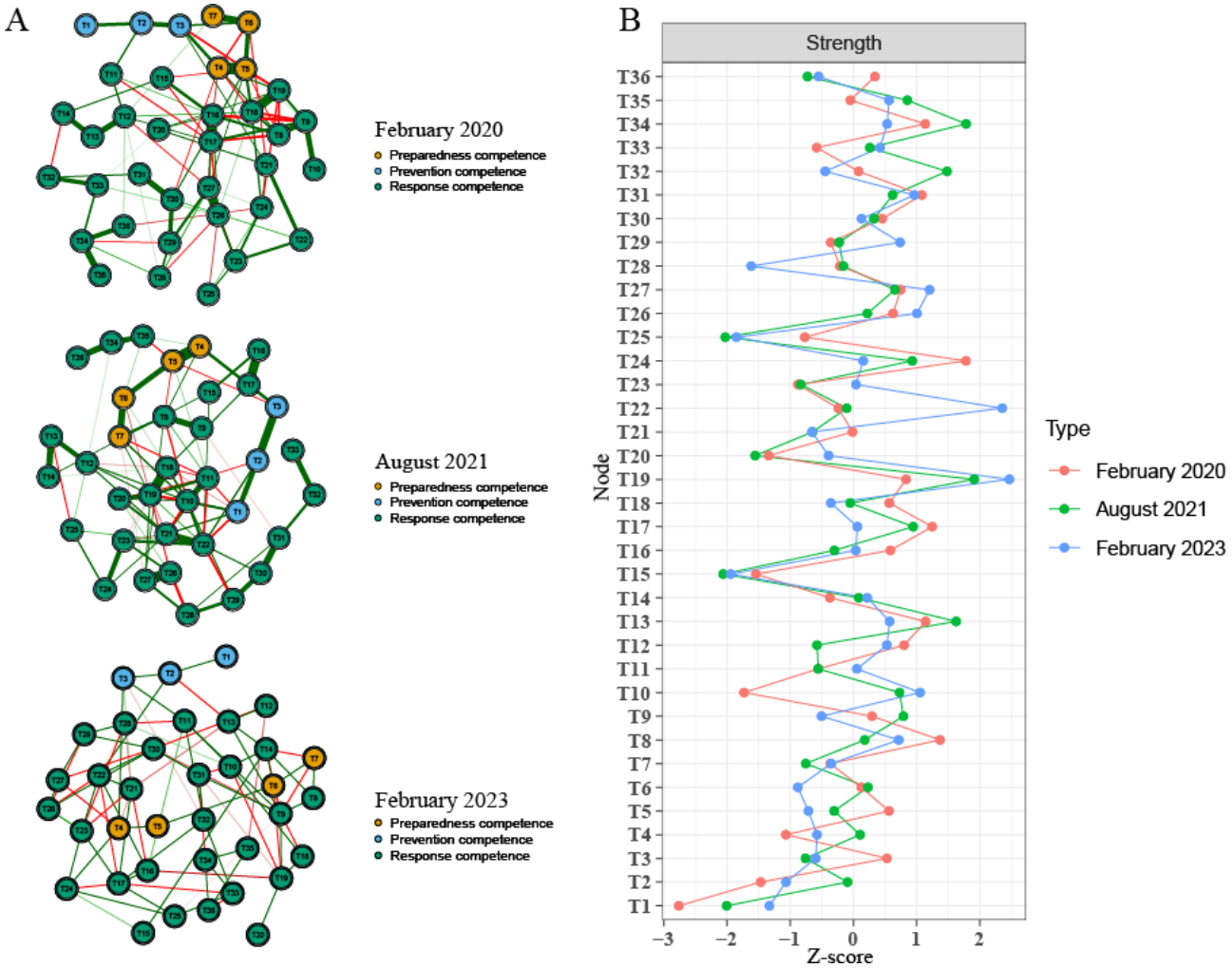


Network Structure and Centrality

On the basis of these stable networks, we examined node centrality and structural patterns. The range of strength centrality values across items was 1.125 to 1.764 in February 2020, 1.249 to 1.759 in August 2021, and 1.239 to 1.852 in February 2023. In February 2020, T24 (emergency management of body fluid exposure) showed the highest strength (1.764), whereas in August 2021 and February 2023, T19 (acquisition of key information on emerging infectious diseases) became the most central node (1.759 and 1.852, respectively). Cross-sectional networks and corresponding strength centralities are shown in Figure 4.

XSL•FO
RenderX

Figure 4. (A) Cross-sectional networks and (B) corresponding strength centralities of the infectious disease emergency response competence questionnaire among propensity score–matched registered nurses in Guangdong Province, China, across 3 COVID-19 pandemic phases (February 2020, August 2021, and February 2023; n=2525 per wave). Networks were estimated using EBICglasso; nodes represent items (T1-T36), and edges represent conditional associations (partial correlations), with thicker edges indicating stronger associations. Item-level strength centrality values for all nodes (T1-T36) across waves are provided in Table S2 in Multimedia Appendix 1.



Network Comparison Tests

Network structure comparisons revealed significant differences between February 2020 and August 2021 (*P*=.01), and between February 2020 and February 2023 (*P*=.01), whereas the difference between August 2021 and February 2023 was not significant (*P*=.07). Global strength invariance tests indicated no significant differences in overall connectivity (*P*>.05). The results suggested that changes were mainly in the pattern or distribution of connections rather than the overall level of connectivity. Network-level summary indices and between-wave comparison metrics (including global strength, network correlations, and the proportion and number of significantly changed edges) are reported in Table S6 in Multimedia Appendix 1.

Discussion

Overview

This repeated cross-sectional study revealed 2 key findings: a statistically significant decline in nurses' infectious disease emergency response competence over successive COVID-19 pandemic phases and a marked reconfiguration of core competence priorities. In contrast to our first hypothesis, overall infectious disease emergency response competence did not increase with pandemic experience; however, our second hypothesis was confirmed by a clear shift in structural centrality. Together, these results highlight systemic fragility in sustaining workforce capacity within the health emergency response system during prolonged crises and underscore the necessity of adaptive, phase-specific reinforcement strategies.

Principal Findings

Contrary to expectations that experiential learning would bolster competence, scores for infectious disease emergency response competence declined steadily from February 2020 to February 2023. Nurses demonstrated above-average competence during the initial outbreak, which is consistent with prior findings [10,11,27]. This initial surge may be associated with short-term intensive training [10], strengthened professional motivation [28], and robust external support from governments and society [11,29]. However, these early gains were not sustained. The subsequent decline observed in later phases corresponded to patterns described in the Selye stress response theory [30], where prolonged exposure to pandemic stressors coincided with

XSL•FO
RenderX

cumulative strain from excessive workloads, escalating burnout rates [31], and progressively diminishing training investments in later phases [8]. These sustained stressors may also contribute to cognitive depletion, reducing attentional and executive resources needed for continuous competence updating in rapidly changing clinical contexts [32]. Collectively, these factors may have counteracted potential competence gains from accumulated experience.

In addition, the observed decline may partly reflect changes in self-evaluation rather than true deterioration in clinical ability. As the pandemic progressed, clinical expectations expanded from basic infection control in the early phase to navigating multiple variants, evolving treatment pathways, vaccination strategies, and long-term management issues, potentially raising the benchmark for what nurses considered "competent" practice [33]. Under such shifting standards, increasing experience may also produce a form of learned humility (an inverse Dunning–Kruger pattern) [34], whereby greater awareness of uncertainty and knowledge gaps leads to more conservative self-assessments [35]. Consistent with this interpretation, our network results suggest a reorientation of competence priorities from technical execution (T24) toward higher-order cognitive competences in information acquisition, appraisal, and integration under uncertainty (T19). Taken together, the decline in self-reported competence likely reflects an interplay of cumulative stress and cognitive depletion, information burden, and recalibrated self-assessment standards rather than a single mechanism.

Beyond the quantitative decline that contradicted hypothesis 1, network analysis validated hypothesis 2 by revealing a structural priority shift—from T24 (emergency management of body fluid exposure) during the initial outbreak to T19 (acquisition of key information on emerging infectious diseases) in the Delta and Omicron phases. This transition signifies a fundamental evolution in competence architecture: from skill-based technical execution (T24) toward higher-order cognitive competence in information synthesis and clinical decision-making under uncertainty (T19).

In the initial phase, limited COVID-19 pandemic transmission knowledge, personal protective equipment shortages, and high mortality mandated rigorous self-protection [36]. The frontline workforce faced heightened infection risk from prolonged contact [37], compounded by fears of stigmatization and acute safety anxiety [38]. Consequently, T24—enhanced personal protective equipment training and exposure protocols—dominated, reflecting a survival-driven response to immediate biological threats. As the pandemic continued, competence demands transcended technical proficiency. Hyperdynamic informational uncertainty from variants and shifting policies created an infodemic [39,40]. Under the Simon bounded rationality model [41], rapidly shifting conditions and frequent guideline revisions strained decision-making. Rising caseloads and ambiguous protocols required real - time information integration [3,42], elevating T19—the ability to rapidly acquire, filter, and contextualize key information—as the new core competence. This shift suggests that prolonged crises reprioritize competencies from procedural skills toward metacognitive capabilities for navigating complexity, and the observed decline in mean scores may partly reflect this reprioritization and a higher benchmark for competence in later phases rather than a uniform deterioration across all competency domains.

To address phase-specific challenges, we propose a tiered intervention framework: in the acute phase, prioritize biosafety through scenario-based training and secure personal protective equipment supply chains [10]. As the pandemic progresses, variant-driven policy turbulence requires the implementation of centralized information platforms (eg, Guangdong's Yuekangma health code system) [43] and the establishment of rapid pandemic information integration systems, such as AI-assisted clinical decision support tools, to help nurses identify key guidance amidst information overload. During the sustained phase, integrate cyclic skill reinforcement programs that bridge the acute and recovery phases [8,44] and offer psychological support to mitigate chronic stress [16]. We provide a phase-specific implementation toolkit with actionable details in Table S7 in Multimedia Appendix 1. This phase-specific approach, aligned with pandemic dynamics, can mitigate competence erosion and increase health care system resilience during prolonged crises.

Limitations

Several limitations warrant consideration. First, although we strengthened internal validity by applying 1:1:1 PSM to balance key staffing characteristics across waves, the repeated cross-sectional design precluded tracking individual nurses over time; residual confounding from time-varying contextual factors or unmeasured characteristics may remain. Second, participants were recruited only from Guangdong Province, where local policies and health system resources may differ from those in other regions, which may limit generalizability. Third, competence was assessed via self-reported measures, which may be influenced by reporting bias or changing self-evaluation standards as clinical expectations evolved. Fourth, most participants were from nonemergency departments; although this reflects workforce composition, the findings may not fully represent nurses in core outbreak-response units (eg, emergency department, intensive care unit, or infection control). Finally, focusing on nurses alone excludes other frontline professions (eg, physicians and public health specialists), constraining inferences about interdisciplinary preparedness.

Recommendations for Further Research

While it is universally desirable that pandemics like the COVID-19 pandemic never recur, preparedness necessitates proactive scholarship. Future research should therefore use longitudinal cohort designs to track the same cohort of nurses across pandemic phases, thereby decoding individual competence trajectories through repeated measures. Integrating mixed methods approaches could reveal how stress, experiential learning, and systemic supports interact to drive competence changes, ultimately informing personalized prevention strategies. Where feasible, future studies should triangulate self-reported competence with objective or performance-based indicators and evaluate the implementation and effectiveness of phase-specific training, information support, and psychological interventions across diverse regions and professional groups.

XSL•FO
RenderX

Conclusions

This study reveals a progressive decline in nurses' infectious disease emergency response competence across 3 years of COVID-19 pandemic response, alongside a shift in core competence from technical infection control tasks to information acquisition and integration. These findings suggest that experience alone may not sustain preparedness during prolonged crises and underscore the need for phase-specific training, robust information infrastructure, and sustained psychological support to mitigate competence erosion and strengthen workforce resilience.

Acknowledgments

The authors would like to thank all the participants for completing the surveys and the nurse directors for supporting data collection.

The authors declare the use of generative AI (GenAI) in the research and writing process. According to the GAIDeT (Generative AI Delegation Taxonomy; 2025), the following tasks were delegated to GenAI tools under full human supervision: proofreading, editing, and translation. The GenAI tool used was ChatGPT (OpenAI). Responsibility for the final manuscript lies entirely with the authors. GenAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Funding

This study was partially supported by the Guangdong Office of Philosophy and Social Science (grant GD23XGL086) and 3 grants from Guangzhou University of Chinese Medicine (GZUCM[2024]/No.278--53, GZUCM[2025]/No.252--134, and GZY2025GB0924). The funding organizations did not participate in the research or the decision to publish.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

JX, QL, and Y Lu contributed to the conception and design of the study and wrote the original draft. JX, ML, J Liu, H Li, JC, and YW contributed to data collection. JX, QL, Y Lu, and H Liu conducted the analyses. ZY and H Liu provided statistical support and helped to interpret the findings. BL, WL, YM, XS, and Y Luo helped to interpret the findings. JX and YT obtained funding. YT, H Liu, J Li, LL, and ZY contributed to methodology, editing, and supervision. All the authors interpreted the data, critically revised the manuscript for important intellectual content, and approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables detailing the COVID-19 pandemic timeline and nursing competence implications, the full 36-item competence questionnaire, participant characteristics before and after propensity score matching, wave-by-wave competence score comparisons, sensitivity analyses, network comparison metrics, and a phase-specific implementation toolkit.
[PDF File, 204 KB - publichealth_v12i1e83281_app1.pdf]

References

1. Su Z, McDonnell D, Ahmad J, Cheshmehzangi A. Disaster preparedness in healthcare professionals amid COVID-19 and beyond: a systematic review of randomized controlled trials. Nurse Educ Pract 2023 May;69:103583. [doi: 10.1016/j.nepr.2023.103583] [Medline: 36996556]
2. Learnings from COVID-19 for future respiratory pathogen pandemic preparedness: a summary of the literature. World Health Organization. 2024. URL: https://www.who.int/publications/i/item/9789240086531 [accessed 2025-12-31]
3. Nie A, Su X, Dong M, Guan W. Are nurses prepared to respond to next infectious disease outbreak: a narrative synthesis. Nurs Open 2022 Mar;9(2):908-919. [doi: 10.1002/nop2.1170] [Medline: 34994079]
4. Fleisher LA, Schreiber M, Cardo D, Srinivasan A. Health care safety during the pandemic and beyond - building a system that ensures resilience. N Engl J Med 2022 Feb 17;386(7):609-611. [doi: 10.1056/NEJMp2118285] [Medline: 35148040]
5. Qiu Q, Liu Y, Meng F, Wang X, Su C. Current status of nurses' public health emergency response capacity for emerging infectious disease: a cross-sectional study. Front Public Health 2025;13:1612790. [doi: 10.3389/fpubh.2025.1612790] [Medline: 40843425]
6. Kan T, Chen C, Huang Y, Liu J, Wang Y, Gui L. Development of a core emergency response competency index system for healthcare workers in infectious disease outbreaks [Article in Chinese]. Chinese Journal of Nursing 2018;53(4):461-466 [FREE Full text]

XSL•FO
RenderX

7. From emergency response to long-term COVID-19 disease management: sustaining gains made during the COVID-19 pandemic. World Health Organization. 2023. URL: https://www.who.int/publications/i/item/WHO-WHE-SPP-2023.1 [accessed 2025-06-23]
8. Loke AY, Guo C, Molassiotis A. Development of disaster nursing education and training programs in the past 20 years (2000-2019): a systematic review. Nurse Educ Today 2021 Apr;99:104809. [doi: 10.1016/j.nedt.2021.104809] [Medline: 33611142]
9. McNeill C, Adams L, Heagele T, Swanson M, Alfred D. Emergency preparedness competencies among nurses: implications for nurse administrators. J Nurs Adm 2020;50(7-8):407-413. [doi: 10.1097/NNA.0000000000000908] [Medline: 32675478]
10. Song S, Li X, Bell SA, Yang X, Zhang W. Emergency response: a cross-sectional study of core competencies for nurses regarding major infectious disease outbreaks. J Emerg Nurs 2021 Nov;47(6):902-913. [doi: 10.1016/j.jen.2021.04.010] [Medline: 34183192]
11. Li H, Dong S, He L, et al. Nurses' core emergency competencies for COVID-19 in China: a cross-sectional study. Int Nurs Rev 2021 Dec;68(4):524-532. [doi: 10.1111/inr.12692] [Medline: 34043839]
12. Liu J, Xiang JG, Yu WQ, Lin MZ, Wu SX. Coping ability to infectious disease emergencies among nurses in TCM hospitals and its influencing factors [Article in Chinese]. J Nurs Sci 2021;36(9):86. [doi: 10.3870/j.issn.1001-4152.2021.09.051]
13. Zhu W, Zhang J, Yang L, Li J, Guo H. Competency in responding to infectious disease outbreaks among nurses in primary healthcare institutions: a quantitative, cross-sectional multicentre study. Front Public Health 2024;12:1406400. [doi: 10.3389/fpubh.2024.1406400] [Medline: 39104898]
14. Zhang D, Chen YJ, Cui T, Zhang J, Chen SY, Zhang YP. Competence and training needs in infectious disease emergency response among Chinese nurses: Cross-sectional study. JMIR Public Health Surveill 2024 Nov 18;10:e62887. [doi: 10.2196/62887] [Medline: 39621872]
15. Hamama L, Marey-Sarwan I, Hamama-Raz Y, Nakad B, Asadi A. Psychological distress and perceived job stressors among hospital nurses and physicians during the COVID-19 outbreak. J Adv Nurs 2022 Jun;78(6):1642-1652. [doi: 10.1111/jan.15041] [Medline: 34532886]
16. Dragioti E, Tsartsalis D, Mentis M, Mantzoukas S, Gouva M. Impact of the COVID-19 pandemic on the mental health of hospital staff: an umbrella review of 44 meta-analyses. Int J Nurs Stud 2022 Jul;131:104272. [doi: 10.1016/j.ijnurstu.2022.104272] [Medline: 35576637]
17. Zhang J, Lin G, Zeng J, Lin J, Tian J, Li G. Challenges of SARS-CoV-2 and lessons learnt from SARS in Guangdong Province, China. J Clin Virol 2020 May;126:104341. [doi: 10.1016/j.jcv.2020.104341] [Medline: 32278300]
18. Li L, Han ZG, Qin PZ, et al. Transmission and containment of the SARS-CoV-2 Delta variant of concern in Guangzhou, China: a population-based study. PLoS Negl Trop Dis 2022 Jan;16(1):e0010048. [doi: 10.1371/journal.pntd.0010048] [Medline: 34986169]
19. Qu L, Xie C, Qiu M, et al. Characterizing infections in two epidemic waves of SARS-CoV-2 Omicron variants: a cohort study in Guangzhou, China. Viruses 2024 Apr 22;16(4):649. [doi: 10.3390/v16040649] [Medline: 38675989]
20. Liu L, Zhang B, Yang Z. Investigation on medical workers' core emergency response competences in epidemics of infectious diseases [Article in Chinese]. J Nurs Sci 2019;34(5):75-77. [doi: 10.3870/j.issn.1001-4152.2019.05.075]
21. Kan T. Development and evaluation of a training program for healthcare workers' response capacity to infectious disease emergencies [dissertation]. : Naval Medical University; 2019 URL: https://xueshu.baidu.com/ndscholar/browse/detail?paperid=122q0ex0hb6j0jv0704n08j018574347 [accessed 2024-05-05]
22. Guo XE, Bian LF, Li Y, Li CY, Lin Y. Common domains of nurses' competencies in public health emergencies: a scoping review. BMC Nurs 2023 Dec 21;22(1):490. [doi: 10.1186/s12912-023-01655-5] [Medline: 38124048]
23. DecisionLinnc. 2023. URL: https://www.statsape.com/ [accessed 2025-12-25]
24. Li J, Li F, Luo D, Liu E, Lan H. Vitamin C supplementation during intensive care unit stay is associated with improved outcomes in critically ill patients with sepsis-induced coagulopathy: a cohort study. Food Sci Nutr 2025 Dec;13(12):e71263. [doi: 10.1002/fsn3.71263] [Medline: 41292678]
25. McNally RJ, Heeren A, Robinaugh DJ. A Bayesian network analysis of posttraumatic stress disorder symptoms in adults reporting childhood sexual abuse. Eur J Psychotraumatol 2017;8(sup3):1341276. [doi: 10.1080/20008198.2017.1341276] [Medline: 29038690]
26. Epskamp S, Borsboom D, Fried EI. Estimating psychological networks and their accuracy: a tutorial paper. Behav Res Methods 2018 Feb;50(1):195-212. [doi: 10.3758/s13428-017-0862-1] [Medline: 28342071]
27. Xu X, Xu J, Kang L, Zhang X, Pan W. Survey and analysis of the emergency response capacity to infectious disease outbreaks among nurses from Shanghai assisting Hubei [Article in Chinese]. Chinese Journal of Nursing Management 2020;20(12):1822-1826. [doi: 10.3969/j.issn.1672-1756.2020.12.014]
28. Xiong Y, Peng L. Focusing on health-care providers' experiences in the COVID-19 crisis. Lancet Glob Health 2020 Jun;8(6):e740-e741. [doi: 10.1016/S2214-109X(20)30214-X] [Medline: 32573442]
29. Alan H, Eskici GT, Sen HT, Bacaksiz FE. Nurses' disaster core competencies and resilience during the COVID-19 pandemic: a cross-sectional study from Turkey. J Nurs Manag 2022 Apr;30(3):622-632. [doi: 10.1111/jonm.13552] [Medline: 35083821]
30. Cunanan AJ, DeWeese BH, Wagle JP, et al. The general adaptation syndrome: a foundation for the concept of periodization. Sports Med 2018 Apr;48(4):787-797. [doi: 10.1007/s40279-017-0855-3] [Medline: 29307100]

XSL•FO
RenderX

31. Ulfa M, Azuma M, Steiner A. Burnout status of healthcare workers in the world during the peak period of the COVID-19 pandemic. Front Psychol 2022;13:952783. [doi: 10.3389/fpsyg.2022.952783] [Medline: 36211838]
32. Cortés-Álvarez NY, Lara-Morales A, Bautista-Rodríguez E, et al. Job burnout, cognitive functioning, and Brain-derived neurotrophic factor expression among hospital Mexican nurses. PLOS ONE 2024;19(5):e0304092. [doi: 10.1371/journal.pone.0304092] [Medline: 38787900]
33. Bhimraj A, Falck-Ytter Y, Kim AY, et al. 2024 Clinical practice guideline update by the Infectious Diseases Society of America on the management of COVID-19: anti-SARS-CoV-2 neutralizing antibody pemivibart for pre-exposure prophylaxis. Clin Infect Dis 2026 Jan 16;81(Supplement_3):i6-i12. [doi: 10.1093/cid/ciae435] [Medline: 39471458]
34. Kruger J, Dunning D. Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. J Pers Soc Psychol 1999 Dec;77(6):1121-1134. [doi: 10.1037//0022-3514.77.6.1121] [Medline: 10626367]
35. Jansen RA, Rafferty AN, Griffiths TL. A rational model of the Dunning-Kruger effect supports insensitivity to evidence in low performers. Nat Hum Behav 2021 Jun;5(6):756-763. [doi: 10.1038/s41562-021-01057-0] [Medline: 33633375]
36. Mason DJ, Friese CR. Protecting health care workers against COVID-19-and being prepared for future pandemics. JAMA Health Forum 2020 Mar 2;1(3):e200353. [doi: 10.1001/jamahealthforum.2020.0353] [Medline: 36218598]
37. Smith GD, Ng F, Ho Cheung Li W. COVID-19: emerging compassion, courage and resilience in the face of misinformation and adversity. J Clin Nurs 2020 May;29(9-10):1425-1428. [doi: 10.1111/jocn.15231] [Medline: 32155302]
38. Nabavian M, Rahmani N, Seyed Nematollah Roshan F, Firouzbakht M. Nurses' experiences of the social stigma caused by the COVID-19 pandemic: a qualitative study. J Res Nurs 2023 Mar;28(2):104-115. [doi: 10.1177/17449871231159604] [Medline: 37152193]
39. Depoux A, Martin S, Karafillakis E, Preet R, Wilder-Smith A, Larson H. The pandemic of social media panic travels faster than the COVID-19 outbreak. J Travel Med 2020 May 18;27(3):taaa031. [doi: 10.1093/jtm/taaa031] [Medline: 32125413]
40. Chong YY, Cheng HY, Chan HYL, Chien WT, Wong SYS. COVID-19 pandemic, infodemic and the role of eHealth literacy. Int J Nurs Stud 2020 Aug;108:103644. [doi: 10.1016/j.ijnurstu.2020.103644] [Medline: 32447127]
41. Simon HA. A behavioral model of rational choice. Q J Econ 1955 Feb;69(1):99. [doi: 10.2307/1884852]
42. Song P, Karako T. COVID-19: Real-time dissemination of scientific information to fight a public health emergency of international concern. Biosci Trends 2020 Mar 16;14(1):1-2. [doi: 10.5582/bst.2020.01056] [Medline: 32092748]
43. Shen T, Wang C. Regulating China's health code system to prepare for future pandemics. Front Public Health 2023;11:1208615. [doi: 10.3389/fpubh.2023.1208615] [Medline: 37841734]
44. Core competencies in disaster nursing: version 2.0. International Council of Nurses. 2019. URL: https://www.icn.ch/resources/publications-and-reports/core-competencies-disaster-nursing-version-20 [accessed 2025-05-26]

Abbreviations

PSM: propensity score matching
SMD: standardized mean difference
CS: correlation stability

Edited by A Mavragani, T Sanchez; submitted 31.Aug.2025; peer-reviewed by LN Mayigane, Z Ma; revised version received 23.Jan.2026; accepted 24.Feb.2026; published 27.Mar.2026.

Please cite as:
Xiang J, Liang Q, Lu Y, Lin M, Liu J, Li H, Cai J, Wu Y, Li B, Liu W, Yang M, Sun X, Luo Y, Tan Y, Liu H, Li J, Lu L, Ye Z
Trends and Core Competence Shifts in Nurses' Infectious Disease Emergency Response Competence Across COVID-19 Pandemic Phases: Repeated Cross-Sectional Survey and Network Analysis
JMIR Public Health Surveill 2026;12:e83281
URL: https://publichealth.jmir.org/2026/1/e83281
doi:10.2196/83281

© Jiagen Xiang, Qiuxia Liang, Yufei Lu, Meizhen Lin, Juan Liu, Hongman Li, Jingjing Cai, Yan Wu, Bingbing Li, Wenting Liu, Ming Yang, Xiaoning Sun, Yan Luo, Yibing Tan, Hairun Liu, Jiaying Li, Liming Lu, Zengjie Ye. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 27.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Gestational Diabetes Mellitus and Associated Factors in Rural Areas of Northern Vietnam: Cross-Sectional Survey

Duc Cuong Le, MD, PhD; Thi Hien Nguyen, MD, PhD; Thi Ly Nguyen, MD; Thi Nuong Tran, MSc, MD; Xuan Hien Luong, PhD; Thi Thu Huong Do, MSc, MD; Xuan Thuy Tran, MD, PhD; Van Manh Ngo, MD, PhD; Thi Huyen Dieu Bui, MD, PhD; Tien Van Nguyen, PhD; Van Thuan Hoang, MD, PhD; Thanh Binh Vu, MD, PhD

Thai Binh University of Medicine and Pharmacy, 373 Ly Bon Street, Thai Binh, Vietnam

Corresponding Author:
Thanh Binh Vu, MD, PhD
Thai Binh University of Medicine and Pharmacy, 373 Ly Bon Street, Thai Binh, Vietnam

Abstract

Background: The prevalence of gestational diabetes mellitus (GDM) and its associated factors have not been investigated in Vietnam, especially in rural areas.

Objective: This study aimed to determine the prevalence of GDM and its associated factors in rural areas of Vietnam.

Methods: This cross-sectional study was conducted among 1003 pregnant women aged ≥18 years in rural areas of northern Vietnam. GDM was identified based on impaired oral glucose tolerance test results according to the guidelines of the International Association of the Diabetes and Pregnancy Study Groups. Associated factors for GDM were analyzed using a modified Poisson regression with robust (sandwich) SE analysis, with maternal age, prepregnancy BMI, and family history of diabetes as explanatory factors, adjusted for parity, method of conception, hormonal therapy for pregnancy maintenance, physical inactivity, and history of chronic medical conditions.

Results: The prevalence of GDM was 26.2%. Patients with GDM were more likely to be older than the control group (odds ratio 3.33, 95% CI 2.31 - 4.78). In the multivariable analysis, maternal age was strongly associated with GDM. Compared with women aged <25 years, those aged 25 to 34 years had a significantly higher prevalence of GDM (adjusted prevalence ratio [PR] 1.50, 95% CI 1.12 - 1.99; *P*=.006). The rate was even higher among women aged ≥35 years (adjusted PR 2.40, 95% CI 1.74 - 3.31; *P*<.001). These associations remained consistent after further adjustment for confounders (25 - 34 years: adjusted PR 1.44, 95% CI 1.07 - 1.95; *P*=.02 and ≥35 years: adjusted PR 2.15, 95% CI 1.49 - 3.11; *P*<.001). Overweight women (BMI 23 to <25 kg/m²) showed a borderline association with GDM, although this did not reach statistical significance (adjusted PR 1.39, 95% CI 0.98 - 1.98; *P*=.06), while women with BMI ≥25 kg/m² had a significantly higher prevalence (adjusted PR 1.58, 95% CI 1.10 - 2.26; *P*=.01). These findings persisted in the adjusted model, with BMI ≥25 kg/m² remaining significantly associated with GDM (adjusted PR 1.54, 95% CI 1.10 - 2.17; *P*=.01). Women with a family history of diabetes had an increased prevalence of GDM in both models, although the association did not reach statistical significance (adjusted PR 1.51, 95% CI 0.85 - 2.68; *P*=.16). Most additional covariates included in the multivariable analysis were not significant. However, women who conceived via in vitro fertilization had a significantly higher prevalence of GDM compared with those who conceived naturally (adjusted PR 1.38, 95% CI 1.04 - 1.85; *P*=.03). The use of hormonal therapy to maintain pregnancy was also associated with an increased risk (adjusted PR 1.32, 95% CI 1.04 - 1.68; *P*=.02).

Conclusions: The findings highlight the need for early screening and counseling before pregnancy. Lifestyle programs should focus on weight control and healthy habits. Future research should test whether these programs can reduce the rate of GDM in rural communities.

(JMIR Public Health Surveill 2026;12:e79688) doi:10.2196/79688

KEYWORDS

gestational diabetes mellitus; obesity; diabetes; risk factors; pregnant women

Introduction

Gestational diabetes mellitus (GDM) is the most common metabolic disorder in pregnancy, affecting approximately 16.5% of pregnancies worldwide, and has become an increasingly major global public health concern because of its adverse implications for maternal and fetal health [1-3]. GDM is defined as any degree of glucose intolerance with the onset of or first recognition during pregnancy [4]. According to the report of the International Diabetes Federation, almost 21.3 million live births to women, or 16% of live births, were affected by some forms of high blood glucose in pregnancy. Moreover, the vast

XSL•FO
RenderX

majority of pregnant women with GDM are in low- and middle-income countries, where access to maternal care is often limited [5,6]. Pregnant women exposed to maternal hyperglycemia from 24 to 28 weeks of gestation onward, if left undiagnosed and untreated, may experience serious short- and long-term outcomes for both mothers and their infants. GDM may cause pregnancy complications, such as preeclampsia or eclampsia or low blood sugar, and tends to be a major risk factor for the development of type 2 diabetes that develops later in life [7-9]. Therefore, identifying the prevalence of GDM and understanding the associated factors for GDM will offer the opportunity to improve pregnancy outcomes and maternal health as well as to drive health-promoting messages about diet and lifestyle.

The prevalence of GDM varies widely across countries, largely due to the use of different diagnostic criteria for GDM. The most recent cross-sectional study and meta-analysis reported that the global prevalence of GDM was 35% based on the new World Health Organization (WHO) diagnostic criteria for GDM and 14.7% based on the International Association of Diabetes and Pregnancy Study Groups (IADPSG) criteria [10,11]. In Lima, Peru, approximately 16% of pregnant women were diagnosed with GDM based on the IADPSG criteria [12]. In 2019, a meta-analysis using the same criteria reported that the prevalence of GDM in China was 14.8% [13]. Another recent systematic review and meta-analysis reported that the pooled prevalence of GDM ranged from 9% to 16%, with higher prevalence observed among pregnant women in urban populations [14].

The development of GDM in pregnant women has been associated with several factors. Previous studies have consistently identified advanced maternal age, obesity, and a family history of type 2 diabetes as major risk factors for GDM compared with women without these characteristics [3,15,16]. Other associated factors for GDM include a history of unexplained stillbirth [16].

Vietnam is undergoing a rapid epidemiologic transition characterized by rising rates of overweight, obesity, and noncommunicable diseases, including diabetes. Several studies from urban or periurban regions of Vietnam have reported GDM prevalence estimates ranging from 7% to 25% depending on diagnostic criteria [11,17]. However, data on GDM burden in rural areas, where access to maternal health services is more limited, remain limited [18,19]. Rural populations face unique challenges, including limited screening capacity, lower health literacy, and substantial socioeconomic disparities, which may contribute to underdiagnosis or delayed detection of GDM. Updated evidence on the magnitude and associated factors of GDM in rural Vietnam is needed to inform targeted interventions and maternal health policies. Therefore, we conducted this study to determine the prevalence of GDM and identify key maternal factors associated with GDM among pregnant women living in rural areas of northern Vietnam.

Methods

Participants

This cross-sectional study was conducted in Thai Binh province from June 2021 to May 2022. It is located approximately 100 km north of Hanoi, the capital city of Vietnam. The economy is mostly agricultural. The total population is 1.86 million people, the gross domestic product per capita is US $1650, and more than 90% of the population lives in rural areas. Thai Binh province is subdivided into 7 districts and 1 provincial city. The study was conducted at 3 hospitals in a provincial city, including Medicine University, Provincial General, and Gyneco-Obstetrics Hospital.

Recruitment Procedures

Recruitment was conducted at the antenatal outpatient clinics of the 3 study hospitals. These clinics provide routine antenatal services, including pregnancy monitoring, screening for obstetric complications, counseling, and laboratory testing, including oral glucose tolerance testing for GDM.

A systematic and consecutive recruitment approach was applied. During the study period, all pregnant women attending antenatal clinics were screened for eligibility. Women were eligible if they were aged 18 years or older and at 24 to 28 weeks of gestation, which corresponds to the routine screening period for GDM in Vietnam. Women with a known diagnosis of diabetes mellitus prior to pregnancy were excluded.

Eligible women were consecutively approached by trained research staff during their routine antenatal visit. Recruitment was conducted on all clinic working days to minimize selection bias. The study objectives, procedures, potential risks, and benefits were explained verbally in Vietnamese, the native language of the participants. Women were given sufficient time to ask questions and consider participation.

Questionnaire

The pregnant women were interviewed using a questionnaire constructed in Vietnamese, the native language of the participants. Information on the following potential associated factors was collected during this survey: women's age in years; women's education level (high school or below or upper high school); women's occupation (housewife, farmer, or small business; worker; state officer; private officer; or others); family history of diabetes (yes or no); parity (nulliparous or multiparous); pregnancy method (natural conception or in vitro fertilization [IVF]); height and weight in the prepregnancy period to calculate BMI; hormonal therapy for pregnancy maintenance; limited movement and exercise, defined as less than 3 hours of physical exercise per week; and medical history of any chronic disease.

GDM Diagnostic Criteria

GDM was diagnosed according to the guidelines of the IADPSG using the 75-g oral glucose tolerance test with 3 venous blood samples. The criteria included diagnosis based on any of the following three values of plasma glucose: (1) fasting blood glucose level of 5.1 to 6.9 mmol/L (92 - 125 mg/dL), (2) 1-hour blood glucose level of ≥10.0 mmol/L (180 mg/dL), and (3)

XSL•FO
RenderX

2-hour blood glucose level of 8.5 to 11.0 mmol/L (153 - 199 mg/dL).

Statistical Analysis

The data were entered using Epidata (version 3.1; EpiData Association) software. Stata (version 15.0; StataCorp LLC) software was used for statistical analysis. General characteristics of the study participants were presented as numbers and percentages. Chi-square or Fisher exact tests were conducted to compare the differences in proportion between 2 groups of patients.

The primary outcome was GDM, defined as a binary variable (yes or no). The main explanatory variables included maternal age (<25, 25 - 34, and ≥35 years), family history of diabetes (yes or no), and prepregnancy BMI, which was categorized according to the classification for Asian populations (<23, 23-<25, and ≥25 kg/m²) [20]. The selection of these variables was based on previous studies [11,12,15,21].

Associations between each explanatory variable and GDM were first explored using univariable analysis to estimate crude prevalence ratios (crude PRs) and 95% CI. Subsequently, a multivariable modified Poisson regression model with robust (sandwich) SEs (model 1) was constructed to assess the association between explanatory factors and GDM. A second model (model 2) was fitted to further adjust for potential confounders identified a priori based on clinical relevance and previous literature. These included parity (nulliparous vs multiparous), method of conception (natural vs IVF), hormonal therapy for pregnancy maintenance (yes or no), physical inactivity (defined as <3 hours of exercise per week), and history of chronic medical conditions (yes or no). The results from both model 1 and model 2 were presented as adjusted prevalence ratios (adjusted PRs) with corresponding 95% CIs. A *P* value of <.05 and a 95% CI were chosen as the statistical significance level and CI, respectively.

Ethical Considerations

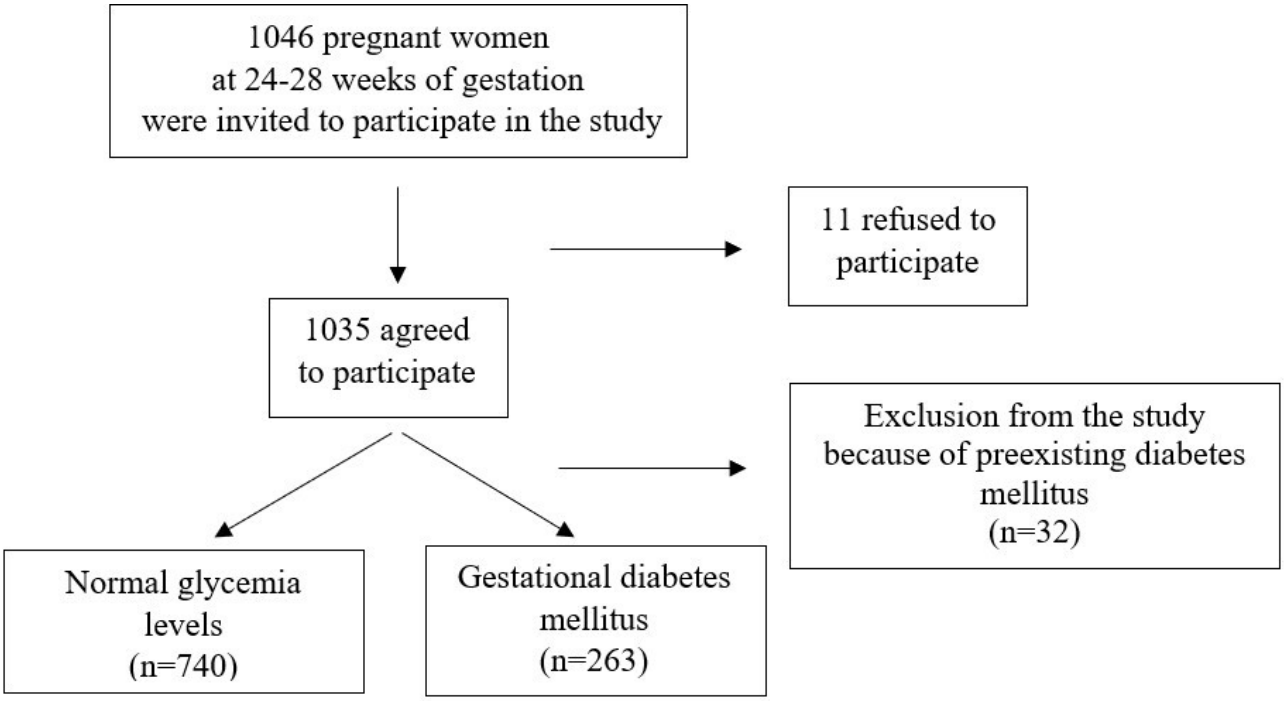
The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of the Thai Binh Department of Science and Technology (protocol code 1369/QĐ-UBND, approved on June 10, 2021) for studies involving humans. All participants provided oral informed consent. Participation in the study was completely voluntary. Participants were free to withdraw from the study at any time without any consequences for their medical care. The data were protected and deidentified before analysis. No financial or material compensation was provided to participants.

Results

Overview

A total of 1046 pregnant women at 24 to 28 weeks of gestation were invited, and 11 (1.1%) refused to participate. In addition, 32 (3.0%) were excluded because of preexisting diabetes mellitus. Of the 1003 (95.9%) included women, 263 were diagnosed with GDM; hence, the GDM proportion was 26.2% (Figure 1).

Figure 1. Study flow diagram.



Characteristics of Participants

Table 1 presents the characteristics of the participants stratified according to GDM status, including sociodemographic characteristics and clinical characteristics of participants. A proportion of 22.8% (60/263) of women with GDM were aged ≥35 years, compared with 10% (74/740) of women without GDM (*P*<.001). In addition, the prevalence of GDM was significantly higher among women with a BMI ≥25 kg/m^2, a family history of diabetes, a history of 2 or more pregnancies, conception via IVF, and use of endocrine medicines, compared

XSL•FO
RenderX

with women without GDM. There was no difference in blood pressure classification between the 2 groups (Table 1).

XSL•FO
RenderX

Table . Background characteristics of participants.

Characteristics	Gestational diabetes mellitus, n (%)		*P* value[a]
	Yes (n=263)	No (n=740)	
Age (years)			<.001
<25	50 (19.0)	240 (32.4)	
25 - 34	153 (58.2)	426 (57.6)	
≥35	60 (22.8)	74 (10.0)	
Educational background (levels)			.16
High school and below	97 (36.9)	310 (41.9)	
Upper high school	166 (63.1)	430 (58.1)	
Occupations			.74
Housewife, farmer, or small business	70 (26.6)	203 (27.4)	
Worker	58 (22.1)	171 (23.1)	
State officer	31 (11.8)	105 (14.2)	
Private officer	90 (34.2)	227 (30.7)	
Others	14 (5.3)	34 (4.6)	
BMI (kg/m^2)			.002
<23	223 (84.8)	682 (92.2)	
23 to <25	23 (8.7)	35 (4.7)	
≥25	17 (6.5)	23 (3.1)	
Family history of diabetes			.02
No	255 (97.0)	731 (99.1)	
Yes	8 (3.0)	7 (0.9)	
Parity			.003
Nulliparous	122 (46.4)	422 (57.0)	
Multiparous	141 (53.6)	318 (43.0)	
Pregnancy method			<.001
Natural	227 (86.3)	684 (94.5)	
In vitro fertilization	36 (13.7)	40 (5.5)	
Hormonal therapy for pregnancy maintenance			<.001
Yes	86 (35.5)	143 (20.7)	
No	156 (64.5)	547 (79.3)	
Limited movement and exercise (<3 hours of physical exercise per week)			.31
Yes	63 (24.0)	155 (21.0)	
No	200 (76.0)	585 (79.0)	
Medical history of chronic diseases			.47
Yes	17 (6.5)	39 (5.3)	
No	246 (93.5)	701 (94.7)	
Blood pressure (systolic and diastolic; mmHg)			.50
<120 and <80	469 (63.4)	1256 (59.3)	
120 - 129 and <80	98 (13.2)	33 (12.6)	
130 - 139 or 80 - 89	151 (20.4)	65 (24.7)	
≥140 or ≥90	22 (3.0)	9 (3.4)	

XSL•FO
RenderX

[a]Fisher exact or chi-square test.

Associated Factors for GDM

In the univariable analysis, maternal age, BMI, family history of diabetes, parity, pregnancy method, and hormonal therapy were significantly associated with GDM. Women aged 25 to 34 years and ≥35 years had higher prevalence of GDM compared with those aged <25 years (crude PR 1.53 and 2.60, respectively). Higher prepregnancy BMI increased the prevalence of GDM, with crude PRs of 1.60 for BMI 23 to <25 kg/m² and 1.72 for BMI ≥25 kg/m². A family history of diabetes was associated with a higher prevalence of GDM (crude PR 2.07). Multiparous women also had an increased prevalence of GDM (crude PR 1.37). IVF and use of hormonal therapy were strongly associated with GDM (crude PR 1.93 and 1.64, respectively). Limited physical activity and a history of chronic disease were not significantly associated with GDM (Table 2).

Table . Associated factors of gestational diabetes mellitus, univariate analysis.

Factors	Crude PR[a] (95% CI)	*P* value
Age (years)		
<25	Reference	
25 - 34	1.53 (1.15 - 2.04)	.003
≥35	2.60 (1.90 - 3.56)	<.001
BMI (kg/m^2)		
<23	Reference	
23 to <25	1.60 (1.15 - 2.26)	.006
≥25	1.72 (1.18 - 2.52)	.005
Family history of diabetes		
No	Reference	
Yes	2.07 (1.27 - 3.36)	.003
Parity		
Nulliparous	Reference	
Multiparous	1.37 (1.11 - 1.69)	.003
Pregnancy method		
Natural	Reference	
In vitro fertilization	1.93 (1.49 - 2.52)	<.001
Hormonal therapy for pregnancy maintenance		
No	Reference	
Yes	1.64 (1.33 - 2.03)	<.001
Limited movement and exercise (<3 hours of physical exercise per week)		
No	Reference	
Yes	1.13 (0.89 - 1.44)	.30
Medical history of chronic diseases		
No	Reference	
Yes	1.17 (0.77 - 1.76)	.46

[a]PR: prevalence ratio.

In the multivariable analysis (model 1), maternal age was strongly associated with GDM. Compared with women aged <25 years, those aged 25 to 34 years had a significantly higher prevalence of GDM (adjusted PR 1.50, 95% CI 1.12 - 1.99; *P*=.006). The prevalence was even higher among women aged ≥35 years (adjusted PR 2.40, 95% CI 1.74 - 3.31; *P*<.001). These associations remained consistent in model 2 after further adjustment for confounders (25 - 34 years: adjusted PR 1.44, 95% CI 1.07 - 1.95; *P*=.02 and ≥35 years: adjusted PR 2.15, 95% CI 1.49 - 3.11; *P*<.001).

Overweight women (BMI 23 to <25 kg/m²) showed a borderline association with GDM, although this did not reach statistical significance (adjusted PR 1.39, 95% CI 0.98 - 1.98; *P*=.06), while women with BMI ≥25 kg/m² had a significantly higher prevalence (adjusted PR 1.58, 95% CI 1.10 - 2.26; *P*=.01). These findings persisted in model 2, with BMI ≥25 kg/m²

XSL•FO
RenderX

remaining significantly associated with GDM (adjusted PR 1.54, 95% CI 1.10 - 2.17; *P*=.01).

Women with a family history of diabetes had an increased prevalence of GDM in both models, although this did not reach statistical significance (model 1: adjusted PR 1.78; 95% CI 1.01 - 3.14; *P*=.04 and model 2: adjusted PR 1.51, 95% CI 0.85 - 2.68; *P*=.16).

Most additional covariates included in model 2 were not significant. However, women who conceived via IVF had a significantly higher prevalence of GDM compared with those who conceived naturally (adjusted PR 1.38, 95% CI 1.04 - 1.85; *P*=.03). The use of hormonal therapy to maintain pregnancy was also associated with an increased prevalence (adjusted PR 1.32, 95% CI 1.04 - 1.68; *P*=.02; Table 3).

Table . Modified Poisson regression model with robust (sandwich) SEs analysis of factors associated with gestational diabetes mellitus.

Associated factors	Model 1[a]		Model 2[b]	
	Adjusted PR[c] (95% CI)	*P* value	Adjusted PR (95% CI)	*P* value
Age (years)				
<25	Reference		Reference	
25 - 34	1.50 (1.12 - 1.99)	.006	1.44 (1.07 - 1.95)	.02
≥35	2.40 (1.74 - 3.31)	<.001	2.15 (1.49 - 3.11)	<.001
BMI (kg/m^2)				
<23	Reference		Reference	
23 to <25	1.39 (0.98 - 1.98)	.06	1.34 (0.94 - 1.92)	>.99
≥25	1.58 (1.10 - 2.26)	.01	1.54 (1.10 - 2.17)	.01
Family history of diabetes				
No	Reference		Reference	
Yes	1.78 (1.01 - 3.14)	.04	1.51 (0.85 - 2.68)	.16
Parity				
Nulliparous	—[d]	—	Reference	
Multiparous	—	—	0.99 (0.78 - 1.25)	.93
Pregnancy method				
Natural	—	—	Reference	
In vitro fertilization	—	—	1.38 (1.04 - 1.85)	.03
Hormonal therapy for pregnancy maintenance				
No	—	—	Reference	
Yes	—	—	1.32 (1.04 - 1.68)	.02
Limited movement and exercise (<3 hours of physical exercise per week)				
No	—	—	Reference	
Yes	—	—	0.94 (0.75 - 1.91)	.62
Medical history of chronic diseases				
No	—	—	Reference	
Yes	—	—	0.96 (0.65 - 1.42)	.83

[a]Model 1 includes maternal age, prepregnancy BMI, and family history of diabetes.

[b]Model 2 includes all variables in model 1 plus parity, pregnancy method, hormonal therapy for pregnancy maintenance, limited physical activity (<3 hours per week), and medical history of chronic diseases.

[c]PR: prevalence ratio.

[d]Not applicable.

XSL•FO
RenderX

Discussion

Principal Findings

This cross-sectional survey provides estimates of the prevalence of GDM among pregnant women, as well as the associations between specific demographic characteristics and GDM, in rural areas of Vietnam. To the best of our knowledge, this is the first study to examine the prevalence of GDM by the IADPSG criteria in Vietnam. In this cross-sectional study of 1003 pregnant women in rural northern Vietnam, we found that the prevalence of GDM was 26.2%. Consistent with the study objectives, several sociodemographic and clinical factors were identified as significantly associated factors for GDM. Older maternal age demonstrated a clear dose-response relationship, with women aged ≥35 years having the highest odds of GDM. Prepregnancy overweight and obesity (BMI ≥23 kg/m²) were also strongly associated with increased GDM risk. A family history of diabetes showed a positive but borderline association.

This prevalence was higher than that reported in previous studies. For instance, the overall prevalence of GDM among study participants in Ghana was 8.5% [22], and in Lima, it was approximately 16% [12]. In Eastern and Southeast Asia, the prevalence of GDM was 11.91% in China, while in Japan, Korea, and Thailand, it was less than 8.0% [23]. A recent prospective cohort study showed that the incidence and age-adjusted incidence of GDM in China were 17.42% and 17.45%, respectively [24].

In this study, we observed that maternal age is an associated factor for GDM. This finding is consistent with studies by Lin et al [25] and Li et al [24], which reported that older maternal age was associated with a higher risk of developing GDM compared with women younger than 35 years. The underlying mechanism for this association may be attributed to hormonal changes that occur as women age, leading to alterations in glucose and insulin processing and consequently raising the risk of GDM during pregnancy. Therefore, this study reinforced the findings of previous studies and indicated that pregnancy is better planned before the age of 35 years, and it is crucial to implement a regular screening plan for GDM to detect and manage the condition effectively, thus ensuring optimal health outcomes for both the mother and the baby.

Our study also showed that pregnant women with BMI ≥25 kg/m² had a significantly higher prevalence of GDM compared with women of normal weight. In addition, those with a BMI of 23 to 25 kg/m² demonstrated a trend toward an increased prevalence of GDM, suggesting a graded association across Asian-specific BMI categories. Although the association in this intermediate BMI category did not reach conventional statistical significance, it represents a clinically important early warning stage, rather than a neutral risk group. This finding is consistent with previous studies [3,12,15,16] and supports the use of lower BMI cutoffs for early risk identification. From a clinical and public health perspective, this finding is particularly important in rural settings, where women with BMI of 23 to 25 kg/m² are often not identified as high risk and may not receive targeted counseling or early screening. Our results suggest that this BMI range could serve as a practical early warning threshold, offering an opportunity for lifestyle modification, dietary counseling, and closer metabolic monitoring to prevent progression to overt obesity and clinically manifest GDM. Overweight and obesity are closely linked to excessive and prolonged calorie intake, which might overwhelm pancreatic β-cell insulin production and insulin signaling pathways. Even independent of BMI, diet and nutrition also play an important role in the development of GDM. Diets high in saturated fat, refined sugar, and red or processed meat have been associated with an increased risk of GDM [3], while diets high in fiber, micronutrients, and polyunsaturated fats are consistently associated with a reduced risk of GDM. Saturated fats directly interfere with insulin signaling. In addition, they can also cause inflammation and endothelial dysfunction. These are the 2 main factors that cause GDM [3].

Women who conceived after IVF had a higher prevalence of GDM in our study. This result suggests that IVF pregnancy is an independent associated factor for GDM in this rural population. The finding is consistent with many previous studies [26-28]. These analyses reported that the proportion of GDM after assisted reproductive technology was often approximately twofold higher. A systematic review and meta-analysis also showed that assisted reproductive technology is associated with about a 1.5-fold increase in GDM [29]. The magnitude of effect in our study is similar to that in the literature. This supports the idea that IVF itself, or factors linked to IVF, may contribute to disturbed glucose metabolism in pregnancy. Several mechanisms may explain this association. Women who need IVF often have older age or long-standing infertility. Many have conditions such as polycystic ovary syndrome or obesity. These factors increase insulin resistance and baseline diabetes risk. Supraphysiologic hormone levels during ovarian stimulation may further worsen insulin resistance [30]. IVF pregnancies also receive closer surveillance. This can increase the detection of milder GDM. Our models were already adjusted for age, BMI, and parity. However, residual confounding by unmeasured fertility-related factors may still be present.

The use of hormonal therapy to maintain pregnancy was also associated with GDM in our study. Indeed, progesterone and other progestogens reduce insulin sensitivity and are part of the hormonal milieu that drives insulin resistance in pregnancy [31]. Meta-analyses on progestogen supplementation to prevent preterm birth suggest a possible increase in GDM risk [32]. In addition, hormonal therapy for pregnancy maintenance likely marks a subgroup of women with threatened miscarriage, luteal phase problems, or other endocrine disorders. These women may already carry a higher metabolic risk.

These findings have important clinical implications. IVF conception and the use of hormonal therapy can be viewed as simple clinical markers for higher GDM risk. Women who receive these treatments may benefit from early glucose screening, closer metabolic follow-up, and strong counseling on weight control and physical activity. Preconception assessment in fertility clinics could also help to identify impaired glucose tolerance before pregnancy.

Although the association was not significant, our study also suggested that a family history of diabetes may be a potential

XSL•FO
RenderX

risk factor for GDM. Similarly, Lin et al [25] and Rhee et al [33] reported that a family history of diabetes is a major risk factor for GDM in Asian populations. This may be explained by the fact that women with GDM often carry susceptibility genes for type 2 diabetes [34]. Therefore, pregnant women with a family history of diabetes should focus on weight management, regular exercise, healthy diet, and monitoring blood sugar levels to prevent GDM. However, the association between obstetric history and family history of diabetes with GDM was only observed in our univariate analysis.

Some limitations must be considered when interpreting the results of our study. First, the use of a cross-sectional study design limits the findings of this study, as causal relationships cannot be established. Second, family history of diabetes was assessed based on self-report; therefore, we cannot rule out recall bias. Additionally, participants in this study were pregnant women living in Thai Binh, which may limit the generalizability of our findings to other populations. Moreover, our study relied on self-reported lifestyle changes and did not assess adherence to these changes, which may have influenced the outcomes. Finally, the study did not include important socioeconomic indicators. We did not collect data on income, living conditions, access to health services, or socioeconomic status. These factors may influence the risk of GDM, especially in rural settings. Their absence may limit the interpretation of our results.

Conclusions

Our study found that maternal age ≥35 years and prepregnancy obesity were associated with an increased prevalence of GDM. Women who conceived through IVF or used hormonal therapy also had higher rates of GDM. These results show that rural women face important metabolic risks during pregnancy. The findings highlight the need for early screening and counseling before pregnancy. Lifestyle programs should focus on weight control and healthy habits. Future research should test whether these programs can reduce the rate of GDM in rural communities.

Acknowledgments

The authors would like to acknowledge all the participants and researchers who contributed to this study. The authors also thank the staff of Medicine University, the Provincial General Hospital, and the Gyneco-Obstetrics Hospital for their support in data collection. Generative artificial intelligence was not used in any part of this work.

Funding

The authors declare that no financial support was received for this work.

Authors' Contributions

All authors contributed to the study design, data collection, and preparation of the manuscript and approved the final version. DCL, VTH, and TVN independently performed the statistical analyses and confirmed the consistency of the results. All authors participated in drafting the article or revising it critically for important intellectual content, approved the final version to be published, and agreed to be accountable for all aspects of the work.

Conflicts of Interest

None declared.

References

1. Zhu Y, Zhang C. Prevalence of gestational diabetes and risk of progression to type 2 diabetes: a global perspective. Curr Diab Rep 2016 Jan;16(1):7. [doi: 10.1007/s11892-015-0699-x] [Medline: 26742932]
2. Lowe WL Jr, Scholtens DM, Lowe LP, et al. Association of gestational diabetes with maternal disorders of glucose metabolism and childhood adiposity. JAMA 2018 Sep 11;320(10):1005-1016. [doi: 10.1001/jama.2018.11628] [Medline: 30208453]
3. Plows JF, Stanley JL, Baker PN, Reynolds CM, Vickers MH. The pathophysiology of gestational diabetes mellitus. Int J Mol Sci 2018 Oct 26;19(11):3342. [doi: 10.3390/ijms19113342] [Medline: 30373146]
4. Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Report of the Expert Committee on the diagnosis and classification of diabetes mellitus. Diabetes Care 2003 Jan;26 Suppl 1:S5-S20. [doi: 10.2337/diacare.26.2007.s5] [Medline: 12502614]
5. Gestational diabetes. International Diabetes Federation. URL: https://idf.org/about-diabetes/types-of-diabetes/gestational-diabetes/ [accessed 2025-11-23]
6. Cho NH, Shaw JE, Karuranga S, et al. IDF Diabetes Atlas: global estimates of diabetes prevalence for 2017 and projections for 2045. Diabetes Res Clin Pract 2018 Apr;138:271-281. [doi: 10.1016/j.diabres.2018.02.023] [Medline: 29496507]
7. HAPO Study Cooperative Research Group, Metzger BE, Lowe LP, et al. Hyperglycemia and adverse pregnancy outcomes. N Engl J Med 2008 May 8;358(19):1991-2002. [doi: 10.1056/NEJMoa0707943] [Medline: 18463375]

XSL•FO
RenderX

8. Najafi L, Abedini A, Kadivar M, et al. Gestational diabetes mellitus: the correlation between umbilical coiling index, and intrapartum as well as neonatal outcomes. J Diabetes Metab Disord 2019 Jun;18(1):51-57. [doi: 10.1007/s40200-019-00389-z] [Medline: 31275874]
9. Kramer CK, Campbell S, Retnakaran R. Gestational diabetes and the risk of cardiovascular disease in women: a systematic review and meta-analysis. Diabetologia 2019 Jun;62(6):905-914. [doi: 10.1007/s00125-019-4840-2] [Medline: 30843102]
10. Giannakou K, Evangelou E, Yiallouros P, et al. Risk factors for gestational diabetes: an umbrella review of meta-analyses of observational studies. PLoS One 2019;14(4):e0215372. [doi: 10.1371/journal.pone.0215372] [Medline: 31002708]
11. Le HM, Nguyen TA, Nguyen DK, et al. Prevalence and risk factors of gestational diabetes mellitus among pregnant women in northern Vietnam: a cross-sectional study. Glob Health Action 2025 Dec;18(1):2460339. [doi: 10.1080/16549716.2025.2460339] [Medline: 39925195]
12. Larrabure-Torrealva GT, Martinez S, Luque-Fernandez MA, et al. Prevalence and risk factors of gestational diabetes mellitus: findings from a universal screening feasibility program in Lima, Peru. BMC Pregnancy Childbirth 2018 Jul 18;18(1):303. [doi: 10.1186/s12884-018-1904-0] [Medline: 30021539]
13. Gao C, Sun X, Lu L, Liu F, Yuan J. Prevalence of gestational diabetes mellitus in mainland China: a systematic review and meta‐analysis. J Diabetes Investig 2019 Jan;10(1):154-162. [doi: 10.1111/jdi.12854] [Medline: 29683557]
14. Mantri N, Goel AD, Patel M, et al. National and regional prevalence of gestational diabetes mellitus in India: a systematic review and meta-analysis. BMC Public Health 2024 Feb 20;24(1):527. [doi: 10.1186/s12889-024-18024-9] [Medline: 38378536]
15. Egbe TO, Tsaku ES, Tchounzou R, Ngowe MN. Prevalence and risk factors of gestational diabetes mellitus in a population of pregnant women attending three health facilities in Limbe, Cameroon: a cross-sectional study. Pan Afr Med J 2018 Nov 20;31:195. [doi: 10.11604/pamj.2018.31.195.17177] [Medline: 31086639]
16. Mwanri AW, Kinabo J, Ramaiya K, Feskens EJM. Gestational diabetes mellitus in sub-Saharan Africa: systematic review and metaregression on prevalence and risk factors. Trop Med Int Health 2015 Aug;20(8):983-1002. [doi: 10.1111/tmi.12521] [Medline: 25877657]
17. Tran NK, Cao TN, Pham LV, et al. Development and multicenter validation of a novel model for selective screening of gestational diabetes mellitus: the Vietnam gestational diabetes mellitus study. BMC Pregnancy Childbirth 2025 Nov 7;25(1):1171. [doi: 10.1186/s12884-025-08249-w] [Medline: 41204205]
18. Nguyen CL, Lee AH, Minh Pham N, et al. Prevalence and pregnancy outcomes of gestational diabetes mellitus by different international diagnostic criteria: a prospective cohort study in Vietnam. J Matern Fetal Neonatal Med 2020 Nov;33(21):3706-3712. [doi: 10.1080/14767058.2019.1583733] [Medline: 30843751]
19. Duong DB, Minh HV, Ngo LH, Ellner AL. Readiness, availability and utilization of rural Vietnamese health facilities for community based primary care of non-communicable diseases: a cross-sectional survey of 3 provinces in northern Vietnam. Int J Health Policy Manag 2019 Mar 1;8(3):150-157. [doi: 10.15171/ijhpm.2018.104] [Medline: 30980631]
20. Haam JH, Kim BT, Kim EM, et al. Diagnosis of obesity: 2022 update of clinical practice guidelines for obesity by the Korean Society for the Study of Obesity. J Obes Metab Syndr 2023 Jun 30;32(2):121-129. [doi: 10.7570/jomes23031] [Medline: 37386771]
21. Mazumder T, Akter E, Rahman SM, Islam MT, Talukder MR. Prevalence and risk factors of gestational diabetes mellitus in Bangladesh: findings from Demographic Health Survey 2017-2018. Int J Environ Res Public Health 2022 Feb 23;19(5):2583. [doi: 10.3390/ijerph19052583] [Medline: 35270274]
22. Boadu WIO, Kugblenu P, Senu E, Opoku S, Anto EO. Prevalence and risk factors associated with gestational diabetes mellitus among pregnant women: a cross-sectional study in Ghana. Front Clin Diabetes Healthc 2022;3:854332. [doi: 10.3389/fcdhc.2022.854332] [Medline: 36992739]
23. Nguyen CL, Pham NM, Binns CW, Duong DV, Lee AH. Prevalence of gestational diabetes mellitus in Eastern and Southeastern Asia: a systematic review and meta-analysis. J Diabetes Res 2018;2018:6536974. [doi: 10.1155/2018/6536974] [Medline: 29675432]
24. Li G, Wei T, Ni W, et al. Incidence and risk factors of gestational diabetes mellitus: a prospective cohort study in Qingdao, China. Front Endocrinol (Lausanne) 2020 Sep 11;11:636. [doi: 10.3389/fendo.2020.00636] [Medline: 33042010]
25. Lin PC, Hung CH, Chan TF, Lin KC, Hsu YY. The risk factors for gestational diabetes mellitus: a retrospective study. Midwifery 2016 Nov;42:16-20. [doi: 10.1016/j.midw.2016.09.008] [Medline: 27705836]
26. Vaajala M, Liukkonen R, Ponkilainen V, Mattila VM, Kekki M, Kuitunen I. In vitro fertilization increases the odds of gestational diabetes: a nationwide register-based cohort study. Acta Diabetol 2023 Feb;60(2):319-321. [doi: 10.1007/s00592-022-01975-z] [Medline: 36271972]
27. Ashrafi M, Gosili R, Hosseini R, Arabipoor A, Ahmadi J, Chehrazi M. Risk of gestational diabetes mellitus in patients undergoing assisted reproductive techniques. Eur J Obstet Gynecol Reprod Biol 2014 May;176:149-152. [doi: 10.1016/j.ejogrb.2014.02.009] [Medline: 24630294]
28. Sachdev D, Sauer MV, Ananth CV. Gestational diabetes mellitus in pregnancies conceived after infertility treatment: a population-based study in the United States, 2015–2020. F S Rep 2024 Mar;5(1):102-110. [doi: 10.1016/j.xfre.2023.11.008] [Medline: 38524205]

XSL•FO
RenderX

29. Mohammadi M, Khedmati Morasae E, Maroufizadeh S, et al. Assisted reproductive technology and the risk of gestational diabetes mellitus: a systematic review and meta-analysis. Middle East Fertil Soc J 2020 Dec;25(1):6. [doi: 10.1186/s43043-020-0018-6]
30. Gerede A, Oikonomou E, Potiris A, et al. A systematic review on biomarkers for gestational diabetes mellitus detection in pregnancies conceived using assisted reproductive technology: current trends and future directions. Int J Mol Sci 2025 Aug 25;26(17):8234. [doi: 10.3390/ijms26178234] [Medline: 40943159]
31. Brănişteanu DD, Mathieu C. Progesterone in gestational diabetes mellitus: guilty or not guilty? Trends Endocrinol Metab 2003 Mar;14(2):54-56. [doi: 10.1016/S1043-2760(03)00003-1] [Medline: 12591170]
32. Pergialiotis V, Bellos I, Hatziagelaki E, Antsaklis A, Loutradis D, Daskalakis G. Progestogens for the prevention of preterm birth and risk of developing gestational diabetes mellitus: a meta-analysis. Am J Obstet Gynecol 2019 Nov;221(5):429-436.e5. [doi: 10.1016/j.ajog.2019.05.033] [Medline: 31132340]
33. Rhee SY, Kim JY, Woo JT, Kim YS, Kim SH. Familial clustering of type 2 diabetes in Korean women with gestational diabetes mellitus. Korean J Intern Med 2010 Sep;25(3):269-272. [doi: 10.3904/kjim.2010.25.3.269] [Medline: 20830223]
34. Wung SF, Lin PC. Shared genomics of type 2 and gestational diabetes mellitus. Annu Rev Nurs Res 2011;29:227-260. [doi: 10.1891/0739-6686.29.227] [Medline: 22891507]

Abbreviations

GDM: gestational diabetes mellitus
IADPSG: International Association of Diabetes and Pregnancy Study Groups
IVF: in vitro fertilization
PR: prevalence ratio
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 25.Jun.2025; peer-reviewed by H Ke, N Goumballa; revised version received 13.Jan.2026; accepted 13.Jan.2026; published 31.Mar.2026.

Please cite as:
Le DC, Nguyen TH, Nguyen TL, Tran TN, Luong XH, Do TTH, Tran XT, Ngo VM, Bui THD, Nguyen TV, Hoang VT, Vu TB
Gestational Diabetes Mellitus and Associated Factors in Rural Areas of Northern Vietnam: Cross-Sectional Survey
JMIR Public Health Surveill 2026;12:e79688
URL: https://publichealth.jmir.org/2026/1/e79688
doi:10.2196/79688

© Duc Cuong Le, Thi Hien Nguyen, Thi Ly Nguyen, Thi Nuong Tran, Xuan Hien Luong, Thi Thu Huong Do, Xuan Thuy Tran, Van Manh Ngo, Thi Huyen Dieu Bui, Tien Van Nguyen, Van Thuan Hoang, Thanh Binh Vu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 31.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Factors Affecting Acceptance of Life Education in Mainland China: National Cross-Sectional Study

Xinyue Zhang[1*], MMed; Nana Huang[2*], BSc; Yuyang Yi[3*], MSc; Yinlin Wang[4], MMed; Wai-Kit Ming[3], MD, PhD; Yibo Wu[5], PhD; Xiaohong Ning[6], PhD; Chen Wang[1,7], MD, PhD

[1]School of Health Policy and Management, Chinese Academy of Medical Sciences & Peking Union Medical College, No. 9, Dongdansantiao Street, Dongcheng District, Beijing, China

[2]Xiangya School of Nursing, Central South University, Changsha, China

[3]Department of Infectious Diseases and Public Health, Jockey Club College of Veterinary Medicine and Life Sciences, City University of Hong Kong, Hong Kong, China (Hong Kong)

[4]Key Laboratory of Oral Biomedical Research of Zhejiang Province, Affiliated Stomatology Hospital, Zhejiang University, Hangzhou, China

[5]Department of Nursing, The Fourth Affiliated Hospital of School of Medicine, and International School of Medicine, International Institutes of Medicine, Zhejiang University, Yiwu, China

[6]Palliative Care Medicine Center, Peking Union Medical College Hospital, Beijing, China

[7]School of Population Medicine and Public Health, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Chen Wang, MD, PhD
School of Health Policy and Management, Chinese Academy of Medical Sciences & Peking Union Medical College, No. 9, Dongdansantiao Street, Dongcheng District, Beijing, China

Abstract

Background: In traditional Chinese culture, discussing death has always been taboo. The social environment characterized by fear, reluctance, and apprehension toward death significantly impedes the development of palliative care. Therefore, establishing a correct view of life and death and promoting life education are prerequisites for the successful implementation of palliative care.

Objective: This study aimed to investigate the public acceptance of life education among individuals in China and analyze the explanatory variables.

Methods: This national cross-sectional study was conducted from June 20 to August 31, 2022, encompassing 23 provinces, 5 autonomous regions, and 4 municipalities directly under the central government. A total of 21,875 participants were included. The generalized linear model was used to analyze influencing factors. Gender, major in medicine, place of residence, education level, family health, media use, etc, were analyzed as the potential variables. Acceptance scores were calculated based on a visual analog scale (VAS). Further subgroup analyses were carried out in different age and major subgroups.

Results: The median (P_{25}, P_{75}) acceptance score for life education in the survey was 71.00 (50.00-95.00) points. Compared with females, males had lower acceptance (β=−2.39; 95% CI −3.08 to −1.69). Respondents who were majoring in medicine (β=3.13; 95% CI 1.11-5.14), residing in urban areas (β=1.25; 95% CI 0.46-2.04), processing a bachelor's degree or higher (β=4.05; 95% CI 2.97-5.12), or having higher scores on the media use (β=0.49; 95% CI 0.41-0.57) had higher acceptance. Compared with people aged 12-17 years, those aged 25-44 years (β=−6.00; 95% CI −7.34 to −4.66), aged 45-64 years (β=−4.55; 95% CI −5.88 to −3.22), and 65 years or older (β=−2.16; 95% CI −3.78 to −0.55) had lower acceptance. For people majoring in medicine, place of residence, family health, and media use were uniquely relevant factors. Higher scores on the Perceived Social Support Scale (PSSS) and Health Literacy Scale-Short Form (HLS-SF) were also significantly associated with greater acceptance of life education.

Conclusions: Gender, place of residence, education level, age, media use, perceived social support, and health literacy were identified as key factors influencing acceptance of life education, providing important evidence to inform targeted policy and educational interventions.

(JMIR Public Health Surveill 2026;12:e78844) doi:10.2196/78844

KEYWORDS

life education; personal intention; public education; affecting factors; cross-sectional study

XSL•FO
RenderX

Introduction

Life education is an educational process that examines the relationship between life and death and encompasses cultural and religious beliefs, as well as attitudes toward death and dying. Its purpose is to cultivate individuals' appreciation for life and encourage individuals to reflect this orientation in everyday life [1]. In addition to an understanding of the process of death, life education also involves the study of attitudes toward death and its significance. In recent years, as life education has become more comprehensive, an increasing number of studies have demonstrated that life education reported beneficial associations in the overall health of individuals throughout the lifespan. In childhood and adolescence, it helps children and teenagers to understand themselves, their emotions, and behaviors, fostering self-esteem and confidence. Furthermore, it provides instruction on the establishment and maintenance of healthy interpersonal relationships, including those with family members, friends, and peers. It fosters respect for life, empathy, and a sense of moral responsibility [2]. In adulthood, individuals are assisted in understanding their interests, abilities, and career choices to facilitate the formulation of suitable career plans. They are furnished with the requisite knowledge and skills about the establishment of a family, reproduction, and child-rearing. Education on maintaining a healthy lifestyle, including diet, exercise, and mental well-being, is provided. In middle age, education is directed toward the assumption of familial responsibilities, including the care of children and older parents. It is recommended that individuals pursue further career development and advancement. Individuals continue to engage in personal growth and development, acquiring new skills and interests [3]. In the geriatric stage, the provision of factual knowledge about life and death enables older adults to confront death realistically, thereby reducing avoidance and encouraging autonomous end-of-life care planning. Furthermore, it enhances awareness of life quality and facilitates the formation of personal views on life and death based on individual experiences and cultural background. This enriches their later life and enhances their sense of well-being [4].

Life education is of significant importance in various aspects of human life, including individual mental health, social-cultural cognition, and health care service quality. Research conducted by Ronconi et al [5] indicates that life education suggests that life education can facilitate open discussions about life, death, and mental health. By participating in life education, individuals can enhance their comprehension of death and cultivate effective coping strategies, which may ultimately reduce fear and anxiety associated with mortality. Concurrently, Lei et al [6] discovered in their research on the educational needs of the older adults regarding death that older adults desire to express their perspectives on the passing of life through life education and mutual exchange to foster gratitude for life and acceptance of the inevitability of death. In health care settings, Lv et al [7] indicate that life education may support the preparation of professionals to care for dying patients by strengthening the competence of medical and nursing staff and promoting more respectful and compassionate attitudes toward dying and bereavement. This is in accordance with the findings of Zhang et al [8], which indicated that life education had a positive impact on the mental health of individual nurses and enhanced their ability to cope with death-related issues. Internationally, life education has been more extensively developed in several countries outside China [9]. From an early age, children are taught about the meaning, value, and dignity of life through school curricula and community activities, which helps them to develop an appropriate understanding of life and death issues [10]. In the United Kingdom, life education plays a pivotal role in health care and social work education, with institutions offering relevant courses and professional organizations such as the Life Education Association and the End-of-Life Care Association providing training and resources. In the United States, the exploration of individual views on death and societal attitudes toward life is a significant focus of university death research institutes or centers. In addition, nonprofit organizations play a role in raising public awareness of death, loss, and palliative care. In Germany, life education is primarily focused on professional training and services in the medical and psychological counseling fields, as well as the provision of life education for children and teenagers. Schools also offer courses on life and death. In Japan, life education is undergoing rapid development in schools and communities. Schools are offering life education courses, communities are promoting understanding and acceptance of death through lectures and activities, and dedicated life education institutions and projects are emerging. Taken together, these practices illustrate how life education can be implemented across the life course, supporting a more holistic understanding of life and death and potentially contributing to quality of life and social cohesion.

However, in part because of the influence of traditional culture, in China, death is frequently treated as a socially delicate topic and is often avoided in everyday conversation, a tendency shaped in part by long-standing cultural traditions. Confucian and Daoist moral frameworks—emphasizing filial obligations, ancestral continuity, and acceptance of the natural order—may contribute to restrained engagement with end-of-life issues and with services that explicitly address dying and bereavement, thereby posing barriers to the broader development of life education. At the same time, demographic aging is intensifying and reconfiguring care demands; the continued growth of the older population, alongside increasing burdens of chronic illness and functional limitations, is placing sustained pressure on health care and long-term care systems and is likely to amplify the need for support related to end-of-life understanding, communication, and planning. Within this policy environment, national initiatives such as the "Silver Economy" agenda—which prioritizes the health and well-being of older adults—create opportunities to strengthen life education as part of a broader strategy to improve quality of life across the life course. Despite these developments, life education in China has often been informed by models and practices developed in other settings, and empirical evidence remains limited regarding its cultural fit, public acceptability, and the conditions under which participation is more likely, especially among older adults. To address these gaps, the present study will examine public acceptance of life education and identify factors associated with receptivity and willingness to engage, with particular attention to older adults' needs and issues of practical access. By

XSL•FO
RenderX

clarifying who is most likely to accept life education and why, the findings are expected to inform the design and delivery of culturally responsive programs and to provide evidence relevant to policy and institutional decision-making on planning, implementation, and refinement; more broadly, the study aims to generate empirically grounded insights into feasible pathways for advancing life education in China in ways that support well-being and contribute to social cohesion.

Methods

Research Instrument

This cross-sectional study aims to describe the acceptance of life education by Chinese residents in an intuitive quantitative way and further reveal the relevant factors affecting their acceptance. The study uses a visual analog scale (VAS) scoring system, ranging from 0 to 100, to quantify participants' level of acceptance of life education. Higher scores indicate a stronger willingness to accept. The fundamental objective of life education is to enhance individuals' comprehension of life and death, instill respect for life, and encourage humane care. This form of education helps individuals to confront the inevitability of death, to overcome their fears of death, and to develop a scientific perspective on life and death. By exploring topics such as death, life education encourages individuals to reflect on the meaning of life, thereby enhancing their composure and rationality when confronted with life challenges. Furthermore, life education offers psychological support and assists individuals in establishing appropriate values to cope with various challenges encountered during the course of life with a more positive outlook (Table S1 in Multimedia Appendix 1).

The questionnaire design of this study is divided into 2 main parts. The initial section of the questionnaire is designed to collect information about the participants' sociodemographic characteristics. This includes data on gender, age, place of residence, marital status, sibling status, income level, medical background, and the presence of chronic diseases. The second part uses a range of scale tools, including the Perceived Social Support Scale (PSSS), Family Health Scale-Short Form (FHS-SF), Health Literacy Scale-Short Form (HLS-SF), Quality of Life Scale (EQ-5D-5L), and Media Exposure Scale, to comprehensively assess the impact of participants' social support, family health status, health literacy, quality of life, and media exposure on their acceptance and understanding of life education.

The Perceived Social Support Scale–Short Form (PSSS-SF) is a psychometric tool designed to assess perceived social support from different sources. We assessed perceived social support using the 3-item PSSS, which covers support from family, friends, and others. Each item is rated on a 7-point Likert scale (1-7), and item scores are summed to yield a total score ranging from 3 to 21, with higher scores indicating greater perceived social support. In our sample, the Cronbach α was 0.923 [11].

The FHS-SF is a multidimensional instrument designed to assess family health. A new International FHS was introduced, covering 4 dimensions, including family socioemotional health processes, family health and social style, family health resources, and external social support for the family. The scale contains a total of ten questions, which are scored using a 5-point Likert scoring method. The scale has scores ranging from 10 to 50, with higher scores indicating better family health functioning. Cronbach α coefficient was found to be 0.84 [12].

The Health Literacy Scale–Short Form (HLS-SF) is designed to assess an individual's ability to access, process, and understand basic health information and services. We used the Chinese 12-item HLS-SF, which comprises 3 dimensions—health care, disease prevention, and health promotion—with 4 items in each dimension. Each item is rated on a 4-point Likert scale. Following the standard scoring procedure, responses are transformed to an index score ranging from 0 to 50, with higher scores indicating higher health literacy. Specifically, the mean item score was converted to a 0 - 50 index using the transformation recommended in the original scale documentation. The reported Cronbach α for the Chinese version is 0.94 [13].

The EQ-5D-5L was used for this study. The EQ-5D-5L, a classic instrument comprising 5 dimensions (mobility, self-care, usual activities, pain and discomfort, and anxiety and depression), was used in this study, with a total of 5 items. The scale is scored using a Likert 5-point rating method. The index score ranges from 0 to 1, with 0 representing a person who has died and 1 representing a completely healthy person. The Cronbach α coefficient was found to be 0.857 [14].

The Media Exposure Scale is a tool designed to assess the extent to which individuals are exposed to various forms of media. The classic Media Exposure Scale, comprising 5 dimensions (seeking information, providing information, self-seeking, social interaction, and relaxation and entertainment), was used in the study. The scale comprises a total of 5 items, which are scored using a Likert 5-point rating method. The scale has scores ranging from 5 to 25, with higher scores indicating higher media exposure. The Cronbach α coefficient is 0.85 [15].

The World Health Organization-Five Well-Being Index (WHO-5) consists of a total of 5 items, covering 3 dimensions, including emotional experience, vitality level, and functional performance. It uses a 6-point Likert scoring method, with a total score ranging from 0 to 25, and the percentage score is calculated by multiplying the total score by 4. In the Chinese sample, the Cronbach α coefficient is 0.83, and both the structural validity and criterion validity are greater than 0.7 [16].

Data Collection

The study was conducted between June 20 and August 31, 2022, encompassing 23 provinces, 5 autonomous regions, and 4 municipalities directly under the central government [17-20]. A stratified sampling technique was used to randomly select 2-6 cities in each province and autonomous region, as well as 2-6 communities in each prefecture-level city. In conclusion, a total of 148 cities were selected as survey targets. The selection was based on the population distribution characteristics obtained from the 7th National Population Census in 2021, which included data on age, gender, and urban-rural distribution [21]. The minimum required sample size was calculated using the following formula for a cross-sectional study design:

XSL•FO
RenderX

$$N=Z\alpha 2\times p\times(1-p)\partial 2$$

where the prevalence refers to the results of the meta-analysis by Krishnamoorthy et al [22]. The given formula establishes the probability of Type I error (α) as .05, while the acceptable error is determined to be .02. Consequently, the minimum sample size is computed as 1752 for each province, and the total sample size is 40,000. A quota sampling method was used to select community residents for the survey. In each selected city, surveyors or survey teams were openly recruited. All surveyors were required to undergo professional training, education, and a rigorous standardized training and process evaluation. The use of high-level, standardized technical terms was maintained throughout. The surveyors obtained informed consent from the interviewees before administering the questionnaire. The questionnaire number was entered into the Wen Juan Xing platform before the questionnaire was made available to the public. Trained surveyors conducted face-to-face interviews in sampled communities and entered participants' responses into the Wen Juan Xing platform in real time (interviewer-administered data collection). Before each interview, surveyors obtained written informed consent. When technical terms appeared, surveyors provided neutral explanations to ensure comprehension, and no leading language was used. If a participant could respond verbally but could not complete the questionnaire due to physical limitations, the surveyor recorded the answers verbatim. Participants could withdraw at any time, and all data were collected anonymously.

The specific quota sampling standards are as follows: (1) age criteria: in the sample calibration stage, we applied age quota control. Based on the questionnaire's age categories, the sample was stratified into 8 age groups, including 12 - 18 years, 19 - 25 years, 26 - 30 years, 31 - 40 years, 41 - 50 years, 51 - 60 years, 61 - 70 years, and ≥71 years. Quota targets were defined as the age distribution of the final analytic sample (n=21,916), expressed as percentages: 14.70% (12 - 18 y), 15.97% (19 - 25 y), 9.75% (26 - 30 y), 13.34% (31 - 40 y), 16.75% (41 - 50 y), 11.68% (51 - 60 y), 10.54% (61 - 70 y), and 7.27% (≥71 y). When the number of respondents in a given age group exceeded the target, we performed random sampling without replacement within that group to reduce the sample size to the target; when it was below or equal to the target, all respondents were retained. (2) Gender criteria: the ratio of males to females is approximately equal. (3) Urban-rural criteria: the ratio of urban to rural samples is approximately 3:2.

The study included individuals who met the following criteria: (1) participants must be at least 12 years of age. (2) They must hold Chinese nationality. (3) They must be Chinese residents who spend no more than one month abroad per year. (4) They must be willing to participate in the survey and complete the informed consent form. (5) They must be able to complete the online questionnaire independently or with the assistance of surveyors. (6) They must be able to understand the meaning of each questionnaire item. The rationale for selecting adolescents aged 12 years and older as research participants is as follows. First, it was found that as a result of improvements in educational quality and social environment, adolescents and adults have acquired knowledge about life education. Furthermore, the majority of adolescents aged 12 years and older possess the requisite conditions and environment to cultivate their distinctive perspectives on how to interact with deceased individuals. Second, it is reasonable to posit that these adolescents are sufficiently capable of understanding the questionnaire and responding through smartphones (if applicable). Based on these considerations, it is generally assumed that participants will engage with the survey objectively. Individuals with mental illnesses or other diseases, those participating in similar research projects, and those unwilling to cooperate are excluded from participation in the study.

Statistical Methods

The data were subjected to descriptive statistical analysis using the Windows version of SPSS 26.0 (IBM Corp). Subsequently, a generalized linear model approach was used to examine the correlation between personal information (gender, region, age, and education level) and family status (marital status, children status, and siblings) with the VAS for acceptance of life education (range: 0 - 100 points, higher scores indicating higher levels of acceptance; Table 1 and Figure S1 and Table S1 in Multimedia Appendix 1). Furthermore, subgroup analyses were conducted by medical background and by age group (Tables 2-7). In addition, sensitivity analyses were conducted for selected measurement instruments to evaluate the robustness of results to alternative scoring and operationalization of key scales.

XSL•FO
RenderX

Table . Multivariable generalized linear model of factors associated with life education acceptance (visual analog scale [VAS] 0 - 100), Mainland China, June-August 2022.

Variate	β	SE	95% CI	*P* value
Gender (reference: female)				
Male	–2.39	0.35	–3.08 to –1.69	<.001
Religion (reference: no)				
Yes	–0.86	0.91	–2.65 to 0.93	.35
Medicine specialty (reference: no)				
Yes	3.13	1.03	1.11 to 5.14	.002
Place of residence (reference: rural)				
Urban	1.25	0.40	0.46 to 2.04	.002
Have debt (reference: no)				
Yes	1.54	0.38	0.80 to 2.28	<.001
Highest educational level (reference: Junior high or below)				
Senior high or specialty	0.33	0.48	–0.62 to 1.27	.50
Undergraduate or above	4.05	0.55	2.97 to 5.12	<.001
Age (years), reference: (12 - 17)				
18 - 24	0.35	0.75	–1.13 to 1.83	.64
25 - 44	–6.00	0.69	–7.34 to –4.66	<.001
45 - 64	–4.55	0.68	–5.88 to –3.22	<.001
≥65	–2.16	0.82	–3.78 to –0.55	.009
Region (reference: Eastern China)				
Central China	–1.98	0.43	–2.83 to –1.14	<.001
Western China	–2.57	0.43	–3.41 to –1.72	<.001
Chronic disease diagnosis (reference: no)				
Yes	1.37	0.45	0.48 to 2.25	.002
PSSS[a]	0.49	0.06	0.38 to 0.61	<.001
FHS-SF[b]	0.46	0.03	0.39 to 0.52	<.001
HLS-SF[c]	0.42	0.04	0.34 to 0.50	<.001
EQ-5D-5L[d]	0.20	0.09	0.02 to 0.37	.03
Media use	0.49	0.04	0.41 to 0.57	<.001

[a]PSSS: Perceived Social Support Scale.

[b]FHS-SF: Family Health Scale-Short Form.

[c]HLS-SF: Health Literacy Scale-Short Form.

[d]EQ-5D-5L: Quality of Life Scale.

XSL•FO
RenderX

Table . Subgroup models by medical background (medical vs nonmedical) for life education acceptance (visual analog scale [VAS] 0 - 100), Mainland China, June−August 2022.

Variate	β	SE	95% CI	*P* value
Medicine specialty				
Gender (reference: female)				
Male	–0.82	1.79	–4.32 to 2.68	.65
Religion (reference: no)				
Yes	–16.73	11.67	–39.60 to 6.14	.15
Place of residence (reference: rural)				
Urban	5.36	1.96	1.51 to 9.21	.006
Have debt (reference: no)				
Yes	0.71	1.73	–2.68 to 4.11	.68
Highest educational level (reference: junior high or below)				
Senior high or specialty	25.89	23.68	–20.51 to 72.29	.27
Undergraduate or above	26.59	23.70	–19.86 to 73.03	.26
Age (years), reference: 12 - 17				
18 - 24	5.22	5.18	–4.93 to 15.37	.31
25 - 44	2.12	5.92	–9.49 to 13.71	.72
45 - 64	3.28	8.18	–12.76 to 19.31	.69
≥65	18.33	23.67	–28.07 to 64.73	.44
Region (reference: Eastern China)				
Central China	–2.30	2.19	–6.59 to 1.98	.29
Western China	–2.18	2.06	–6.23 to 1.86	.29
Chronic disease diagnosis (reference: no)				
Yes	1.62	2.82	–3.90 to 7.13	.57
PSSS[a]	0.48	0.30	–0.11 to 1.07	.11
FHS−SF[b]	0.36	0.16	0.05 to 0.67	.02
HLS−SF[c]	0.26	0.22	–0.17 to 0.69	.24
EQ−5D-5L[d]	0.51	0.45	–0.38 to 1.40	.26
Media use	1.28	0.24	0.81 to 1.75	<.001
Nonmedicine specialty				
Gender (reference: female)				
Male	–2.45	0.36	–3.16 to –1.75	<.001
Religion (reference: no)				
Yes	–0.77	0.92	–2.57 to 1.03	.40

XSL•FO
RenderX

Variate	β	SE	95% CI	*P* value
Place of residence (reference: rural)				
Urban	1.10	0.41	0.29 to 1.91	.008
Have debt (reference: no)				
Yes	1.60	0.39	0.84 to 2.35	<.001
Highest educational level (reference: junior high or below)				
Senior high or specialty	0.35	0.49	–0.61 to 1.30	.48
Undergraduate or above	4.21	0.56	3.12 to 5.30	<.001
Age (years) (reference: 12 - 17)				
18 - 24	0.28	0.77	–1.22 to 1.78	.71
25 - 44	–6.13	0.69	–7.49 to –4.77	<.001
45 - 64	–4.65	0.69	–6.00 to –3.31	<.001
≥65	–2.31	0.83	–3.93 to –0.68	.006
Region (reference: Eastern China)				
Central China	–1.94	0.44	–2.80 to –1.09	<.001
Western China	–2.53	0.44	–3.40 to –1.67	<.001
Chronic disease diagnosis (reference: no)				
Yes	1.36	0.46	0.46 to 2.25	.003
PSSS	0.49	0.06	0.37 to 0.61	<.001
FHS−SF	0.46	0.03	0.39 to 0.53	<.001
HLS−SF	0.42	0.04	0.34 to 0.51	<.001
EQ−5D-5L	0.18	0.09	0.00 to 0.36	.05
Media use	0.47	0.04	0.38 to 0.55	<.001

[a]PSSS: Perceived Social Support Scale.

[b]FHS-SF: Family Health Scale-Short Form.

[c]HLS-SF: Health Literacy Scale-Short Form.

[d]EQ-5D-5L: Quality of Life Scale.

XSL•FO
RenderX

Table . Subgroup model for life education acceptance (visual analog scale [VAS] 0-100) among participants aged 12-17 years, Mainland China, June-August 2022

Variable	β	SE	95% CI	*P* value
Gender (reference: female)				
Male	–0.39	1.16	–2.66 to 1.88	.74
Religion (reference: no)				
Yes	4.45	3.15	–1.73 to 10.63	.16
Medicine specialty (reference: no)				
Yes	0.19	5.80	–11.17 to 11.55	.97
Place of residence (reference: Rural)				
Urban	4.08	1.24	1.66 to 6.51	.001
Have debt (reference: No)				
Yes	3.95	1.24	1.52 to 6.39	.001
Highest educational level (reference: junior high or below)				
Senior high or specialty	2.19	1.21	–0.18 to 4.56	.07
Undergraduate or above	–0.36	3.00	–6.24 to 5.51	.90
Region (reference: Eastern China)				
Central China	–3.11	1.40	–5.86 to –0.37	.03
Western China	–4.77	1.40	–7.51 to –2.04	.001
Chronic disease diagnosis (reference: No)				
Yes	–0.17	2.36	–4.79 to 4.45	.94
PSSS[a]	0.26	0.18	–0.09 to 0.60	.15
FHS-SF[b]	0.40	0.11	0.18 to 0.61	<.001
HLS-SF[c]	0.46	0.13	0.21 to 0.71	<.001
EQ-5D-5L[d]	0.13	0.27	−0.40 to 0.66	.63
Media use	0.55	0.13	0.30 to 0.80	<.001

[a]PSSS: Perceived Social Support Scale.

[b]FHS-SF: Family Health Scale-Short Form.

[c]HLS-SF: Health Literacy Scale-Short Form.

[d]EQ-5D-5L: Quality of Life Scale.

XSL•FO
RenderX

Table . Subgroup model for life education acceptance (visual analog scale [VAS] 0-100) among participants aged 18-24 years.

Variable	β	SE	95% CI	*P* value
Gender (reference: female)				
Male	–4.38	0.77	–5.88 to –2.87	<.001
Religion (reference: no)				
Yes	–2.90	3.39	–9.54 to 3.74	.39
Medicine specialty (reference: no)				
Yes	2.88	1.08	0.77 to 5.00	.008
Place of residence (reference: rural)				
Urban	4.99	0.86	3.30 to 6.67	<.001
Have debt (reference: no)				
Yes	1.93	0.77	0.43 to 3.44	.01
Highest educational level (reference: junior high or below)				
Senior high or specialty	8.20	3.08	2.16 to 14.24	.008
Undergraduate or above	8.84	3.07	2.82 to 14.85	.004
Region (reference: Eastern China)				
Central China	–3.45	0.92	–5.25 to –1.65	<.001
Western China	–3.97	0.92	–5.77 to –2.18	<.001
Chronic disease diagnosis (reference: no)				
Yes	3.83	1.21	1.45 to 6.21	.002
PSSS[a]	0.22	0.13	–0.03 to 0.48	.09
FHS-SF[b]	0.59	0.07	0.45 to 0.73	<.001
HLS-SF[c]	0.31	0.09	0.12 to 0.49	.001
EQ-5D-5L[d]	0.76	0.22	0.33 to 1.19	.001
Media use	0.78	0.10	0.58 to 0.98	<.001

[a]PSSS: Perceived Social Support Scale.

[b]FHS-SF: Family Health Scale-Short Form.

[c]HLS-SF: Health Literacy Scale-Short Form:

[d]EQ-5D-5L: Quality of Life Scale.

XSL•FO
RenderX

Table . Subgroup model for life education acceptance (visual analog scale [VAS] 0-100) among participants aged 25-44 years.

Variable	β	SE	95% CI	*P* value
Gender (reference: female)				
Male	–2.46	0.66	–3.75 to –1.17	<.001
Religion (reference: no)				
Yes	3.04	1.85	–0.59 to 6.67	.10
Medicine specialty (reference: no)				
Yes	5.01	3.42	–1.70 to 11.72	.14
Place of residence (reference: rural)				
Urban	0.06	0.81	–1.53 to 1.65	.94
Have debt (reference: no)				
Yes	1.32	0.67	0.01 to 2.62	.05
Highest educational level (reference: junior high or below)				
Senior high or specialty	0.80	0.99	–1.14 to 2.74	.42
Undergraduate or above	7.09	0.97	5.19 to 8.99	<.001
Region (reference: Eastern China)				
Central China	1.27	0.80	–0.29 to 2.83	.11
Western China	–0.80	0.82	–2.41 to 0.81	.33
Chronic disease diagnosis (reference: no)				
Yes	3.69	0.91	1.90 to 5.47	<.001
PSSS[a]	0.36	0.11	0.14 to 0.58	.001
FHS-SF[b]	0.45	0.06	0.33 to 0.57	<.001
HLS-SF[c]	0.46	0.08	0.31 to 0.61	<.001
EQ-5D-5L[d]	0.17	0.17	–0.16 to 0.51	.32
Media use	0.52	0.09	0.35 to 0.69	<.001

[a]PSSS: Perceived Social Support Scale.

[b]FHS-SF: Family Health Scale-Short Form.

[c]HLS-SF: Health Literacy Scale-Short Form.

[d]EQ-5D-5L: Quality of Life Scale.

XSL•FO
RenderX

Table . Subgroup model for life education acceptance (visual analog scale [VAS] 0-100) among participants aged 45-64 years.

Variable	β	SE	95% CI	*P* value
Gender (reference: female)				
Male	–2.31	0.68	–3.64 to –0.99	.001
Religion (reference: no)				
Yes	–3.32	1.57	–6.40 to –0.25	.03
Medicine specialty (reference: no)				
Yes	5.73	7.21	–8.41 to 19.86	.43
Place of residence (reference: rural)				
Urban	–1.32	0.79	–2.86 to 0.22	.09
Have debt (reference: no)				
Yes	0.37	0.74	–1.08 to 1.82	.62
Highest educational level (reference: junior high or below)				
Senior high or specialty	0.12	0.82	–1.49 to 1.72	.89
Undergraduate or above	4.50	1.00	2.54 to 6.46	<.001
Region (reference: Eastern China)				
Central China	–1.78	0.83	–3.40 to –0.16	.03
Western China	–1.08	0.82	–2.68 to 0.53	.19
Chronic disease diagnosis (reference: No)				
Yes	0.86	0.72	–0.55 to 2.26	.23
PSSS[a]	0.79	0.11	0.57 to 1.01	<.001
FHS-SF[b]	0.49	0.06	0.36 to 0.62	<.001
HLS-SF[c]	0.39	0.08	0.23 to 0.55	<.001
EQ-5D-5L[d]	–0.03	0.19	–0.40 to 0.35	.89
Media use	0.22	0.08	0.06 to 0.37	.006

[a]PSSS: Perceived Social Support Scale.

[b]FHS-SF: Family Health Scale-Short Form.

[c]HLS-SF: Health Literacy Scale-Short Form.

[d]EQ-5D-5L: Quality of Life Scale.

XSL•FO
RenderX

Table . Subgroup model for life education acceptance (visual analog scale [VAS] 0-100) among participants aged ≥65 years.

Variable	β	SE	95% CI	*P* value
Gender (reference: female)				
Male	–0.95	0.97	–2.86 to 0.96	.33
Religion (reference: no)				
Yes	–2.76	1.77	–6.23 to 0.71	.12
Medicine specialty (reference: no)				
Yes	12.42	25.37	–37.31 to 62.15	.62
Place of residence (reference: rural)				
Urban	0.84	1.03	–1.19 to 2.86	.42
Have debt (reference: no)				
Yes	3.03	1.37	0.34 to 5.72	.03
Highest educational level (reference: junior high or below)				
Senior high or specialty	–1.06	1.26	–3.53 to 1.41	.40
Undergraduate or above	1.36	2.04	–2.63 to 5.36	.50
Region (reference: Eastern China)				
Central China	–5.95	1.21	–8.31 to –3.58	<.001
Western China	–4.27	1.16	–6.55 to –1.99	<.001
Chronic disease diagnosis (reference: no)				
Yes	–1.44	1.05	–3.49 to 0.61	.17
PSSS[a]	0.94	0.18	0.59 to 1.29	<.001
FHS-SF[b]	0.17	0.10	–0.03 to 0.36	.10
HLS-SF[c]	0.53	0.11	0.31 to 0.75	<.001
EQ-5D-5L[d]	0.11	0.19	–0.27 to 0.49	.57
Media use	0.44	0.10	0.24 to 0.63	<.001

[a]PSSS: Perceived Social Support Scale.

[b]FHS-SF: Family Health Scale-Short Form.

[c]HLS-SF: Health Literacy Scale-Short Form.

[d]EQ-5D-5L: Quality of Life Scale.

Ethical Considerations

This study was approved by the Ethics Review Committee of Shaanxi Health Culture Research Center (Approval No.: JKWH-2022‐02). All participants signed an electronic informed consent form before participating, which clearly stated the research objectives, procedures, and the right to withdraw without consequences and were informed again by dedicated staff before starting the questionnaire. For respondents aged 12‐18 years, their guardians signed informed consent forms and completed the questionnaire survey with the accompaniment of their guardians. To ensure the privacy and confidentiality of participants, all collected data are fully anonymized. The research dataset does not retain any identifiable personal information, and access is limited to authorized researchers only. The participants did not receive any economic compensation.

Results

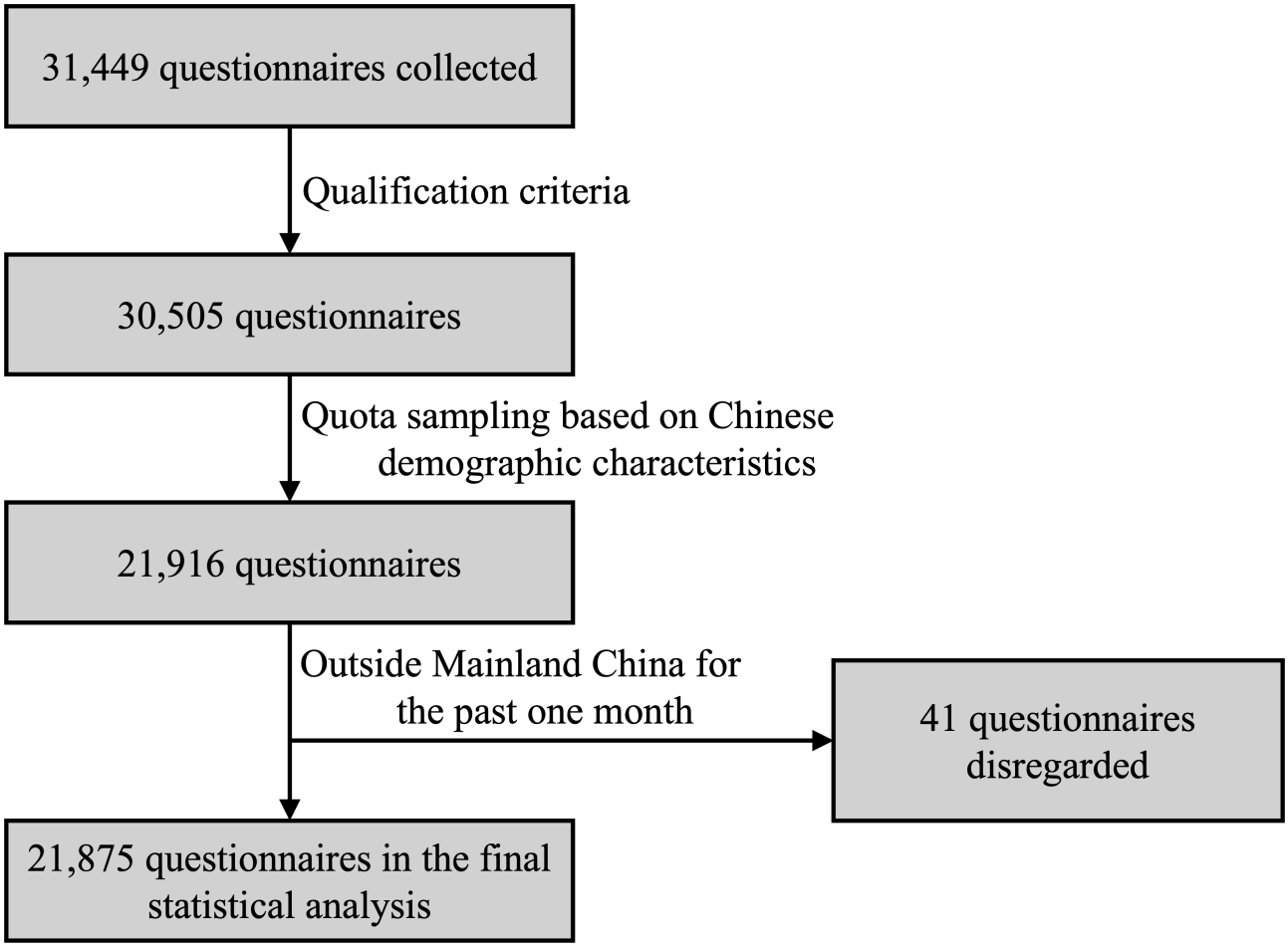
Sociological Characteristics of the Population

A total of 21,875 participants were analyzed in this study (Figure 1). Among the 21,875 respondents, 10,942 (50.2%) were female, 19,934 (91.13%) were ethnic Han, 21,005 (96.02%) were nonreligious believers, 15,155 (69.28%) resided in urban areas, and 21,137 (96.63%) were nonmedical majors. In terms of educational level, 6960 out of 21,875 (31.82%) had junior high or below, 7687 out of 21,875 (35.14%) had senior high or

XSL•FO
RenderX

specialty education, and 7228 out of 21,875 (33.04%) had undergraduate or above education. Out of 21,875, a total of 8083 (36.95%) participants were from eastern China, 6905 (31.57%) participants were from central China, and 6887 (31.48%) were from western China. A total of 5652 out of 21,875 (25.84%) participants had a diagnosis of a chronic disease, and 8138 (37.20%) participants had a debt. Out of 21,875, 2170 (9.92%) participants were aged 12 - 17 years, 4382 (20.03%) participants were aged 18 - 24 years, 6604 (30.19%) were aged 25 - 44 years, 5934 (27.13%) were aged 45 - 64 years, and 2785 (12.73%) were aged 65 years or older (Table S1 in Multimedia Appendix 1).

Figure 1. Participant low for a national cross-sectional survey of community residents in Mainland China, June-August 2022.



Acceptance of Life Education and Associated Factors

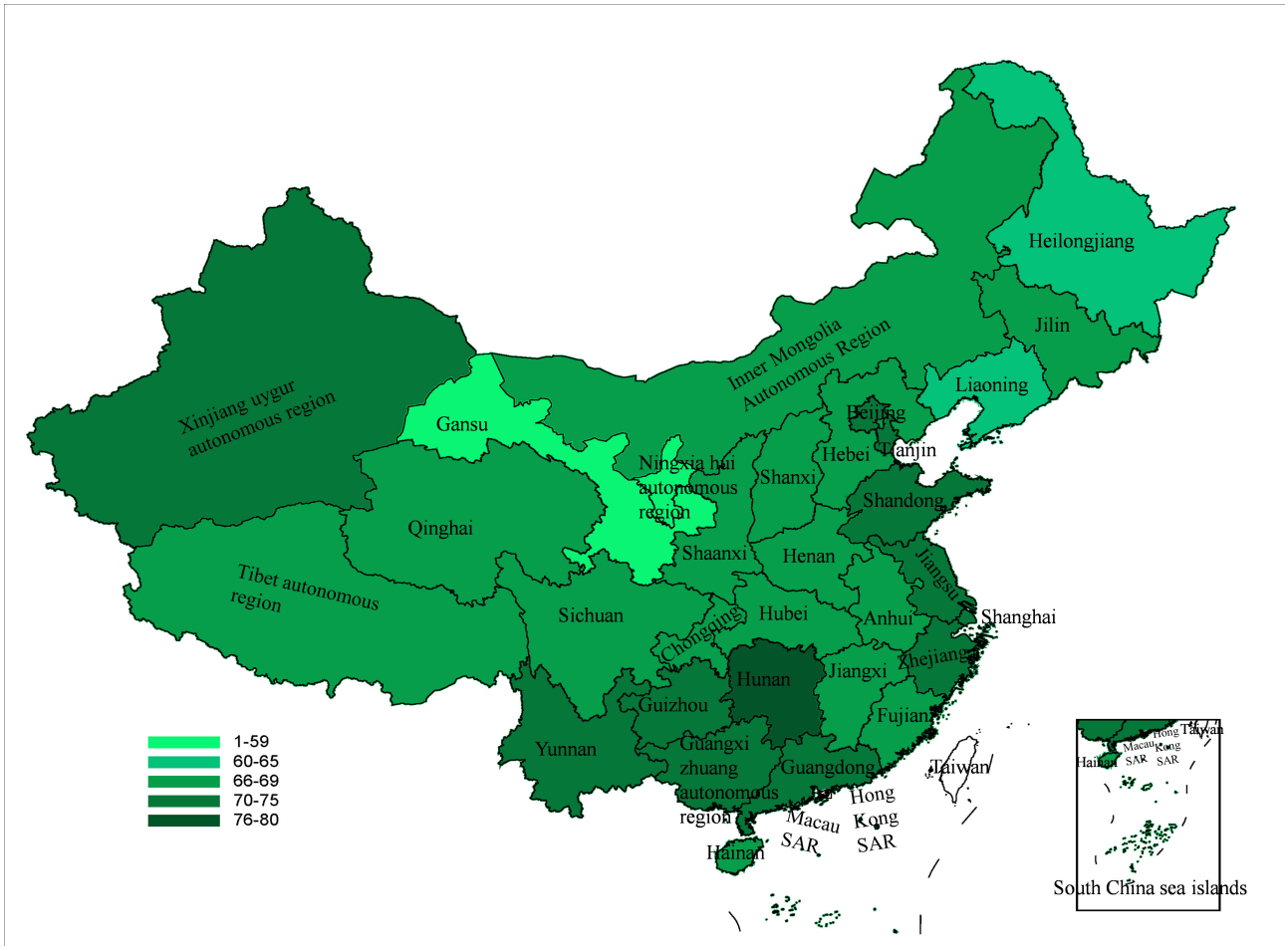
Among the 21,149 participants, the distribution of life education acceptance scores (VAS 0–100) was as follows: 80 (0.4%) scored 0-9; 558 (2.6%) scored 10-19; 918 (4.3%) scored 20-29; 1006 (4.8%) scored 30-39; 1802 (8.5%) scored 40-49; 2640 (12.5%) scored 50-59; 2,24 (12.9%) scored 60-69; 2347 (11.1%) scored 70-79; 2742 (13.0%) scored 80-89; and 6332 (29.9%) scored 90-100. The distribution showed a progressive increase in participant counts across higher score intervals, with the highest proportion concentrated in the 90–100 range, indicating a strong overall acceptance of life education in this population. The median (P_{25}, P_{75}) acceptance score for life education in this study was 71.00 (50.00-95.00), 76.7% of participants exhibited an acceptance level of 50 or above toward life education, and 6332 (28.9%) participants were listed at a 90 - 100 score, which was the largest number of all people. The acceptance of life education was different in each province (Figure S2 in Multimedia Appendix 1). And the score of most provinces was more than 60, except Ningxia province and Gansu province (Figure 2). The results showed that gender, medicine specialty, place of residence, debt status, highest level of education, age, region, chronic disease diagnosis, and media use were all significant factors associated with life education acceptance (Table 1).

Compared with females, males had lower acceptance of life education (β=−2.39; 95% CI −3.08 to −1.69; Figure 3). Compared with the eastern region, the central region (β=−1.98; 95% CI −2.83 to −1.14) and western region (β=−2.57; 95% CI −3.41 to −1.72) both had lower acceptance of life education. Respondents who were majoring in medicine (β=3.13; 95% CI 1.11-5.14), living in an urban area (β=1.25; 95% CI 0.46-2.04), having debt (β=1.54; 95% CI 0.80-2.28), holding a bachelor's degree or higher (β=4.05; 95% CI 2.97-5.12), having a chronic disease (β=1.37; 95% CI 0.48-2.25), or having higher scores on the Perceived Social Support Scale (β=0.49; 95% CI 0.38-0.61), the FHS-SF (β=0.46; 95% CI 0.39-0.52), the Short-Form Health Literacy Questionnaire (β=0.42; 95% CI 0.34-0.50), the Euro-Qol 5 Dimensions Questionnaire (β=0.20; 95% CI 0.02-0.37), or media use (β=0.49; 95% CI 0.41-0.57) had higher acceptance of life education. Compared with people aged 12-17 years, those aged 25-44 years (β=−6.00; 95% CI

XSL•FO
RenderX

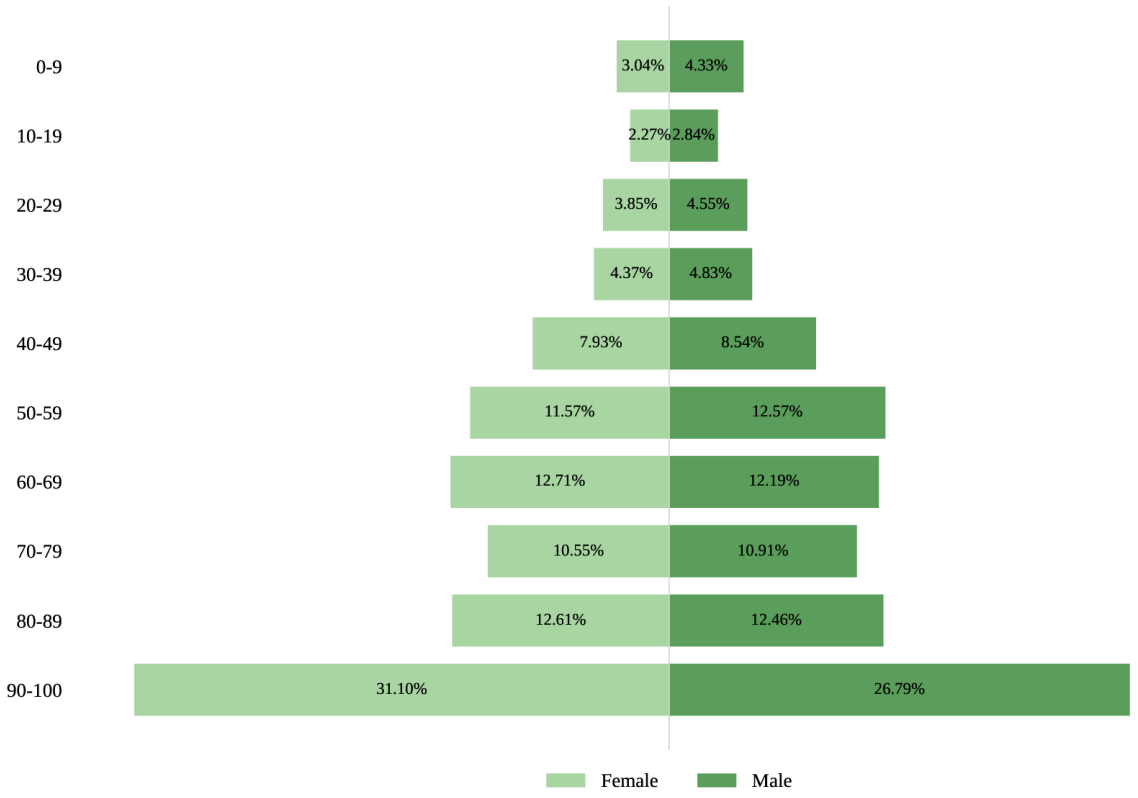
−7.34 to −4.66), aged 45-64 years (β=−4.55; 95% CI −5.88 to −3.22), and 65 years or older (β=−2.16; 95% CI −3.78 to −0.55) had lower acceptance of life education.

Figure 2. Provincial mean life education acceptance (visual analog scale [VAS] 0-100) in Mainland China, national cross-sectional survey, June-August 2022.



XSL•FO
RenderX

Figure 3. Life education acceptance (visual analog scale [VAS] 0-100) by gender in Mainland China, national cross-sectional survey, June-August 2022.



Subgroup Analysis of Life Education Acceptance in Medical and Nonmedical Majors

Understanding the factors influencing the level of life education among medical professionals can guide the development of palliative care and hospice medicine, which has important practical significance. Therefore, we further analyzed medicine specialty and nonmedicine specialty subgroups (Table 2). Medicine specialty and non-medicine specialty had different acceptance levels of life education. In the medicine specialty group, only place of residence, family health (FHS-SF), and media use had statistical significance. Compared to rural areas, people from urban areas had a greater acceptance of life education (β=5.36; 95% CI 1.51-9.21). Higher scores on the FHS-SF (β=0.36; 95% CI 0.05-0.67) and media use (β=1.28; 95% CI 0.81-1.75) had more acceptance of life education. For the nonmedicine specialty group, most of the influencing factors were the same as the whole population.

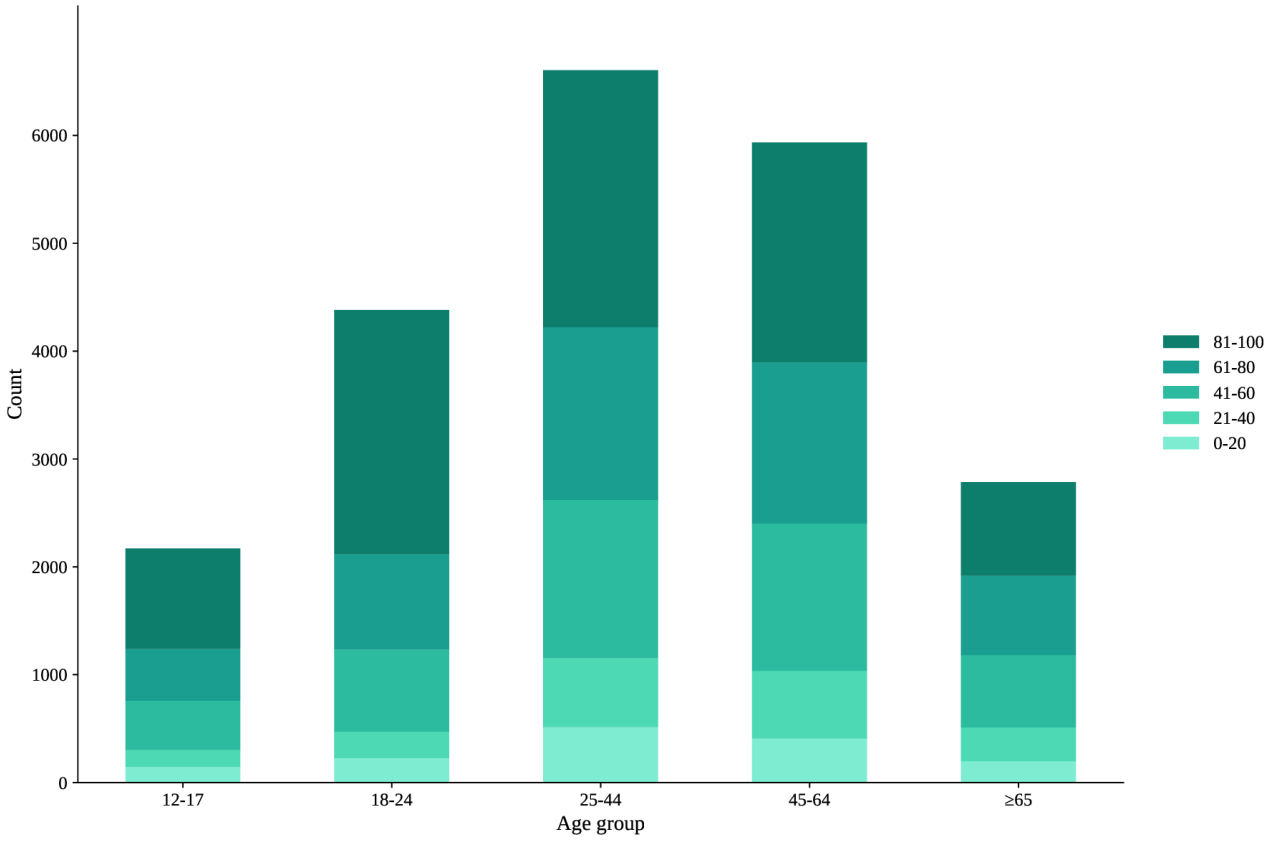
Subgroup Analysis of Life Education Acceptance of Different Ages

The essence of life education is the education of the whole life cycle, so age is a worthwhile variable to pay attention to (Table 3). Therefore, we further analyzed the influencing factors of different age subgroups. People of different ages had different acceptance levels of life education (Figure 4). HLS-SF and media use were the common influencing factors of all ages and were positively correlated with the acceptance of life education. Besides, those aged 12-17 years, who resided in an urban area (β=4.08; 95% CI 1.66-6.51), and had debt (β=3.95; 95% CI 1.52-6.39) had a higher acceptance of life education. The higher the score of the FHS-SF (β=0.40; 95% CI 0.18-0.61), the higher the acceptance level of life education they had. People aged 18-24 years had approximately the same acceptance level as all ages. However, perceived social support (PSSS) had no significance, and the higher the score of the EQ-5D-5L (β=0.76; 95% CI 0.33-1.19), the higher the acceptance level of life education they had, which was different across all ages. For those aged 25 -44 years, having debt (β=1.32; 95% CI 0.01-2.62), an educational level of undergraduate or above (β=7.09; 95% CI 5.19-8.99), having chronic disease (β=3.69; 95% CI 1.90-5.47), a higher score on the PSSS (β=0.36; 95% CI 0.14-0.58) and a higher score on the FHS-SF (β=0.45; 95% CI 0.33-0.57) exhibited higher acceptance of life education. However, medicine specialty, place of residence, and region had no significance, which was different from the findings in the overall sample. For those aged 45-64 years, those with an undergraduate degree or above (β=4.50; 95% CI 2.54-6.46) had higher acceptance of life education. Compared to eastern China, central China (β=−1.78; 95% CI −3.40 to −0.16) had lower acceptance of life education. Medicine specialty, place of residence, having debt, and chronic disease diagnosis had no significance, which was inconsistent with the results observed in the overall sample. For those aged older years, the higher score on the PSSS (β=0.94; 95% CI 0.59-1.29), the higher acceptance level of life education they had. Compared to eastern China, central China (β=−5.95; 95% CI −8.31 to −3.58), and western China (β=−4.27; 95% CI −6.55 to −1.99) had lower acceptance of life education. Gender, medicine specialty, place of residence, the highest level of education, chronic disease

XSL•FO
RenderX

diagnosis, and family health (FHS-SF) had no significance, which were differences across all ages.

Figure 4. Life education acceptance (visual analog scale [VAS] 0-100) by age group in Mainland China: national cross-sectional survey, June-August 2022.



Discussion

Principal Findings

The definition and essence of life education should be education for all and education throughout the life cycle [23]. This study was a large-sample cross-sectional survey on the acceptance of life education and its influencing factors among Chinese residents. It is important for promoting the acceptance of life education among the general public and the development of life education.

This study found that, overall, residents had a relatively high level of acceptance of life education, and acceptance was notably high among medical professionals. However, acceptance levels were lower among older residents, males, those with lower levels of education, and rural residents. While the findings regarding religious belief were nonsignificant. Media use, gender, place of residence, education level, and age were identified as the main factors influencing the level of acceptance of life education.

The subgroup analysis in medicine indicated that place of residence, family health, and media use were factors influencing the acceptance level of life education among medical professionals. Age subgroup analyses revealed that the influences on the acceptance of life education varied across age groups. Among them, health literacy and media use were common influencing factors across age groups. Besides, the age group of 12 - 17 years was also influenced by residence, debt situation, and family health. Additionally, religious beliefs were a relatively special influencing factor among residents aged 45 - 64 years. For the group aged 65 years and older, the debt situation, region, and social support were the influencing factors.

Life education had lower acceptance among the older adult population, possibly due to the influence of traditional cultural beliefs. Older adult individuals may avoid discussing topics related to death, which inhibits their ability to face and accept death calmly. Consequently, they may struggle to fully comprehend the true essence of life [4,24]. Therefore, it is important to actively provide end-of-life education to the older adult population. Helping them understand that the end of life is a natural process can assist in fostering a healthy perspective on life and death. In contrast, the younger generation shows higher acceptance, primarily due to school-based curricula targeting students [21,22]. However, further engagement is still needed to cultivate early awareness of health and interpersonal relationships, fostering positive perspectives on life and death for their future. Another important objective is that young people, as children in the family, are important decision-makers for the affairs of older adults [25]. Only when the young people have a correct understanding of life and death can they help the older adults cope with death in a better way. This study also found that the acceptance of life education among medical professionals was high, reaching a score of 85, which is 14 points higher than the average and should be maintained or further improved in the future. This is because medical

XSL•FO
RenderX

professionals, as a special group closely associated with life and death, undertake various tasks such as saving lives, providing life education, and offering end-of-life care [26-28]. Therefore, this group must possess a deep understanding of life and assume the responsibility of public education. Enhancing their capabilities is crucial for advancing hospice care and ensuring a meaningful life journey for patients.

In terms of gender, females are more receptive to life education than males, which is consistent with the findings of several previous studies [29-31]. Traditional Chinese gender roles influence the acceptance of life education. As primary caregivers for older adults, females are more exposed to the end-of-life process, enabling them to be better prepared for death than males, who typically focus on providing for the family [32]. Females build resilience by viewing death as a natural process and engaging in open family discussions. Conversely, males, burdened by societal expectations as primary providers, often perceive death as a heavy subject. They tend to avoid such discussions to maintain a positive, strong image for their children [33]. Therefore, females tend to have a higher acceptance of life education.

Individuals with higher media use have a higher level of acceptance of life education. Media use refers to the use of the internet, mobile phones, and television as channels for accessing desired information [34]. The widespread use of the internet provides individuals with greater access to knowledge about life education. Therefore, the higher the level of media exposure, the higher the acceptance of life education. The findings of previous studies [35,36] have consistently demonstrated that media usage can effectively enhance individuals' access to safety and health information, foster interpersonal connections, facilitate knowledge acquisition and sharing, and serve as a valuable tool for health education. Therefore, it is a good way to carry out life education through the internet media. But since the application of internet media is a double-edged sword [37,38], caution must be exercised regarding its method and extent of application.

The findings regarding religious belief in this study were nonsignificant, deviating from prior research outcomes. Previous research [39,40] indicated that within religious contexts, death is often perceived as a natural occurrence, prompting individuals to turn to religious beliefs for solace, purpose, and inner peace. Consequently, individuals with religious convictions tended to be more receptive to life education. However, in China's general environment, irreligious beliefs prevail [41], so the results of this study were not significant. Other studies have shown that traditional Chinese culture, influenced by Confucianism, Taoism, and Buddhism, regards life as a sacred and noble symbol, advocates caring for and cherishing life, and the majority of people show a psychology of denial, avoidance, and rejection of death, and therefore have a low degree of acceptance of life education [42]. Another interesting finding of this study was that the acceptance of life education among individuals in debt is unexpectedly high, contrary to previous research [43,44]. This may be related to the types of individuals' debts and individuals' perceptions of debt. For example, compared to other types of debts, student loans often cause greater stress and impact on mental and physical health [45]. However, when people do not perceive loans as a source of stress, but rather as a motivation for life's endeavors [46], they tend to be more positive, striving to live actively, cherishing life, embracing new experiences, and thereby exhibiting a higher acceptance of life education.

The results of the medical subgroup showed that only place of residence, family health, and media use were significant, which differs from previous studies. This may be due to the smaller sample size of the medical subgroup. In previous studies [47-50], there has been a greater focus on the acceptance of life education among medical professionals. It has been found that factors such as participation in palliative care activities, level of education, and attendance at training sessions related to death all have significant impacts. There is a close relationship between better acceptance and mastery of death-related knowledge and the sustained and healthy development of palliative and hospice care. To enhance professional capacity, schools should integrate life education and palliative care into curricula, emphasizing clinical practice to deepen medical students' understanding. Similarly, hospitals must provide continuing education to keep staff engaged in these topics. Furthermore, prioritizing cultural sensitivity, personalized care, and interdisciplinary collaboration is essential to improve patient services and advance the field of palliative medicine [51,52].

In terms of age subgroups, the influences on the acceptance of life education varied by age group, with health literacy and media use as influences common to all age groups, with higher levels being associated with higher acceptance of life education. Health literacy is defined as a cognitive and social skill that determines an individual's motivation and ability to acquire, understand, and use information in ways that promote and maintain health [53]. Individuals with high health literacy are more receptive to health information, making them more likely to accept life education [54-56]. For aging groups 12 - 17 years, the higher the level of family health, the higher the acceptance of life education [57,58]. Therefore, family-oriented life education can be conducted to enhance the acceptance level of life education among adolescent groups [59]. In addition, this study segmented the age group of young adults (18-24 years) for research because young adults are transitioning from school to society, experiencing corresponding changes in their thoughts and beliefs [60]. Therefore, life education for this age group holds significant importance. For this group, we should emphasize the dissemination of the meaning of life and knowledge about life safety. Focusing on themes such as "appreciating life," "respecting life," "cherishing life," and "nurturing life," the education should progress step by step, from shallow to deep, from theory to practice. This allows young people to appreciate the value of life and cherish it. For the group aged 65 years and above, individuals with higher levels of social support tend to have a higher acceptance level. Therefore, for the older population, we should continually strengthen social support and provide education on death-related knowledge to help them develop a correct perspective on life and death. First, it is necessary to create a societal atmosphere where discussions about death can be openly held. Building on this foundation, we can assist the older adults in understanding the concepts of life and death, recognizing the life cycle and

XSL•FO
RenderX

the aging process. Gradually, as they come to understand that death is a natural part of life, we can help them initiate rational "end-of-life" planning. This involves informing them that palliative care is a scientific approach to ensuring a peaceful end, reducing fear, and helping them to face the final stages of life calmly [4].

This shift in societal development patterns and the deepening of population aging have brought to light various issues regarding life education in our country. These issues primarily manifest in the following aspects: First, this target group of life education is mainly students [29,32]. Second, most life education primarily focuses on topics related to "life," with fewer discussions on topics related to death [61]. Third, there is a lack of specialized legal regulations and policy support specifically targeting life education (summary of existing legislation, see Table 8). The existing policies are broad in scope and lack specificity, making it difficult to directly focus on the essence of life and education. Therefore, they are unable to provide direct guidance for the promotion and development of life education. Fourth, there is insufficient interdisciplinary collaboration in the development of life education. The connections between disciplines such as education, medicine, and psychology are not sufficiently tight, making it difficult to establish comprehensive life education programs.

XSL•FO
RenderX

Table . National laws and policies related to life education in China: domains, key content, and source documents.

Category	Content of the law	Source
Educational programs and pathways	• The state integrates health education into the national education system: schools use various forms to implement health education and offer health courses [62]. • Encourage citizens to establish and practice the health management concept of being responsible for their own health, actively learn health knowledge, improve health literacy, and strengthen health management. Encourage family members to care for each other and form a healthy lifestyle that suits themselves and their family [62]. • The state encourages the development of various forms of continuing education so that citizens can receive appropriate forms of education in politics, economics, culture, science, technology, business, etc, promotes mutual recognition and connection of different types of learning results, and promotes lifelong learning for all [63]. • The state encourages social groups, social and cultural institutions, and other social organizations and individuals to carry out social and cultural education activities that are beneficial to the physical and mental health of the educated [63]. • Integrating comprehensive life safety and health education into primary and secondary school curriculum and textbooks, fostering a consciousness of cherishing life and self-protection at the primary school stage, teaching objective self-awareness and self-treatment at the middle school stage, and delving into the significance and value of healthy life at the high school stage [64]. • Establish and improve a health promotion and education system, improve health education service capabilities, and popularize health science knowledge from an early age. All types of media at all levels should increase the publicity of health science knowledge, actively build and standardize various radio and television and other health columns, and use new media to expand health education [65]. • Increase publicity on mental health among the people and improve their mental health literacy. Strengthen intervention for common mental disorders and psychological behavioral problems such as depression and anxiety, and increase early detection and timely intervention of psychological problems among key groups [65].	• Basic Medical Care and Health Promotion Law of the People's Republic of China (December 28, 2019) Chapter 6, Article 68 • Basic Medical Care and Health Promotion Law of the People's Republic of China (December 28, 2019) Chapter 6, Article 69 • Education Law of the People's Republic of China (April 29, 2021) Chapter 2, Article 20 • Education Law of the People's Republic of China (April 29, 2021) Chapter 6, Article 53 • Guidelines on Life Safety and Health Education in Primary and Secondary School Curricula and Materials (October 26, 2021) • "Healthy China 2030" Plan Outline (October 25, 2016) Chapter 4, Article 1 • "Healthy China 2030" Plan Outline (October 25, 2016) Chapter 5, Article 3

XSL•FO
RenderX

Category	Content of the law	Source
Legislative and policy guarantees	• The state and society respect and protect the right to health of citizens: the state establishes a health education system to ensure citizens' right to access health education [62]. • Governments at all levels prioritize the health of the people in their development strategies [62]. • The State Council and governments at all levels lead the work of medical and health care and health promotion [62]. • The state strengthens basic scientific research, supports the development of clinical medicine, and promotes the transformation and application of medical scientific and technological achievements [62]. • National health is the fundamental goal of building a Healthy China. Focusing on the entire population and across the entire lifespan, efforts are made to provide equitable and accessible, systematic and continuous health services, achieving a higher level of national health [65]. • Medical and health institutions at all levels should collaborate and cooperate to provide citizens with comprehensive and continuous medical and health services throughout the entire spectrum of prevention, health care, treatment, nursing, rehabilitation, and palliative care [62].	• Basic Medical Care and Health Promotion Law of the People's Republic of China (December 28, 2019) Chapter 1, Article 4 • Basic Medical Care and Health Promotion Law of the People's Republic of China (December 28, 2019) Chapter 1, Article 6 • Basic Medical Care and Health Promotion Law of the People's Republic of China (December 28, 2019) Chapter 1, Article 7 • Basic Medical Care and Health Promotion Law of the People's Republic of China (December 28, 2019) Chapter 1, Article 8 • "Healthy China 2030" Plan Outline (October 25, 2016) Chapter 2 • Basic Medical Care and Health Promotion Law of the People's Republic of China (December 28, 2019) Chapter 3, Article 36
Resource allocation	• The state increases its financial investment in medical and health care and health services [62]. • . The state promotes national health informatization and accelerates the construction of medical and health information infrastructure [62]. • Establish a mechanism for investment in medical and health care and health services that is commensurate with economic and social development, fiscal conditions, and health indicators [62]. • The state establishes a system that mainly focuses on financial allocations and is supplemented by raising education funds through various other channels [63].	• Basic Medical Care and Health Promotion Law of the People's Republic of China (December 28, 2019) Chapter 1, Article 11 • Basic Medical Care and Health Promotion Law of the People's Republic of China (December 28, 2019) Chapter 3, Article 49 • Basic Medical Care and Health Promotion Law of the People's Republic of China (December 28, 2019) Chapter 7, Article 80 • Education Law of the People's Republic of China (April 29, 2021) Chapter 7, Article 54
Training of professionals		• Basic Medical Care and Health Promotion Law of the People's Republic of China (December 28, 2019) Chapter 1, Article 8 • Basic Medical Care and Health Promotion Law of the People's Republic of China (December 28, 2019) Chapter 4, Article 52 • Basic Medical Care and Health Promotion Law of the People's Republic of China (December 8, 2019) Chapter 6, Article 67 • "Healthy China 2030" Plan Outline (October 25, 2016) Chapter 22, Article 1

XSL•FO
RenderX

Category	Content of the law	Source
	• The state develops medical education, improves the medical education system, and vigorously cultivates medical and health talents [62]. • The state formulates a training plan for medical and health personnel, establishes a medical and health personnel training mechanism and a supply and demand balance mechanism that adapt to industry characteristics and social needs, and improves medical school education, post-graduation education and continuing education systems [62]. • Governments at all levels should strengthen health education work and the training of health education professionals, establish systems for disseminating core information on health knowledge and skills [62]. • Strengthen the coordination of medicine and education and establish and improve the supply and demand balance mechanism for medical talent training. Reform the medical education system and accelerate the establishment of a medical talent cultivation and training system that is organically connected with the three stages of institutional education, post-graduation education, and continuing education that adapts to the characteristics of the industry [65].	

Therefore, in response to these issues with life education in China and considering its influencing factors, the following improvement measures are proposed: First, tailor personalized educational programs [4,29]. Develop life education plans tailored to different groups, taking into account their cultural backgrounds, regional characteristics, and economic conditions, especially targeting young people, older adults, and male populations. Second, diverse learning pathways should be established [66,67]. Community outreach via radio, bulletin boards, and workshops can enhance awareness. Additionally, flexible online courses should be developed to complement traditional offline methods, catering to different demographics. Third, a comprehensive support system involving families, communities, and government agencies should be established [68]. This network facilitates resource sharing and fosters mutual assistance at the family and community levels. Furthermore, the government should increase comprehensive resource allocation to ensure the sustainability and expansion of life education programs. Additionally, regular policy assessments are essential to optimize their effectiveness and maximize impact. Fourth, train promoters and cultivate medical professional teams [69]. Train community workers and volunteers to become advocates and educators for life education. Disseminate health knowledge in communities to enhance the acceptance and effectiveness of life education. At the same time, it is necessary to strengthen the construction of palliative care disciplines and the professional team of palliative care professionals so as to make a good manpower reserve for hospice.

This study has several strengths. First, the large sample size reduces random sampling error and increases the precision of estimates. Second, by examining factors associated with acceptance of life education, our findings may inform the design of public education strategies and support the integration of hospice and palliative care into broader health promotion efforts. This study also has limitations. The cross-sectional design limits causal inference: the observed associations do not establish temporality (ie, whether exposures preceded acceptance), and reverse causation is possible. Although we adjusted for multiple covariates, residual confounding from unmeasured factors may remain. Future research using longitudinal designs and, where feasible, intervention-based studies is warranted to clarify causal pathways and to test whether changes in modifiable factors (eg, health literacy or media exposure) lead to changes in acceptance of life education.

Conclusion

This study found that gender, age, media use, place of residence, level of education, perceived social support (PSSS), and health literacy (HLS-SF) are key factors influencing the acceptance of life education among Chinese residents. These findings suggest that both structural factors (sociodemographic characteristics) and modifiable psychosocial factors (social support and health literacy) should be considered when designing targeted improvement measures. Accordingly, reinforced life education interventions are particularly needed for key populations such as older adults, youth, and medical professionals. These results have important practical and contemporary value for advancing life education across the life

XSL•FO
RenderX

course, improving population quality of life, and supporting the development of palliative care.

Funding

The study is supported by the Non-profit Central Research Institute Fund of Chinese Academy of Medical Sciences (Grant No. 2022-ZHCH330-01).

Data Availability

The data used in this study are not currently available for public release.

Authors' Contributions

Conceptualization: XZ, NH, YY, CW
Methodology: XZ, NH, YY, Y Wang, WKM, CW
Validation: XZ, NH, YY, Y Wang
Formal analysis: XZ, NH, YY, Y Wang
Investigation: XZ, NH, YY, Y Wang
Resources: CW, WKM, XN, Y Wu
Data curation: CW, XN, Y Wu
Writing – original draft: XZ, NH, YY, Y Wang
Writing – review & editing: All authors
Visualization: XZ, NH
Supervision: CW, XN, Y Wu
Project administration: XZ, CW, XN, Y Wu
Additional statements: CW, XN, and Y Wu had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
XZ, NH, and YY contributed equally to this work and should be considered co-first authors.
Y Wu (email: wuyiboism@zju.edu.cn) and XN (email: ningxh1973@foxmail.com) are co-corresponding authors with CW.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary materials from the national cross-sectional survey on life education acceptance in mainland China, including provincial variation, descriptive statistics of acceptance scores, and measurement instruments with scoring and reliability information.
[DOCX File, 324 KB - publichealth_v12i1e78844_app1.docx]

Checklist 1
STROBE checklist.
[DOCX File, 38 KB - publichealth_v12i1e78844_app2.docx]

References

1. Phan HP, Ngu BH, Chen SC, Wu L, Lin WW, Hsu CS. Introducing the study of life and death education to support the importance of positive psychology: an integrated model of philosophical beliefs, religious faith, and spirituality. Front Psychol 2020;11:580186. [doi: 10.3389/fpsyg.2020.580186] [Medline: 33117246]
2. Shih YH. Improving the learning in life education for young children aged 3 to 6 years: a review on the research themes of early childhood life education in Taiwan. Children (Basel) 2022 Oct 9;9(10):10. [doi: 10.3390/children9101538] [Medline: 36291473]
3. Phan HP, Chen SC, Ngu BH, Hsu CS. Advancing the study of life and death education: theoretical framework and research inquiries for further development. Front Psychol 2023;14:1212223. [doi: 10.3389/fpsyg.2023.1212223] [Medline: 37575440]
4. Lei L, Lu Y, Zhao H, Tan J, Luo Y. Construction of life-and-death education contents for the elderly: a Delphi study. BMC Public Health 2022 Apr 21;22(1):802. [doi: 10.1186/s12889-022-13197-7] [Medline: 35449042]
5. Ronconi L, Biancalani G, Medesi GA, Orkibi H, Testoni I. Death education for palliative psychology: the impact of a death education course for Italian university students. Behav Sci (Basel) 2023 Feb 16;13(2):182. [doi: 10.3390/bs13020182] [Medline: 36829411]

XSL•FO
RenderX

6. Lei L, Zhao H, Ran L, Wang L, Luo Y. Influencing factors of attitudes towards death and demands for death education among community-dwelling Chinese older adults: a cross-sectional study. BMC Public Health 2022 Jun 23;22(1):1242. [doi: 10.1186/s12889-022-13655-2] [Medline: 35733112]
7. Lv T, Li L, Wang H, et al. Relationship between death coping and death cognition and meaning in life among nurses: a cross-ectional study. Omega (Westport) 2025 Aug;91(3):1528-1547. [doi: 10.1177/00302228231158911] [Medline: 36803151]
8. Zhang H, Lv T, Li L, Chen F, He S, Ding D. Nurse life-and-death education from the perspective of Chinese traditional culture. Omega (Westport) 2024 Mar 4:302228241236981. [doi: 10.1177/00302228241236981] [Medline: 38438064]
9. Cheng H. Seeing life through life: unpacking life education in intergenerational learning in China. Humanit Soc Sci Commun 2024;11(1):1048. [doi: 10.1057/s41599-024-03551-y]
10. Fu F, Jiao K, Luo C. Circle of light education model: death education for children in end of life care. Child Youth Serv Rev 2026 Jan;180:108678. [doi: 10.1016/j.childyouth.2025.108678]
11. Wu Y, Tang J, Du Z, et al. Development of a short version of the perceived social support scale: based on classical test theory and ant colony optimization. BMC Public Health 2025;25(1):232. [doi: 10.1186/s12889-025-21399-y]
12. Crandall A, Weiss-Laxer NS, Broadbent E, et al. The family health scale: reliability and validity of a short- and long-form. Front Public Health 2020;8:587125. [doi: 10.3389/fpubh.2020.587125] [Medline: 33330329]
13. Sun X, Lv K, Wang F, et al. Validity and reliability of the Chinese version of the Health Literacy Scale Short-Form in the Chinese population. BMC Public Health 2023 Feb 23;23(1):385. [doi: 10.1186/s12889-023-15237-2] [Medline: 36823591]
14. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec;20(10):1727-1736. [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
15. den Hamer AH, Konijn EA, Plaisier XS, Keijer MG, Krabbendam LC, Bushman BJ. The Content-based Media Exposure Scale (C-ME): development and validation. Comput Human Behav 2017 Jul;72:549-557. [doi: 10.1016/j.chb.2017.02.050]
16. Fung SF, Kong CYW, Liu YM, et al. Validity and psychometric evaluation of the Chinese Version of the 5-Item WHO Well-Being Index. Front Public Health 2022;10:872436. [doi: 10.3389/fpubh.2022.872436] [Medline: 35433612]
17. Liu D, Fan S, Huang X, et al. Study protocol: a national cross-sectional study on psychology and behavior investigation of Chinese residents in 2023. Health Care Sci 2024 Dec;3(6):475-492. [doi: 10.1002/hcs2.125] [Medline: 39735279]
18. Wu Y, Fan S, Liu D, Sun X. Psychological and behavior investigation of Chinese residents: concepts, practices, and prospects. Chin Gen Pract J 2024 Sep;1(3):149-156. [doi: 10.1016/j.cgpj.2024.07.006]
19. Yang Y, Fan S, Chen W, Wu Y. Broader open data needed in psychiatry: practice from the psychology and behavior investigation of Chinese residents. Alpha Psychiatry 2024 Aug;25(4):564-565. [doi: 10.5152/alphapsychiatry.2024.241804] [Medline: 39360297]
20. Zhang Y, Fan S, Hui H, et al. Privacy protection for open sharing of psychiatric and behavioral research data: ethical considerations and recommendations. Alpha Psychiatry 2025 Feb;26(1):38759. [doi: 10.31083/AP38759] [Medline: 40110382]
21. Jizhe N. Main data of the Seventh National Population Census. : National Bureau of Statistics; 2021 URL: https://www.stats.gov.cn/english/PressRelease/202105/t20210510_1817185.html [accessed 2026-04-06]
22. Krishnamoorthy Y, Nagarajan R, Saya GK, Menon V. Prevalence of psychological morbidities among general population, healthcare workers and COVID-19 patients amidst the COVID-19 pandemic: a systematic review and meta-analysis. Psychiatry Res 2020 Nov;293:113382. [doi: 10.1016/j.psychres.2020.113382] [Medline: 32829073]
23. Gao SH, Liu XM, Tian ZP. Research on the path of life education in higher continuing education. Adult Educ 2023;01:14-17 [FREE Full text]
24. Wysokiński M, Fidecki W, Jarosz M. Elderly people's acceptance of death: a study of a Polish cohort. Int J Environ Res Public Health 2019 Sep 12;16(18):3374. [doi: 10.3390/ijerph16183374] [Medline: 31547290]
25. Yu X, Wang Q. A study on the factors influencing old age identity among the Chinese elderly. Front Public Health 2022;10:1027678. [doi: 10.3389/fpubh.2022.1027678]
26. Gocmen Baykara Z, Keles S, Demir Karabulut S, et al. The effect of professional education on medical and nursing students' attitudes toward death. Death Stud 2022;46(7):1728-1740. [doi: 10.1080/07481187.2020.1850546] [Medline: 33287666]
27. Oba K, Tanimura C, Noguchi Y, Chujo M, Yoshioka SI. Perspectives of nursing students on end-of-life nurse education: a qualitative study of the guided death experience. Nurse Educ Today 2023 Jul;126:105834. [doi: 10.1016/j.nedt.2023.105834] [Medline: 37167831]
28. Tajiri T, Matoba S, Kuwabara M, et al. Death education for future physicians. Geriatr Gerontol Int 2020 Jan;20(1):5-6. [doi: 10.1111/ggi.13807] [Medline: 31916694]
29. Han H, Ye Y, Xie Y, et al. The impact of death attitudes on death education needs among medical and nursing students. Nurse Educ Today 2023 Mar;122:105738. [doi: 10.1016/j.nedt.2023.105738] [Medline: 36731246]
30. Zhang H, Cai X, Peng LY. An analysis of the status quo of college students' life and health education investigation and countermeasures. Chin J Health Educ 2020;36(5):441-445. [doi: 10.16168/j.cnki.issn.1002-9982.2020.05.012]
31. Lin PJ, Zhu Y, Olchanski N, et al. Racial and ethnic differences in hospice use and hospitalizations at end-of-life among medicare beneficiaries with dementia. JAMA Netw Open 2022 Jun 1;5(6):e2216260. [doi: 10.1001/jamanetworkopen.2022.16260] [Medline: 35679046]

XSL•FO
RenderX

32. Wang Y, Tang S, Hu X, Qin C, Khoshnood K, Sun M. Gender differences in attitudes toward death among Chinese college students and the implications for death education courses. Omega (Westport) 2022 May;85(1):59-74. [doi: 10.1177/0030222820934944] [Medline: 32580650]
33. Xie L, Li Y, Ge W, Lin Z, Xing B, Miao Q. The relationship between death attitude and professional identity in nursing students from mainland China. Nurse Educ Today 2021 Dec;107:105150. [doi: 10.1016/j.nedt.2021.105150] [Medline: 34560392]
34. Thomson K, Hunter SC, Butler SH, Robertson DJ. Social media "addiction": the absence of an attentional bias to social media stimuli. J Behav Addict 2021 Jul 15;10(2):302-313. [doi: 10.1556/2006.2021.00011] [Medline: 33852419]
35. Mendoza-Herrera K, Valero-Morales I, Ocampo-Granados ME, Reyes-Morales H, Arce-Amaré F, Barquera S. An overview of social media use in the field of public health nutrition: benefits, scope, limitations, and a Latin American Experience. Prev Chronic Dis 2020 Aug 6;17:E76. [doi: 10.5888/pcd17.200047] [Medline: 32762809]
36. Dunn PH, Woo BKP. Social media's role in the dissemination of health information. J Am Geriatr Soc 2019 Sep;67(9):1989-1990. [doi: 10.1111/jgs.16070] [Medline: 31317540]
37. Lee S, Tandoc EC Jr, Lee EWJ. Social media may hinder learning about science; social media's role in learning about COVID-19. Comput Human Behav 2023 Jan;138:107487. [doi: 10.1016/j.chb.2022.107487] [Medline: 36124311]
38. Sharma MK, John N, Sahu M. Influence of social media on mental health: a systematic review. Curr Opin Psychiatry 2020 Sep;33(5):467-475. [doi: 10.1097/YCO.0000000000000631] [Medline: 32639362]
39. de la Herrán Gascón A, Rodríguez Herrero P, Peralta IR, Rodríguez JJM. The pedagogy of death and special needs education. A phenomenological study. Eur J Spec Needs Educ 2022 Sep 3;37(5):747-760. [doi: 10.1080/08856257.2021.1943269]
40. Scherer JS, Milazzo KC, Hebert PL, et al. Association between self-reported importance of religious or spiritual beliefs and end-of-life care preferences among people receiving dialysis. JAMA Netw Open 2021 Aug 2;4(8):e2119355. [doi: 10.1001/jamanetworkopen.2021.19355] [Medline: 34347059]
41. Zhang J, Lew B, Liu Y, Chistopolskaya K, Zhao S. Religion, psychological strain, and suicidality in China: a preliminary study. Omega (Westport) 2024 May;89(1):275-291. [doi: 10.1177/00302228211072985] [Medline: 35098786]
42. Liu Y, Zhang X, Wang S, Shi F, Guo M. Effect of death education training based on "Thousand Chats live" on hospice care attitudes and death attitudes among emergency department nurses. Chin Nurs Res 2023;05:922-925. [doi: 10.12102/j.issn.1009-6493.2023.05.030]
43. Batomen B, Sweet E, Nandi A. Social inequalities, debt, and health in the United States. SSM Popul Health 2021 Mar;13:100736. [doi: 10.1016/j.ssmph.2021.100736] [Medline: 33553568]
44. Stickley A, Shirama A, Sumiyoshi T. Financial debt, worry about debt and mental health in Japan. BMC Psychiatry 2023 Oct 17;23(1):761. [doi: 10.1186/s12888-023-05235-4] [Medline: 37848860]
45. Froidevaux A, Koopmann J, Wang M, Bamberger P. Is student loan debt good or bad for full-time employment upon graduation from college? J Appl Psychol 2020 Nov;105(11):1246-1261. [doi: 10.1037/apl0000487] [Medline: 32105095]
46. Nau M, Dwyer RE, Hodson R. Can't afford a baby? Debt and young Americans. Res Soc Stratif Mobil 2015 Dec;42:114-122. [doi: 10.1016/j.rssm.2015.05.003] [Medline: 28090131]
47. Park SK, Lee CK, Kim H. Suicide mortality and marital status for specific ages, genders, and education levels in South Korea: using a virtually individualized dataset from national aggregate data. J Affect Disord 2018 Sep;237:87-93. [doi: 10.1016/j.jad.2018.05.003] [Medline: 29803100]
48. Duran S, Polat S. Nurses' attitudes towards death and its relationship with anxiety levels. Omega (Westport) 2024 Mar;88(4):1530-1544. [doi: 10.1177/00302228211065963] [Medline: 34982589]
49. Chang WP, Lin YK. Influence of basic attributes and attitudes of nurses toward death on nurse turnover: a prospective study. Int Nurs Rev 2023 Dec;70(4):476-484. [doi: 10.1111/inr.12781] [Medline: 35768904]
50. Blaschke SM, Schofield P, Taylor K, Ugalde A. Common dedication to facilitating good dying experiences: qualitative study of end-of-life care professionals' attitudes towards voluntary assisted dying. Palliat Med 2019 Jun;33(6):562-569. [doi: 10.1177/0269216318824276] [Medline: 30688145]
51. Su FJ, Zhao HY, Wang TL, Zhang LJ, Shi GF, Li Y. Death education for undergraduate nursing students in the China Midwest region: an exploratory analysis. Nurs Open 2023 Dec;10(12):7780-7787. [doi: 10.1002/nop2.2025] [Medline: 37823440]
52. Young J, Lyons A, Egan R, Dew K. Embodied decisions unfolding over time: a meta-ethnography systematic review of people with cancer's reasons for delaying or declining end-of-life care. BMC Palliat Care 2024 Feb 19;23(1):45. [doi: 10.1186/s12904-024-01342-5] [Medline: 38369452]
53. Li H, Wu Y, Bai Z, et al. The association between family health and frailty with the mediation role of health literacy and health behavior among older adults in China: nationwide cross-sectional study. JMIR Public Health Surveill 2023 Jun 27;9:e44486. [doi: 10.2196/44486] [Medline: 37368463]
54. Chen X, Giles J, Yao Y, et al. The path to healthy ageing in China: a Peking University-Lancet Commission. Lancet 2022 Dec 3;400(10367):1967-2006. [doi: 10.1016/S0140-6736(22)01546-X] [Medline: 36423650]
55. Nutbeam D, Lloyd JE. Understanding and responding to health literacy as a social determinant of health. Annu Rev Public Health 2021 Apr 1;42:159-173. [doi: 10.1146/annurev-publhealth-090419-102529] [Medline: 33035427]

XSL•FO
RenderX

56. Shin HR, Choi EY, Kim SK, Lee HY, Kim YS. Health literacy and frailty in community-dwelling older adults: evidence from a Nationwide Cohort Study in South Korea. Int J Environ Res Public Health 2021 Jul 27;18(15):15. [doi: 10.3390/ijerph18157918] [Medline: 34360213]
57. Goodman DM. Family health is child health. J Pediatr 2017 Feb;181:1-2. [doi: 10.1016/j.jpeds.2016.12.009] [Medline: 28129867]
58. Weiss-Laxer NS, Crandall A, Okano L, Riley AW. Building a foundation for family health measurement in national surveys: a modified Delphi expert process. Matern Child Health J 2020 Mar;24(3):259-266. [doi: 10.1007/s10995-019-02870-w] [Medline: 31912378]
59. Barnes MD, Hanson CL, Novilla LB, Magnusson BM, Crandall AC, Bradford G. Family-centered health promotion: perspectives for engaging families and achieving better health outcomes. Inquiry 2020;57:46958020923537. [doi: 10.1177/0046958020923537] [Medline: 32500768]
60. Huang Y, Heflin CM, Validova A. Material hardship, perceived stress, and health in early adulthood. Ann Epidemiol 2021 Jan;53:69-75. [doi: 10.1016/j.annepidem.2020.08.017] [Medline: 32949721]
61. Ruíz-Fernández MD, Fernández-Medina IM, Granero-Molina J, Hernández-Padilla JM, Correa-Casado M, Fernández-Sola C. Social acceptance of death and its implication for end-of-life care. J Adv Nurs 2021 Jul;77(7):3132-3141. [doi: 10.1111/jan.14836] [Medline: 33755231]
62. Basic medical care and health promotion law of the People's Republic of China. Standing Committee of the National People's Congress. URL: https://www.nhc.gov.cn/fzs/c100048/202007/e1192329ea4d45ef892892cd8232d385.shtml [accessed 2026-01-15]
63. Education law of the People's Republic of China. Ministry of Education of the People's Republic of China. URL: http://www.moe.gov.cn/jyb_sjzl/sjzl_zcfg/zcfg_jyfl/202107/t20210730_547843.html [accessed 2026-01-15]
64. Notice of the Ministry of Education on issuing the "Guidelines on life safety and health education in primary and secondary school curricula and materials". Ministry of Education of the People's Republic of China. URL: http://www.moe.gov.cn/srcsite/A26/s8001/202111/t20211115_579815.html [accessed 2026-01-15]
65. The Central Committee of the Communist Party of China and the State Council issue the "Healthy China 2030" plan outline. The Central Committee of the Communist Party of China and the General Office of the State Council. URL: https://www.gov.cn/zhengce/2016-10/25/content_5124174.htm [accessed 2026-01-15]
66. Suarez-Balcazar Y. Meaningful engagement in research: community residents as co-creators of knowledge. Am J Community Psychol 2020 Jun;65(3-4):261-271. [doi: 10.1002/ajcp.12414] [Medline: 31907950]
67. Zimitat C. Designing effective on-line continuing medical education. Med Teach 2001 Mar;23(2):117-122. [doi: 10.1080/01421590120036538] [Medline: 11371287]
68. Weiss JB, Berner ES, Johnson KB, Giuse DA, Murphy BA, Lorenzi NM. Recommendations for the design, implementation and evaluation of social support in online communities, networks, and groups. J Biomed Inform 2013 Dec;46(6):970-976. [doi: 10.1016/j.jbi.2013.04.004] [Medline: 23583424]
69. Pelayo M, Cebrián D, Areosa A, Agra Y, Izquierdo JV, Buendía F. Effects of online palliative care training on knowledge, attitude and satisfaction of primary care physicians. BMC Fam Pract 2011 May 23;12:37. [doi: 10.1186/1471-2296-12-37] [Medline: 21605381]

Abbreviations

EQ-5D-5L: Quality of Life Scale
FHS-SF: Family Health Scale-Short Form
HLS-SF: Health Literacy Scale-Short Form
PSSS: Perceived Social Support Scale
VAS: visual analog scale
WHO-5: World Health Organization-Five Well-Being Index

Edited by A Mavragani, T Sanchez; submitted 10.Jun.2025; peer-reviewed by D Dinh, RHM Macena, S Saritas, Z Xiang; revised version received 15.Feb.2026; accepted 24.Feb.2026; published 21.Apr.2026.

Please cite as:
Zhang X, Huang N, Yi Y, Wang Y, Ming WK, Wu Y, Ning X, Wang C
Factors Affecting Acceptance of Life Education in Mainland China: National Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e78844
URL: https://publichealth.jmir.org/2026/1/e78844
doi:10.2196/78844

XSL•FO
RenderX

© Xinyue Zhang, Nana Huang, Yuyang Yi, Yinlin Wang, Wai-Kit Ming, Yibo Wu, Xiaohong Ning, Chen Wang. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 21.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Prevalence and Factors Associated With Acute Stress Disorder Among Adults Ever Infected With COVID-19 During the Ending Phase of the Pandemic in 7 Chinese Cities: Cross-Sectional Study

Ziying Yang[1], PhD; Yanqiu Yu[2], PhD; Hui Lu[3], MS; Xu Wang[1], PhD; Yong Xu[4], BS; Junqiang Ying[4], BS; Xianying Wen[5], MS; Lei Luo[5], BS; Meng Wang[6], MS; Muwen Liu[6], MS; Xingyi Geng[7], MS; Xuchong Zhao[7], BS; Biyu He[8], PhD; Tao Liu[9], MS; Remina Maimaitijiang[10], MS; Jing Gu[1,11,12,13], PhD; Joseph T F Lau[14,15], PhD

[1]Department of Medical Statistics, School of Public Health, Sun Yat-sen University, No. 74, Zhongshan 2nd Road, Nonglin Street, Yuexiu District, Guangzhou, China

[10]Department of Public Health, School of Medicine, Shihezi University, Shihezi, China

[11]Guangzhou Joint Research Center for Disease Surveillance, Early Warning, and Risk Assessment, Guangzhou, China

[12]Sun Yat-Sen University Global Health Institute, School of Public Health and Institute of State Governance, Sun Yat-Sen University, Guangzhou, China

[13]Guangdong Key Laboratory of Health Informatics, Guangzhou, China

[14]Zhejiang Provincial Clinical Research Center for Mental Disorders, The Affiliated Wenzhou Kangning Hospital, Wenzhou Medical University, 1 Shengjing Road, Lucheng District, Wenzhou, China

[15]School of Mental Health, Wenzhou Medical University, Wenzhou, China

[2]School of Public Health, Fudan University, Shanghai, China

[3]Oujiang Laboratory (Zhejiang Lab for Regenerative Medicine, Vision, and Brain Health), Postgraduate Training Base Alliance of Wenzhou Medical University, Wenzhou, China

[4]Neijiang Center for Disease Control and Prevention, Neijiang, China

[5]Mianyang Center for Disease Control and Prevention, Mianyang, China

[6]Hangzhou Center for Disease Control and Prevention, Hangzhou, China

[7]Jinan Center for Disease Control and Prevention, Jinan, China

[8]Shanghai Municipal Center for Health Promotion, Shanghai, China

[9]Community Health Service Center of Friendship Street, Shanghai, China

Corresponding Author:
Joseph T F Lau, PhD
Zhejiang Provincial Clinical Research Center for Mental Disorders, The Affiliated Wenzhou Kangning Hospital, Wenzhou Medical University, 1 Shengjing Road, Lucheng District, Wenzhou, China

Abstract

Background: Acute stress disorder (ASD) among people ever infected with COVID-19 is prevalent and may lead to posttraumatic stress disorder. Soon after China relaxed their COVID-19 control measures in November 2022 or December 2022, the infection rate surged rapidly, creating huge uncertainty and stressful situations. Little is known about situations regarding ASD at the ending phase of the pandemic.

Objective: The study aimed to investigate the potential of personal cognitive or emotional factors and environmental factors of ASD.

Methods: A cross-sectional study was conducted among 5545 people ever infected with COVID-19 aged 18 - 60 years from December 27, 2022, to January 9, 2023, living in 7 cities of China. The 5-item Chinese version of the Primary Care PTSD Screen was used to assess ASD. Multiple logistic regression analyses were performed to identify factors of ASD.

Results: The prevalence of ASD was 21.2% (1174/5545). Adjusted for the background variables, significant personal risk factors (COVID-19 infection severity, cognitions including perceived high reinfection risk and perceived weak acquired natural immunity, and emotions including worry about the long-term physical harms and panic about infection of older or younger family members), and significant environmental risk factors (difficulties in getting information and medical supplies, having unvaccinated older or younger family members, and having significant others with severe COVID-19 symptoms) were identified.

Conclusions: The prevalence of ASD among people ever infected with COVID-19 was noticeable. It is warranted to identify those at high risk of developing ASD and provide them with care and early interventions to prevent deterioration. Such programs may consider targeting the modifiable risk factors found in this study.

XSL•FO
RenderX

(JMIR Public Health Surveill 2026;12:e73002) doi:10.2196/73002

KEYWORDS

COVID-19; China; acute stress disorder; perceptions; emotions; risk factors

Introduction

As of September 29, 2023, COVID-19 had caused approximately 770 million confirmed cases and over 6 million deaths globally [1]. The Chinese government responded immediately to implement the zero-COVID-19 strategy after the outbreak in Wuhan, including strict measures such as lockdown, quarantine, travel restrictions, and regular mass testing [2-4]. By November 2022, the country had recorded about 9.3 million COVID-19 cases and 29,000 related deaths [5], respectively. Subsequently, the "Twenty Measures" were announced to ease certain controls (eg, quarantine measures and testing requirements for entry to public venues). On December 7, 2022, the "New Ten Measures" removed all control measures in China [6]. Due to the previously low population immunity to COVID-19 in China and the high transmissibility of the virus, infections surged rapidly across the country [7,8]. For instance, a seroprevalence study detecting open reading frame 8 antigen reported infection rates from 61.5% (922/1500) to 80.7% (1210/1500) in Guangzhou, China, during January 5, 2023, and January 14, 2023 [9].

The pandemic has been associated with an elevated prevalence of depression in both infected and uninfected individuals [10,11]. For example, the global prevalence of depression increased by 25% during the pandemic [12]. Meta-analyses have reported depression prevalence rates from 18% to 33% in the general population [10] and 37% to 54% among people ever infected with COVID-19 during that time period [13]. People ever infected with COVID-19 may encounter stronger stressors and hence experience a higher prevalence of mental distress than their uninfected counterparts [14-20].

People ever infected with COVID-19 are at risk of developing acute stress disorder (ASD) and posttraumatic stress disorder (PTSD). ASD is a mental health disorder that can emerge shortly after a traumatic event. If the symptoms persist beyond 1 month, the diagnosis may progress to PTSD [21]. PTSD was significantly associated with depression and suicidal ideation [22,23]. Furthermore, ASD is a known predictor of PTSD [24,25] and other harms such as memory impairment [26]. Studies in China, Italy, and Singapore reported PTSD prevalence among people ever infected with COVID-19 ranging from 9.3% to 43.3% [27]. The prevalence of ASD was 25.0%, 40.0%, and 24.1% in New York City, Romania, and Spain, respectively [28-30]. In China, only 1 study has reported ASD in people ever infected with COVID-19 (71%) during the early outbreak phase [31]. The unprecedented acceleration of COVID-19 spread in the final phase of the pandemic in China created an extremely stressful social context. People ever infected with COVID-19 may worry about transmitting the virus to their family members, which can worsen their mental distress [32-34]. Given that hundreds of millions of Chinese people were infected within a few months, a substantial number were exposed to a high risk of ASD. To inform the design of prevention programs, it is crucial to understand the risk factors for ASD among people ever infected with COVID-19.

According to the diathesis-stress model and its contemporary refinements (eg, differential susceptibility), the development of mental disorders is jointly shaped by personal predispositions and contextual or environmental factors [35-38]. Based on this framework, severity of symptoms, negative cognitions, and negative emotions related to COVID-19 have been identified as personal risk factors of mental distress. For example, severe symptoms experienced during the acute phase of infection have been shown to increase the risk of long-term depression among people ever infected with COVID-19 [14]. Negative cognitions such as a perceived risk of infection were associated with negative mental health outcomes among health care workers [39,40]. Personal emotional factors commonly studied included concern about family members' infection [41] and worry about the long-term negative effects of COVID-19 [42]. Contextual risk factors included limited access to accurate disease-related information [43], necessary supplies [44], masks [45], medicine [46], and medical care [47]. Additionally, COVID-19 infection among family members and colleagues was identified as an environmental risk factor of depression, anxiety, and insomnia among people ever infected with COVID-19 [19].

Only 5 studies were identified in our literature review that specifically investigated risk factors for ASD among people ever infected with COVID-19. A survey conducted in Pakistan (n=114) identified several risk factors of ASD, including inadequate communication, poor ward facilities, worries about family members, and financial problems [48]. A Spanish study (n=90) found risk factors, including younger age, female sex, obesity, history of psychiatric diseases, and disease severity at intensive care unit (ICU) admission [49]. Another Italian study found that hypoxemia at ICU admission was negatively associated with ASD [30]. An Indian study found that smell or taste disturbance was associated with ASD [50]. An American study found that neither demographic characteristics nor clinical characteristics (eg, ICU admission and inflammatory markers) were associated with ASD [29]. These findings indicate that a relatively narrow range of potential risk factors has been examined. To the best of our knowledge, no such ASD studies were conducted among people ever infected with COVID-19 in China, although several reported factors of ASD in populations, including both people ever infected with COVID-19 and non–people ever infected with COVID-19 [31,51-54]. Furthermore, most of such studies were conducted during the early phase of the pandemic, whereas the pandemic's ending phase in China presented a distinct social and psychological context with its own uncertainties.

The present study investigated the prevalence of ASD and associated personal factors and contextual or environmental factors of ASD among people ever infected with COVID-19 during the ending phase of the COVID-19 pandemic (from December 2022 to January 2023) in mainland China. It is

XSL•FO
RenderX

hypothesized that (1) the following personal factors will be positively associated with ASD: (a) severity of COVID-19 infection, (b) cognitive variables, including perceived risk of reinfection and perceived strong acquired natural immunity within 6 months since infection, and (c) emotional factors, including worry about long-term harms of COVID-19 infection to oneself and panic about risks of infection of one's older or younger family members; (2) the following contextual or environmental factors will be positively associated with ASD: (a) difficulties in obtaining COVID-19 related information, (b) difficulties in obtaining drugs rapid antigen test (RAT) supplies, (c) presence of unvaccinated older or younger family members, and (d) having relatives or friends with severe COVID-19 symptoms. Most of these factors of ASD among people ever infected with COVID-19 have not been studied.

Methods

Participants and Data Collection

An anonymous cross-sectional study was conducted in 7 Chinese cities during December 27, 2022, and January 9, 2023. Eligible participants were adults aged 18 to 60 years who received a COVID-19 diagnosis between December 1, 2022 (when the control measures started to be loosened or removed) and the survey date. The 7 cities—Hangzhou and Shanghai (east), Guangzhou (south), Jinan (northeast), Neijiang and Mianyang (west), and Shihezi (northwest)—were purposively selected to reflect different geographic regions.

Participants were recruited using a multistage, cluster-based recruitment across the above selected cities. Administratively, Chinese cities are divided into districts, which are further subdivided into streets, and in turn, into communities. With support from local Centers for Disease Control and Prevention (CDC), at least 3 districts were selected in each city. Within each selected district, at least 1 street was randomly selected using simple random sampling from the administrative list. Within each selected street, at least 6 communities with more than 100 residents and an existing WeChat group were randomly selected from the administrative list. WeChat groups are widely used communication platforms in communities in China [55]. Centers for Disease Control and Prevention representatives or community leaders then posted the survey link to the WeChat group of each community and invited only 1 adult per household to complete the questionnaire via the Wenjuanxing platform (Changsha Ranxing Information Technology Co, Ltd) [56].

A total of 6028 completed questionnaires met the inclusion criteria. For quality control, 483 participants were excluded (completion time > 4 min: n=445; repetitive answer patterns and substantial missingness, n=38), yielding an analytic sample size of 5545.

Ethical Considerations

The participation was voluntary and anonymous, and returning the completed questionnaire indicated informed consent. No incentives were given to the participants. Personal identifiers were not collected, and all analyses were performed on deidentified data to ensure participant privacy and confidentiality. Ethics approval was obtained from the ethics committee of the School of Public Health, Zhejiang University (ZLG202301-01).

Measures

The overall survey comprised modules on illness perception, protective behaviors, and mental health. This study focuses on the prevalence of ASD and its associated personal and environmental factors. Findings from the other modules are reported in published papers [57].

Background Information

Such information included the study city, community type (urban or rural area), sex, age, education level, current marital status, employment status, and chronic disease status (no chronic disease, chronic disease well controlled, and chronic disease not well controlled).

Acute Stress Disorder Assessment

ASD was assessed by the 5-item Chinese version of the Primary Care PTSD Screen (Primary Care PTSD Screen for *Diagnostic and Statistical Manual of Mental Disorders, Fifth Revision* [PC-PTSD-5]) [58], which was also validated in a Chinese population [59]. The items have binary responses (0=no, 1=yes). Probable ASD was defined as PC-PTSD-5 total scores greater than or equal to 3 [58]. The Cronbach α was .808 in this study.

Personal Factors

Severity of COVID-19 Infection

The item was as follows: "How severe were your COVID-19 symptoms?" It was rated with a 4-point Likert scale (asymptomatic, mild, severe and not hospitalized, and severe and hospitalized).

Cognitive Factors

The cognitive factors were as follows:

1. Perceived risk of reinfection was assessed by the item: "In the next month, there is a high possibility that I might contract or experience a reinfection of COVID-19" ("1=strongly disagree," "2=disagree," "3=neutral," "4=agree," and "5=strongly agree"). The responses were recoded into high (4 and 5), average (3), and low (1 and 2).
2. Perceived natural immunity level within the 6 months since COVID-19 infection was assessed by the item: "With COVID-19 infection, I have acquired natural immunity and would not be re-infected within the next six months" ("1=strongly disagree," "2=disagree," "3=neutral," "4=agree," and "5 =strongly agree"). The responses were recoded into strong (4 and 5), average (3), and weak (1 and 2).

Emotional Factors

The emotional factors were as follows:

1. Worry about long-term physical harms of COVID-19: the item was as follows: "What are the chances that the COVID-19 infection would cause serious long-term physical harms to you?" (0=no chance to 10=extremely great chances).

XSL•FO
RenderX

2. Panic about infection of older or younger family members: the item was as follows: “Do you feel panic about the high risk that older or younger family members (age >65 years and aged 3‐11 years) would contract or recontract COVID-19?” (0=no, 1=yes).

Contextual or Environmental Factors

Information-Related Difficulty

Difficulties in obtaining high-quality information related to COVID-19 were assessed by the summative scores of 3 items: “Have you frequently received inconsistent or confusing COVID-19-related information from external sources?,” “Do you find it hard to determine the accuracy of the COVID-19-related information you have received?,” and “Do you feel overwhelmed and uncertain about what to do due to the rapid changes in COVID-19 related information” (0=no, 1=yes). Higher scores indicate more difficulties in obtaining high-quality information.

Difficulty in Obtaining Drugs or RAT Supplies

It was assessed by the item: “How difficult was it to obtain medications (eg, for fever or cough) and RAT currently?” (1=not difficult at all to 7=impossible to obtain).

Vaccination and Infection Status of Family Members, Relatives, or Friends

The vaccination and infection status were as follows:

1. Having unvaccinated family members aged greater than or equal to 65 years: the item was as follows: “Do you have any unvaccinated family members aged greater than or equal to 65 years?” (0=no, 1=yes).
2. Having unvaccinated family members aged 3‐11 years: the item asked: “Do you have any unvaccinated family members aged 3‐11 years?” (0=no, 1=yes).
3. Having relatives or friends having severe COVID-19 symptoms: the item was as follows: “Do you know of any relatives or friends having developed severe COVID-19 symptoms?” A 4-point Likert scale was used (none, only a few, some, and many).

Statistical Analysis

The distributions of studied variables are presented. Univariate logistic regression models were used to test the associations between the background factors and ASD; univariate odds ratio and 95% CIs were estimated. Multivariable logistic regression models were used to identify independent correlates of probable ASD. For each of the considered personal and environmental factors, adjusted odds ratio (ORa) and their 95% CI were respectively estimated with adjustment for all background variables. Statistical significance was reached when the 2-tailed $P<.05$. The analyses were performed using *R* Software, version 4.2.3 (R Foundation for Statistical Computing) [60].

Results

Descriptive Statistics

The characteristics of the study participants are summarized in Table 1. Of the 5545 participants, the majority were female participants (n=1590, 71.3%), living in an urban area (n=4569, 82.4%), aged below 40 years (n=3422, 61.7%), had attained an undergraduate degree or above (n=3095, 55.8%), working full-time (n=4065, 73.3%), and did not have chronic diseases (n=4643, 83.7%). More than half of them had mild COVID-19 symptoms (n=2883, 52%); 0.7% (n=37) had been hospitalized. Among the participants, 25.8% (n=1433), 45.4% (n=2515), or 28.8% (n=1597) perceived a high, average, or low risk of reinfection; 21.6% (n=1197), 50.2% (n=2782), or 28.2% (n=1566) perceived strong, average, or weak natural immunity within 6 months since COVID-19 infection; 75.7% (n=4195) felt panic about infection of their older or younger family members. The majority did not have unvaccinated family members aged greater than or equal to 65 years (n=4789, 86.4%) and 3 to 11 years (n=5216, 94.1%); 52.2% (n=2897) had a few relatives or friends having severe COVID-19 symptoms. The mean scores (SD; range) of worry about long-term physical harms of COVID-19, information-related difficulty, and difficulty in obtaining drugs or RAT supplies were 6.39 (2.44; 0‐10), 1.42 (1.14; 0‐3), and 4.44 (1.30; 1‐7), respectively. The prevalence of ASD was 21.2% (n=1174; 95% CI 20.1%-22.2%).

XSL•FO
RenderX

Table . Descriptive statistics among people ever infected with COVID-19 aged 18 - 60 years (N=5545).[a]

	Value
Background variables, n (%)	
City	
Guangzhou	414 (7.5)
Hangzhou	1790 (32.3)
Jinan	668 (12.0)
Mianyang	1176 (21.2)
Neijiang	1090 (19.7)
Shanghai	243 (4.3)
Shihezi	164 (3.0)
Community type	
Rural	976 (17.6)
Urban	4569 (82.4)
Sex	
Male	1590 (28.7)
Female	3955 (71.3)
Age group (y)	
18-30	1465 (26.4)
31-40	1957 (35.3)
41-50	1493 (26.9)
51-60	630 (11.4)
Education level	
Below undergraduate degree	2450 (44.2)
Undergraduate degree and above	3095 (55.8)
Current marital status	
Unmarried	1296 (23.4)
Married	4249 (76.6)
Employment status	
Not being employed	1321 (23.8)
Full-time job	4062 (73.3)
Part-time job	162 (2.9)
Chronic disease status	
None	4643 (83.7)
Yes, extremely well controlled	126 (2.3)
Yes, well controlled	564 (10.2)
Yes, not well controlled	212 (3.8)
Personal factors, n (%)	
Severity of COVID-19 infection	
Asymptomatic	65 (1.1)
Mild	2883 (52)
Severe and not hospitalized	2560 (46.2)
Severe and hospitalized	37 (0.7)
Cognitive factors, n (%)	

XSL•FO
RenderX

	Value
Perceived risk of reinfection	
Low	1597 (28.8)
Average	2515 (45.4)
High	1433 (25.8)
Perceived natural immunity level within 6 mo since COVID-19 infection	
Strong	1197 (21.6)
Average	2782 (50.2)
Weak	1566 (28.2)
Emotional factors	
Worry about long-term physical harms of COVID-19, mean (SD)	6.39 (2.44)
Panic about infection of older or younger family members, n (%)	
No	1350 (24.3)
Yes	4195 (75.7)
Contextual/environmental factors	
Information-related difficulty, mean (SD)	1.42 (1.14)
Difficulty in obtaining drugs/RAT[b] supplies, mean (SD)	4.44 (1.3)
Having unvaccinated family members aged ≥65 years, n (%)	
No	4789 (86.4)
Yes	756 (13.6)
Having unvaccinated younger family members aged 3-11 years, n (%)	
No	5216 (94.1)
Yes	329 (5.9)
Having relatives/friends with severe COVID-19 symptoms, n (%)	
None	884 (15.9)
Only a few	2897 (52.2)
Some	1377 (24.8)
Many	387 (7.0)
Dependent variable, n (%)	
Acute stress disorder	
No	4371 (78.8)
Yes	1174 (21.2)

[a]Chronic disease includes any of hypertension, diabetes, chronic lung disease, myocardial infarction, heart failure, cerebrovascular disease, dementia, and ulcerative diseases such as gastric ulcer, liver disease, and tumor. Range of worry for long-term impact of COVID-19, information difficulty, and difficulty in obtaining drugs or rapid antigen test supplies are 0 - 10, 0 - 3, and 1 - 7, respectively. Acute stress disorder is defined using a cutoff of 3 or higher.

[b]RAT: rapid antigen test.

Factors of ASD

Background Factors

Univariate logistic regression analyses showed that living in Neijiang (vs Guangzhou), female sex, ages greater than 30 years (vs 18 - 30 y), being currently married, having a part-time job (vs not being employed), and having chronic diseases (both well controlled and not well controlled vs no chronic disease) were more likely than others to develop ASD. Conversely, university or above education and full-time employment (vs not being employed) were negatively associated with ASD (Table 2).

XSL•FO
RenderX

Table . The association between background factors and acute stress disorder among people ever infected with COVID-19 aged 18 - 60 years (N=5545).

	ORu[a] (95% CI)
City	
Guangzhou	Ref[b]
Hangzhou	1.11 (0.84-1.46)
Jinan	1.29 (0.95-1.75)
Mianyang	1.15 (0.87-1.53)
Neijiang	1.45 (1.09-1.93)[c]
Shanghai	1.43 (0.97-2.09)
Shihezi	0.65 (0.39-1.10)
Community type	
Rural	Ref
Urban	0.97 (0.82-1.15)
Sex	
Male	Ref
Female	1.41 (1.21-1.63)[d]
Age group (y)	
18 - 30	Ref
31 - 40	1.24 (1.05-1.48)[e]
41 - 50	1.25 (1.05-1.50)[e]
51 - 60	1.38 (1.10-1.73)[e]
Education level	
Below undergraduate degree-	Ref
Undergraduate degree and above	0.85 (0.74-0.96)[c]
Current marital status	
Unmarried	Ref
Married	1.28 (1.09-1.50)
Employment status	
Not being employed	Ref
Full-time job	0.81 (0.70-0.94)
Part-time job	1.49 (1.04-2.12)[c]
Chronic disease status[f]	
None	Ref
Yes, extremely well controlled	1.08 (0.70-1.67)
Yes, well controlled	1.48 (1.21-1.81)[d]
Yes, not well controlled	3.71 (2.81-4.91)[d]

[a]ORu: univariate odds ratio.
[b]Ref: reference group.
[c]$P<.05$.
[d]$P<.001$.
[e]$P<.01$.
[f]Chronic disease includes any of hypertension, diabetes, chronic lung disease, myocardial infarction, heart failure, cerebrovascular disease, dementia, and ulcerative diseases such as gastric ulcer, liver disease, and tumor.

XSL•FO
RenderX

Personal Factors

The results of the adjusted logistic regression analyses are shown in Table 3. Regarding the personal factors, having severe symptoms that did not hospitalize (ORa=5.56, 95% CI 2.00-15.47; reference group: asymptomatic) and a condition that hospitalized (ORa=9.74, 95% CI 2.88-32.93; reference group: asymptomatic) were both positively associated with ASD. Regarding the cognitive factors, both higher perceived risk of reinfection (average vs low: ORa=1.68, 95% CI 1.40-2.02; high versus low: ORa=3.62, 95% CI 3.00-4.38) and lower perceived natural immunity level within 6 months since infection (weak vs strong: ORa=1.51, 95% CI 1.25-1.82) were positively associated with ASD. In terms of emotional factors, higher levels of worry about long-term physical harms of COVID-19 (ORa=1.47, 95% CI 1.42-1.52) and panic about risk of infection of older or younger family members (ORa=2.00, 95% CI 1.68-2.39) were positively associated with ASD.

XSL•FO
RenderX

Table . Logistic regression on personal/environmental factors and acute stress disorder among people ever infected with COVID-19 aged 18 - 60 y (N=5545). The models were adjusted for city, urban/rural area, sex, age group, education level, current marital status, employment status, and chronic disease status.

	ORa[a] (95% CI)
Personal factors	
Severity of COVID-19 infection	
Asymptomatic	Ref[b]
Mild	2.48 (0.89-6.90)
Severe and not hospitalized	5.56 (2.00-15.47)[c]
Severe and hospitalized	9.74 (2.88-32.93)[d]
Cognitive factors	
Perceived risk of reinfection	
Low	Ref
Average	1.68 (1.40-2.02)[d]
High	3.62 (3.00-4.38)[d]
Perceived natural immunity level within 6 months since COVID-19 infection	
Strong	Ref
Average	1.13 (0.95-1.35)
Weak	1.51 (1.25-1.82)[d]
Emotional factors	
Worry about long-term physical harms of COVID-19	1.47 (1.42-1.52)[d]
Panic about infection of older or younger family members	
No	Ref
Yes	2.00 (1.68-2.39)[d]
Contextual/environmental factors	
Information-related difficulty	1.42 (1.34-1.51)[d]
Difficulty in obtaining drugs/RAT[e] supplies	1.35 (1.28-1.44)[d]
Having unvaccinated family members aged ≥65 years	
No	Ref
Yes	1.48 (1.23-1.78)[d]
Having unvaccinated family members aged 3-11 years	
No	Ref
Yes	1.35 (1.04-1.75)[f]
Having relatives/friends with severe COVID-19 symptoms?	
None	Ref
Only a few	1.65 (1.29, 2.10)[d]
Some	3.99 (3.10, 5.13)[d]
Many	6.95 (5.13, 9.41)[d]

[a]ORa: adjusted odds ratio.

[b]Ref: reference group.

[c]$P<.01$.

[d]$P<.001$.

XSL•FO
RenderX

[e]RAT: rapid antigen test.
[f]$P<.05$.

Contextual or Environmental Factors

Significant contextual or environmental factors included higher levels of information-related difficulty (ORa=1.42, 95% CI 1.34-1.51), difficulty in obtaining drugs or RAT supplies (ORa=1.35, 95% CI 1.28-1.44), having unvaccinated family members aged greater than or equal to 65 years (ORa=1.48, 95% CI 1.23-1.78), having unvaccinated family member aged 3 - 11 years (ORa=1.35, 95% CI 1.04-1.75), and having relatives or friends having severe COVID-19 symptoms (only a few versus none: ORa=1.65, 95% CI 1.29-2.10; some versus none: ORa=3.99, 95% CI 3.10-5.13; many versus none: ORa=6.95, 95% CI 5.13-9.41).

Discussion

Substantially, about one-fifth of the sampled people ever infected with COVID-19 showed ASD, especially among those who were female, older, less educated, married, not in full-time employment, and having chronic disease conditions. Supporting the initial hypotheses, significant personal risk factors of ASD (severity of infection, perceived high reinfection risk, perceived low natural immunity level postinfection, worry about long-term physical harms of COVID-19, and panic about infection of older or younger family members) and contextual or environmental risk factors of ASD (difficulties in getting information and medical supplies, the unvaccinated status of older and younger family members, and having relatives or friends with severe COVID-19 symptoms) were identified.

The ASD prevalence in our study was lower than the rate of 71% reported in another Chinese study [31] and that of some other countries [28-30]. Such discrepancy could be attributed to variations in the stages of the COVID-19 pandemic as the other studies were conducted during the early phases. The early outbreak was characterized by fear of a novel and more severe virus, coupled with the experience of strict lockdowns—both of which are established population-level stressors [10,13,61]. In contrast, our study was conducted during the ending phase, which was dominated by the milder Omicron variant and followed the lifting of all restrictive measures. Given that disease severity is a key risk factor for ASD [14], the population-level shift toward less severe clinical manifestations, together with the removal of lockdown-related stressors, provides a parsimonious explanation for the observed difference in ASD prevalence.

Among the personal factors, the severity of COVID-19 infection was positively associated with ASD, consistent with previous research [14]. This relationship can be understood through several pathways. First, the direct physiological impact of a severe infection, including high fever, intense bodily pain, and dyspnea, can be perceived as a life-threatening physical trauma, which constitutes a core diagnostic criterion for stress-related disorders [62-65]. Second, the management of severe symptoms often necessitates intensive efforts, including seeking urgent medical care and securing necessary supplies, thereby amplifying situational stressors and potentially exacerbating the perception of threat [62,64,65]. Importantly, the markedly elevated ASD risk observed in the “severe and hospitalized” group likely extends beyond a simple gradient of symptom intensity. During the study period in China, the decision to hospitalize a patient was shaped by a complex interplay of clinical necessity, health care resource availability, and local pandemic policies [66,67]. Therefore, hospitalization itself constituted a distinct and potent psychosocial stressor, characterized by forced isolation from one’s daily milieu, reliance on an overburdened emergency health care system, anxieties related to treatment costs, and direct confrontation with the potential for severe health outcomes [68,69]. These factors—stemming from health care system interaction and institutionalization—may operate independently of, or synergistically with, the physiological distress of the infection to jointly exacerbate acute stress responses. Consequently, our findings highlight that individuals who experienced more severe COVID-19 illness, particularly those who required hospitalization, constitute a critical high-risk subgroup warranting prioritized screening and early psychological intervention to mitigate the potential progression to PTSD.

Corroborating previous studies among health care workers, a higher perceived risk of reinfection was positively associated with ASD [39,40]. This empirical relationship underscores the significant role of threat appraisal in the development of stress-related disorders. Within the framework of fear appeal theories, our results highlight the public health imperative of calibrating risk communication: excessively elevated perceptions of susceptibility may precipitate maladaptive emotional outcomes such as ASD [70,71], whereas unrealistically low perceptions can undermine protective behaviors [72]. In the context of an evolving pathogen like SARS-CoV-2, this underscores the critical public health imperative to disseminate timely and accurate information. For instance, during Omicron-predominant waves, public communication should concurrently reinforce the high transmissibility (to motivate booster vaccination) and the substantially reduced severity (to alleviate excessive fear and prevent ASD), thereby fostering both protection and psychological resilience. Conversely, a stronger perceived level of natural immunity 6 months post-infection served as a protective factor against ASD, likely mediated by a reduction in perceived susceptibility. However, the durability of infection-induced immunity is finite, waning over time [73,74], and public awareness regarding its extent and duration is often incomplete or inaccurate [75]. This knowledge gap underscores the necessity for clear communication that not only delineates the properties and limitations of natural immunity but also proactively promotes booster vaccination and the superior, more robust protection conferred by hybrid immunity [76].

We postulate that confidence in vaccine-induced immunity may buffer against the anxiety stemming from waning natural immunity. Given the demonstrated efficacy of social media in shaping pandemic-related perceptions [77], these platforms represent an indispensable channel for the prompt and accessible

XSL•FO
RenderX

dissemination of evolving epidemiological and clinical evidence to the public.

Persistent worry is a recognized precursor to anxiety [78] and ASD [79]. In line with this, our study identified a significant positive association between worry about the long-term physical harms of COVID-19 and ASD, a finding consistent with prior research [41,42]. The prevalence of such concerns is understandable given evidence on long COVID; a meta-analysis reported a pooled prevalence of 43% [80], encompassing symptoms such as fatigue, cognitive impairment, and cardiovascular sequelae [81,82]. As the majority of the Chinese population was infected during our survey period, long COVID emerged as a widespread public health concern, thereby elevating the population-level risk for ASD—a dynamic potentially applicable to other global populations. The persistence of these concerns may be partly attributable to limited public understanding of the long-term effects of COVID-19 [83]. Therefore, evidence-based information on the prevalence and severity of long COVID is warranted to mitigate excessive worry. Furthermore, our results indicate that panic about the risk of infection for vulnerable family members (the older adults and children) was independently associated with ASD, echoing its link to depression [19]. The increased odds of ASD among participants with unvaccinated older or younger family members substantiates this concern, as unvaccinated infected older adults face a significantly elevated risk of severe outcomes and mortality [84]. This underscores the continued importance of prompting vaccination in these vulnerable demographics. Collectively, these findings highlight the central role of emotional disturbance in ASD. Consequently, fostering adaptive emotional regulation should be a cornerstone of secondary prevention strategies for ASD. Maladaptive regulation strategies, such as rumination and catastrophizing, were robustly linked to mental distress during the pandemic [85]. Evidence-based brief interventions, including cognitive behavioral therapy [86] and mindfulness-based programs, which have demonstrated efficacy in ameliorating maladaptive emotional regulation [87], should be integrated into such preventive frameworks.

Regarding environmental factors, both difficulties in accessing timely and consistent COVID-19 information and in obtaining essential medical supplies were significantly associated with ASD. Information-related difficulties (eg, inconsistent messaging and information overload) may generate profound uncertainty and hinder effective coping, thereby elevating the risk of anxiety and ASD [53,88]. Furthermore, the significant association between difficulties in obtaining drug or RAT supplies and ASD underscores a critical and tangible dimension of the pandemic's psychological impact. Unlike informational uncertainty, the scarcity of essential medical supplies posed a direct threat to personal and family health, leaving individuals feeling vulnerable and powerless to manage their illness [89,90]. These findings collectively demonstrate that ensuring both informational clarity and physical access to key medical resources is critical for public mental health protection. Future epidemic responses must therefore prioritize robust supply chains and equitable distribution systems to mitigate these potential stressors.

Similarly, other studies have found that limited access to accurate COVID-19–related information was associated with mental health problems [43]. To our knowledge, the relationship between difficulty in accessing related information and ASD specifically has been underreported. The lack of clear information generates uncertainty or misconceptions, which can increase anxiety and hence the risk of ASD [91]. This risk was exacerbated during the unprecedented surge in infections, which created an urgent public demand for updated, authoritative information regarding transmission dynamics, disease severity, access to health care, and government policies [92-94]. As supported by previous findings, infected individuals with limited understanding of the situation can experience significant uncertainty and panic [93]. Therefore, ensuring public access to regular, accurate, and authoritative information from official sources is a critical public health strategy to mitigate ASD during crises. This also necessitates active efforts to combat the spread of misinformation on social media [95].

Our study further identified that having relatives or friends who experienced severe COVID-19 symptoms was positively associated with ASD among people ever infected with COVID-19. This aligns with existing evidence linking such exposure to mental distress [96]. This association may be explained by several mechanisms. First, emotional contagion [97] could play a role, whereby the psychological distress of close contacts adversely affects the participant's own mental state. This suggests that people ever infected with COVID-19 should be counseled on the potential mental health impact their illness disclosures may have on significant others. Second, providing emotional and instrumental support to infected significant others may impose a caregiver burden on participants [98], contributing to their distress. This burden can be particularly prolonged and severe when significant others develop long COVID, necessitating ongoing monitoring for ASD not only in the patients themselves but also in their caregivers [99]. Community-based programs that foster mutual aid networks could help alleviate this burden. In addition, integrating interventions designed to bolster psychological resources such as resilience and optimism—which are known to buffer against stress [100]—into support services for people ever infected with COVID-19 could represent an effective secondary prevention strategy.

As a universal public health strategy, it is recommended to raise awareness among stakeholders regarding the signs and consequences and preventive measures of ASD or PTSD, including the use of evidence-based brief positive psychology interventions [101]. The establishment of a comprehensive, one-stop digital platform for ASD screening and support—integrating self-assessment tools, online or printable self-help materials, emotional support hotlines, and facilitated referral pathways—could be highly beneficial. Individuals screening positive for ASD should be proactively monitored for the potential development of PTSD in the subsequent months. Furthermore, the risk factors identified in this study could be operationalized into a structured checklist to facilitate the provision of personalized interventions. The development of algorithm-driven digital tools to select specific brief interventions from a pre-established library, based on an

XSL•FO
RenderX

individual's unique risk profile, represents a promising avenue for scalable and efficient mental health support. Proactive development of such systems is crucial, as public health capacity is often strained during pandemics, underscoring the importance of preparedness during interpandemic periods.

The present study has interesting findings but also several limitations. First, our multistage, cluster-based recruitment across purposively selected cities, and voluntary, online participation means the sample is not strictly probability based; response metrics or weights were unavailable, so generalizability should be interpreted with caution, and WeChat-based recruitment may have introduced coverage and self-selection bias. Second, as this is a cross-sectional study, causal relationships cannot be inferred. Third, this study did not assess several important factors, including interpersonal factors (eg, perceived social support [52]) and baseline dispositional traits (eg, emotional stability and emotion-regulation skills [102,103]), which are known predictors of ASD. Additionally, potential confounders such as a history of previous COVID-19 infection were not measured. Fourth, the variable "having unvaccinated family members aged greater than or equal to 65" measured a composite household stressor that conflates 2 concerns. Future research would benefit from measuring these 2 factors—family age structure and vaccination status within the household—separately to elucidate their unique and interactive contributions to mental distress. At last, we used logistic regression for risk identification, which does not assess mediation or moderation. Structural equation modeling or related methods based on multi-item scales are needed to clarify pathways and interconnections in future work.

To summarize, this study found noticeable ASD prevalence among people ever infected with COVID-19 at the explosive ending phase of the COVID-19 pandemic in China and identified some significant personal and environmental factors. The findings are implicative. First, the risk of ASD and PTSD is noticeable. Preparations for universal prevention, screening, brief secondary intervention, stocking of key materials, information campaign, and follow-ups are greatly warranted to prevent ASD and PTSD. Second, health promotion may consider modifying the identified risk factors of ASD. Longitudinal studies are required in future pandemics to better understand the relationship between ASD and PTSD. It is also important to investigate the level of PTSD during the postpandemic stage. The findings may improve prevention of ASD and PTSD in future pandemics.

Acknowledgments

The authors sincerely appreciate participants for their contributions to this study.

Data Availability

The data presented in this study are available upon reasonable request from the corresponding author.

Authors' Contributions

JG contributed to this study as a co-corresponding author, providing supervision and critical revision of the manuscript, and can be reached via email at gujing5@mail.sysu.edu.cn.

Conflicts of Interest

None declared.

References

1. COVID-19 epidemiological update. World Heath Organization. URL: https://www.who.int/publications/m/item/covid-19-epidemiological-update---29-september-2023 [accessed 2023-09-29]
2. Chen S, Yang J, Yang W, Wang C, Bärnighausen T. COVID-19 control in China during mass population movements at new year. The Lancet 2020 Mar;395(10226):764-766. [doi: 10.1016/S0140-6736(20)30421-9]
3. Lin H, Guo C, Hu Y, et al. COVID-19 control strategies in Taizhou city, China. Bull World Health Organ 2020 Sep 1;98(9):632-637. [doi: 10.2471/BLT.20.255778] [Medline: 33012863]
4. Wang J, Liao Y, Wang X, et al. Incidence of novel coronavirus (2019-nCoV) infection among people under home quarantine in Shenzhen, China. Travel Med Infect Dis 2020;37:101660. [doi: 10.1016/j.tmaid.2020.101660] [Medline: 32247931]
5. COVID-19 situation in WHO - Western Pacific region. World Health Organization. 2023. URL: https://who.maps.arcgis.com/apps/dashboards/345dfdc82b5c4f6a815f1d54a05d18ec [accessed 2024-01-30]
6. COVID-19 response further optimized with 10 new measures. The State Council of the People's Republic of China. 2022. URL: https://english.www.gov.cn/statecouncil/ministries/202212/08/content_WS63913c92c6d0a757729e4135.html [accessed 2024-10-30]
7. Leung K, Lau EHY, Wong CKH, Leung GM, Wu JT. Estimating the transmission dynamics of SARS-CoV-2 Omicron BF.7 in Beijing after adjustment of the zero-COVID policy in November-December 2022. Nat Med 2023 Mar;29(3):579-582. [doi: 10.1038/s41591-023-02212-y] [Medline: 36638825]

XSL•FO
RenderX

8. Liang J, Liu R, He W, et al. Infection rates of 70% of the population observed within 3 weeks after release of COVID-19 restrictions in Macao, China. J Infect 2023 Apr;86(4):402-404. [doi: 10.1016/j.jinf.2023.01.029] [Medline: 36731635]
9. Huang J, Zhao S, Chong KC, et al. Infection rate in Guangzhou after easing the zero-COVID policy: seroprevalence results to ORF8 antigen. Lancet Infect Dis 2023 Apr;23(4):403-404. [doi: 10.1016/S1473-3099(23)00112-3] [Medline: 36803917]
10. Bueno-Notivol J, Gracia-García P, Olaya B, Lasheras I, López-Antón R, Santabárbara J. Prevalence of depression during the COVID-19 outbreak: a meta-analysis of community-based studies. Int J Clin Health Psychol 2021;21(1):100196. [doi: 10.1016/j.ijchp.2020.07.007] [Medline: 32904715]
11. Wu T, Jia X, Shi H, et al. Prevalence of mental health problems during the COVID-19 pandemic: a systematic review and meta-analysis. J Affect Disord 2021 Feb 15;281:91-98. [doi: 10.1016/j.jad.2020.11.117] [Medline: 33310451]
12. COVID-19 pandemic triggers 25% increase in prevalence of anxiety and depression worldwide. World Heath Organization. URL: https://www.who.int/news/item/02-03-2022-covid-19-pandemic-triggers-25-increase-in-prevalence-of-anxiety-and-depression-worldwide [accessed 2024-01-30]
13. Deng J, Zhou F, Hou W, et al. The prevalence of depression, anxiety, and sleep disturbances in COVID-19 patients: a meta-analysis. Ann N Y Acad Sci 2021 Feb;1486(1):90-111. [doi: 10.1111/nyas.14506] [Medline: 33009668]
14. Magnúsdóttir I, Lovik A, Unnarsdóttir AB, et al. Acute COVID-19 severity and mental health morbidity trajectories in patient populations of six nations: an observational study. Lancet Public Health 2022 May;7(5):e406-e416. [doi: 10.1016/S2468-2667(22)00042-1] [Medline: 35298894]
15. Meng ZQ, Leng Q, Wang X, Xu CH, Xu J, Liu Y. Comparative analysis of anxiety/depression scores in COVID-19 patients with disease severity, sleep patterns, and certain laboratory test results. Arch Med Sci 2022;18(5):1262-1270. [doi: 10.5114/aoms/136111] [Medline: 36160342]
16. Zahangir MS, Rokonuzzaman M. Depression, anxiety and stress among people infected with COVID-19 in Dhaka and Chittagong cities. Heliyon 2022 Sep;8(9):e10415. [doi: 10.1016/j.heliyon.2022.e10415] [Medline: 36060465]
17. Liu D, Epskamp S, Isvoranu AM, Chen C, Liu W, Hong X. Network analysis of physical and psychiatric symptoms of hospital discharged patients infected with COVID-19. J Affect Disord 2021 Nov 1;294:707-713. [doi: 10.1016/j.jad.2021.07.043] [Medline: 34343929]
18. Liu D, Baumeister RF, Veilleux JC, et al. Risk factors associated with mental illness in hospital discharged patients infected with COVID-19 in Wuhan, China. Psychiatry Res 2020 Oct;292:113297. [doi: 10.1016/j.psychres.2020.113297] [Medline: 32707218]
19. Yuan K, Zheng YB, Wang YJ, et al. A systematic review and meta-analysis on prevalence of and risk factors associated with depression, anxiety and insomnia in infectious diseases, including COVID-19: a call to action. Mol Psychiatry 2022 Aug;27(8):3214-3222. [doi: 10.1038/s41380-022-01638-z] [Medline: 35668158]
20. Islam MS, Ferdous MZ, Islam US, Mosaddek ASM, Potenza MN, Pardhan S. Treatment, persistent symptoms, and depression in people infected with COVID-19 in Bangladesh. Int J Environ Res Public Health 2021 Feb 5;18(4):1453. [doi: 10.3390/ijerph18041453] [Medline: 33562427]
21. Riggs DS, Foa EB. Posttraumatic disorders. In: Spielberger CD, editor. Encyclopedia of Applied Psychology: Elsevier; 2004:83-90. [doi: 10.1016/B0-12-657410-3/00158-6]
22. Chen S, Bi K, Lyu S, Sun P, Bonanno GA. Depression and PTSD in the aftermath of strict COVID-19 lockdowns: a cross-sectional and longitudinal network analysis. Eur J Psychotraumatol 2022;13(2):2115635. [doi: 10.1080/20008066.2022.2115635] [Medline: 36186164]
23. Stanley IH. Advancements in the understanding of PTSD and suicide risk: introduction to a special section. Psychol Trauma 2021 Oct;13(7):723-724. [doi: 10.1037/tra0001121] [Medline: 34723566]
24. Bryant RA, Harvey AG. Gender differences in the relationship between acute stress disorder and posttraumatic stress disorder following motor vehicle accidents. Aust N Z J Psychiatry 2003 Apr;37(2):226-229. [doi: 10.1046/j.1440-1614.2003.01130.x] [Medline: 12656964]
25. Shalev A, Liberzon I, Marmar C. Post-traumatic stress disorder. N Engl J Med 2017 Jun 22;376(25):2459-2469. [doi: 10.1056/NEJMra1612499] [Medline: 28636846]
26. Guez J, Cohen J, Naveh-Benjamin M, et al. Associative memory impairment in acute stress disorder: characteristics and time course. Psychiatry Res 2013 Oct 30;209(3):479-484. [doi: 10.1016/j.psychres.2012.12.013] [Medline: 23312478]
27. Cénat JM, Blais-Rochette C, Kokou-Kpolou CK, et al. Prevalence of symptoms of depression, anxiety, insomnia, posttraumatic stress disorder, and psychological distress among populations affected by the COVID-19 pandemic: a systematic review and meta-analysis. Psychiatry Res 2021 Jan;295:113599. [doi: 10.1016/j.psychres.2020.113599] [Medline: 33285346]
28. Dehelean L, Papava I, Musat MI, et al. Coping strategies and stress related disorders in patients with COVID-19. Brain Sci 2021 Sep 28;11(10):1287. [doi: 10.3390/brainsci11101287] [Medline: 34679351]
29. Parker C, Shalev D, Hsu I, et al. Depression, anxiety, and acute stress disorder among patients hospitalized with COVID-19: a prospective cohort study. J Acad Consult Liaison Psychiatry 2021;62(2):211-219. [doi: 10.1016/j.psym.2020.10.001] [Medline: 33198962]

XSL•FO
RenderX

30. Mongodi S, Salve G, Tavazzi G, et al. High prevalence of acute stress disorder and persisting symptoms in ICU survivors after COVID-19. Intensive Care Med 2021 May;47(5):616-618. [doi: 10.1007/s00134-021-06349-7] [Medline: 33730197]
31. Shi L, Lu ZA, Que JY, et al. Prevalence of and risk factors associated with mental health symptoms among the general population in China during the coronavirus disease 2019 pandemic. JAMA Netw Open 2020 Jul 1;3(7):e2014053. [doi: 10.1001/jamanetworkopen.2020.14053] [Medline: 32609353]
32. Torales J, O'Higgins M, Castaldelli-Maia JM, Ventriglio A. The outbreak of COVID-19 coronavirus and its impact on global mental health. Int J Soc Psychiatry 2020 Jun;66(4):317-320. [doi: 10.1177/0020764020915212] [Medline: 32233719]
33. Wang XL, Gao LY, Miu QF, et al. Perceived uncertainty stress and its predictors among residents in China during the COVID-19 pandemic. Psychol Health Med 2022 Jan;27(1):265-279. [doi: 10.1080/13548506.2021.1883692] [Medline: 33573426]
34. Barzilay R, Moore TM, Greenberg DM, et al. Resilience, COVID-19-related stress, anxiety and depression during the pandemic in a large population enriched for healthcare providers. Transl Psychiatry 2020 Aug 20;10(1):291. [doi: 10.1038/s41398-020-00982-4] [Medline: 32820171]
35. Ellis BJ, Sheridan MA, Belsky J, McLaughlin KA. Why and how does early adversity influence development? Toward an integrated model of dimensions of environmental experience. Dev Psychopathol 2022 May;34(2):447-471. [doi: 10.1017/S0954579421001838] [Medline: 35285791]
36. Jolicoeur-Martineau A, Belsky J, Szekely E, et al. Distinguishing differential susceptibility, diathesis-stress, and vantage sensitivity: beyond the single gene and environment model. Dev Psychopathol 2020 Feb;32(1):73-83. [doi: 10.1017/S0954579418001438] [Medline: 30626458]
37. Monroe SM, Simons AD. Diathesis-stress theories in the context of life stress research: implications for the depressive disorders. Psychol Bull 1991 Nov;110(3):406-425. [doi: 10.1037/0033-2909.110.3.406] [Medline: 1758917]
38. Sawyer CR. Developmental psychophysiology and the human stress response during communication. In: The Handbook of Communication Science and Biology: Routledge; 2020:332-343. [doi: 10.4324/9781351235587-28]
39. Londoño-Ramírez AC, García-Pla S, Bernabeu-Juan P, Pérez-Martínez E, Rodríguez-Marín J, van-der Hofstadt-Román CJ. Impact of COVID-19 on the anxiety perceived by healthcare professionals: differences between primary care and hospital care. Int J Environ Res Public Health 2021 Mar 22;18(6):3277. [doi: 10.3390/ijerph18063277] [Medline: 33810004]
40. Silva-Costa A, Griep RH, Rotenberg L. Perceived risk from COVID-19 and depression, anxiety, and stress among workers in healthcare units. Cad Saude Publica 2022;38(3):e00198321. [doi: 10.1590/0102-311X00198321] [Medline: 35293539]
41. Li Y, Zhao J, Ma Z, et al. Mental health among college students during the COVID-19 pandemic in China: a 2-wave longitudinal survey. J Affect Disord 2021 Feb;281:597-604. [doi: 10.1016/j.jad.2020.11.109]
42. Sharpe D, Rajabi M, Chileshe C, et al. Mental health and wellbeing implications of the COVID-19 quarantine for disabled and disadvantaged children and young people: evidence from a cross-cultural study in Zambia and Sierra Leone. BMC Psychol 2021 May 15;9(1):79. [doi: 10.1186/s40359-021-00583-w] [Medline: 33992113]
43. Bassey EE, Gupta A, Kapoor A, Bansal A. COVID-19 and poverty in South America: the mental health implications. Int J Ment Health Addict 2022 Jan 24;21:1-7. [doi: 10.1007/s11469-022-00765-6] [Medline: 35095354]
44. Vogel EA, Chieng A, Robinson A, Pajarito S, Prochaska JJ. Associations between substance use problems and stress during COVID-19. J Stud Alcohol Drugs 2021 Nov;82(6):776-781. [doi: 10.15288/jsad.2021.82.776] [Medline: 34762037]
45. Chen H, Gao J, Dai J, et al. Generalized anxiety disorder and resilience during the COVID-19 pandemic: evidence from China during the early rapid outbreak. BMC Public Health 2021 Oct 9;21(1):1830. [doi: 10.1186/s12889-021-11877-4] [Medline: 34627208]
46. Mistry SK, Ali ARMM, Irfan NM, et al. Prevalence and correlates of depressive symptoms among Rohingya (forcibly displaced Myanmar nationals or FDMNs) older adults in Bangladesh amid the COVID-19 pandemic. Glob Ment Health 2021;8:e23. [doi: 10.1017/gmh.2021.24]
47. De PK, Sun R. Impacts of COVID-19 on mental health in the US: evidence from a national survey. J Ment Health 2023 Oct;32(5):910-919. [doi: 10.1080/09638237.2023.2210651] [Medline: 37194622]
48. Imran N, Aamer I, Afzal H, Hashmi A, Shabbir B, Asif A. Psychiatric impact on COVID-19 patients isolated in a tertiary care hospital in Pakistan. East Mediterr Health J 2022 Jan 31;28(1):5-13. [doi: 10.26719/emhj.21.062] [Medline: 35165873]
49. Gil D, Tiscar C, Gómez M, et al. Health-related quality of life and stress-related disorders in COVID-19 ICU survivors: are they worse than with other causes of ARDS? J Intensive Med 2022 Apr;2(2):103-109. [doi: 10.1016/j.jointm.2022.02.002] [Medline: 36785778]
50. Shankar S, Rahamathulla MMI. Incidence of psychiatric illness among COVID-19-positive individuals with and without loss of smell or taste symptoms in a tertiary care hospital in South India: a prospective cohort study. Annals of Indian Psychiatry 2022;6(2):169-175. [doi: 10.4103/aip.aip_39_22]
51. Zhang L, Qi H, Wang L, et al. Effects of the COVID‐19 pandemic on acute stress disorder and career planning among healthcare students. Int J Mental Health Nurs 2021 Aug;30(4):907-916. [doi: 10.1111/inm.12839]
52. Liu S, Xu R, Liu L. Influencing factors of acute stress disorder among frontline nurses in Wuhan, China. J Psychosoc Nurs Ment Health Serv 2021 Sep;59(9):38-47. [doi: 10.3928/02793695-20210324-03] [Medline: 34110951]

XSL•FO
RenderX

53. Lin D, Friedman DB, Qiao S, Tam CC, Li X, Li X. Information uncertainty: a correlate for acute stress disorder during the COVID-19 outbreak in China. BMC Public Health 2020 Dec 7;20(1):1867. [doi: 10.1186/s12889-020-09952-3] [Medline: 33287780]
54. Luo Y, He X, Wang S, Li J, Zhang Y. Media exposure predicts acute stress and probable acute stress disorder during the early COVID-19 outbreak in China. PeerJ 2021;9:e11407. [doi: 10.7717/peerj.11407] [Medline: 34026362]
55. Lin W. On interactive teaching model of translation course based on Wechat. Engl Lang Teach 2017;10(3):21. [doi: 10.5539/elt.v10n3p21]
56. Shan Y, Ji M, Dong Z, Xing Z, Xu X. Assessing Patients' critical health literacy and identifying associated factors: cross-sectional study. J Med Internet Res 2023 Apr 5;25:e43342. [doi: 10.2196/43342] [Medline: 37018027]
57. Zhang X, Lu H, Xu Y, et al. Multi-dimensional factors of behavioral intention of COVID-19 booster vaccination among people having contracted COVID-19: a 7-city study conducted during the last national major outbreak in China. BMC Public Health 2025;25(1):4391. [doi: 10.1186/s12889-025-23915-6]
58. Prins A, Bovin MJ, Smolenski DJ, et al. The primary care PTSD screen for DSM-5 (PC-PTSD-5): development and evaluation within a veteran primary care sample. J Gen Intern Med 2016 Oct;31(10):1206-1211. [doi: 10.1007/s11606-016-3703-5] [Medline: 27170304]
59. Cheng P, Jasinski N, Zheng W, et al. Psychometric properties of the primary care PTSD screen for DSM-5: findings from family members of Chinese healthcare workers during the outbreak of COVID-19. Front Psychiatry 2021;12:695678. [doi: 10.3389/fpsyt.2021.695678] [Medline: 34594246]
60. The R project for statistical computing. : R Foundation for Statistical Computing; 2023 URL: https://www.r-project.org [accessed 2026-02-28]
61. Friis NU, Martin-Bertelsen T, Pedersen RK, Nielsen J, Krause TG, Andreasen V, et al. COVID-19 mortality attenuated during widespread Omicron transmission. Euro Surveill 2020;28:2200547. [doi: 10.2807/1560-7917.ES.2023.28.3.2200547]
62. Brooks SK, Webster RK, Smith LE, et al. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. The Lancet 2020 Mar;395(10227):912-920. [doi: 10.1016/S0140-6736(20)30460-8]
63. Chung MC. Posttraumatic Stress in Physical Illness: Oxford University Press; 2024. [doi: 10.1093/oso/9780198727323.001.0001]
64. Kassir G, El Hayek S, Zalzale H, Orsolini L, Bizri M. Psychological distress experienced by self-quarantined undergraduate university students in Lebanon during the COVID-19 outbreak. Int J Psychiatry Clin Pract 2021 Jun;25(2):172-179. [doi: 10.1080/13651501.2021.1900872] [Medline: 33775208]
65. Wang Y, Shi L, Que J, et al. The impact of quarantine on mental health status among general population in China during the COVID-19 pandemic. Mol Psychiatry 2021 Sep;26(9):4813-4822. [doi: 10.1038/s41380-021-01019-y]
66. Isasi F, Naylor MD, Skorton D, Grabowski DC, Hernández S, Rice VM. Patients, families, and communities COVID-19 impact assessment: lessons learned and compelling needs. NAM Perspect 2021;2021:10-31478. [doi: 10.31478/202111c] [Medline: 35118349]
67. Ko JY, Danielson ML, Town M, et al. Risk factors for COVID-19-associated hospitalization: COVID-19-associated hospitalization surveillance network and behavioral risk factor surveillance system. Clin Infect Dis 2020:ciaa1419. [doi: 10.1093/cid/ciaa1419]
68. De Berardis G, D'Ettorre A, Graziano G, et al. The burden of hospitalization related to diabetes mellitus: a population-based study. Nutr Metab Cardiovasc Dis 2012 Jul;22(7):605-612. [doi: 10.1016/j.numecd.2010.10.016] [Medline: 21333508]
69. Di Fusco M, Shea KM, Lin J, et al. Health outcomes and economic burden of hospitalized COVID-19 patients in the United States. J Med Econ 2021;24(1):308-317. [doi: 10.1080/13696998.2021.1886109] [Medline: 33555956]
70. Dillard JP. Rethinkin the study of fear appeals: an emotional perspective. Commun Theory 1994 Nov;4(4):295-323. [doi: 10.1111/j.1468-2885.1994.tb00094.x]
71. Mello S, Glowacki E, Fuentes I, Seabolt J. Communicating COVID-19 risk on Instagram: a content analysis of official public health messaging during the first year of the pandemic. J Health Commun 2023 Jan 2;28(1):38-52. [doi: 10.1080/10810730.2023.2175278] [Medline: 36760119]
72. Brewer NT, Weinstein ND, Cuite CL, Herrington JE. Risk perceptions and their relation to risk behavior. Ann Behav Med 2004 Apr;27(2):125-130. [doi: 10.1207/s15324796abm2702_7] [Medline: 15026296]
73. Stein C, Nassereldine H, Sorensen RJD, et al. Past SARS-CoV-2 infection protection against re-infection: a systematic review and meta-analysis. The Lancet 2023 Mar;401(10379):833-842. [doi: 10.1016/S0140-6736(22)02465-5]
74. Goldberg Y, Mandel M, Bar-On YM, et al. Protection and waning of natural and hybrid immunity to SARS-CoV-2. N Engl J Med 2022 Jun 9;386(23):2201-2212. [doi: 10.1056/NEJMoa2118946] [Medline: 35613036]
75. Song N, Chen K, Peng J, Zhao Y, Wang J. Misleading and correcting: characteristics of online misinformation and refutations in China during the COVID-19 pandemic. Aslib J Inf Manag 2026 Jan 8;78(1):214-236. [doi: 10.1108/AJIM-08-2023-0287]
76. Nordström P, Ballin M, Nordström A. Risk of SARS-CoV-2 reinfection and COVID-19 hospitalisation in individuals with natural and hybrid immunity: a retrospective, total population cohort study in Sweden. Lancet Infect Dis 2022 Jun;22(6):781-790. [doi: 10.1016/S1473-3099(22)00143-8] [Medline: 35366962]

XSL•FO
RenderX

77. Tsoy D, Tirasawasdichai T, Ivanovich Kurpayanidi K. Role of social media in shaping public risk perception during COVID-19 pandemic: a theoretical review. Int J Manag Sci Bus Adm 2021;7(2):35-41. [doi: 10.18775/ijmsba.1849-5664-5419.2014.72.1005]
78. Worry and anxiety. Centers for Disease Control and Prevention. 2023. URL: https://www.cdc.gov/howrightnow/emotion/worry/index.html [accessed 2024-01-30]
79. Spinhoven P, Penninx BW, Krempeniou A, van Hemert AM, Elzinga B. Trait rumination predicts onset of post-traumatic stress disorder through trauma-related cognitive appraisals: a 4-year longitudinal study. Behav Res Ther 2015 Aug;71:101-109. [doi: 10.1016/j.brat.2015.06.004] [Medline: 26093467]
80. Chen C, Haupert SR, Zimmermann L, Shi X, Fritsche LG, Mukherjee B. Global prevalence of post-coronavirus disease 2019 (COVID-19) condition or long COVID: a meta-analysis and systematic review. J Infect Dis 2022 Nov 1;226(9):1593-1607. [doi: 10.1093/infdis/jiac136] [Medline: 35429399]
81. Karakayalı M, Artac I, Ilis D, et al. Evaluation of outpatients in the post-COVID-19 period in terms of autonomic dysfunction and silent ischemia. Cureus 2023 Jun;15(6):e40256. [doi: 10.7759/cureus.40256] [Medline: 37440812]
82. Karakayalı M, Artac I, Ilis D, et al. Evaluation of symptomatic and asymptomatic outpatients in the post-COVID-19 period with electrocardiographic ventricular depolarization and repolarization parameters. Int J Cardiovasc Sci 2024 Feb 29;37:e20230105. [doi: 10.36660/ijcs.20230105]
83. Yang C, Zhao H, Espín E, Tebbutt SJ. Association of SARS-CoV-2 infection and persistence with long COVID. Lancet Respir Med 2023 Jun;11(6):504-506. [doi: 10.1016/S2213-2600(23)00142-X] [Medline: 37178694]
84. Sanecka M, Youssef M, Abdulsalam M, et al. Hospital outcomes in patients hospitalized for COVID-19 pneumonia: the effect of SARS-CoV-2 vaccination and vitamin D status. Nutrients 2023 Jun 30;15(13):2976. [doi: 10.3390/nu15132976] [Medline: 37447302]
85. Yang X, Xu Y, Tan R, Zhou X. Event centrality and post-traumatic stress symptoms among college students during the COVID-19 pandemic: the roles of attention to negative information, catastrophizing, and rumination. Eur J Psychotraumatol 2022;13(1):2078563. [doi: 10.1080/20008198.2022.2078563] [Medline: 35695844]
86. Goldin PR, Ziv M, Jazaieri H, Weeks J, Heimberg RG, Gross JJ. Impact of cognitive-behavioral therapy for social anxiety disorder on the neural bases of emotional reactivity to and regulation of social evaluation. Behav Res Ther 2014 Nov;62:97-106. [doi: 10.1016/j.brat.2014.08.005] [Medline: 25193002]
87. Metz SM, Frank JL, Reibel D, Cantrell T, Sanders R, Broderick PC. The effectiveness of the learning to BREATHE program on adolescent emotion regulation. Res Hum Dev 2013 Jul;10(3):252-272. [doi: 10.1080/15427609.2013.818488]
88. Millroth P, Frey R. Fear and anxiety in the face of COVID-19: negative dispositions towards risk and uncertainty as vulnerability factors. J Anxiety Disord 2021 Oct;83:102454. [doi: 10.1016/j.janxdis.2021.102454] [Medline: 34298237]
89. Pfefferbaum B, North CS. Mental health and the COVID-19 pandemic. N Engl J Med 2020 Aug 6;383(6):510-512. [doi: 10.1056/NEJMp2008017] [Medline: 32283003]
90. Tomko C, Sisson LN, Haney K, et al. Multiple material need insecurities and severity of psychological distress during the COVID-19 pandemic among women who use drugs. BMC Public Health 2024 Dec 19;24(1):3541. [doi: 10.1186/s12889-024-21104-5] [Medline: 39702174]
91. Sarry SM, Anggreiny N. Responding to uncertainty caused by COVID-19. IOP Conf Ser: Earth Environ Sci 2021 Apr 1;708(1):012095. [doi: 10.1088/1755-1315/708/1/012095]
92. Kim H, Oh S. Everyday life information seeking in South Korea during the COVID-19 pandemic: daily topics of information needs in social Q&A. OIR 2023 Mar 30;47(2):414-430. [doi: 10.1108/OIR-10-2021-0547]
93. Juanamasta IG, Aungsuroch Y, Gunawan J. Experiences of patients with coronavirus in the COVID-19 pandemic era in Indonesia. Asian J Public Opin Res 2020;8:377-392. [doi: 10.15206/ajpor.2020.8.3.377]
94. Lu H, Xie J, Cheng Y, Chen Y. Modeling the formation of the public's information needs in the early stage of the COVID-19 outbreak. Aslib J Inf Manag 2023 Aug 30;75(5):961-984. [doi: 10.1108/AJIM-04-2022-0167]
95. Al-Zaman MS. Prevalence and source analysis of COVID-19 misinformation in 138 countries. IFLA J 2022 Mar;48(1):189-204. [doi: 10.1177/03400352211041135]
96. Koçak O, Koçak Ö, Younis MZ. The psychological consequences of COVID-19 fear and the moderator effects of individuals' underlying illness and witnessing infected friends and family. Int J Environ Res Public Health 2021 Feb 13;18(4):1836. [doi: 10.3390/ijerph18041836] [Medline: 33668655]
97. Zillmann D. Theory of affective dynamics: emotions and moods. In: Communication and Emotion: Essays in Honor of Dolf Zillmann: Routledge; 2003:533-567.
98. Russell D, Spence NJ, Chase JAD, Schwartz T, Tumminello CM, Bouldin E. Support amid uncertainty: long COVID illness experiences and the role of online communities. SSM Qual Res Health 2022 Dec;2:100177. [doi: 10.1016/j.ssmqr.2022.100177] [Medline: 36212783]
99. Baz SA, Fang C, Carpentieri JD, Sheard L. "I don't know what to do or where to go". Experiences of accessing healthcare support from the perspectives of people living with Long COVID and healthcare professionals: a qualitative study in Bradford, UK. Health Expect 2023 Feb;26(1):542-554. [doi: 10.1111/hex.13687] [Medline: 36512382]
100. Waters L, Algoe SB, Dutton J, et al. Positive psychology in a pandemic: buffering, bolstering, and building mental health. J Posit Psychol 2022 May 4;17(3):303-323. [doi: 10.1080/17439760.2021.1871945]

XSL•FO
RenderX

101. Ivandic I, Freeman A, Birner U, Nowak D, Sabariego C. A systematic review of brief mental health and well-being interventions in organizational settings. Scand J Work Environ Health 2017 Mar 1;43(2):99-108. [doi: 10.5271/sjweh.3616] [Medline: 28042963]
102. Shangguan C, Wang Y, Wang X, Zhang L, Zhou B. The more anxious, the less flexible? Association of trait anxiety with expressive flexibility and the mediating role of emotion regulation strategies. Curr Psychol 2024 Aug;43(32):26121-26133. [doi: 10.1007/s12144-024-06291-z]
103. Tyra AT, Griffin SM, Fergus TA, Ginty AT. Individual differences in emotion regulation prospectively predict early COVID-19 related acute stress. J Anxiety Disord 2021 Jun;81:102411. [doi: 10.1016/j.janxdis.2021.102411] [Medline: 33962141]

Abbreviations

ASD: acute stress disorder
ICU: intensive care unit
ORa: adjusted odds ratio
PC-PTSD-5: Primary Care PTSD Screen for *Diagnostic and Statistical Manual of Mental Disorders, Fifth Revision*
PTSD: posttraumatic stress disorder
RAT: rapid antigen test

Edited by B Biswas; submitted 23.Feb.2025; peer-reviewed by X Chen, YQ Gan; revised version received 17.Dec.2025; accepted 29.Dec.2025; published 11.Mar.2026.

Please cite as:
Yang Z, Yu Y, Lu H, Wang X, Xu Y, Ying J, Wen X, Luo L, Wang M, Liu M, Geng X, Zhao X, He B, Liu T, Maimaitijiang R, Gu J, Lau JTF
Prevalence and Factors Associated With Acute Stress Disorder Among Adults Ever Infected With COVID-19 During the Ending Phase of the Pandemic in 7 Chinese Cities: Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e73002
URL: https://publichealth.jmir.org/2026/1/e73002
doi:10.2196/73002

© Ziying Yang, Yanqiu Yu, Hui Lu, Xu Wang, Yong Xu, Junqiang Ying, Xianying Wen, Lei Luo, Meng Wang, Muwen Liu, Xingyi Geng, Xuchong Zhao, Biyu He, Tao Liu, Remina Maimaitijiang, Jing Gu, Joseph T F Lau. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 11.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Regional Disparities in Factors Associated With Subjective Health Among Older Adults in Aging and Super-Aged Areas of Korea: Nationwide Cross-Sectional Study

Gi-Moon Nam[1], BS; Ye-Rin Lee[1], BS; Ji-Hye Mun[2], MA; Su-Hyeun Cho[2], MS; Hee-Sook Lim[1], PhD

[1]AgeTech-Service Convergence Major, Department of Gerontology, Graduate School of East-West Medical Science, Kyung Hee University, Deogyeong-daero 1732, Giheung-gu, Yongin-si, Republic of Korea

[2]Dietary and Nutritional Safety Policy Division, Food and Consumer Safety Bureau, Ministry of Food and Drug Safety, Cheongju, Republic of Korea

Corresponding Author:
Hee-Sook Lim, PhD
AgeTech-Service Convergence Major, Department of Gerontology, Graduate School of East-West Medical Science, Kyung Hee University, Deogyeong-daero 1732, Giheung-gu, Yongin-si, Republic of Korea

Abstract

Background: As South Korea transitions into a super-aged society, understanding regional disparities in subjective health among older adults is critical to addressing health inequalities and supporting healthy aging.

Objective: This study aimed to compare determinants of subjective health between aging and super-aged areas in South Korea and identify region-specific characteristics contributing to disparities among older adults.

Methods: A cross-sectional analysis was conducted using data from the Korea Community Health Survey (2020 - 2023), a nationwide population-based survey at the city, county, and district levels. Adults aged 65 years and older (n=179,571) were categorized into aging (n=19,759) or super-aged (n=159,782) areas based on regional aging rates. Propensity score matching was applied to adjust for demographic differences, yielding 18,574 matched participants in each group. Subjective health was assessed using a 5-point Likert scale. Ordinal logistic regression was used to examine associations between subjective health and various exposures, including demographic characteristics, health behaviors, physical and mental health status, and health literacy indicators such as nutrition label recognition and reading.

Results: Older adults in super-aged areas reported poorer subjective health than those in aging regions. Physical activity and mental health were consistently associated with better subjective health in both regions, region-specific patterns were observed. In aging regions, nutrition label recognition was significantly associated with better subjective health, whereas in super-aged areas, nutrition label reading showed a stronger association. The negative impact of hypertension and diabetes on subjective health was more pronounced in super-aged areas.

Conclusions: Although key determinants of subjective health were similar across regions, regional differences underscore the importance of tailored public health strategies. Interventions that strengthen health literacy and provide nutrition education focused on disease-related nutrients may help mitigate disparities and enhance subjective health among older adults in aging and super-aged areas.

(JMIR Public Health Surveill 2026;12:e80189) doi:10.2196/80189

KEYWORDS

subjective health status; health literacy; older adult; chronic disease; health disparities

Introduction

South Korea's older population has increased by approximately 4.4% per year in recent years, reflecting one of the fastest aging trends among Organization for Economic Co-operation and Development (OECD) countries [1]. In December 2024, the proportion of individuals aged 65 and older exceeded 20%, and South Korea was classified as a super-aged society [2]. This transition occurred just 7 years after entering an aged society, a steeper trajectory than Japan's 12-year process [3], and highlights the need for rapid societal adaptation.

Among older adults, the focus has shifted from simply extending life expectancy to increasing healthy life expectancy, meaning living longer in good health [4]. Reducing the gap between these 2 indicators by minimizing illness periods has become a key public health issue. In 2019, South Korea's life expectancy was 83.3 years, and healthy life expectancy was 73.1 years, indicating a 10.2-year gap, exceeding the OECD average of 9.6 years [5]. A 2-year difference between the top 20% and bottom 20% of regions within the country further underscores the need to reduce regional disparities [6].

XSL•FO
RenderX

The health status of older adults, a key determinant of healthy life expectancy, is influenced by a range of factors, including socioeconomic and behavioral factors [7]. Objective health is typically defined as measurable clinical indicators, such as clinical diagnoses or physiological functioning, whereas subjective health refers to individuals' self-assessment of their overall health status. Because objective health is difficult to measure directly, subjective health is often used as a proxy indicator [8]. South Korea scored 52.4 in subjective health, ranking 28th among 38 OECD countries and below the OECD average of 67.6 [9]. Similar to objective health, subjective health is associated with sociodemographic characteristics, physical and mental health, and health behaviors. Individuals who live with others, participate in economic activity, earn higher incomes, and have higher education levels tend to report better subjective health [10-13]. Smoking, low physical activity, depression, and chronic disease diagnoses have also been associated with poorer subjective health [14-16].

Environmental factors associated with older adults' health include regional attributes such as access to medical services, economic conditions, and housing, as well as social environments including neighborhood relationships and health-supportive community settings [17-19]. These factors may be related to or shape health behaviors within everyday living contexts. However, most studies rely on broad regional classifications, such as metropolitan versus nonmetropolitan or urban versus rural, which limit the analysis of local variation.

Using data from the Korea Community Health Survey (KCHS), which produces community-level health statistics, this study aimed to analyze factors influencing the subjective health of older adults by comparing aging and super-aged areas and considering relevant environmental determinants.

Methods

Study Design and Setting

This study employed a cross-sectional design using nationally representative data from the KCHS, conducted annually by the Korea Disease Control and Prevention Agency (KDCA). The KCHS is the largest nationwide health survey in Korea, including approximately 230,000 participants each year and representing about 0.5% of the country's adult population. It provides regionally disaggregated data that enable community-level health analyses. The sampling frame was established by linking resident registration data from the Ministry of the Interior and Safety with housing data from the Ministry of Land, Infrastructure and Transport to construct a complete list of households.

The survey used a multistage, stratified probability sampling method. In the first stage, primary sampling units (PSUs) were selected based on administrative districts (tong/ban/ri) with probabilities proportional to the number of households by housing type. In the second stage, households were systematically sampled within each PSU. The overall sample size was predetermined by the national survey design, and approximately 10 households per PSU in urban areas and 8 households per PSU in rural areas were selected to account for regional variations in health indicators and sampling adequacy.

Trained interviewers visited selected households and collected data through computer-assisted personal interviewing using tablet-based electronic questionnaires between August and October each year.

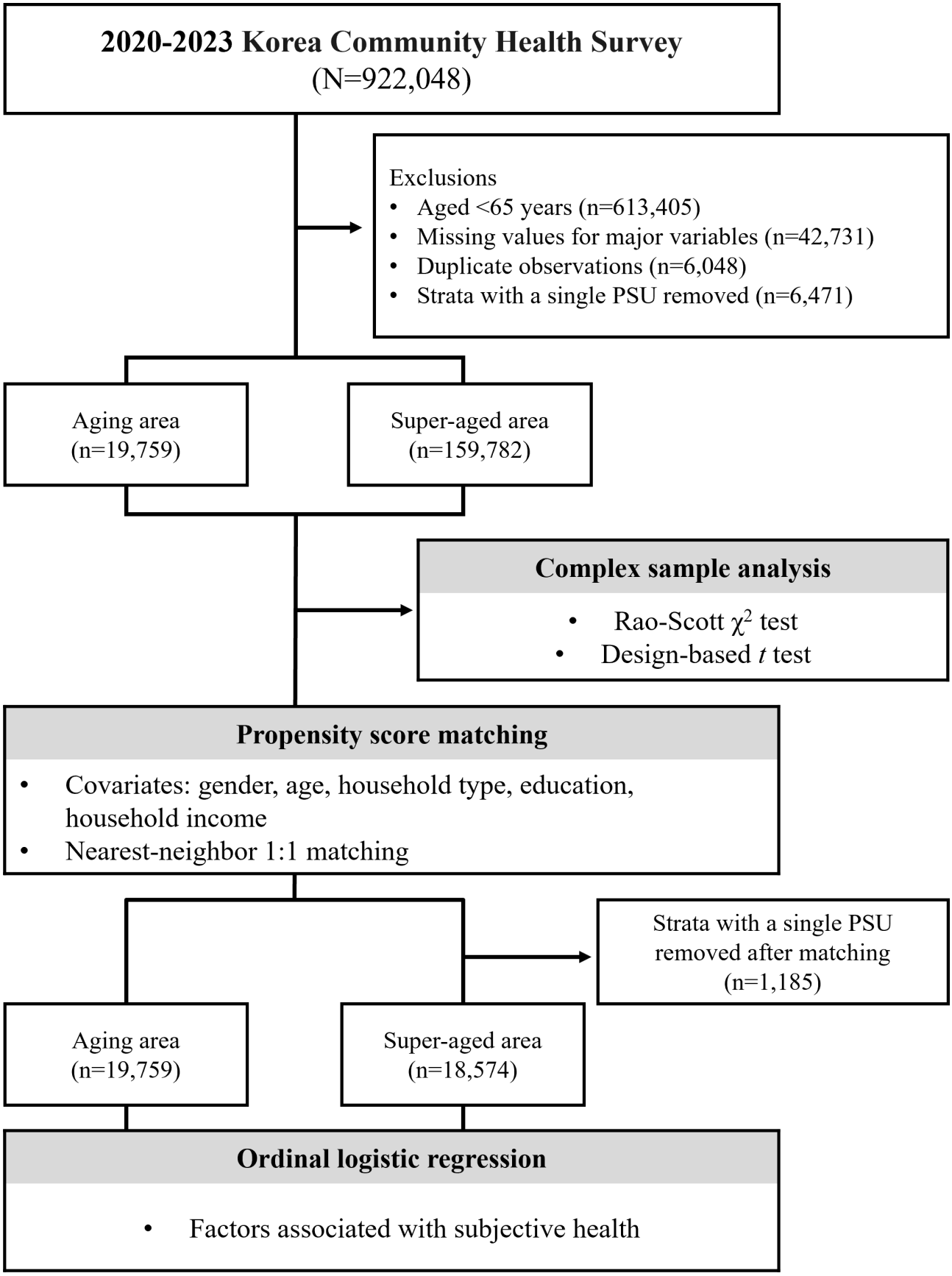
Since 2010, the KCHS has been organized into 4-year survey cycles, with items categorized according to importance, applicability, and sampling size. Accordingly, certain items are assessed every 1, 2, or 4 years. The total number of questionnaire items ranged from approximately 140 to 160 per year (142 in 2020, 163 in 2021, 138 in 2022, and 145 in 2023), with minor revisions made annually to reflect changes in public health priorities and regional conditions.

Subjects and Data Collection

This study used merged KCHS data from 2020 to 2023, including 922,048 participants, to examine factors influencing subjective health by regional aging levels. Exclusions included 613,405 individuals aged 65 years or younger, 42,731 with missing or outlier values, 6048 with duplicated identifiers, and 6471 from strata with only 1 PSU, resulting in a final sample of 179,571. Regions were classified by linking each respondent's location to aging rate data from Statistics Korea at the Si-Gun-Gu (city-county-district) level. Areas with aging rates of 14% or lower were defined as "Aging area" (n=19,759), and those with aging rates of 20% or higher were defined as "Super-aged area" (n=159,782), based on commonly used demographic thresholds that have been applied in national-level classifications as well as regional analyses in Korea [20]. The study design and analytical procedures are summarized in Figure 1.

XSL•FO
RenderX

Figure 1. Conceptual framework of the study. Under "Exclusions," major variables included demographic characteristics, health behavior characteristics, and health literacy variables (see Tables 1 and 2). Aging area has an aging rate ≤14%, while super-aged area has an aging rate ≥20%. PSU: primary sampling unit.



XSL•FO
RenderX

Research Variables

Subjective health was assessed using a 5-point Likert scale ranging from "Very bad" to "Very good" (1=Very bad, 2=Bad, 3=Fair, 4=Good, 5=Very good). Respondents were asked the question, "In general, how would you rate your health?." To identify the factors influencing subjective health, items from the KCHS related to demographic characteristics, health behavior characteristics, health status characteristics, and health literacy were used.

Demographic Characteristics

Demographic variables included gender, age (years), household type, marital status, education, economic activity, and household income (monthly). Specifically, household type was categorized into "living alone," "with a spouse or other (single-generation household)," "two generations (couple with unmarried children, single parent with unmarried children, or other 2-generation households)," and "three generations." Marital status was divided into "yes, cohabitation," "separated or divorced," "bereavement," and "no." Education was classified into "below elementary (including no formal education or traditional education)," "elementary," "middle school," "high school," and "college or higher (2-y or 3-y college, 4-y university, graduate school or higher)." Monthly household income was reported as a continuous variable.

Health Behavior Characteristics

Health behavioral characteristics included smoking (yes), drinking frequency, active physical activity (weekly, min), moderate physical activity (weekly, min), walking (weekly, min), subjective body shape, and weight control experience. Smoking was categorized into "yes" and "nonsmoker (smoked fewer than 5 packs or never smoked)." Drinking frequency was classified into "almost never (no drinking experience or none within the past year)," "less than once a month," "1 - 4 times a month (once or 2 - 4 times per month)," and "more than 2 times a week (2 - 3 times or more than 4 times per week)." Physical activity (active, moderate, and walking) was calculated as a continuous variable by multiplying the average daily minutes spent on each activity by the number of days it was performed per week, resulting in total weekly minutes.

Health Status Characteristics

Health-related characteristics included subjective stress, depression, hypertension, diabetes, subjective health, chewing discomfort, and subjective oral health. Specifically, chewing discomfort was classified into "uncomfortable (very uncomfortable, uncomfortable)," "neutral," and "comfortable (slightly uncomfortable, not uncomfortable at all)."

Health Literacy

Health literacy variables included nutrition label recognition, nutrition label reading, nutrition label utilization, unmet health care needs, and written health literacy. Nutrition label items were structured hierarchically to represent sequential levels of label literacy: recognition (awareness of labels on food packaging), reading (understanding the information displayed), and utilization (using label information in dietary decisions). Unmet health care needs were categorized into 2 groups: "yes" (excluding cases indicating no need for care) and "no."

Statistical Analysis

Statistical analyses were performed using RStudio (version 4.3.3; Posit, PBC) and IBM SPSS Statistics (version 25.0; IBM Corp.). A complex sample survey analysis method was applied considering the individual weights, strata, and cluster variables provided by the KCHS. These weights were calculated as the inverse of the selection probability and adjusted for response rates and population structure to represent the Korean adult population. Categorical variables were summarized as frequencies and percentages, and group differences were tested using the Rao-Scott chi-square test. Continuous variables were presented as mean (SD) and compared using design-based *t* tests. Statistical significance was defined as $P<.05$.

Propensity scores were calculated using 5 variables: gender, age, household type, education, and household income (monthly). Based on these propensity scores, data from aging and super-aged areas were matched using a 1:1 nearest-neighbor matching method. The covariate balance between the matched groups was confirmed by comparing standardized mean differences before and after matching. Postmatching weights were calculated using poststratification adjustment based on original KCHS sampling weights and applied in subsequent analyses [21].

Ordinal logistic regression analysis was conducted using subjective health as the dependent variable. The backward elimination method of stepwise selection was applied, starting with a full model containing all predictors and sequentially removing variables with the highest nonsignificant *P* values. The threshold for variable removal was set at $P>.10$. After each removal, the model was refitted and the *P* values were recalculated. This process was repeated until all the remaining variables reached statistical significance.

To identify significant regional differences in determinants of subjective health, subgroup analyses were performed separately for aging and super-aged areas using the same model structure. Odds ratios (OR) and 95% CI were estimated for each group. Subsequently, a *z* test was performed to compare the OR and SDs between the regions. The significance level used to identify trends in regional differences was set at 90% ($P<.10$).

Ethical Considerations

This study was conducted using secondary data from KCHS provided by the KDCA. In accordance with Article 2(2) of the Enforcement Rule of the Bioethics and Safety Act [22], the study was exempt from institutional review board review because it did not involve direct human participation. The original data were collected under the Regional Public Health Act and its Enforcement Decree to support evidence-based community health programs. All participants in the KCHS provided written informed consent that included information about the purpose of data collection, the types of personal information gathered, and the period of data retention and use.

The dataset supplied by the KDCA was deidentified in accordance with the Personal Information Protection Act, using

XSL•FO
RenderX

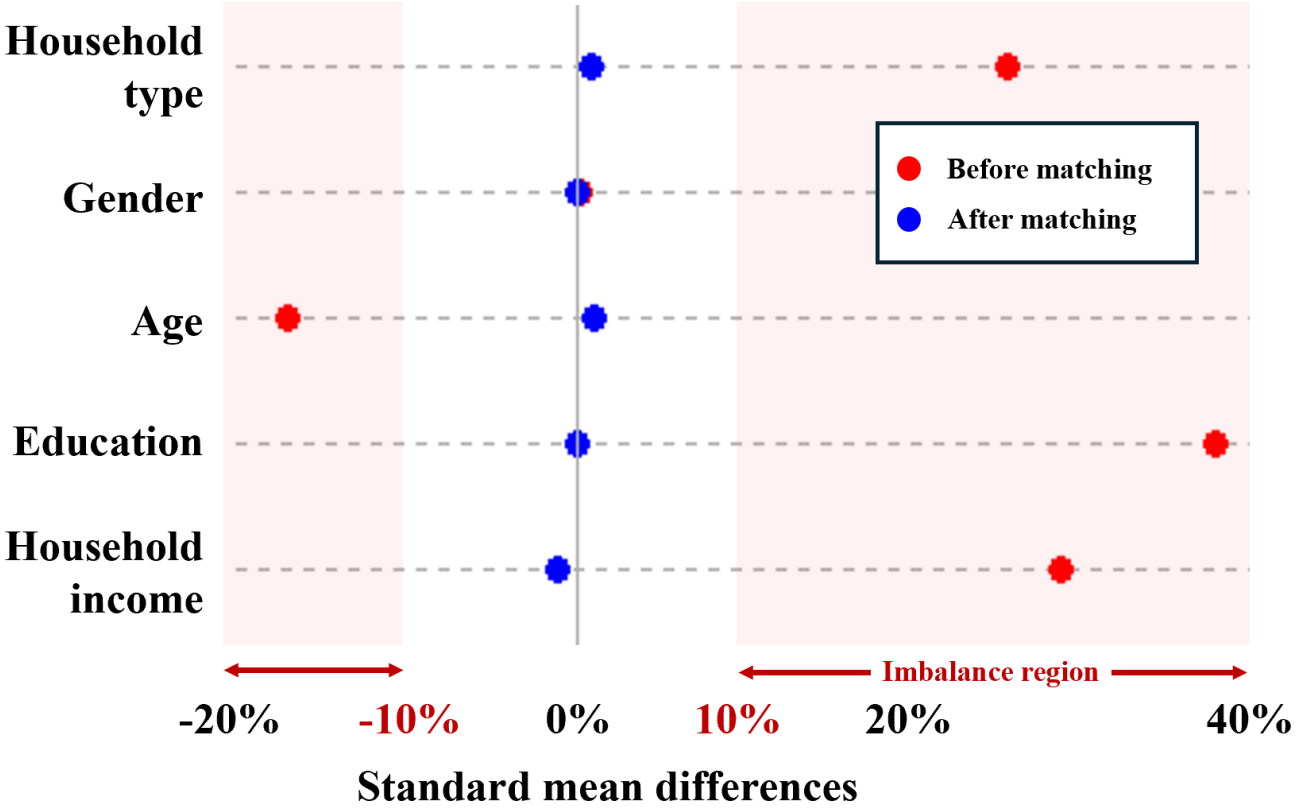
pseudonymization procedures in which identifiers such as sample point numbers and household numbers were replaced with numerical codes to prevent personal identification. Participants in the original survey received small incentives, including a gift voucher worth approximately KRW 10,000 (US $7) and additional daily necessities, provided by local community health centers. No identifiable images or materials that could reveal participants' identities were used in this study.

Results

Propensity Score Matching

Of the 159,782 respondents in super-aged areas, 19,759 were matched 1:1 to those in aging areas using nearest-neighbor matching. After excluding strata with a single PSU, the final matched dataset included 18,574 individuals. All covariates achieved standardized mean differences below 0.10 [23], including gender (<0.001), age (0.01), household type (0.009), education (<0.001), and household income (−0.01), confirming an adequate covariate balance between groups (Figure 2).

Figure 2. Covariate changes before and after propensity score matching. An absolute standardized mean difference (|SMD|>0.10) indicates covariate imbalance (highlighted in red). SMD was used to assess covariate balance after matching.



Comparison of Demographic Characteristics After Matching

After matching, demographic characteristics were compared between aging and super-aged areas, including both the variables used for matching and additional variables not used in the matching process, such as age groups, marital status, and economic activity (Table 1). Comparisons of the same variables between aging and super-aged areas before propensity score matching are provided in Table S1 in Multimedia Appendix 1. Marital status and participation in economic activity differed significantly between regions (both P<.001).

XSL•FO
RenderX

Table . Demographic characteristics of older adults (≥65 y) in aging and super-aged areas of Korea after propensity score matching.

Variable	Matched data[a]		
	Aging area[b] (n=19,759)	Super-aged area (n=18,574)	*P* value[c] (SMD)[d]
Gender, n (%)			.11 (<.001)[d]
Male	2,216,305 (45.7)[3]	1,169,928 (44.5)	
Female	2,654,450 (54.3)	1,457,693 (55.5)	
Age (y), mean (SD)	73.3 (6.6)	73.2 (6.5)	.29 (.01)[d]
Age (y), n (%)			.36
65 - 74	3,016,953 (61.7)	1,641,793 (62.5)	
Over 75	1,871,773 (38.3)	985,828 (37.5)	
Household type, n (%)			<.001 (.009)[d]
Living alone	956,634 (19.6)	509,564 (19.4)	
With a spouse or other	2,357,156 (48.2)	1,368,420 (52.1)	
2 generations	1,110,306 (22.7)	538,923 (20.5)	
3 generations	464,630 (9.5)	210,713 (8.0)	
Marital status, n (%)			<.001
Yes, cohabitation	3,158,471 (64.6)	1,763,107 (67.1)	
Separation or divorced	390,972 (8.0)	158,176 (6.0)	
Bereavement	1,303,745 (26.7)	689,221 (26.2)	
No	35,538 (0.7)	17,117 (0.7)	
Education, n (%)			.006 (<.001)[d]
Under elementary	379,104 (7.8)	223,548 (8.5)	
Elementary	1,415,418 (29.0)	821,941 (31.3)	
Middle school	1,005,594 (20.6)	513,614 (19.5)	
High school	1,300,113 (26.6)	679,110 (25.8)	
College or over	788,496 (16.1)	389,409 (14.8)	
Economic activity, yes, n (%)	1,316,003 (26.9)	987,473 (37.6)	<.001
Household income, 1 million KRW[e], mean (SD)	262.7 (247.3)	242.8 (236.0)	<.001 (–.01)[d]

[a]Propensity score matching was performed using gender, age, household type, education, and household income.

[b]Aging areas were defined as regions with aging rates ≤14%, and super-aged areas as those with ≥20%.

[c]Rao-Scott chi-square tests for categorical variables, independent sample *t* tests for continuous variables.

[d]Standardized mean differences (SMD) for variables used in propensity score matching.

[e]1 million KRW=US $700.

Comparison of Variables Between 2 Groups

Table 2 summarizes regional differences in health behaviors, health status, and health literacy. Among health behaviors, drinking frequency differed significantly between regions, with respondents in the super-aged areas showing more polarized drinking patterns and slightly higher rates of frequent drinking (*P*=.02). Smoking prevalence was similar across regions and was not significantly different. Regarding chronic conditions, the prevalence of depression was significantly lower in the super-aged areas (*P*=.003), whereas hypertension and diabetes did not differ notably between regions. Subjective health status and subjective oral health were poorer among respondents in the super-aged areas (*P*<.001 and *P*=.004, respectively). For health literacy, recognition of nutrition labels was higher in the super-aged areas (*P*=.002), but reading and utilization rates were lower (*P*<.001 and *P*=.03, respectively), suggesting difficulties in applying nutrition information in daily practice.

XSL•FO
RenderX

Table . Health-related characteristics of aging areas and super-aged areas after matching.

Variable	Matched data[a]		
	Aging area[b] (n=19,759)	Super-aged area (n=18,574)	*P* value[c]
Smoking, n (%)			.97
Yes	1,763,544 (36.1)	947,086 (36.0)	
Nonsmoker	3,125,182 (63.9)	1,680,535 (64.0)	
Drinking, n (%)			.02
Almost never	2,986,748 (61.1)	1,621,490 (61.7)	
Less than once a month	581,809 (11.9)	283,298 (10.8)	
1 - 4 times a month	762,930 (15.6)	391,855 (14.9)	
More than 2 times a week	557,239 (11.4)	330,977 (12.6)	
Active physical activity (weekly, min), mean (SD)	38.8 (182.1)	47.0 (216.6)	.02
Moderate physical activity (weekly, min), mean (SD)	84.0 (263.1)	133.3 (385.1)	<.001
Walking (weekly, min), mean (SD)	273.7 (334.3)	250.9 (347.0)	<.001
Subjective body shape, n (%)			<.001
Very thin	128,784 (4.2)	104,497 (4.8)	
Slightly thin	462,204 (15.1)	376,489 (17.3)	
Average	1,493,374 (48.6)	1,074,713 (49.4)	
Slightly obese	885,780 (28.9)	558,236 (25.7)	
Very obese	99,544 (3.2)	59,644 (2.7)	
Weight control experience, n (%)			<.001
Tried to lose weight	1,349,787 (27.6)	607,446 (23.1)	
Tried to maintain	979,810 (20.0)	463,944 (17.7)	
Tried to gain weight	274,146 (5.6)	179,628 (6.8)	
Never tried	2,284,983 (46.7)	1,376,603 (52.4)	
Subjective stress, n (%)			.84
Extremely stressed	91,665 (1.9)	46,217 (1.8)	
Much stressed	642,962 (13.2)	340,560 (13.0)	
Little stressed	2,353,430 (48.1)	1,257,394 (47.9)	
Stressless	1,800,669 (36.8)	983,450 (37.4)	
Depression, yes, n (%)	420,954 (8.6)	191,963 (7.3)	.003
Hypertension, yes, n (%)	2,548,124 (52.1)	1,377,820 (52.4)	.70
Diabetes, yes, n (%)	1,171,928 (24.0)	648,287 (24.7)	.33
Chewing discomfort, n (%)			.11
Uncomfortable	1,478,303 (30.2)	798,880 (30.4)	
Neutral	935,398 (19.1)	467,012 (17.8)	
Comfortable	2,475,025 (50.6)	1,361,728 (51.8)	
Subjective oral health, n (%)			.004
Very good	101,933 (2.1)	58,141 (2.2)	
Good	866,235 (17.7)	456,730 (17.4)	
Fair	1,885,743 (38.6)	956,906 (36.4)	
Bad	1,580,120 (32.3)	925,295 (35.2)	

XSL•FO
RenderX

Variable	Matched data[a]		
	Aging area[b] (n=19,759)	Super-aged area (n=18,574)	*P* value[c]
Very bad	454,695 (9.3)	230,549 (8.8)	
Subjective health, n (%)			<.001
Very good	209,405 (4.3)	84,084 (3.2)	
Good	1,329,045 (27.2)	674,556 (25.7)	
Fair	2,057,748 (42.1)	1,078,442 (41.0)	
Bad	1,025,139 (21.0)	622,846 (23.7)	
Very bad	267,389 (5.5)	167,692 (6.4)	
Nutrition label recognition, n (%)	1,891,633 (38.7)	1,086,721 (41.4)	.002
Nutrition label reading, n (%)	759,477 (15.5)	375,564 (14.3)	<.001
Nutrition label utilization, n (%)	612,869 (12.5)	292,871 (11.1)	.03
Unmet health care needs, n (%)			.26
Yes (unmet)	189,944 (4.1)	110,665 (4.5)	
No (met)	4,489,698 (95.9)	2,374,057 (95.5)	
Written health literacy, n (%)			<.001
Very easy	106,457 (8.0)	62,850 (10.5)	
Somewhat easy	572,592 (43.1)	223,136 (37.1)	
Somewhat difficult	289,349 (21.8)	125,325 (20.8)	
Very difficult	82,198 (6.2)	31,535 (5.2)	
No interest	279,095 (21.0)	158,437 (26.3)	

[a]Propensity score matching was performed using gender, age, household type, education, and household income.

[b]Aging areas were defined as regions with aging rates ≤14%, and super-aged areas as those with ≥20%.

[c]Rao-Scott chi-square tests for categorical variable, independent sample *t* tests for continuous variable.

Results of Ordinal Logistic Regression

Ordinal logistic regression was first performed on the overall matched data to identify factors associated with subjective health. The results, presented in Table S2 in Multimedia Appendix 2, showed consistent patterns with prior evidence. Positive associations were observed for male gender, younger age, higher socioeconomic status, nonsmoking, and greater physical activity (moderate), whereas higher stress, depression, and chronic conditions were negatively associated with subjective health. Subsequently, subgroup analyses were conducted for the aging and super-aged areas using identical models.

Comparison of Factors Between the 2 Groups

Using matched datasets for aging and super-aged areas, ordinal logistic regression was conducted with identical independent variables and ORs were derived to compare the factors influencing subjective health between the 2 groups (Table 3). Most variables showed significant associations consistent with the overall dataset, and the direction of these associations was largely similar across both regions. However, some differences emerged. Smoking status was not significantly associated with subjective health in the super-aged area (*P*=.06). Regarding nutrition label variables, label recognition was significant only in the aging area (*P*=.04), whereas label reading was significant only in the super-aged area (*P*=.05).

XSL•FO
RenderX

Table . Subgroup ordinal logistic regression analysis of factors associated with subjective health among older adults (≥65 y) in aging and super-aged areas of Korea.

Variable	Aging area		Super-aged area		*P* value[c]
	OR[a] (95% CI)	*P* value[b]	Or (95% CI)	*P* value[b]	
Gender					
Male	1.29 (1.16 - 1.42)	<.001	1.22 (1.05 - 1.42)	.01	.56
Female	Reference	__[d]	Reference	—	—
Age (y)	0.99 (0.98 - 0.99)	<.001	0.98 (0.97 - 0.99)	<.001	.27
Education					
Under elementary	0.43 (0.36 - 0.50)	<.001	0.45 (0.34 - 0.58)	<.001	.81
Elementary school	0.51 (0.45 - 0.57)	<.001	0.51 (0.44 - 0.60)	<.001	.93
Middle school	0.60 (0.54 - 0.67)	<.001	0.59 (0.50 - 0.69)	<.001	.83
High school	0.76 (0.68 - 0.84)	<.001	0.83 (0.72 - 0.96)	.01	.30
College or over	Reference	—	Reference	—	—
Economic activity					
Yes	1.77 (1.65 - 1.91)	<.001	1.63 (1.46 - 1.82)	<.001	.22
No	Reference	—	Reference	—	—
Household income (1 million KRW)[e]	1.03 (1.01 - 1.04)	<.001	1.03 (1.01 - 1.05)	.006	.78
Smoking					
Yes	0.82 (0.74 - 0.90)	<.001	0.86 (0.74 - 1.01)	.06	.52
Nonsmoker	Reference	—	Reference	—	—
Drinking					
Almost never	0.62 (0.56 - 0.69)	<.001	0.60 (0.50 - 0.71)	<.001	.70
Less than once a month	0.87 (0.77 - 0.99)	.03	0.76 (0.62 - 0.95)	.01	.30
2 - 4 times a month	0.90 (0.80 - 1.02)	.11	0.91 (0.76 - 1.09)	.31	.95
More than 2 times a week	Reference	—	Reference	—	—
Moderate physical activity (h)	1.02 (1.01 - 1.03)	<.001	1.02 (1.01 - 1.03)	<.001	.43
Walking (h)	1.04 (1.03 - 1.05)	<.001	1.04 (1.02 - 1.05)	<.001	.68
Subjective body shape					
Very thin	0.47 (0.39 - 0.56)	<.001	0.41 (0.31 - 0.55)	<.001	.49
Slightly thin	0.75 (0.69 - 0.83)	<.001	0.77 (0.65 - 0.91)	.002	.81
Average	Reference	—	Reference	—	—
Slightly obese	0.77 (0.72 - 0.83)	<.001	0.84 (0.75 - 0.96)	.008	.24
Very obese	0.40 (0.33 - 0.47)	<.001	0.64 (0.48 - 0.86)	.002	.005
Weight control experience					
Tried to lose weight	1.17 (1.07 - 1.27)	<.001	1.04 (0.90 - 1.20)	.60	.17
Tried to maintain	1.18 (1.09 - 1.28)	<.001	1.25 (1.10 - 1.42)	.001	.48
Tried to gain weight	0.92 (0.79 - 1.07)	.26	1.17 (0.94 - 1.44)	.16	.07[†f]
Never tried	Reference	—	Reference	—	—
Subjective stress					

XSL•FO
RenderX

Variable	Aging area		Super-aged area		*P* value[c]
	OR[a] (95% CI)	*P* value[b]	Or (95% CI)	*P* value[b]	
Extremely stressed	0.18 (0.13 - 0.24)	<.001	0.15 (0.10 - 0.22)	<.001	.56
Much stressed	0.34 (0.31 - 0.37)	<.001	0.34 (0.29 - 0.40)	<.001	.97
Little stressed	0.63 (0.59 - 0.67)	<.001	0.62 (0.56 - 0.70)	<.001	.94
Stressless	Reference	—	Reference	—	—
Depression					
Yes	0.62 (0.55 - 0.69)	<.001	0.56 (0.46 - 0.67)	<.001	.39
No	Reference	—	Reference	—	—
Hypertension					
Yes	0.72 (0.68 - 0.77)	<.001	0.65 (0.59 - 0.72)	<.001	.09[†f]
No	Reference	—	Reference	—	—
Diabetes					
Yes	0.65 (0.61 - 0.69)	<.001	0.58 (0.52 - 0.65)	<.001	.10[†f]
No	Reference	—	Reference	—	—
Chewing difficulty					
Uncomfortable	0.68 (0.62 - 0.74)	<.001	0.78 (0.68 - 0.90)	.001	.11
Neutral	0.81 (0.74 - 0.88)	<.001	0.99 (0.87 - 1.13)	.89	.01
Comfortable	Reference	—	Reference	—	—
Subjective oral health					
Very good	8.50 (6.11 - 11.82)	<.001	8.13 (5.47 - 12.09)	<.001	.87
Good	3.82 (3.28 - 4.45)	<.001	4.80 (3.67 - 6.27)	<.001	.15
Fair	2.21 (1.93 - 2.52)	<.001	2.76 (2.15 - 3.55)	<.001	.12
Bad	1.73 (1.54 - 1.96)	<.001	2.02 (1.59 - 2.56)	<.001	.27
Very bad	Reference	—	Reference	—	—
Nutrition label recognition					
Yes	1.08 (1.00 - 1.16)	.04	1.07 (0.96 - 1.20)	.19	.93
No	Reference	—	Reference	—	—
Nutrition label reading					
Yes	1.09 (0.99 - 1.19)	.09	1.16 (1.00 - 1.35)	.046	.44
No	Reference	—	Reference	—	—
Unmet medical needs					
Yes (unmet)	0.55 (0.47 - 0.64)	<.001	0.57 (0.46 - 0.72)	<.001	.78
No (met)	Reference	—	Reference	—	—

[a]OR: odds ratio.
[b]*P* value within one region.
[c]*P* value between 2 region groups.
[d]Not applicable.
[e]1 million KRW=US $700.
[f]† P < 0.10

Comparison of ORs Between the 2 Groups

To assess regional differences in the impact of variables influencing subjective health, *z* tests comparing ORs between regions showed that perceiving oneself as "very obese" had a significantly weaker negative association with subjective health in the super-aged area (OR 0.64) than in the aging area (OR 0.40; *P*=.005). Hypertension (*P*=.09) and diabetes (*P*=.10) also

XSL•FO
RenderX

showed marginally significant trends, with both conditions presenting lower ORs in the super-aged area.

Discussion

Key Findings

In this study, the proportion of respondents from super-aged areas who positively evaluated their subjective health ("very good" or "good") was 28.9%. Furthermore, the proportion of positive subjective health evaluations in aging areas was 31.5%, higher than that in super-aged areas.

For drinking frequency, those consuming alcohol more frequently ("more than 2 times a week") reported higher odds of positive health assessments compared to those drinking infrequently. Although the proportion of smokers was similar across regions, smoking status was not significantly associated with subjective health in super-aged areas.

Nutrition label recognition and reading exhibited distinct patterns across regions. In aging areas, nutrition label recognition was significantly associated with positive subjective health, whereas nutrition label reading was not. Conversely, in super-aged areas, despite higher recognition rates, recognition alone was not significantly related to subjective health, whereas nutrition label reading significantly influenced positive evaluations.

Hypertension and diabetes did not differ significantly between the 2 regions; both conditions tended to reduce the likelihood of positive subjective health evaluations more strongly in super-aged areas than in aging areas [24].

Comparison With Previous Studies

The regional disparities observed in this study are consistent with prior research showing that rural regions tend to report lower subjective health than urban areas [25,26], likely due to the higher proportion of rural communities within super-aged areas [27].

Demographic characteristics, health behaviors, health status, and unmet health care needs were significantly associated with subjective health, consistent with previous studies [11,15,28-30]. Similar patterns were also observed for chronic conditions such as hypertension and diabetes, consistent with prior findings [31].

Although research investigating the association between nutrition labels and subjective health is limited, Dumotier et al [32] reported that recognizing and using nutrition labels significantly influenced dietary habits and obesity. Similarly, Jang's study [33] indicated that health literacy significantly improved subjective health status.

Plausible Underlying Mechanisms

Although previous studies generally suggest that higher drinking frequency is negatively associated with health [14], some studies have found that moderate drinking frequency is positively linked to subjective health [34], possibly due to the social interactions accompanying drinking [35]. The lack of association between smoking and subjective health may indicate that factors beyond individual smoking behavior, such as government tobacco control policies or the socioeconomic characteristics of smokers in different regions, could have influenced self-reported health, potentially contributing to regional differences [36,37].

In Korea, approximately 84% of older adults have at least one chronic condition, many of which require dietary modifications, making it particularly important for this population to utilize information on food labels [38]. However, a high proportion of older adults report difficulties in reading and interpreting food labels [39]. Such challenges may stem from age-related declines in numeracy and attention, which can reduce the ability to process and apply nutritional information effectively [40].

Nutrition knowledge plays a key role as a predictor of both comprehension and utilization of nutrition labels [41]. In Korea, nutrition education for older adults is mainly provided through public health centers and senior welfare centers, yet access to such programs remains limited in super-aged areas. These areas are disproportionately rural, where older residents often face longer travel distances to health care facilities and have fewer community-based resources, leading to reduced participation in preventive programs and health education initiatives.

Limitations

This study has several limitations. First, although the KCHS is a nationally representative dataset, its cross-sectional design allows for the assessment of associations but not causality between subjective health and its related factors. Longitudinal studies are needed to establish causal relationships and to address potential reverse causation. Second, regional classification based solely on aging rates may not fully reflect the complex social and environmental characteristics of each community. However, by applying this classification at the district level rather than using broad administrative boundaries, the study was able to capture substantial variations in aging rates and subjective health. This finer granularity likely improved the sensitivity of the analysis and better reflected Korea's ongoing demographic transitions.

Policy Implications

Improving subjective health is essential for increasing healthy life expectancy and enhancing quality of life in South Korea, a nation experiencing rapid population aging. It has also been designated as a specific indicator within Korea's 5th National Health Plan (HP2030), highlighting it as a national priority requiring focused intervention [42]. The lower proportion of older adults reporting positive subjective health in super-aged areas, compared with both the 2023 KCHS national average (47.6%) and the HP2030 target (34.7%), underscores the need for targeted public health interventions.

Given that 77% of households with older adults in Korea report that older adults are the primary purchasers of food [43], the ability to accurately read and interpret nutrition labels is particularly important for this population. In Korea, front-of-package (FOP) nutrition labeling has not yet been implemented, whereas countries such as Canada have adopted standardized FOP symbols to help consumers quickly identify foods high in sodium, sugar, or saturated fat [44]. Such government-endorsed labeling has been shown to improve the accuracy and speed of consumers' food choices and is

XSL•FO
RenderX

particularly beneficial for populations with lower health literacy, including older adults [45]. Introducing a similar labeling policy in Korea could therefore promote informed food selection and facilitate the effective use of nutrition labels in daily dietary decisions.

Complementary to such policy measures, nutrition education remains essential, particularly in super-aged areas where accessibility to educational programs remains limited. Previous research has shown that when older adults focus on specific nutrients, their accuracy in interpreting nutrition labels can be comparable to that of younger adults [40]. This suggests that targeted and nutrient-specific education can effectively enhance comprehension among older adults. Notably, in this study, hypertension and diabetes, although similarly prevalent across regions, showed a stronger negative association with subjective health in super-aged areas than in aging areas [24], indicating that nutrition label education emphasizing disease-related nutrients may be particularly beneficial in super-aged areas, as supported by previous evidence that tailored nutrition education significantly improved dietary knowledge and behaviors among the elderly [46].

When combined with FOP labeling and community-based education through local health centers, such tailored nutrition programs could strengthen dietary awareness, promote active label use, and help reduce health disparities among older adults [47].

Conclusions

This study revealed regional differences in factors influencing subjective health among older adults in Korea. In super-aged areas, reading and understanding nutrition labels, rather than merely recognizing them, were more strongly associated with positive subjective health, while the negative effects of hypertension and diabetes were also more pronounced. These findings highlight the need for region-specific public health strategies that include tailored nutrition education focusing on disease-related nutrients to improve subjective health among older adults.

Funding

This work was supported by the Ministry of Food and Drug Safety, Republic of Korea (contract 20240411E1A-00), through the research project “Healthy Aging MAP Table Program.” The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

The dataset is available from the Korea Disease Control and Prevention Agency through a data use agreement.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Demographic characteristics of older adults (≥65 years) in aging and super-aged areas of Korea before propensity-score matching.
[DOCX File, 20 KB - publichealth_v12i1e80189_app1.docx]

Multimedia Appendix 2
Ordinal logistic regression analysis of factors associated with subjective health among older adults (≥65 y) in Korea.
[DOCX File, 20 KB - publichealth_v12i1e80189_app2.docx]

Checklist 1
STROBE checklist.
[PDF File, 195 KB - publichealth_v12i1e80189_app3.pdf]

References

1. Population projections: estimated population by major age groups (working-age population, older population, etc.), national data [Web page in Korean]. Korean Statistical Information Service (KOSIS). URL: https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1BPA003&conn_path=I2 [accessed 2025-03-26]
2. Koohsari MJ, Nakaya T, Oka K. Activity-friendly built environments in a super-aged society, Japan: current challenges and toward a research agenda. Int J Environ Res Public Health 2018 Sep 19;15(9):2054. [doi: 10.3390/ijerph15092054] [Medline: 30235862]
3. Jones RS. Addressing demographic headwinds in Japan: a long-term perspective. : Organisation for Economic Co-operation and Development; 2024 URL: https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/04/addressing-demographic-headwinds-in-japan-a-long-term-perspective_85b9a67f/96648955-en.pdf [accessed 2026-02-12] [doi: 10.1787/96648955-en]

XSL•FO
RenderX

4. Robine JM, Saito Y, Jagger C. The relationship between longevity and healthy life expectancy. Qual Ageing Older Adults 2009 Jun 14;10(2):5-14. [doi: 10.1108/14717794200900012]
5. World health statistics 2023: monitoring health for the SDGs, sustainable development goals. : World Health Organization; 2023 URL: https://www.who.int/publications/i/item/9789240074323 [accessed 2025-03-26]
6. Healthy life expectancy statistics: healthy life expectancy at a glance, 2022 [Report in Korean]. : Korea Health Promotion Institute; 2023 URL: https://www.khepi.or.kr/board/view?linkId=1007632&menuId=MENU01592&no1=2&pageNum=1&rowCnt=10 [accessed 2025-03-26]
7. Browning CJ, Leith KH, Thomas SA. Editorial: economic and social factors affecting the health of older adults. Front Public Health 2023;11:1323030. [doi: 10.3389/fpubh.2023.1323030] [Medline: 38074752]
8. Is subjective health reliable as a proxy variable for true health? A comparison of self-rated health and self-assessed change in health among middle-aged and older South Koreans [Article in Korean]. J Health Soc Sci 2016 Dec;36(4):431-459. [doi: 10.15709/hswr.2016.36.4.431]
9. OECD health statistics 2024 (summary) booklet [Report in Korean]. : Korea Institute for Health and Social Affairs; 2024 URL: https://www.mohw.go.kr/board.es?mid=a10411010100&bid=0019&act=view&list_no=1483195 [accessed 2025-03-26]
10. Wang J, Zhang L, Wang S, Zhang L. Living arrangements, health lifestyles, and health outcomes among Chinese oldest-old. Front Public Health 2023;11:1235768. [doi: 10.3389/fpubh.2023.1235768]
11. Kwon MJ. Factors influencing convergence quality of life of the elderly according to economic activity [Article in Korean]. J Korea Converg Soc 2021;12(5):345-354. [doi: 10.15207/JKCS.2021.12.5.345]
12. Barsha RAA, Najand B, Zare H, Assari S. Immigration, educational attainment, and subjective health in the United States. J Ment Health Clin Psychol 2024;8(1):16-25. [doi: 10.29245/2578-2959/2024/1.1299] [Medline: 38455255]
13. Lin CC, Kuo CT, Tsai MR. Association of functional, interactive, and critical health literacy with good self-rated health among Taiwanese community-dwelling older adults. Geriatr Nurs 2022;43:91-96. [doi: 10.1016/j.gerinurse.2021.11.007] [Medline: 34844130]
14. Lee JC, Park JS, Kim GH. The effects of stress, quality of life and family relationship of smokers and drinkers on tobacco and alcohol use: focusing on the mediating effects of self-rated health [Article in Korean]. Korean Public Health Res 2011;37:29-43. [doi: 10.22900/kphr.2011.37.1.003]
15. Cho JO, Park SH. Relationship between exercise practice and subjective health status in the elderly. Soc Occup Ther Aged Dement 2022 Dec 31;16(2):63-70. [doi: 10.34263/jsotad.2022.16.2.63]
16. Chun HJ, Kim HY, Chun KK. The effects of objective characteristics and subjective evaluation of community on depression of adults - the case of Seoul city. J Community Welf 2019 Nov 30;71:1-39. [doi: 10.15300/jcw.2019.71.4.1]
17. Lee MS. Health inequalities among Korean adults: socioeconomic status and residential area differences [Article in Korean]. Korean J Sociol 2005;39:183-209 [FREE Full text]
18. Liao Y, Cheng X, Li Z, Li Y. The mediating role of physical activity and health status between a health-supportive environment and well-being: a cross-sectional study. Front Public Health 2023;11:1233970. [doi: 10.3389/fpubh.2023.1233970] [Medline: 37927875]
19. Yen IH, Michael YL, Perdue L. Neighborhood environment in studies of health of older adults: a systematic review. Am J Prev Med 2009 Nov;37(5):455-463. [doi: 10.1016/j.amepre.2009.06.022] [Medline: 19840702]
20. Age-friendly baseline assessment of Taebaek city. : World Health Organization; 2025 URL: https://extranet.who.int/agefriendlyworld/wp-content/uploads/2025/09/Age-freindly-Baseline-Assessment-of-Taebaek-City-1.pdf [accessed 2026-02-12]
21. Jang EJ, Ahn J, Jung SY, et al. Methods for the Control of Measured Confounders in Outcomes Research: National Evidence-based Healthcare Collaborating Agency; 2013.
22. Enforcement Rule of the Bioethics and Safety Act (Ordinance of the Ministry of Health and Welfare no1115) [Web page in Korean]. Ministry of Health and Welfare, Republic of Korea. URL: https://www.law.go.kr/%EB%B2%95%EB%A0%B9/%C3%B6%C3%B6%C3%B6

29. Lee JW. Effect of unmet healthcare needs on quality of life [Article in Korean]. J Korea Acad Ind Coop Soc 2020;21:283-290. [doi: 10.5762/KAIS.2020.21.9.283]
30. Venegas-Sanabria LC, Moreno-Echeverry MM, Borda MG, Chavarro-Carvajal DA, Cano-Gutierrez CA. Oral health and self-rated health in community-dwelling older adults in Colombia. BMC Oral Health 2023 Oct 20;23(1):772. [doi: 10.1186/s12903-023-03401-4] [Medline: 37858108]
31. Sahril N, Chan YM, Chan YY, et al. Poor self-rated health and associated factors among older persons in Malaysia: a population-based study. Int J Environ Res Public Health 2023 Feb 28;20(5):4342. [doi: 10.3390/ijerph20054342] [Medline: 36901351]
32. Dumoitier A, Abbo V, Neuhofer ZT, McFadden BR. A review of nutrition labeling and food choice in the United States. Obes Sci Pract 2019 Dec;5(6):581-591. [doi: 10.1002/osp4.374] [Medline: 31890249]
33. Jang YE. Influence of the health literacy on subjective health status [Article in Korean]. J Korea Acad Ind Cooper Soc 2023 Nov 30;24(11):691-699. [doi: 10.5762/KAIS.2023.24.11.691]
34. Nam YH, Nam JR. A study of the factors affecting the subjective health status of elderly people in Korea [Article in Korean]. Korean J Fam Walf 2011;16:145-162 [FREE Full text]
35. Lim JS, Kim D, Yoo J, Jung H, Park JH. A longitudinal study of the impact of social activity on drinking in older adults [Article in Korean]. J Ind Converg 2024 Feb 28;22(2):105-117. [doi: 10.22678/JIC.2024.22.2.105]
36. Park S, Ahn J, Lee BK. Self-rated subjective health status Is strongly associated with sociodemographic factors, lifestyle, nutrient intakes, and biochemical indices, but not smoking status: KNHANES 2007-2012. J Korean Med Sci 2015 Sep;30(9):1279-1287. [doi: 10.3346/jkms.2015.30.9.1279] [Medline: 26339168]
37. Verropoulou G. Key elements composing self-rated health in older adults: a comparative study of 11 European countries. Eur J Ageing 2009 Sep;6(3):213-226. [doi: 10.1007/s10433-009-0125-9] [Medline: 28798605]
38. Post RE, Mainous AG 3rd, Diaz VA, Matheson EM, Everett CJ. Use of the nutrition facts label in chronic disease management: results from the National Health and Nutrition Examination Survey. J Am Diet Assoc 2010 Apr;110(4):628-632. [doi: 10.1016/j.jada.2009.12.015] [Medline: 20338291]
39. Deakin TA. Consumers find food labels confusing and too small to read. Pract Diab Int 2011 Jul;28(6):261-264c. [doi: 10.1002/pdi.1611]
40. Miller LMS, Applegate E, Beckett LA, Wilson MD, Gibson TN. Age differences in the use of serving size information on food labels: numeracy or attention? Public Health Nutr 2017;20(5):786-796. [doi: 10.1017/S1368980016003219]
41. Miller LMS, Cassady DL. The effects of nutrition knowledge on food label use. a review of the literature. Appetite 2015 Sep;92:207-216. [doi: 10.1016/j.appet.2015.05.029] [Medline: 26025086]
42. 5th comprehensive national health promotion plan (health plan 2030, 2021~2030) [Report in Korean]. : Korea Health Promotion Institute; 2021 URL: https://www.khepi.or.kr/board/view?boardStyle=Gallery&categoryId=&contents1=&continent=&country=&linkId=1002769&menuId=MENU01320&no1=27&pageNum=1&rowCnt=8 [accessed 2025-03-26]
43. Food consumption behavior survey 2023 [Report in Korean]. : Korea Rural Economic Institute; 2024 URL: https://repository.krei.re.kr/bitstream/2018.oak/31053/1/E16-2023.pdf [accessed 2025-10-25]
44. Regulations amending the food and drug regulations (nutrition symbols, other labelling provisions, vitamin D, and hydrogenated fats or oils), SOR/2022-168. Government of Canada. 2022. URL: https://canadagazette.gc.ca/rp-pr/p2/2022/2022-07-20/html/sor-dors168-eng.html [accessed 2025-11-12]
45. Consumer research on front-of-package nutrition labeling. : Léger; 2018 URL: https://publications.gc.ca/collections/collection_2018/sc-hc/H164-252-2018-eng.pdf [accessed 2025-11-12]
46. Wallace R, Lo J, Devine A. Tailored nutrition education in the elderly can lead to sustained dietary behaviour change. J Nutr Health Aging 2016 Jan;20(1):8-15. [doi: 10.1007/s12603-016-0669-2] [Medline: 26728927]
47. Fan L, Wang Z, Zhao Y, Ma Y. Urban-rural disparities in knowledge, use and perceived benefits of nutrition labels in China: evidence from 10 provinces. Nutrients 2023 Feb 26;15(5):1171. [doi: 10.3390/nu15051171] [Medline: 36904170]

Abbreviations

FOP: front-of-package
KCHS: Korea Community Health Survey
KDCA: Korea Disease Control and Prevention Agency
OECD: Organization for Economic Co-operation and Development
OR: odds ratio
PSU: primary sampling unit

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 08.Jul.2025; peer-reviewed by HM Wang, SK Sharma; revised version received 31.Dec.2025; accepted 06.Jan.2026; published 02.Mar.2026.

Please cite as:
Nam GM, Lee YR, Mun JH, Cho SH, Lim HS
Regional Disparities in Factors Associated With Subjective Health Among Older Adults in Aging and Super-Aged Areas of Korea: Nationwide Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e80189
URL: https://publichealth.jmir.org/2026/1/e80189
doi:10.2196/80189

© Gi-Moon Nam, Ye-Rin Lee, Ji-Hye Mun, Su-Hyeun Cho, Hee-Sook Lim. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 2.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

HIV Risk Profiles and Pre-Exposure Prophylaxis and Postexposure Prophylaxis Uptake Among Men Who Have Sex With Men in China: Cross-Sectional Latent Class Analysis

Ziwei Wu[1], MPH; Xinrui Zhang[1], BS; Xingliang Zhang[2*], PhD; Junfang Xu[1*], PhD

[1]School of Public Health, Second Affiliated Hospital, School of Medicine, Zhejiang University, 866 Yuhangtang Road, Xihu District, Hangzhou, China

[2]Hangzhou Center for Disease Control and Prevention, Hangzhou Health Supervision Institution, Hangzhou, China

[*]these authors contributed equally

Corresponding Author:
Junfang Xu, PhD
School of Public Health, Second Affiliated Hospital, School of Medicine, Zhejiang University, 866 Yuhangtang Road, Xihu District, Hangzhou, China

Abstract

Background: Men who have sex with men remain disproportionately affected by HIV globally and in China. Despite the availability of pre-exposure prophylaxis (PrEP) and postexposure prophylaxis (PEP), their uptake remains suboptimal. Previous studies have rarely integrated both sexual behavioral factors and prevention-related cognitive factors. A clearer understanding of heterogeneity in HIV exposure and prevention literacy is needed to inform targeted HIV prevention strategies.

Objective: This study aimed to identify latent subgroups of men who have sex with men based on sexual behaviors and prevention-related cognition and to examine differences in PrEP and PEP uptake across these subgroups.

Methods: An online community–recruited cross-sectional survey was conducted among men who have sex with men in Hangzhou, China, from January 2024 to August 2024. A total of 3267 eligible participants (male at birth, aged ≥16 years, and history of sex with men) completed the structured self-administered questionnaires. Measures included sociodemographic characteristics, sexual behaviors within the past 3 months, exposure to HIV-related information, HIV testing history, and PrEP or PEP awareness and use. Latent class analysis was used to identify behavioral-cognitive subgroups. Multinomial logistic regression examined sociodemographic predictors of class membership, and binary logistic regression assessed associations between latent classes and PrEP or PEP use. A 2-sided α of .05 was used.

Results: Four latent subgroups were identified: low-risk and well-informed (n=1219, 37.3%), high-risk and well-informed (n=735, 22.5%), moderate-risk and poorly informed (n=830, 25.4%), and high-risk and poorly informed (n=483, 14.8%). Overall, PrEP use was 7.81% (n=255; 95% CI 6.93% - 8.78%) and PEP use was 4.74% (n=155; 95% CI 4.07% - 5.53%). Compared with the low-risk, well-informed group, the high-risk, well-informed group had higher PrEP use (n=99, 13.48%; odds ratio [OR] 4.89, 95% CI 3.22 - 7.42) and PEP use (n=69, 9.40%; OR 3.04, 95% CI 1.93 - 4.77). The high-risk, poorly informed group also showed elevated PrEP use (n=65, 13.55%; OR 5.42, 95% CI 11.52% - 15.87%) and PEP use (n=31, 6.41%; OR 2.57, 95% CI 1.67 - 3.98). The moderate-risk, poorly informed subgroup had the lowest uptake of PrEP (n=19, 2.30%; OR 0.89, 95% CI 1.35% - 3.90%) and PEP (n=9, 1.06%; OR 0.39, 95% CI 0.49% - 2.30%). Membership in poorly informed subgroups was associated with lower education, lower income, living in a rural area, and short-term residence.

Conclusions: Distinct behavioral–cognitive typologies exist among men who have sex with men in Hangzhou, with substantial gaps between HIV risk exposure and prevention uptake. A large subgroup characterized by moderate behavioral risk but very low prevention literacy showed minimal PrEP or PEP use. These findings suggest that knowledge alone does not determine prevention utilization and highlight the need for integrated education, low-threshold counseling, and improved service access for underserved men who have sex with men. Linking behavioral–cognitive profiles to prevention outcomes provides actionable evidence for designing targeted and equitable HIV prevention strategies in real-world public health settings.

(JMIR Public Health Surveill 2026;12:e81405) doi:10.2196/81405

KEYWORDS

men who have sex with men; postexposure prophylaxis; pre-exposure prophylaxis; PrEP or PEP uptake; latent class analysis; LCA; sexual behavioral factors; cognition of HIV preventive measures

XSL•FO
RenderX

Introduction

The Joint United Nations Program on HIV/AIDS reported that an estimated 39 million people were living with HIV by the end of 2023 in the whole world, with approximately 1.3 million new infections occurring in that year [1]. Although the overall global incidence has dropped by 59% since its peak in 1995, some key populations, especially men who have sex with men, continue to bear a disproportionate burden of new infections. For example, men who have sex with men accounted for an estimated 23% of new HIV infections in 2023 [1,2]. Notably, the proportion of new HIV infections attributed to men who have sex with men increased from approximately 11% in 2010 to 20% in 2022, reflecting a growing concentration of HIV transmission within this population [3]. In China, men who have sex with men accounted for 25.2% of newly reported HIV diagnoses nationwide, and in some cities, the proportion of newly notified HIV cases attributable to men who have sex with men approaches 50% in 2024 [4,5]. Such as, previous studies showed that the proportion of men who have sex with men among new HIV cases was 54.3% in Zhejiang, 74% in Hebei, and 76.6% in Tianjin [5]. Despite efforts including safer sex education, voluntary HIV counseling and testing, and peer-led interventions, the HIV epidemic persists among men who have sex with men in China [6]. These data highlight the urgent need to strengthen HIV prevention measures, with particular focus on the high-risk men who have sex with men groups.

Pre-exposure prophylaxis (PrEP) and postexposure prophylaxis (PEP) are proven to be effective ways to reduce the infection risk of HIV for men who have sex with men and are recommended by both World Health Organization (WHO) and the Chinese Center for Disease Control and Prevention [7,8]. While PEP has been available for occupational and nonoccupational exposure prevention since the late 1990s, PrEP represents a newer approach for ongoing protection before potential exposure. However, the uptake of PrEP and PEP remains limited in China, with only 20% of men who have sex with men aware of these options and fewer than 1% initiating PrEP since 2019 [9]. A multicity survey of men who have sex with men in 2021 reported that the usage rates remained low at 7.8% for PrEP and 9.5% for PEP, although with awareness rates of 65.9% and 78.7% respectively [10]. In China, PrEP has not yet been incorporated into the national health insurance system, and no formal subsidized program exists. Nevertheless, generic tenofovir disoproxil fumarate or emtricitabine formulations approved by the National Medical Products Administration in 2020 have lowered medication costs [9]. Men who have sex with men can access PrEP through infectious disease clinics, designated hospitals, and community-based programs in major cities, or seek information from local centers for disease control and prevention and online platforms, with professional consultation before initiation [11].

Empirical evidence indicates that sexual orientation and receptive versus insertive roles during anal intercourse contribute to differential biological vulnerability and network-mediated exposure to HIV [12,13]. Risk is further elevated by behaviors such as multiple sexual partnerships [14,15], group sex [16,17], and lack of awareness of partners' HIV status [18,19]. Additionally, engagement in commercial sex, condomless anal intercourse, and substance use during sex are consistently associated with both increased HIV acquisition risk and higher likelihood of PrEP or PEP use [18,20-22]. A history of sexually transmitted infections (STIs) also serves as a clinical marker of biological susceptibility and behavioral risk clustering [23,24]. Cognitive and informational factors are equally critical in shaping prevention behavior. Exposure to HIV-related public health messaging enhances prevention literacy, while awareness of biomedical interventions (such as PrEP and PEP) and recent HIV testing experience are among the strongest predictors of uptake [23,25]. However, existing research has only examined sexual behavioral factors separately without accounting for the synergistic interactions or cumulative impacts of sexual behaviors as well as cognition of HIV preventive measures [12]. A more holistic evaluation of risk behaviors and prevention engagement is crucial for developing multidimensional vulnerability profiles among men who have sex with men.

In this context, we conducted a population-based cross-sectional study among men who have sex with men in China to examine associations between different HIV risk profiles (by considering both HIV-related risk behaviors and cognition of HIV preventive measures) and PrEP or PEP use. We aim to provide critical evidence for developing targeted HIV prevention interventions that address the diverse behavioral and cognitive profiles of men who have sex with men populations in China and comparable epidemiological settings.

Methods

Study Design and Participants

This study was designed, conducted, and reported in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines for cross-sectional studies (Checklist 1) [26]. The survey was implemented in Hangzhou, Zhejiang Province, China. Participants were recruited via digital outreach of the Sunshine Coast Public Welfare Center (a community-based local nongovernmental organization [NGO] providing HIV-related education and services to men who have sex with men) between January 2024 and August 2024. The inclusion criteria for men who have sex with men were male at birth, aged 16 years or older, and reporting a history of anal or oral sex with a man. The link was disseminated on the NGO's official WeChat account, community WeChat groups, and WeChat Mini Program, and the participants completed a structured, anonymous, self-administered questionnaire online. No a priori sample size calculation was performed because the study used community-based recruitment and aimed to characterize latent subgroups rather than estimate a single parameter.

Data Measurement

Sample and Data Collection

The sociodemographic characteristics (eg, age, registered residence location and type, duration of residence in Hangzhou, marital status, employment status, educational attainment, and monthly income), sexual behavioral factors, cognition of HIV

XSL•FO
RenderX

preventive measures, and experiences related to PrEP and PEP use were collected.

In our study, sexual behaviors over the preceding 3 months included sexual orientation (homosexual or bisexual), anal sex role (insertive, receptive, or versatile), number of male sex partners (0, 1, 2 - 4, or ≥5), condom use, knowledge of partner's HIV status, engagement in commercial sex (yes or no), unprotected sex (yes or no), drug use during intercourse (never or sometimes or every time), and any self-reported diagnosis of a STI within the past 3 months (yes or no). Knowledge of a partner's HIV status was classified as "unknown" (participant did not know partner's status), "partially known" (participant knew status of some but not all partners in the past 3 months, or partner reported but not documented), and "fully known" (participant reported that all recent partners' HIV status was known, based on partner disclosure or test results). Self-reported STI diagnosis referred to any clinician-diagnosed STI within the past 3 months, including syphilis, gonorrhea, genital herpes, genital warts, chlamydia trachomatis, and other self-reported infections.

Cognition of HIV preventive measures included whether they had been exposed to HIV-related information, undergone HIV testing, and their awareness and uptake of biomedical HIV prevention strategies (ie, PrEP and PEP). Awareness of medication-based HIV prevention was measured using a single yes or no composite item: "Have you ever heard of HIV prevention medications such as PrEP or PEP?" This composite question was used to minimize respondent burden and to improve comprehension among participants with varying biomedical literacy. While some studies assess PrEP and PEP awareness separately, the combined phrasing has also been used in prior research and surveys assessing baseline chemoprophylaxis awareness [27]. PrEP and PEP uptake were assessed using 2 separate yes or no items: "Have you used PrEP in the past 3 months?" and "Have you used PEP in the past 3 months?" These outcomes were modeled using binary logistic regression, with latent sexual behavior class membership as the primary independent variable and sociodemographic characteristics as covariates.

Most sexual-behavior items were single-item behavioral self-reports (number of partners, condom use, and drug use during sex) rather than multi-item psychometric scales. For cognition or awareness items, we used direct single-item indicators (exposure to HIV publicity, awareness of PrEP, and awareness of PEP); these were face-valid and widely used in similar men who have sex with men surveys [28,29].

There were no missing data for variables included in the analysis because the online survey platform required all items to be answered before submission (forced-response design). Therefore, no imputation procedures were necessary. Variables of both sexual behavioral factors and cognition of HIV preventive measures were considered to identify latent risk profiles of men who have sex with men.

Bias

To enhance data quality and reduce potential sources of bias, several procedural safeguards were implemented during data collection. The survey was administered anonymously online without collecting any personal identifiers, which minimized interviewer influence and encouraged candid reporting. All items were mandatory within the survey system to ensure completeness of responses. A 6-month reference frame was used for behavioral questions to improve accuracy and consistency of recall. Recruitment was conducted broadly through a major men who have sex with men community-based organization to maximize reach within the target population. These methodological steps were designed to strengthen internal validity and reduce measurement error.

Ethical Considerations

The study protocol and consent procedure were approved by the Medical Ethics Committee of the School of Public Health, Zhejiang University (project ZGL202306-9). Electronic informed consent was obtained at the beginning of the online survey; participants proceeded only after checking an "I agree" box. Survey responses were anonymous; no personal identifiers were collected. Data were stored on a password-protected server accessible only to study staff. No compensation was provided. No images containing identifiable participant information were included.

Statistical Analyses

Descriptive statistics were used to describe the sociodemographic characteristics and sexual behavior profiles of the incorporated men who have sex with men participants. Latent class analysis (LCA) was conducted using the *poLCA* package (1.6.0.1) in R (version 4.3.3; R Foundation for Statistical Computing) to identify distinct behavioral and cognitive subgroups based on the main factors influencing the use of PrEP and PEP (ie, sexual behaviors and cognition of HIV preventive measures). LCA models with 2 to 5 latent classes were estimated, and model fit was assessed using multiple indices, including the Akaike information criterion (AIC), Bayesian information criterion (BIC), adjusted BIC (aBIC), entropy, Vuong-Lo-Mendell-Rubin likelihood ratio test (VLMR-LRT), bootstrapped likelihood ratio test (BLRT) [30], and average latent class posterior probability (ALCPP) [31]. Lower AIC, BIC, and aBIC values, and higher entropy values, indicate better model fit. Entropy values of more than 0.60 were considered indicative of acceptable model fit [32], while ALCPP values of more than 0.70 indicated reliable classification [31]. The VLMR-LRT and BLRT were used to compare the relative fit of K-class versus (K–1)-class models. Prior research suggests that among fit indices, BLRT exhibits the highest performance, followed by BIC and aBIC [33]. Additionally, selection of the optimal model also considered the interpretability, meaningfulness, and proportional size of latent classes [34]. In case of conflicting fit indices, we prioritized BLRT and BIC or aBIC for statistical fit, and then considered entropy, ALCPP, minimum class size (>5% a priori), and interpretability.

Following the identification of latent sexual behavior subgroups via LCA, further analyses were conducted to examine differences across classes. Multinomial logistic regression models were used to assess associations between sociodemographic variables and latent sexual behavior class membership. In addition, latent classes were used as the main

XSL•FO
RenderX

independent variable in binary logistic regression models to explore their association with the PrEP and PEP uptake, while adjusting for the same set of sociodemographic covariates. All statistical analyses were conducted using SPSS (Statistical Package for Social Sciences; version 29.0; IBM Corp). A 2-sided *P* value <.05 was considered statistically significant. Results were reported as odds ratios (ORs) with corresponding 95% CIs.

Results

Characteristics of Men Who Have Sex with Men

The Sunshine Coast Public Welfare Center's online platform serves approximately 10,700 men who have sex with men residing in Hangzhou. A total of 3267 participants were recruited (from January 2024 to August 2024), yielding an estimated response rate of 30.52%. The majority of participants were young, with 58.4% aged 16 to 29 years (Table 1). Most participants (n=2179, 66.7%) had resided in the city for more than 2 years. Educational attainment was generally high, with 46.8% (n=1529) holding a bachelor's degree and 12.5% (n=407) attaining a master's degree or higher. The income distribution suggested a moderate socioeconomic profile, with 45.1% (n=1473) earning between 5000 and 10,000 RMB (1 RMB=US $0.14) per month, and 29.2% (n=1043) reporting monthly earnings above 10,000 RMB.

XSL•FO
RenderX

Table . Sociodemographic and behavioral characteristics of men who have sex with men participants, Hangzhou, China, January 2024 to August 2024 (N=3267).

Items	Frequency, n (%)
Age (y)	
Approximately 16-29	1908 (58.4)
30 - 49	1236 (37.8)
≥50	123 (3.8)
Duration of residence in Hangzhou	
<6 months	323 (9.9)
Approximately 6-12 months	257 (7.9)
1-2 years	477 (14.6)
>2 years	2179 (66.7)
Not in Hangzhou	31 (0.9)
Registered residence	
Zhejiang province	1496 (45.8)
Other provinces	1771 (54.2)
Type of registered residence	
Urban	1487 (45.5)
Rural	1780 (54.5)
Occupation	
Student	432 (13.2)
Full-time (public sector)	228 (7)
Full-time (nonpublic sector)	1865 (57.1)
Part-time	70(2.1)
Unemployed	177 (5.4)
Retired	21 (0.6)
Other	474 (14.5)
Marital status	
Single/never married	2696 (82.5)
Cohabiting	99 (3)
Divorced or widowed	96 (2.9)
Married	376 (11.5)
Education level	
Primary school or below	29 (0.9)
Junior high school	171 (5.2)
Senior high school	258 (7.9)
Vocational or technical school	119 (3.6)
Associate degree	754 (23.1)
Bachelor's degree	1529 (46.8)
Master's degree or higher	407 (12.5)
Monthly income (RMB[a])	
<3000	415 (12.7)
3000-4999	424 (13)
5000-10,000	1473 (45.1)

XSL•FO
RenderX

Items	Frequency, n (%)
>10,000	955 (29.2)
Sexual orientation	
Homosexual	2224 (68.1)
Bisexual	1043 (31.9)
Sexual role during sexual behaviors	
Receptive (0)	680 (20.8)
Versatile (0.5)	1529 (46.8)
Insertive (1)	1058 (32.4)
Sexual partners[b]	
0	745 (22.8)
1	1455 (44.5)
Approximately 2-4	939 (28.7)
≥5	128 (3.9)
Awareness of partners' HIV status	
Not at all aware	494 (15.1)
Partially aware	1187 (36.3)
Fully aware	1586 (48.5)
Engaged in commercial sex[b]	
Yes	89 (2.7)
No	3178 (97.3)
Had unprotected sex[b]	
Yes	1044 (32)
No	2223 (68)
Used drugs during sex[b]	
Never	2462 (75.4)
Sometimes	649 (19.9)
Every time	156 (4.8)
Diagnosed with STI[b,c]	
Syphilis	34 (1)
Gonorrhea	2 (0.1)
Genital herpes	1 (0)
Genital warts (HPV[d])	10 (0.3)
Chlamydia trachomatis	0 (0)
Other	19 (0.6)
None	3201 (98)
Saw HIV-related publicity[b]	
Yes	2496 (76.4)
No	771 (23.6)
Awareness of PrEP[e] and PEP[f]	
Not at all aware	311 (9.5)
Partially aware	1791 (54.8)

XSL•FO
RenderX

Items	Frequency, n (%)
Fully aware	1165 (35.7)
Used PrEP[b]	
Yes	255 (7.8)
No	3012 (92.2)
Used PEP[b]	
Yes	155 (4.7)
No	3112 (95.3)
Tested for HIV in the past 6 months	
Yes	2470 (75.6)
No	797 (24.4)

[a]1 RMB=US $0.14.

[b]Within 3 months before this survey.

[c]STI: sexually transmitted infection.

[d]HPV: human papillomavirus.

[e]PrEP: pre-exposure prophylaxis.

[f]PEP: postexposure prophylaxis.

A considerable proportion (68.1%) self-identified as homosexual, whereas 31.9% reported being bisexual. Nearly half of the respondents (44.5%) reported having one male sexual partner in the past 3 months, while 28.7% had 2 to 4 partners, and 3.9% reported 5 or more. Risk-related sexual behaviors were prominent, with 32% engaging in unprotected intercourse and 2.7% reporting engagement in commercial sex. Drug use during sexual activity was reported by 4.8% of participants. Only 48.5% of respondents reported being fully informed of their partners' status. Only 2% of men who have sex with men reported being diagnosed with STIs, with 1% with syphilis. Most participants (75.6%) had received HIV testing within the previous six months.

Despite high exposure to HIV-related education campaigns (76.4%), only 35.7% demonstrated awareness of PrEP and PEP, with only 7.8% having used PrEP and 4.7% having used PEP.

Model Fitting and Selection

LCA was conducted to compare models comprising 2 to 5 latent classes (Table 2). The ALCPP declined from 0.892 in the 2-class model to 0.757 in the 5-class model, and entropy values were consistently low—particularly for the 3-class (0.591) and 5-class (0.619) models—indicating reduced classification precision with increasing numbers of classes. Both the VLMR-LRT and the BLRT yielded statistically significant results ($P<.001$), indicating improved model fit with the inclusion of additional classes. Specifically, the 2-class model exhibited the most distinct classification structure but demonstrated poor overall model fit, whereas the 3-class model showed improved fit but suffered from the lowest entropy value. The 4-class model provided the optimal tradeoff among model fit (as reflected by lower AIC and BIC values), entropy (0.630), and classification reliability (ALCPP=0.787), and was therefore selected as the final solution. The 5-class model demonstrated signs of overfitting and diminished clarity in latent class differentiation.

XSL•FO
RenderX

Table . Fit statistics for latent class solutions by number of classes.

Classes	AIC[a]	BIC[b]	aBIC[c]	Log-likelihood	ALCPP[d]	Entropy	Minimum sample size for the class (n)	VLMR-LRT[e]	BLRT[f]
2	49363.0	49637.1	49858.2	−24636.5	0.892	0.633	45	__[g]	—
3	48726.5	49140.7	49474.8	−24295.2	0.804	0.591	68	<.001	<.001
4	48496.4	49050.7	49497.9	−24157.2	0.787	0.630	91	<.001	<.001
5	48324.6	49019.1	49579.2	−24048.3	0.757	0.619	114	<.001	<.001

[a]AIC: Akaike information criteria.

[b]BIC: Bayesian information criteria.

[c]aBIC: adjusted Bayesian information criteria.

[d]ALCPP: average latent class posterior probability.

[e]VLMR-LRT: Vuong-Lo-Mendell-Rubin likelihood ratio test.

[f]BLRT: bootstrapped likelihood ratio test.

[g]Not applicable.

Accordingly, 4 distinct latent classes (ie, low-risk and well-informed class, high-risk and well-informed class, moderate-risk and poorly informed class, and high-risk and poorly informed class) were identified, characterized by varying sexual behaviors and cognition of HIV preventive measures (Table 3). *The low-risk, well-informed group* exhibited the most protective behavioral profile: only 2.4% reported having 2 or more partners in the past 3 months, no participants engaged in condomless sex, and 74.9% had undergone HIV testing within that period. Awareness of PrEP and PEP was moderate (42.3%), and a substantial proportion (87.4%) reported recent exposure to HIV-related public health messaging. *The high-risk, well-informed group* reported elevated sexual risk behaviors—38.1% had two or more partners and 67.2% engaged in condomless sex—yet demonstrated strong prevention engagement: 77.3% were fully aware of PrEP and PEP, 88.5% had undergone recent HIV testing, and 97% knew their partners' HIV status. *The moderate-risk, poorly informed group* displayed intermediate levels of behavioral risk—24.1% reported multiple partners and 23.5% reported condomless sex—but exhibited the lowest levels of prevention knowledge: 99.8% lacked full awareness of PrEP or PEP, 80.4% had not encountered HIV-related messaging, and only 48.5% had recently undergone testing. *The high-risk, poorly informed group* represented the most vulnerable subgroup: 72% reported multiple partners, 55.9% engaged in condomless sex, and 76.3% lacked comprehensive awareness of PrEP and PEP. Despite a relatively high rate of HIV testing (85%), prevention literacy remained consistently low across this group.

XSL•FO
RenderX

Table . HIV risk-clustering characteristics of men who have sex with men participants, Hangzhou, China, January 2024 to August 2024 (N=3267).

Characteristic	Class 1	Class 2	Class 3	Class 4
Participants, n (%)	1186 (36.3)	564 (17.3)	565 (17.3)	952 (29.1)
95% CI	34.7-38.0	16.0-18.6	16.0-18.6	27.6-30.7
Sexual orientation, n (%)				
Homosexual	783 (66)	439 (77.8)	290 (51.3)	712 (74.8)
Bisexual	403 (34)	125 (22.2)	275 (48.7)	240 (25.2)
Sexual role during anal sex with a man, n (%)				
Receptive (0)	215 (18.1)	129 (22.9)	108 (19.1)	228 (23.9)
Versatile (0.5)	571 (48.1)	213 (37.8)	297 (52.6)	44 8(47.1)
Insertive (1)	400 (33.7)	222 (39.4)	160 (38.3)	276 (29)
Number of sexual partners, n (%)[a]				
0	527 (44.4)	42 (7.4)	162 (28.7)	14 (1.5)
1	631 (53.2)	307 (54.4)	267 (47.3)	250 (26.3)
Approximately 2-4	28 (2.4)	202 (35.8)	128 (22.7)	581 (61)
≥5	0 (0)	13 (2.3)	8 (1.4)	107 (11.2)
Awareness of partners' HIV status, n (%)				
Not at all aware	124 (10.5)	17 (3)	216 (38.2)	137 (14.4)
Partially aware	298 (25.1)	0 (0)	203 (35.9)	686 (72.1)
Fully aware	764 (64.4)	547(97)	146 (25.8)	129 (13.6)
Engaged in commercial sex, n (%)[a]				
No	1178 (99.3)	519 (92)	558 (98.8)	923 (97)
Yes	8 (0.7)	45 (8)	7 (1.2)	29 (3)
Had unprotected sex, n (%)[a]				
No	1186 (100)	185 (32.8)	432 (76.5)	420 (44.1)
Yes	0 (0)	379 (67.2)	133 (23.5)	532 (55.9)
Used drugs during sex, n (%)[a]				
Never	1146 (96.6)	348 (61.7)	474 (83.9)	494 (51.9)
Sometimes	25 (2.1)	157 (27.8)	63 (11.2)	404 (42.4)
Every time	15 (1.3)	59 (10.5)	28 (5)	54 (5.7)
Diagnosed with STIs[b], n (%)[a]				
None	1159 (97.7)	552 (97.9)	560 (99.1)	930 (97.7)
Syphilis	14 (1.2)	4 (0.7)	5 (0.9)	11 (1.2)
Gonorrhea	1 (0.1)	0 (0)	0 (0)	1 (0.1)
Genital warts (HPV[c])	0 (0)	0 (0)	0 (0)	1 (0.1)
Chlamydia trachomatis	0 (0)	6 (1.1)	0 (0)	4 (0.4)
Other	12 (1)	2 (0.4)	0 (0)	5 (0.5)
Saw HIV-related publicity, n (%)[a]				
No	149 (12.6)	29 (5.1)	454 (80.4)	139 (14.6)
Yes	1037 (87.4)	535 (94.9)	111 (19.6)	813 (85.4)
Awareness of PrEP[d] and PEP, n (%)[e]				
Not at all aware	0 (0)	9 (1.6)	297 (52.6)	5 (0.5)

XSL•FO
RenderX

Characteristic	Class 1	Class 2	Class 3	Class 4
Partially aware	684 (57.7)	119 (21.1)	267 (47.3)	721 (75.7)
Fully aware	502 (42.3)	436 (77.3)	1 (0.2)	226 (23.7)
Tested for HIV in the past 6 months, n (%)				
No	298 (25.1)	65 (11.5)	291 (51.5)	143 (15)
Yes	888 (74.9)	499 (88.5)	274 (48.5)	809 (85)

[a]Within 3 months before this survey.

[b]STI: sexually transmitted infection.

[c]HPV: human papillomavirus.

[d]PrEP: pre-exposure prophylaxis.

[e]PEP: postexposure prophylaxis.

Composition of 4 HIV Risk-Related Behavior Patterns

As shown in Table 4, men who have sex with men aged 16 to 29 years had significantly lower odds of belonging to the moderate-risk, poorly informed class compared with those aged 50 years or older (OR 0.452, 95% CI 0.255-0.798; *P*=.006). Household registration characteristics were significantly associated with latent class membership. Participants with registered residence in Zhejiang province had lower odds of being in the moderate-risk, poorly informed class (OR 0.745, 95% CI 0.596-0.931; *P*=.01) and the high-risk, poorly informed class (OR 0.759, 95% CI 0.631-0.912; *P*=.003) compared with those from other provinces. Similarly, individuals from urban areas had reduced odds of being classified into the high-risk, poorly informed class compared with those from rural areas (OR 0.808, 95% CI 0.671-0.972; *P*=.02). Full-time employment in the public sector was associated with lower odds of membership in the moderate-risk, poorly informed class compared with other occupational categories (OR 0.559, 95% CI 0.317-0.987; *P*=.045). Educational attainment was strongly associated with class membership. Compared with those holding a master's degree or above, participants with primary school or below (OR 3.496, 95% CI 1.268-9.639; *P*=0.02), junior high school (OR 2.482, 95% CI 1.380-4.463; *P*=0.002), senior high school (OR 2.279, 95% CI 1.379-3.766; *P*=0.001), vocational or technical education (OR 2.922, 95% CI 1.551-5.504; *P*<0.001), and associate degree (OR 1.880, 95% CI 1.253-2.822; *P*=0.002) had higher odds of belonging to the moderate-risk, poorly informed class. Monthly income was inversely associated with membership in the high-risk, poorly informed class. Compared with those earning more than 10,000 RMB per month, participants earning less than 3000 RMB (OR 0.620, 95% CI 0.416-0.923; *P*=.02), 3000 to 4999 RMB (OR 0.726, 95% CI 0.529-0.996; *P*=.047), and 5000 to 10,000 RMB (OR 0.682, 95% CI 0.550-0.847; *P*<.001) had lower odds of belonging to the high-risk, poorly informed class.

XSL•FO
RenderX

Table . Association between sociodemographic characteristics and HIV risk patterns of men who have sex with men participants, Hangzhou, China, January 2024 to August 2024 (N=3267)[a].

Characteristics	High-risk, well-informed class		Moderate-risk, poorly informed class		High-risk, poorly informed class	
	OR[b] (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Age (y)						
Approximately 16-29	1.686 (0.717-3.967)	.23	0.452 (0.255-0.798)	.006	1.854 (0.949-3.623)	.07
30 - 49	1.261 (0.548-2.901)	.59	0.510 (0.302-0.862)	.01	1.756 (0.919-3.354)	.09
≥50	1.000		1.000		1.000	
Duration of residence in Hangzhou						
<6 months	2.002 (0.399-10.047)	.40	0.412 (0.148-1.149)	.09	0.577 (0.208-1.601)	.29
Approximately 6-12 months	1.686 (0.333-8.522)	.53	0.392 (0.139-1.105)	.08	0.469 (0.168-1.315)	.15
Approximately 1-2 years	1.672 (0.336-8.312)	.53	0.342 (0.124-0.939)	.04	0.488 (0.178-1.335)	.16
>2 years	1.648 (0.336-8.078)	.54	0.364 (0.137-0.973)	.04	0.545 (0.203-1.464)	.23
Not in Hangzhou	1.000		1.000		1.000	
Registered residence						
Zhejiang province	1.057 (0.853-1.309)	.61	0.745 (0.596-0.931)	.01	0.759 (0.631-0.912)	.003
Other provinces	1.000		1.000		1.000	
Type of registered residence						
Urban	1.119 (0.902-1.388)	.31	0.932 (0.746-1.164)	.54	0.808 (0.671-0.972)	.02
Rural	1.000		1.000		1.000	
Occupation						
Student	1.358 (0.853-2.162)	.20	0.986 (0.604-1.611)	.96	0.799 (0.525-1.215)	.29
Full-time (public sector)	1.449 (0.912-2.303)	.12	0.559 (0.317-0.987)	.045	1.069 (0.718-1.592)	.74
Full-time (non–public sector)	1.073 (0.775-1.486)	.67	1.169 (0.865-1.580)	.31	1.039 (0.801-1.348)	.77
Part-time	1.678 (0.785-3.587)	.18	1.282 (0.591-2.779)	.53	1.406 (0.731-2.704)	.31
Unemployed	1.003 (0.579-1.740)	.99	1.179 (0.700-1.984)	.54	1.118 (0.716-1.746)	.62
Retired	3.183 (0.611-16.567)	.17	1.385 (0.367-5.226)	.63	3.567 (0.928-13.707)	.06
Other	1.000		1.000		1.000	
Marital status						
Single or never married	1.362 (0.876-2.118)	.17	0.617 (0.431-0.882)	.008	1.029 (0.734-1.443)	.87
Cohabiting	1.942 (0.947-3.979)	.07	0.980 (0.491-1.956)	.95	1.513 (0.825-2.773)	.18
Divorced or widowed	0.652 (0.264-1.608)	.35	0.973 (0.542-1.749)	.93	1.087 (0.605-1.953)	.78
Married	1.000		1.000		1.000	
Education level						
Primary school or lower	2.969 (1.053-8.375)	.04	3.496 (1.268-9.639)	.02	0.204 (0.025-1.678)	.14
Junior high school	1.351 (0.655-2.786)	.42	2.482 (1.380-4.463)	.002	1.851 (1.093-3.133)	.02

XSL•FO
RenderX

Characteristics	High-risk, well-informed class		Moderate-risk, poorly informed class		High-risk, poorly informed class	
	OR[b] (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Senior high school	1.651 (0.994-2.742)	.05	2.279 (1.379-3.766)	.001	1.246 (0.801-1.939)	.33
Vocational or technical school	2.186 (1.140-4.189)	.09	2.922 (1.551-5.504)	<.001	1.942 (1.109-3.401)	.02
Associate degree	1.501 (1.036-2.174)	.03	1.880 (1.253-2.822)	.002	1.344 (0.973-1.857)	.07
Bachelor's degree	1.062 (0.773-1.460)	.71	1.050 (0.726-1.519)	.80	1.027 (0.777-1.357)	.85
Master's degree or higher	1.000		1.000		1.000	
Monthly income (RMB[c])						
<3000	0.720 (0.465-1.114)	.14	1.172 (0.741-1.855)	.50	0.620 (0.416-0.923)	.02
Approximately 3000-4999	0.562 (0.381-0.830)	.004	0.823 (0.560-1.210)	.32	0.726 (0.529-0.996)	.047
Approximately 5000-10,000	0.718 (0.556-0.927)	.01	0.902 (0.687-1.185)	.46	0.682 (0.550-0.847)	<.001
>10,000	1.000		1.000		1.000	

[a]The "low-risk, well-informed class" was included in the model as the reference group.

[b]OR: odds ratio.

[c]1 RMB=US $0.14.

Uptake of PEP and PrEP Across HIV Risk Patterns

Table 5 revealed substantial disparities in the uptake of PrEP and PEP across distinct HIV risk profiles. Overall, PrEP use was 7.81% (255/3267; 95% CI 6.93%-8.78%) and PEP use was 4.74% (155/3267; 95% CI 4.07%-5.53%). Compared with the low-risk, well-informed reference group (PrEP: 37/1186, 3.12%; PEP: 35/1186, 2.95%), individuals in the high-risk, well-informed class reported significantly higher use of PrEP (76/564, 13.48%; OR 4.890, 95% CI 3.224-7.418, *P*<.001) and PEP (53/564, 9.40%; OR 3.035, 95% CI 1.930-4.771, *P*<.001). Similarly, participants in the high-risk, poorly informed class also showed higher PrEP uptake (129/952, 13.55%; OR 5.419, 95% CI 3.690-7.959, *P*<.001) and higher PEP use (61/952, 6.41%; OR 2.574, 95% CI 1.666-3.978, *P*<.001) compared with the reference group. In contrast, the moderate-risk, poorly informed group showed no significant difference in PrEP use (13/565, 2.30%; OR 0.891, 95% CI 0.463-1.713, *P*=.73) and had significantly lower PEP uptake (6/565, 1.06%; OR 0.385, 95% CI 0.158-0.939, *P*=.04) relative to the reference group.

Table . Association between HIV risk patterns and the use of pre-exposure prophylaxis (PrEP) and postexposure prophylaxis (PEP) by men who have sex with men participants, Hangzhou, China, January 2024 to August 2024 (N=3267).

Class	PrEP			PEP		
	Used, n (%; 95% CI)	Odds ratio (95% CI)	*P* value	Used, n (%; 95% CI)	Odds ratio (95% CI)	*P* value
Overall	255 (7.81; 6.93-8.78)	__[c]	—	155 (4.74; 4.07-5.53)	—	—
Low-risk, well-informed class (n=1186)	37 (3.12; 2.27-4.27)	1.000	<.001	35 (2.95; 2.13-4.08)	1.000	<.001
High-risk, well-informed class (n=564)	76 (13.48; 10.90-16.54)	4.890 (3.224-7.418)	<.001	53 (9.40; 7.26-12.09)	3.035 (1.930-4.771)	<.001
Moderate-risk, poorly informed class (n=565)	13 (2.30; 1.35-3.90)	0.891 (0.463-1.713)	.73	6 (1.06; 0.49-2.30)	0.385 (0.158-0.939)	.04
High-risk, poorly informed class (n=952)	129 (13.55; 11.52-15.87)	5.419 (3.690-7.959)	<.001	61 (6.41; 5.02-8.15)	2.574 (1.666-3.978)	<.001

[a]Not applicable.

[b]The "high-risk, well-informed" class was the reference group.

XSL•FO
RenderX

In contrast, the moderate-risk, poorly informed group showed no significant difference in PrEP use (2.30%; OR 0.891, 95% CI 0.463 - 1.713; *P*=.729) and had significantly lower PEP uptake (1.06%; OR 0.385, 95% CI 0.158 - 0.939; *P*=.036) compared to the reference group.

Discussion

Principal Findings

This study identified 4 latent subgroups among men who have sex with men in Hangzhou based on sexual behaviors and prevention-related awareness (low-risk, well-informed; high-risk, well-informed; moderate-risk, poorly informed; and high-risk, poorly informed groups). PrEP and PEP use were primarily concentrated among the 2 high-risk groups, whereas the moderate-risk but poorly informed group exhibited particularly low uptake (PrEP: 2.30%; PEP: 1.06%), indicating a substantial mismatch between behavioral exposure and coverage of HIV prevention services.

Determinants of Behavioral-Cognitive Profiles

Sociodemographic factors play a pivotal structural role in differentiating men who have sex with men subpopulations according to their HIV risk profiles. Education and income emerged as the most consistent predictors among the high-risk and moderate-risk but poorly informed groups. Lower educational attainment—particularly junior high school and vocational training—was significantly associated with membership in the high-risk, poorly informed group, while individuals with unstable residence and nonlocal household registration were more likely to fall into the moderate-risk, poorly informed group. These findings align with existing literature indicating that lower socioeconomic status is strongly associated with reduced access to HIV-related knowledge, rather than knowledge-use behaviors. For example, research from Brazil and China suggests that men who have sex with men with lower income and education levels tend to have lower HIV knowledge scores and weaker risk perception [35,36]. This phenomenon may be explained by structural factors such as informational exclusion, institutional discrimination, and unequal access to community-based education programs. Studies from both high- and middle-income countries also indicate that health communication tends to favor those with higher educational attainment, reinforcing knowledge gaps within socioeconomically marginalized men who have sex with men subgroups [37,38]. These findings highlight the need to design differentiated health promotion strategies that address the cognitive disadvantage created by broader structural inequities.

Structural Inequalities and Implications for HIV Prevention

Both high-risk subgroups—irrespective of their cognition of HIV preventive measures—exhibited significantly elevated rates of PrEP and PEP use. This aligns with well-documented associations between high-risk sexual behaviors and the uptake of biomedical prevention tools observed among men who have sex with men globally, including in China [39,40]. Notably, PrEP usage among the high-risk, poorly informed group was comparable to that of their well-informed counterparts, despite markedly lower levels of HIV-related knowledge. This pattern suggests that PrEP uptake in this subgroup may be less influenced by comprehensive prevention literacy and more reflective of reactive, event-driven decision-making—for example, seeking PEP following perceived exposure, or initiating PrEP due to a partner's HIV status or prior testing experiences. Similar behavioral dynamics have been reported in other men who have sex with men populations, where PrEP use is often shaped more by perceived vulnerability and episodic risk than by an in-depth understanding of biomedical interventions. For instance, a study in Belgium found that men who have sex with men modulated PrEP adherence based on fluctuating risk exposure, indicating a largely reactive usage model [41]. Likewise, research in India reported that men who have sex with men with higher perceived HIV risk were more willing to use PrEP, regardless of their actual knowledge of the intervention [42]. However, the rate of PEP use in the high-risk, poorly informed group (6.41%) was substantially lower than in the high-risk, well-informed group (9.40%). This discrepancy likely reflects the time-sensitive nature of PEP initiation—typically within 72 hours of exposure—which demands a higher degree of health awareness. The comparatively lower uptake in the low-literacy group underscores the urgent need for targeted health education to bridge this critical knowledge-to-action gap.

Interestingly, no significant difference in PrEP or PEP uptake was observed between the high-risk, well-informed and high-risk, poorly informed groups, challenging the assumption that cognition of HIV prevention measures alone directly translates into preventive behavior. Several factors may account for this disconnect. First, both groups may encounter similar structural barriers, such as limited access, financial constraints, and stigma [43-45], that hinder service use regardless of knowledge levels. Second, general awareness of HIV prevention may not equate to a heightened perception of personal risk, thereby weakening the motivation to initiate PrEP or PEP [46,47]. Third, the knowledge indicators used in the LCA may capture general informational exposure rather than specific, service-oriented literacy or navigational capacity. Fourth, men who have sex with men populations often demonstrate event-driven, reactive patterns of biomedical prevention use, such as initiating PEP following perceived exposure, rather than sustained, proactive engagement [48]. Finally, social norms and peer influences may exert a stronger effect than individual knowledge [43,45,49], especially in the first context.

The moderate-risk, poorly informed group emerged as a notably underserved population in this study. Despite reporting non-negligible behavioral risks—such as a 24.1% prevalence of multiple concurrent sexual partnerships—this group exhibited the lowest uptake of PrEP (2.30%) and PEP (1.06%). Strikingly, 99.8% of individuals in this class lacked adequate knowledge of PrEP or PEP (Table 3), revealing a profound gap in biomedical prevention coverage for what may be termed an "invisible risk group." This misalignment between behavioral vulnerability and prevention engagement echoes findings from prior research in China. For instance, Wang et al [9] reported rising PEP usage in Shenzhen—from 3.92% in 2018 to 10.29% in 2020—yet noted that such uptake remained concentrated

XSL•FO
RenderX

among high-risk populations, with limited reach among moderately at-risk individuals with low prevention literacy [50]. Importantly, the disconnect between moderate-risk individuals and biomedical prevention tools reflects not only gaps in personal awareness but also deeper structural and informational asymmetries. As Lei [51] highlights in a qualitative study, knowledge dissemination about PrEP and PEP within men who have sex with men communities in China is highly networked and stratified. Information circulates primarily through specific digital platforms and community-based networks, often excluding individuals from rural backgrounds, those with limited education, or recent migrants [51]. These mechanisms reinforce their "invisibility" in existing prevention infrastructures, rendering them unable to access timely information or interventions despite engaging in risk behaviors. This phenomenon is supported by quantitative evidence: studies in China have shown that low health literacy, rural or nonlocal household registration, recent migration, and limited access to public sector resources are independently associated with restricted access to HIV prevention services [52-54]. Traditional intervention strategies, which often prioritize those at the highest observable risk, may inadvertently neglect populations with moderate but sustained behavioral risk. For this structurally disconnected group, a multilevel intervention framework is urgently needed. This includes peer-led community education, embedding PrEP or PEP counseling into routine HIV testing, and implementing culturally responsive outreach for rural and migrant populations [52,55]. Enhancing both institutional design and the communication ecology will be essential for achieving equitable, targeted service delivery across diverse men who have sex with men subpopulations.

Limitations

This study has several limitations. First, its cross-sectional design limits the ability to infer causal relationships between HIV risk patterns and PrEP or PEP uptake. Second, self-reported data may be subject to recall bias or social desirability bias, potentially affecting the accuracy of behavioral and prevention-related indicators. Third, participant recruitment was conducted exclusively online through a NGO's digital platforms, which may have excluded men who have sex with men who are less digitally connected or more structurally marginalized. Fourth, PrEP and PEP usage were measured as binary variables, without capturing duration, adherence, or regimen type, thereby restricting the interpretability of biomedical prevention engagement. Finally, this study did not collect or report participants' HIV serological results, and thus the HIV status of respondents could not be objectively verified, which may introduce uncertainty regarding the actual infection status. Despite these limitations, the identification of distinct behavioral-knowledge typologies offers meaningful insights for refining HIV prevention strategies.

Conclusions

This study identified that the 4 distinct subgroups of HIV-related risk profiles among men who have sex with men and PrEP or PEP use were concentrated among high-risk individuals. By integrating behavioral and prevention-cognition indicators through LCA, the study revealed an underserved subgroup, men who have sex with men with moderate behavioral risk but very low prevention literacy, which traditional behavior–only frameworks fail to detect. These findings demonstrate that knowledge alone is insufficient to drive PrEP or PEP uptake and highlight the need for differentiated, equity-focused strategies that pair education with improved access and supportive service environments. Targeted community-based education, embedding PrEP or PEP counseling into low-threshold HIV testing, and outreach to men who have sex with men with lower socioeconomic status emerge as practical priorities to close prevention gaps. By offering a multidimensional, person-centered perspective, this study contributes novel evidence to inform real-world HIV prevention planning in China and other middle-income settings facing similar structural challenges.

Acknowledgments

The authors are grateful to all the study participants for their valuable participation. The authors confirm that no generative artificial intelligence tools were used to write, analyze, or interpret data in this manuscript.

Funding

This work was supported by Fundamental Research Funds for the Central Universities and Zhejiang Science and Technology Plan for Disease Prevention and Control (project 2025JK246). The funder had no role in study design, data collection, analysis, interpretation, or manuscript writing.

Data Availability

The de-identified participant-level dataset that underlies the findings of this study is available from the corresponding author upon reasonable request and subject to institutional data-sharing agreements to protect participant confidentiality.

Authors' Contributions

Conceptualization: ZW, JX
Data curation: ZW, Xinrui Z
Formal analysis: ZW
Investigation: Xinrui Z

XSL•FO
RenderX

Methodology: ZW
Supervision: Xinliang Z, JX
Writing – original draft: ZW, Xinrui Z, Xinliang Z, JX
Writing – review & editing: ZW, Xingliang Z, JX
Xingliang Zhang is co-corresponding author (xingliangyufang@163.com)

Conflicts of Interest

None declared.

Checklist 1
STROBE checklist.
[DOCX File, 76 KB - publichealth_v12i1e81405_app1.docx]

References

1. Global HIV & AIDS statistics—fact sheet. Joint United Nations Programme on HIV/AIDS. 2024. URL: https://www.unaids.org/en/resources/fact-sheet [accessed 2026-02-13]
2. Understanding fast-track: accelerating action to end the AIDS epidemic by 2030. Joint United Nations Programme on HIV/AIDS. 2023. URL: https://www.unaids.org/en/resources/documents/2015/201506_JC2743_Understanding_FastTrack [accessed 2026-02-13]
3. Korenromp EL, Sabin K, Stover J, et al. New HIV infections among key populations and their partners in 2010 and 2022, by world region: a multisources estimation. J Acquir Immune Defic Syndr 2024 Jan 1;95(1S):e34-e45. [doi: 10.1097/QAI.0000000000003340] [Medline: 38180737]
4. Editorial Office of Chinese Journal of AIDS & STD. National AIDS and sexually transmitted disease epidemic in December 2024. Chin J AIDS STD 2025;31(3):225. [doi: 10.13419/j.cnki.aids.2025.03.01]
5. Yin Y, Liu Y, Zhu J, et al. The prevalence, temporal trends, and geographical distribution of HIV-1 subtypes among men who have sex with men in China: a systematic review and meta-analysis. Epidemiol Infect 2019 Jan;147:e83. [doi: 10.1017/S0950268818003400] [Medline: 30869019]
6. Ma Y, Dou Z, Guo W, et al. The human immunodeficiency virus care continuum in China: 1985–2015. Clin Infect Dis 2018 Mar 5;66(6):833-839. [doi: 10.1093/cid/cix911] [Medline: 29216405]
7. Consolidated guidelines on HIV prevention, testing, treatment, service delivery and monitoring. World Health Organization. 2021. URL: https://www.who.int/publications/i/item/9789240031593 [accessed 2026-02-13]
8. Acquired Immunodeficiency Syndrome Professional Group, Society of Infectious Diseases, Chinese Medical Association; Chinese Center for Disease Control and Prevention. Chinese guidelines for the diagnosis and treatment of human immunodeficiency virus infection/acquired immunodeficiency syndrome (2024 edition). Chin Med J 2024;137(22):2654-2680. [doi: 10.1097/CM9.0000000000003383]
9. Wang H, Tang W, Shang H. Expansion of PrEP and PEP services in China. Lancet HIV 2022 Jul;9(7):e455-e457. [doi: 10.1016/S2352-3018(22)00138-2] [Medline: 35688167]
10. Liu S, Yu F, Xue H, et al. Factors influencing awareness and use of HIV prophylaxis medications among men who have sex with men in seven cities. China AIDS & STD 2022;28(7):810-815. [doi: 10.13419/j.cnki.aids.2022.07.11]
11. Jiang Z, Wang Q, Liang J, et al. Exploration of PrEP/PEP service delivery model in China: a pilot in eastern, central and western region. Glob Health Med 2024 Oct 31;6(5):295-303. [doi: 10.35772/ghm.2024.01048] [Medline: 39483454]
12. Liu Y, Liu X, Wei S, et al. Identifying patterns of sexual behaviors and PrEP uptake characteristics among MSM who were eligible for PrEP: a national cross-section study. J Virus Erad 2024 Jun;10(2):100382. [doi: 10.1016/j.jve.2024.100382] [Medline: 39015215]
13. Baggaley RF, White RG, Boily MC. HIV transmission risk through anal intercourse: systematic review, meta-analysis and implications for HIV prevention. Int J Epidemiol 2010 Aug;39(4):1048-1063. [doi: 10.1093/ije/dyq057] [Medline: 20406794]
14. Okafor CN, Gorbach PM, Ragsdale A, Quinn B, Shoptaw S. Correlates of preexposure prophylaxis (PrEP) use among men who have sex with men (MSM) in Los Angeles, California. J Urban Health 2017 Oct;94(5):710-715. [doi: 10.1007/s11524-017-0172-z] [Medline: 28600749]
15. Guan Y, Qi T, Liao Q, et al. Multi-dimensional mismatch and barriers for promoting PrEP among men who have sex with men in China: a cross sectional survey from the Demand-side. AIDS Res Ther 2023 Feb 13;20(1):11. [doi: 10.1186/s12981-022-00497-6] [Medline: 36782323]
16. de Sousa AFL, Lima SVMA, Ribeiro CJN, et al. Adherence to pre-exposure prophylaxis (PrEP) among men who have sex with men (MSM) in Portuguese-speaking countries. Int J Environ Res Public Health 2023 Mar 10;20(6):4881. [doi: 10.3390/ijerph20064881] [Medline: 36981792]
17. He Y, Chu X, Li N, et al. Meta-analysis of factors influencing willingness to accept HIV post-exposure prophylaxis among Chinese men who have sex with men. Mod Prev Med 2025;52(1):167-173. [doi: 10.20043/j.cnki.MPM.202409285]

XSL•FO
RenderX

18. Kulie P, Castel AD, Zheng Z, et al. Targeted screening for HIV pre-exposure prophylaxis eligibility in two emergency departments in Washington, DC. AIDS Patient Care STDS 2020 Dec;34(12):516-522. [doi: 10.1089/apc.2020.0228] [Medline: 33296271]
19. Ma Q, Jiang T, Chen W, et al. Awareness of sexual partner's HIV status among men who have sex with men in China: cross-sectional survey study. JMIR Public Health Surveill 2025 Jan 13;11:e66139. [doi: 10.2196/66139] [Medline: 39810334]
20. Xu JF, Wang PC, Cheng F. Health related behaviors among HIV-infected people who are successfully linked to care: an institutional-based cross-sectional study. Infect Dis Poverty 2020 Mar 10;9(1):28. [doi: 10.1186/s40249-020-00642-1] [Medline: 32169118]
21. Mgbako O, Park SH, Mayer KH, et al. Transactional sex and preferences for pre-exposure prophylaxis (PrEP) administration modalities among men who have sex with men (MSM). J Sex Res 2019;56(4-5):650-658. [doi: 10.1080/00224499.2018.1449190] [Medline: 29634362]
22. Wang Z, Mo PKH, Ip M, Fang Y, Lau JTF. Uptake and willingness to use PrEP among Chinese gay, bisexual and other men who have sex with men with experience of sexualized drug use in the past year. BMC Infect Dis 2020 Apr 22;20(1):299. [doi: 10.1186/s12879-020-05024-4] [Medline: 32321442]
23. Lin H, Li JH, Yang X, et al. Discrepancy between behavioral-indicated and perceived candidacy for HIV pre-exposure prophylaxis among men who have sex with men in Chengdu, China. Beijing Da Xue Xue Bao Yi Xue Ban 2023 Jun 18;55(3):511-520. [doi: 10.19723/j.issn.1671-167X.2023.03.018] [Medline: 37291928]
24. Tang V, Montemayor BN, Owens C. Factors influencing oral pre-, post-, and doxycycline post-exposure prophylaxis uptake among substance-using men who have sex with men in the rural southern US. AIDS Patient Care STDS 2025 Jun;39(6):233-244. [doi: 10.1089/apc.2025.0051] [Medline: 40323716]
25. Dai S, Zhang D, Zhang J, et al. Survey on the current status of HIV pre- and post-exposure prophylaxis awareness and service usage among men who have sex with men (MSM). Chin J Dis Control Prev 2024;28(12):1454-1462. [doi: 10.16462/j.cnki.zhjbkz.2024.12.014]
26. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med 2007 Oct 16;4(10):e296. [doi: 10.1371/journal.pmed.0040296] [Medline: 17941714]
27. Zhang X, Qi SZ, Du FZ, et al. Awareness and willingness to accept syphilis chemoprophylaxis among men who have sex with men from three cities in China: a cross-sectional study. BMC Public Health 2022 Oct 15;22(1):1926. [doi: 10.1186/s12889-022-14323-1] [Medline: 36243692]
28. Chen N. A study on pre- and post-exposure prophylaxis medication for HIV among men who have sex with men based on social network analysis [Master's Thesis]. : Youjiang Medical University for Nationalities; 2024. [doi: 10.27908/d.cnki.gymzy.2024.000200]
29. Liu Z. Study on the Current Status and Related Factors Inself-Protective Behaviors for the Prevention of HIV Infection Among Men Who Have Sex With Men [Master's Thesis]. : Central South University; 2023. [doi: 10.27661/d.cnki.gzhnu.2023.002913]
30. Hoijtink H. Confirmatory latent class analysis: model selection using Bayes factors and (pseudo) likelihood ratio statistics. Multivariate Behav Res 2001;36:563-588. [doi: 10.1207/S15327906MBR3604_04]
31. Masyn KE. Latent class analysis and finite mixture modeling. In: The Oxford Handbook of Quantitative Methods in Psychology: Statistical Analysis: Oxford University Press; 2013:551-611. [doi: 10.1093/oxfordhb/9780199934898.013.0025]
32. Asparouhov T, Muthén B. Auxiliary variables in mixture modeling: three-step approaches using Mplus. Struct Equ Modeling 2014 Jul 3;21(3):329-341. [doi: 10.1080/10705511.2014.915181]
33. Nylund KL, Asparouhov T, Muthén BO. Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. Struct Equ Modeling 2007 Oct 23;14(4):535-569. [doi: 10.1080/10705510701575396]
34. Nylund-Gibson K, Choi AY. Ten frequently asked questions about latent class analysis. Transl Issues Psychol Sci 2018;4(4):440-461. [doi: 10.1037/tps0000176]
35. Guimarães MDC, Magno L, Ceccato MGB, et al. HIV/AIDS knowledge among MSM in Brazil: a challenge for public policies. Rev Bras Epidemiol 2019;22(Suppl 1):e190005. [doi: 10.1590/1980-549720190005.supl.1] [Medline: 31576981]
36. Fan W, Yin L, Qian HZ, et al. HIV risk perception among HIV negative or status-unknown men who have sex with men in China. Biomed Res Int 2014;2014:232451. [doi: 10.1155/2014/232451] [Medline: 24795880]
37. Kiviniemi MT, Orom H, Waters EA, McKillip M, Hay JL. Education-based disparities in knowledge of novel health risks: the case of knowledge gaps in HIV risk perceptions. Br J Health Psychol 2018 May;23(2):420-435. [doi: 10.1111/bjhp.12297] [Medline: 29388364]
38. Philbin MM, Hirsch JS, Wilson PA, Ly AT, Giang LM, Parker RG. Structural barriers to HIV prevention among men who have sex with men (MSM) in Vietnam: diversity, stigma, and healthcare access. PLoS ONE 2018;13(4):e0195000. [doi: 10.1371/journal.pone.0195000] [Medline: 29614104]
39. Jin J, Sun R, Mu T, et al. Awareness and use of post-exposure prophylaxis for HIV prevention among men who have sex with men: a systematic review and meta-analysis. Front Med (Lausanne) 2022;8:783626. [doi: 10.3389/fmed.2021.783626] [Medline: 35083243]

XSL•FO
RenderX

40. Du J, Wang S, Zhang H, et al. Pre-exposure Prophylaxis (PrEP) awareness and engagement among MSM at high risk of HIV infection in China: a multi-city cross-sectional survey. AIDS Behav 2025 May;29(5):1629-1639. [doi: 10.1007/s10461-025-04633-x] [Medline: 39883368]
41. Vuylsteke B, Reyniers T, De Baetselier I, et al. Daily and event-driven pre-exposure prophylaxis for men who have sex with men in Belgium: results of a prospective cohort measuring adherence, sexual behaviour and STI incidence. J Int AIDS Soc 2019 Oct;22(10):e25407. [doi: 10.1002/jia2.25407] [Medline: 31663257]
42. Sun Z, Gu Q, Dai Y, et al. Increasing awareness of HIV pre-exposure prophylaxis (PrEP) and willingness to use HIV PrEP among men who have sex with men: a systematic review and meta-analysis of global data. J Int AIDS Soc 2022 Mar;25(3):e25883. [doi: 10.1002/jia2.25883] [Medline: 35255193]
43. Sun Y, Lu H, Ye J, Li D, Li G. Awareness and use of HIV pre-exposure prophylaxis and factors associated with awareness among MSM in Beijing, China. Sci Rep 2023 Jan 11;13(1):554. [doi: 10.1038/s41598-023-27485-8] [Medline: 36631515]
44. Mayer KH, Agwu A, Malebranche D. Barriers to the wider use of pre-exposure prophylaxis in the United States: a narrative review. Adv Ther 2020 May;37(5):1778-1811. [doi: 10.1007/s12325-020-01295-0] [Medline: 32232664]
45. Guo JH, Zhang G, Qin QQ, Chen HJ, Wang L, Lyu F. Progress in research of knowledge, attitude and practice of pre-exposure prophylaxis in men who have sex with men and its influencing factors. Zhonghua Liu Xing Bing Xue Za Zhi 2022 Nov 10;43(11):1854-1859. [doi: 10.3760/cma.j.cn112338-20220427-00351] [Medline: 36444473]
46. Blumenthal J, Jain S, Mulvihill E, et al. Perceived versus calculated HIV risk: implications for pre-exposure prophylaxis uptake in a randomized trial of men who have sex with men. J Acquir Immune Defic Syndr 2019 Feb 1;80(2):e23-e29. [doi: 10.1097/QAI.0000000000001888] [Medline: 30422909]
47. Kota KK, Gelaude D, Carnes N, et al. Low self-perceived need for PrEP and behavioral indications of MSM who recently refused daily PrEP: a mixed methods study in three US cities. AIDS Behav 2024 Jun;28(6):1845-1857. [doi: 10.1007/s10461-024-04276-4] [Medline: 38457051]
48. Beymer MR, Bolan RK, Flynn RP, et al. Uptake and repeat use of postexposure prophylaxis in a community-based clinic in Los Angeles, California. AIDS Res Hum Retroviruses 2014 Sep;30(9):848-855. [doi: 10.1089/AID.2014.0017] [Medline: 24970113]
49. Quinn KG, Christenson E, Spector A, Amirkhanian Y, Kelly JA. The influence of peers on PrEP perceptions and use among young black gay, bisexual, and other men who have sex with men: a qualitative examination. Arch Sex Behav 2020 Aug;49(6):2129-2143. [doi: 10.1007/s10508-019-01593-x] [Medline: 32016815]
50. Wang Y, Liu S, Zhang Y, et al. Use of HIV post-exposure prophylaxis among men who have sex with men in Shenzhen, China: a serial cross-sectional study. AIDS Behav 2022 Oct;26(10):3231-3241. [doi: 10.1007/s10461-022-03673-x] [Medline: 35380286]
51. Lei Z. Accessibility, utilization and governmentality: a qualitative research on prep in MSM community [Master's Thesis]. : East China Normal University; 2023. [doi: 10.27149/d.cnki.ghdsu.2023.003549]
52. Chen W, Chen L, He L, Chai C. Urban-rural disparity in risky sexual behavior, HIV knowledge, and healthy practices among men who have sex with men: a cross-sectional study in Southeast China. PLoS ONE 2024;19(11):e0312006. [doi: 10.1371/journal.pone.0312006] [Medline: 39527536]
53. Wu J, Wu H, Li P, Lu C. HIV/STIs risks between migrant MSM and local MSM: a cross-sectional comparison study in China. PeerJ 2016;4:e2169. [doi: 10.7717/peerj.2169] [Medline: 27478695]
54. Jaiswal J, Griffin M, Singer SN, et al. Structural barriers to pre-exposure prophylaxis use among young sexual minority men: the P18 cohort study. Curr HIV Res 2018;16(3):237-249. [doi: 10.2174/1570162X16666180730144455] [Medline: 30062970]
55. Li C, Xiong Y, Muessig KE, et al. Community-engaged mHealth intervention to increase uptake of HIV pre-exposure prophylaxis (PrEP) among gay, bisexual and other men who have sex with men in China: study protocol for a pilot randomised controlled trial. BMJ Open 2022 May 10;12(5):e055899. [doi: 10.1136/bmjopen-2021-055899] [Medline: 35537794]

Abbreviations

aBIC: adjusted Bayesian information criterion
AIC: Akaike information criterion
ALCPP: average latent class posterior probability
BIC: Bayesian information criterion
BLRT: bootstrapped likelihood ratio test
LCA: latent class analysis
NGO: nongovernmental organization
PEP: postexposure prophylaxis
PrEP: pre-exposure prophylaxis
SPSS: Statistical Package for Social Sciences
STI: sexually transmitted infection
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

XSL•FO
RenderX

VLMR–LRT: Vuong-Lo-Mendell-Rubin likelihood ratio test
WHO: World Health Organization

Edited by A Mavragani, S Brini; submitted 28.Jul.2025; peer-reviewed by J Jones, NS Wong; revised version received 12.Dec.2025; accepted 29.Dec.2025; published 06.Mar.2026.

Please cite as:
Wu Z, Zhang X, Zhang X, Xu J
HIV Risk Profiles and Pre-Exposure Prophylaxis and Postexposure Prophylaxis Uptake Among Men Who Have Sex With Men in China: Cross-Sectional Latent Class Analysis
JMIR Public Health Surveill 2026;12:e81405
URL: https://publichealth.jmir.org/2026/1/e81405
doi:10.2196/81405

© Ziwei Wu, Xinrui Zhang, Xingliang Zhang, Junfang Xu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 6.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Association Between Health Literacy and Prehypertension in South Korean Adults: Cross-Sectional Study Using the 2023 Korea National Health and Nutrition Examination Survey

Jiyeon Chun[1,2], MPH; Dan Bi Kim[1,2], BSS; Suk-Yong Jang[2,3,4], JD, MD, PhD; Eun-Cheol Park[1,5], MD, PhD

[1]Department of Public Health, Graduate School, Yonsei University, Seoul, Republic of Korea

[2]Institute of Health Services Research, Yonsei University, Seoul, Republic of Korea

[3]Department of Biohealth Industry, Policy Analysis Division, Graduate School of Transdisciplinary Health Science, Seoul, Republic of Korea

[4]Department of Health Policy & Management, Graduate School of Public Health, Yonsei University, Seoul, Republic of Korea

[5]Department of Preventive Medicine, College of Medicine, Yonsei University, Seodaemun-gu, 50-1 Yonsei-ro, Seoul, Republic of Korea

Corresponding Author:
Eun-Cheol Park, MD, PhD
Department of Public Health, Graduate School, Yonsei University, Seoul, Republic of Korea

Abstract

Background: Hypertension represents an important global health challenge, closely linked to cardiovascular diseases and elevated premature mortality rates. Prehypertension, defined as elevated blood pressure not meeting the diagnostic criteria for hypertension, necessitates early intervention to prevent disease progression. Health literacy, defined as the capacity to comprehend and use health-related information, is a key determinant of health outcomes but has rarely been studied in the context of prehypertension prevention.

Objective: This study investigated the association between health literacy and prehypertension in South Korean adults. Unlike prior research focusing on treatment adherence in diagnosed patients, this study used the most recent nationally representative data to explore how domain-specific health literacy is associated with prehypertension across various subgroups, identifying potential mechanisms for intervention.

Methods: Data were obtained from the 2023 Korea National Health and Nutrition Examination Survey, a nationally representative cross-sectional study. A stratified, multistage clustered sampling design was used to invite participants. Adults aged 19 years and older (N=1873) who completed the Korean Health Literacy Index were included. Prehypertension was defined as a systolic blood pressure of 130 to 139 mm Hg or a diastolic blood pressure of 80 to 89 mm Hg. A multivariable survey-weighted logistic regression model was used to assess the associations between health literacy and prehypertension, adjusting for sociodemographic and health-related covariates.

Results: Of the 1873 participants, 319 (17.0%) had prehypertension, and 1098 (58.6%) showed low health literacy. After adjustment, those with low health literacy had a 43% higher likelihood of prehypertension (odds ratio 1.43, 95% CI 1.07 - 1.91) than those with high health literacy. Subgroup analyses revealed that the protective impact of health literacy is not uniform but is modulated by demographic contexts.

Conclusions: The observed patterns may reflect three potential mechanisms: (1) motivation for and dependency on health information (eg, in women, middle-aged, lower education, and unemployed groups), (2) synergy between health literacy and resources (eg, in high-income, urban, married, and employer-insured groups), and (3) preventive efficacy in low-risk populations. Low health literacy was significantly associated with prehypertension, with variations across subgroups suggesting context-dependent mechanisms. Health literacy may serve as a modifiable determinant and compensatory resource for cardiovascular risk prevention, particularly in populations with limited access to health care. Targeted interventions that address domain-specific health literacy deficits are needed to reduce the prehypertension burden.

(JMIR Public Health Surveill 2026;12:e82684) doi:10.2196/82684

KEYWORDS

hypertension; prehypertension; cardiovascular diseases; CVD; public health; prevention; health literacy; awareness; social determinant of health

XSL•FO
RenderX

Introduction

Hypertension is often referred to as “the silent killer” and is recognized as a leading global risk factor for cardiovascular morbidity and mortality, contributing to stroke, heart attack, heart failure, and kidney failure [1]. It is the primary individual risk factor for premature mortality, resulting in approximately 10.8 million preventable deaths worldwide each year. The global prevalence of hypertension has doubled from 650 million to 1.3 billion individuals over the past 3 decades since 1990 [1], with a similar trend observed in South Korea, where the number of patients undergoing hypertension treatment steadily increased from 6.32 million in 2018 to 7.27 million in 2022 [2].

Considering that hypertension risk increases with cumulative exposure [3], proactive care during the prehypertensive stage is crucial. Individuals in the prehypertensive stage face a higher risk of developing hypertension and cardiovascular diseases (CVDs) than those with normal blood pressure [4-6]. Recognizing this, the 2017 American College of Cardiology and the American Heart Association redefined blood pressure readings of 130 to 139 mmHg systolic or 80 to 89 mm Hg diastolic—previously termed prehypertension—as stage 1 hypertension, emphasizing the necessity for proactive intervention [7,8].

Health literacy—the capacity of individuals to obtain, comprehend, and use health information to promote and maintain optimal health—has emerged as a critical modifiable determinant in disease prevention and health-related decision-making. People with higher health literacy are less likely to experience medication issues or system process errors, and they tend to engage in healthy behaviors and actively participate in accessible health care services [9,10]. Numerous studies have indicated that low health literacy correlates with hypertension and CVD [11-14], prompting governments worldwide to acknowledge the importance of enhancing health literacy for the prevention and management of CVD and other non-communicable diseases following the 2014 United Nations General Assembly high-level meeting [15]. In the public health and clinical context, health literacy is considered a crucial concept closely linked to patient safety and is emphasized as a major health determinant [10,12,16-19].

Previous studies examining the relationship between health literacy and cardiovascular outcomes have predominantly focused on treatment adherence, behavior improvement, and patient prognosis [13,20-22]. However, the role of health literacy in the prehypertensive stage remains poorly understood, where lifestyle modification alone can prevent disease progression without pharmacological intervention. The health literacy measurement tools used in prior research were mostly designed for specific diseases or disadvantaged populations, typically focusing on the functional comprehension of health literacy related to medical terminology or hospital-centered health care systems [23]. Furthermore, using clinically measured blood pressure data alongside a comprehensive, multidimensional health literacy measure to investigate how health literacy functions across different sociodemographic contexts and which specific literacy domains are most relevant to early-stage blood pressure management would inform the design of targeted and resource-efficient health interventions.

The recent development of a comprehensive health literacy measurement tool by the Korea Disease Control and Prevention Agency, which aligns with Korea’s 5th Health Plan 2030 and reflects the Korean health environment [24], provides an opportunity to address these questions. The tool was included in the health survey section of the 2023 Korea National Health and Nutrition Examination Survey (KNHANES), where it measured the health literacy of participants alongside clinical health examinations, including blood pressure measurements, and the results were publicly released in 2025. This enables the investigation of the association between health literacy and objectively measured prehypertension in a general population context using a culturally relevant, multidimensional measurement tool.

Therefore, this study aimed to investigate the association between health literacy and prehypertension in the Korean adult population using data from the 2023 KNHANES, to identify sociodemographic and health-related subgroups most substantially affected by low health literacy in relation to prehypertension prevalence, and to examine which specific domains of health literacy deficiency are most strongly associated with the likelihood of prehypertension. This study focuses on the prehypertensive stage in a population free from chronic diseases, using clinically measured blood pressure data with a multidimensional health literacy measure.

Methods

Ethical Considerations

As this study is a secondary analysis of publicly available data from the KNHANES, individual informed consent was waived by the institutional review board of the KDCA (2022-11-16-R-A). No additional consent, compensation, or confidentiality agreements were required, as all data were fully anonymized prior to public release.

Study Design

This cross-sectional study was conducted using data from the 2023 KNHANES, collected by the KDCA between January 2023 and December 2023. The study followed the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guideline for cross-sectional studies (Checklist 1) [25]. The institutional review board of the KDCA granted approval of the study protocol (2022-11-16-R-A), and the requirement for informed consent for secondary data analyses was waived under the same approval.

Data Source

The KNHANES employs a stratified, multistage clustered probability sampling design based on the 2021 Population and Housing Census. In 2023, 192 primary sampling units were identified nationwide, with 25 households systematically selected within each unit. Trained field staff visited selected households in advance to verify eligibility and arrange participation at mobile examination centers. The survey comprised Health Interview, Health Examination, and Nutrition

XSL•FO
RenderX

surveys. Of 9825 sampled individuals, 6929 participated (response rate: 70.5%).

Study Sample

Of the 6929 individuals invited to participate in the KNHANES, this study primarily included respondents aged 19 years and older who completed all health literacy questionnaires. Subsequently, individuals with hypertension (systolic blood pressure [SBP] ≥140 mm Hg, diastolic blood pressure [DBP] ≥90 mm Hg, or those on antihypertensive medication), diabetes (fasting blood glucose level ≥126 mg/dL, those on antidiabetes medication or insulin, medically diagnosed cases, or those with a glycated hemoglobin level ≥6.5%), or dyslipidemia (8-h fasting total cholesterol level ≥240 mg/dL, low-density lipoprotein cholesterol level ≥160 mg/dL, high-density lipoprotein cholesterol level <40 mg/dL, 12-h fasting triglyceride level ≥200 mg/dL, or those on cholesterol-lowering medication) were excluded from the study. In particular, individuals with diabetes or dyslipidemia were excluded to ensure internal validity by eliminating the confounding effects of metabolic comorbidity [26] and to minimize potential bias from disease management for these comorbidities [27,28]. Such management, including lifestyle modifications and pharmacological treatment, may independently influence blood pressure levels, thus obscuring the true, independent association between health literacy and blood pressure. Furthermore, individuals with missing values for covariates incorporated into the study model were excluded. A total of 1873 participants were included in the final study sample.

Variables

Prehypertension

Prehypertension was defined as an SBP of 130 to 139 mmHg or a DBP of 80 to 89 mmHg. The KNHANES used an oscillometric blood pressure monitor (WatchBP Office; Microlife AG) for blood pressure measurements. The results and subsequent adjustments were evaluated and published by a panel of experts from the Korean Society of Hypertension and associated disciplines.

Health Literacy

Health literacy was measured using the Health Literacy Index [24], which was developed by adapting the integrated health literacy model proposed by Sørensen et al [29], covering the competencies of accessing, understanding, appraising, and applying health information to reflect the Korean context [24]. The measurement comprises 10 items across 4 key domains: disease prevention, which involves the ability to manage risk factors like smoking, alcohol consumption, and exercise; health care, defined as the ability to access, understand, interpret, and apply information on medical or clinical issues, such as adhering to medication and medical instructions; technology and resources, which entails the ability to access and evaluate the reliability of health information obtained through the internet or media; and health promotion, referring to the ability to make informed decisions on everyday health behaviors based on social and physical environmental determinants [24]. The 10 self-reported items were scored on a scale from 10 to 40 points. Scores of 31 or higher were categorized as "high health literacy," whereas scores below 31 were classified as "low health literacy" in accordance with the criteria established in the research that aided in the development of the Health Literacy Index.

Covariates

Demographic and socioeconomic variables included sex; age (categorized in 10-year intervals from the 20s to 70s and above); highest educational attainment (below elementary school, middle school, high school, or college and higher); household income (high, medium, or low); employment status; residential area (metropolitan, city, or rural); marital status; and type of health care coverage (employer-insured, self-employed insured, or receiving medical aid). Health status and lifestyle variables included BMI (≥25 kg/m^2 or <25 kg/m^2); comorbidities (clinically verified elevated cholesterol level or borderline diabetes); smoking status (smoker or nonsmoker); high-risk alcohol consumption—defined as an average intake of 7 or more drinks per occasion for men and 5 or more for women, with consumption occurring more than twice per week; and physical activity—defined as walking for at least 30 minutes on 5 or more days per week or completing strength training on at least 2 days per week.

Statistical Analysis

Participant characteristics are presented as frequency and percentage for categorical variables, and group differences were evaluated using the Rao-Scott chi-square test. Continuous variables are presented as mean and SD, and comparisons between groups were made using 2-tailed *t* tests. A multivariable survey-weighted logistic regression analysis was conducted to evaluate the relationship between health literacy and prehypertension status and to perform subgroup analyses by independent variables and health literacy domains. All statistical analyses accounted for the complex sampling design of the KNHANES by applying health examination weights provided by the KDCA. Survey-weighted analyses were performed using SAS (version 9.4; SAS Institute Inc) survey procedures (PROC SURVEYFREQ and PROC SURVEYLOGISTIC) to ensure nationally representative estimates. The same regression model was applied to assess dose-response trends (*P* for trend). A survey-weighted multinomial logistic regression model was used to evaluate the relative likelihood of differing blood pressure levels between low and high health literacy groups, adjusting for demographic, socioeconomic, and health-related covariates.

All estimates were reported as odds ratios (ORs) with corresponding 95% CIs. A *P* value of less than .05 was defined as statistically significant, and all analyses were performed using SAS. Multicollinearity was verified using variance inflation factors, confirming no significant multicollinearity among the independent variables included in the model.

Results

Table 1 outlines the general characteristics of the study population, revealing that 319 (17%) participants had prehypertension among the 1873 surveyed (Figure 1). A total of 1098 (58.6%) participants were classified as having low health literacy. The distribution of health literacy differed

XSL•FO
RenderX

significantly between the groups; the proportion of participants with low health literacy was notably higher in the prehypertension group than in the normal blood pressure group, as confirmed by the Rao-Scott chi-square test. Moreover, significant differences in the distribution of characteristics were observed across various factors, including sex, age, highest level of education attained, household income, marital status, type of health care coverage, BMI, presence of prechronic disease stages, smoking status, and high-risk drinking behavior.

XSL•FO
RenderX

Table . General characteristics of the study sample stratified by prehypertension status. Cross-sectional analysis of 1873 Korean adults aged 19 years or older without hypertension, diabetes, or dyslipidemia from the 2023 Korea National Health and Nutrition Examination Survey (KNHANES) conducted in South Korea. Prehypertension was defined as systolic blood pressure (SBP) of 130 to 139 mm Hg or diastolic blood pressure (DBP) of 80 to 89 mm Hg.

Variables	Total (N=1873)	Prehypertension		*P* values[a]
		Yes (n=319)	No (n=1554)	
Health literacy[b], n (%)				<.001
Low	1098 (58.6)	228 (71.5)	870 (56)	
High	775 (41.4)	91 (28.5)	684 (44)	
Sex, n (%)				<.001
Male	659 (35.2)	158 (49.5)	501 (32.2)	
Female	1214 (64.8)	161 (50.5)	1053 (67.8)	
Age (y), n (%)				<.001
19-29	394 (21)	28 (8.8)	366 (23.6)	
30-39	374 (20)	41 (12.9)	333 (21.4)	
40-49	446 (23.8)	74 (23.2)	372 (23.9)	
50-59	275 (14.7)	56 (17.6)	219 (14.1)	
60-69	254 (13.6)	74 (23.2)	180 (11.6)	
70≤	130 (6.9)	46 (14.4)	84 (5.4)	
Educational level, n (%)				<.001
Middle school graduate or lower	199 (10.6)	63 (19.7)	136 (8.8)	
High school graduate	625 (33.4)	110 (34.5)	515 (33.1)	
College graduate or higher	1049 (56)	146 (45.8)	903 (58.1)	
Household income, n (%)				.02
High	651 (34.8)	89 (27.9)	562 (36.2)	
Middle	1024 (54.7)	179 (56.1)	845 (54.4)	
Low	198 (10.6)	51 (16)	147 (9.5)	
Work status, n (%)				.05
Employed	1249 (66.7)	196 (61.4)	1053 (67.8)	
Unemployed	624 (33.3)	123 (38.6)	501 (32.2)	
Residence, n (%)				.11
Metropolitan or city	1578 (84.2)	259 (81.2)	1319 (84.9)	
Rural	295 (15.8)	60 (18.8)	235 (15.1)	
Marital status, n (%)				.<.001
Married	1277 (68.2)	251 (78.7)	1026 (66.0)	
Single	596 (31.8)	68 (21.3)	528 (34.0)	
Health care coverage, n (%)				.009
Employee insured	536 (28.6)	111 (34.8)	425 (27.3)	
Self-employed insured	1295 (69.1)	199 (62.4)	1096 (70.5)	
Medical aid	42 (2.2)	9 (2.8)	33 (2.1)	
BMI (kg/m^2), mean (SD)	22.8 (3.4)	24.2 (3.6)	22.6 (3.3)	<.001
Borderline dyslipidemia, n (%)				<.001

XSL•FO
RenderX

Variables	Total (N=1873)	Prehypertension		*P* values[a]
		Yes (n=319)	No (n=1554)	
Yes	756 (40.4)	169 (53)	587 (37.8)	
No	1117 (59.6)	150 (47)	967 (62.2)	
Prediabetes, n (%)				<.001
Yes	399 (21.3)	117 (36.7)	282 (18.1)	
No	1474 (78.7)	202 (63.3)	1272 (81.9)	
Smoking status, n (%)				.048
Current smoker	234 (12.5)	50 (15.7)	184 (11.8)	
Never or former smoker	1639 (87.5)	269 (84.3)	1370 (88.2)	
High-risk drinking, n (%)				.003
Yes	183 (9.8)	46 (14.4)	137 (8.8)	
No	1690 (90.2)	273 (85.6)	1417 (91.2)	
Physical activity, n (%)				.09
Yes	1075 (57.4)	167 (52.4)	908 (58.4)	
No	798 (42.6)	152 (47.6)	646 (41.6)	

[a]The Rao-Scott chi-square test for categorical variables and the 2-sample *t* test for continuous variables were performed at a .05 significance level.

[b]Health literacy was measured using the Korean Health Literacy Index developed by Yoon et al [24]. Following the recommendations presented in the paper, health literacy scores were classified as low if <31 and as high if ≥31.

XSL•FO
RenderX

Figure 1. Study sample selection. Flowchart depicting the criteria for identifying the study sample for this cross-sectional study. Data were obtained from the 2023 Korea National Health and Nutrition Examination Survey (KNHANES) conducted in South Korea. The final sample consisted of 1873 participants aged 19 years or older who completed the health literacy survey; had no history of hypertension, diabetes, or dyslipidemia; and had complete covariate data.

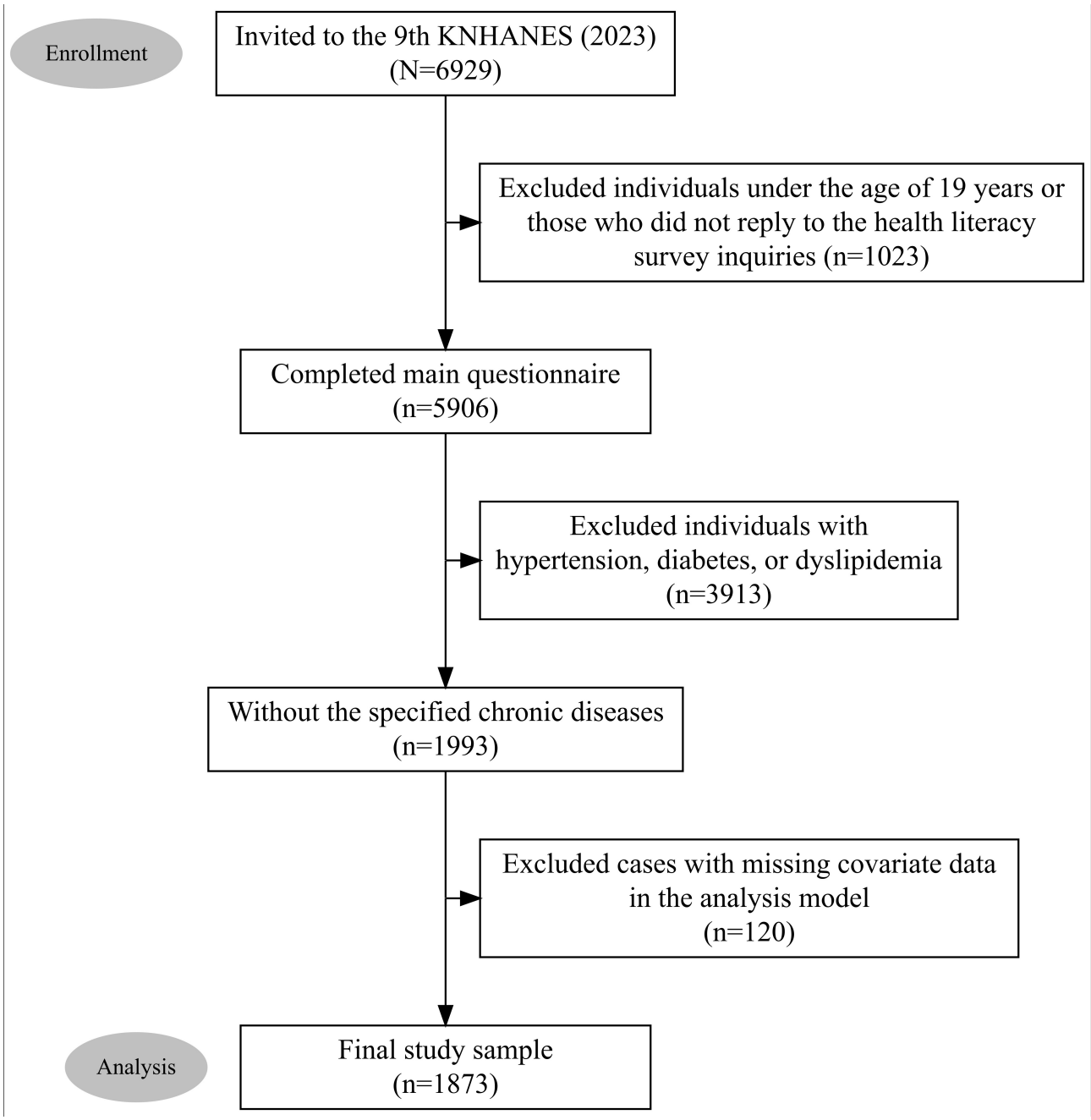


Table 2 presents the results of the multivariable survey-weighted logistic regression analysis examining the association between health literacy and prehypertension. In the crude model, individuals with low health literacy showed a significantly higher likelihood of prehypertension than those with higher health literacy. After adjusting for sociodemographic and health-related factors, this association remained statistically significant, though somewhat attenuated. The main analysis indicated that individuals with low health literacy had a 43% increased likelihood of prehypertension compared with those with higher health literacy.

XSL•FO
RenderX

Table . Association between health literacy and prehypertension in Korean adults. Cross-sectional survey-weighted logistic regression analysis of 1873 adults without chronic diseases from the 2023 Korea National Health and Nutrition Examination Survey (KNHANES) conducted in South Korea. Crude and adjusted ORs and 95% CIs are presented, with adjustments for sociodemographic and health-related covariates.

Variables	Prehypertension	
	Crude OR[a] (95% CI)	Adjusted OR[b] (95% CI)
Health Literacy		
Low	1.75 (1.34-2.28)	1.43 (1.07-1.91)
High	1.00	1.00
Sex		
Male	2.00 (1.55-2.59)	1.86 (1.39-2.51)
Female	1.00	1.00
Age (y)		
19 - 29	1.00	1.00
30 - 39	1.71 (0.94-3.09)	2.07 (1.12-3.83)
40 - 49	2.64 (1.54-4.52)	3.82 (2.10-6.97)
50 - 59	3.32 (1.86-5.92)	4.45 (2.28-8.70)
60 - 69	5.38 (3.13-9.25)	6.62 (3.31-13.24)
70≤	7.07 (3.86-12.98)	6.86 (3.07-15.32)
Educational level		
Middle school graduate or lower	3.19 (2.23-4.55)	1.26 (0.73-2.17)
High school graduate	1.16 (0.85-1.58)	1.04 (0.74-1.47)
College graduate or higher	1.00	1.00
Household income		
High	1.00	1.00
Middle	1.22 (0.84-1.77)	1.09 (0.73-1.62)
Low	2.13 (1.27-3.57)	1.57 (0.79-3.15)
Work status		
Employed	1.00	1.00
Unemployed	1.33 (1.00-1.77)	1.56 (1.13-2.15)
Residence		
Metropolitan or city	1.00	1.00
Rural	1.31 (0.94-1.84)	0.83 (0.55-1.26)
Marital status		
Married	1.00	1.00
Single	0.57 (0.41-0.78)	1.36 (0.89-2.08)
Health care coverage		
Employee insured	1.00	1.00
Self-employed insured	0.68 (0.52-0.90)	0.87 (0.63-1.19)
Medical aid	0.90 (0.44-1.86)	0.48 (0.19-1.26)

XSL•FO
RenderX

Variables	Prehypertension	
	Crude OR[a] (95% CI)	Adjusted OR[b] (95% CI)
BMI (kg/m^2)	1.13 (1.09-1.17)	1.13 (1.09-1.18)
Borderline dyslipidemia		
Yes	1.91 (1.46-2.51)	1.50 (1.12-2.00)
No	1.00	1.00
Prediabetes		
Yes	2.50 (1.86-3.37)	1.20 (0.87-1.67)
No	1.00	1.00
Smoking status		
Current smoker	1.45 (1.00-2.11)	1.16 (0.76-1.76)
Never or former smoker	1.00	1.00
High-risk drinking		
Yes	1.75 (1.21-2.55)	1.86 (1.21-2.87)
No	1.00	1.00
Physical activity		
Yes	1.00	1.00
No	1.24 (0.96-1.60)	1.25 (0.92-1.70)

[a]OR: odds ratio.

[b]Adjusted for sex, age, education, income, employment, residence, marriage, health care coverage, BMI, borderline dyslipidemia, prediabetes, smoking, high-risk drinking, and physical activity.

Table 3 summarizes the results of the analysis of the correlation between health literacy and prehypertension by subgroup of independent variables. The findings revealed that the association between low health literacy and prehypertension was more pronounced in certain subgroups, including women, middle-aged adults (adults in their 40s and 50s), individuals with a high school education or less, high-income households, unemployed individuals, residents of metropolitan or city areas, married individuals, individuals with employee health insurance, individuals with a BMI less than 25 kg/m^2, nonsmokers, and non–high-risk drinkers.

XSL•FO
RenderX

Table . Subgroup analysis of the association between health literacy and prehypertension stratified by sociodemographic and health-related characteristics. Cross-sectional analysis of 1873 Korean adults without chronic diseases from the 2023 Korea National Health and Nutrition Examination Survey (KNHANES) conducted in South Korea. Adjusted odds ratios (ORs) and 95% CIs comparing low and high health literacy are presented for each subgroup.

Variables	Prehypertension: health literacy	
	High	Low, OR[a] (95% CI)
Sex		
Male	1.00	1.04 (0.67-1.63)
Female	1.00	2.27 (1.48-3.47)
Age (y)		
19 - 29	1.00	1.20 (0.45-3.20)
30 - 39	1.00	0.73 (0.35-1.56)
40 - 49	1.00	2.19 (1.22-3.94)
50 - 59	1.00	3.20 (1.28-8.05)
60 - 69	1.00	1.63 (0.73-3.65)
70≤	1.00	0.82 (0.27-2.48)
Educational level		
High school graduate or lower	1.00	1.59 (1.01-2.53)
College graduate or higher	1.00	1.38 (0.91-2.08)
Household income		
High	1.00	2.21 (1.29-3.79)
Middle	1.00	1.18 (0.79-1.75)
Low	1.00	1.33 (0.47-3.78)
Work status		
Employed	1.00	1.29 (0.90-1.84)
Unemployed	1.00	1.81 (1.12-2.93)
Residence		
Metropolitan or city	1.00	1.38 (1.02-1.87)
Rural	1.00	1.83 (0.86-3.87)
Marital status		
Married	1.00	1.66 (1.20-2.28)
Single	1.00	1.10 (0.59-2.04)
Health care coverage		
Employed insured	1.00	2.15 (1.33-3.47)
Self-employed insured	1.00	1.23 (0.84-1.78)
BMI (kg/m^2)		
≥25	1.00	0.95 (0.56-1.62)
<25	1.00	1.74 (1.17-2.60)
Borderline dyslipidemia		
Yes	1.00	1.33 (0.86-2.05)
No	1.00	1.49 (0.98-2.25)
Prediabetes		
Yes	1.00	1.60 (0.92-2.78)

XSL•FO
RenderX

Variables	Prehypertension: health literacy	
	High	Low, OR[a] (95% CI)
No	1.00	1.34 (0.93-1.93)
Smoking status		
Current smoker	1.00	0.82 (0.39-1.73)
Never or former smoker	1.00	1.61 (1.18-2.20)
High-risk drinking		
Yes	1.00	1.19 (0.54-2.62)
No	1.00	1.49 (1.08-2.08)
Physical activity		
Yes	1.00	1.50 (0.98-2.30)
No	1.00	1.41 (0.94-2.12)

[a]Adjusted for sex, age, education, income, employment, residence, marriage, health care coverage, BMI, borderline dyslipidemia, prediabetes, smoking, high-risk drinking, and physical activity.

Table 4 presents the results of the subgroup analyses by health literacy domain, along with the outcomes of the response trend test. Our analysis demonstrated that specific domains of health literacy, including disease prevention and health care, show a significant association with prehypertension even after adjusting for socioeconomic and health-related variables. Additionally, lower health literacy scores correlate with an increased likelihood of prehypertension, confirming statistical significance for the trend.

Table . Sensitivity analysis of health literacy domains and categorized health literacy scores for prehypertension. Cross-sectional analysis of 1873 Korean adults without chronic diseases from the 2023 Korea National Health and Nutrition Examination Survey (KNHANES) conducted in South Korea. Adjusted odds ratios (ORs) and 95% CIs are presented for health literacy domains and categorized health literacy scores. *P* value for trend is .02.

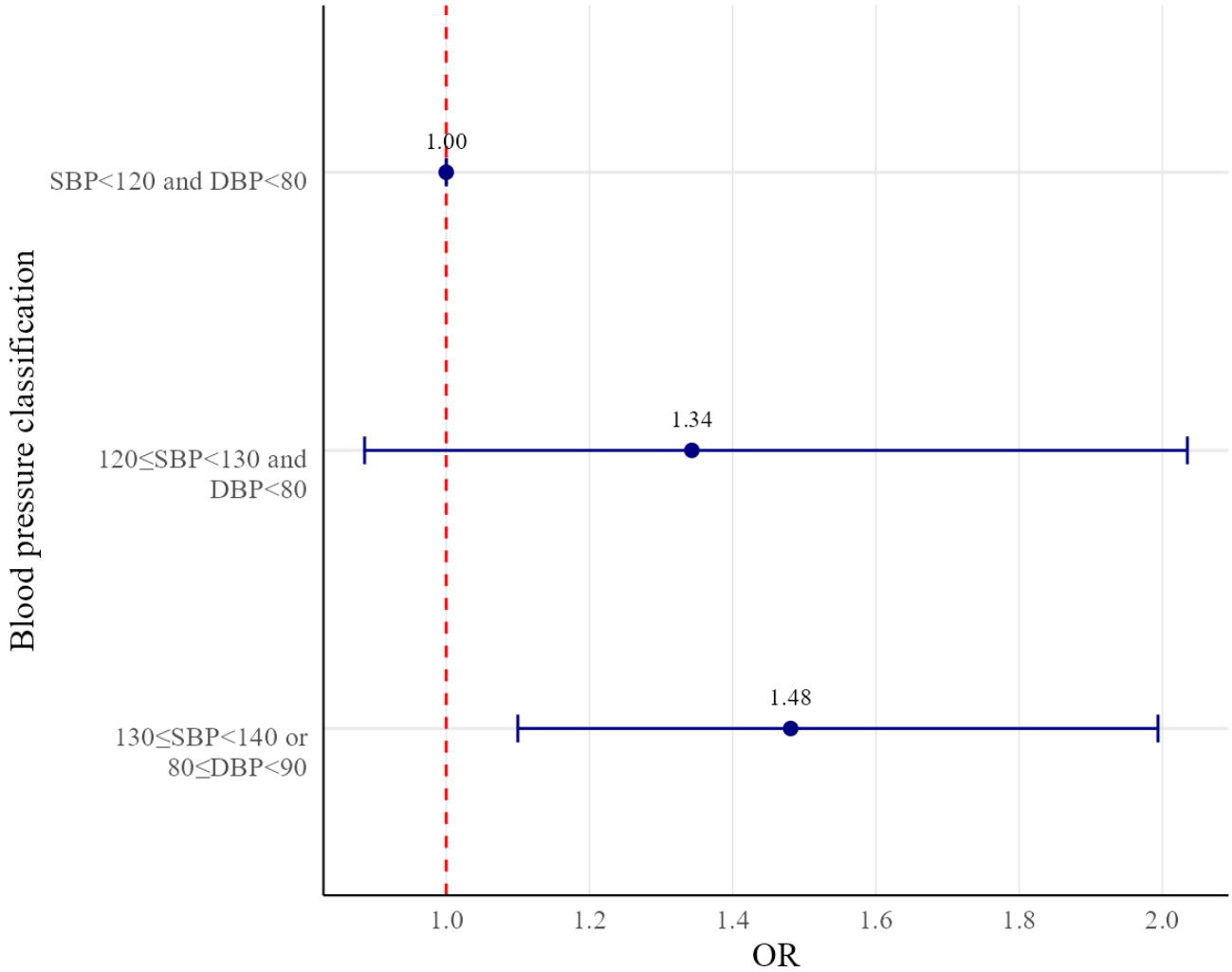
Variables	Prehypertension, OR[a] (95% CI)
Domains of Health Literacy Index (reference: high)	
Disease prevention	1.54 (1.10-2.17)
Health care	1.43 (1.05-1.96)
Technology and resources	1.38 (0.93-2.06)
Health promotion	1.33 (0.91-1.95)
Categorized health literacy scores	
≤23 (10th percentile)	2.10 (0.97-4.53)
24 - 32 (3rd quartile)	1.62 (0.95-2.75)
33 - 37 (90th percentile)	1.26 (0.92-2.63)
38+	1.00

[a]Adjusted for sex, age, education, income, employment, residence, marriage, health care coverage, BMI, borderline dyslipidemia, prediabetes, smoking, high-risk drinking, and physical activity.

In addition, Figure 2 illustrates the results of the multinomial survey-weighted logistic regression analysis, with the dependent variable categorized according to standard blood pressure classifications. Categorizing blood pressure measurements into three groups revealed that individuals with lower health literacy had a 48% higher likelihood of being classified in the prehypertension stage than those with normal blood pressure. Specifically, the ORs were as follows: normal group (SBP<120 mm Hg and DBP<80 mm Hg; OR 1.00, reference group), caution group (120 mm Hg ≤SBP<130 mm Hg and DBP<80 mm Hg; OR 1.34, 95% CI 0.89 - 2.04), and prehypertension group (130 mm Hg ≤SBP<140 mmHg and 80 mm Hg ≤DBP<90 mm Hg; OR 1.48, 95% CI 1.10 - 1.99).

XSL•FO
RenderX

Figure 2. Association between health literacy and blood pressure classification in Korean adults. Forest plot derived from a cross-sectional multinomial logistic regression analysis of 1873 adults without chronic diseases from the 2023 Korea National Health and Nutrition Examination Survey (KNHANES) conducted in South Korea. Adjusted odds ratios (ORs) and 95% CIs comparing low and high health literacy are presented for blood pressure categories: normal (systolic blood pressure [SBP]<120 mm Hg and diastolic blood pressure [DBP]<80 mm Hg), caution (120≤SBP<130 mm Hg and DBP<80 mm Hg), and prehypertension (130≤SBP<140 or 80≤DBP<90 mm Hg). Models were adjusted for sociodemographic and health-related covariates.



Discussion

Principal Findings

This cross-sectional, population-based study revealed a significant association between low health literacy and prehypertension, thereby confirming a negative correlation between health literacy and blood pressure levels. This association remained statistically significant even after adjusting for sociodemographic and health-related factors, indicating a tendency for the strength of the relationship to increase as health literacy decreased and blood pressure increased. The relationship was particularly significant among specific subgroups, including women, middle-aged adults, individuals with a high school education or less, those in high-income groups, unemployed individuals, urban residents, persons with a BMI of 25 kg/m^2 or lower, nonsmokers, and individuals who were not high-risk drinkers. This study provides evidence for preventive intervention strategies in the initial phase of blood pressure elevation.

Previous studies have investigated the association between health literacy and CVD, including hypertension, from multiple perspectives. Studies involving patients with hypertension have consistently shown that low health literacy is associated with greater difficulty in controlling blood pressure, often resulting in inadequate self-management behaviors, such as poor medication adherence and unhealthy lifestyle habits [11,12,30-34]. Furthermore, studies focused on the general adult population have indicated that higher health literacy correlates with improved management and prevention of cardiovascular risk factors including blood pressure [13,14,35]. Additionally, objective CVD risk indicators, such as the Framingham risk score, have been shown to vary based on health literacy levels [14]. Health literacy has also been recognized as playing a crucial role in medication adherence, self-management, and effective communication with health care providers [31,33,36,37].

Domestic studies using population-based data, including health literacy measurement tools such as the Korean Health Panel, have so far been limited by the absence of clinical examinations during data collection [38,39]. This has resulted in discrepancies from actual diagnostic criteria, as disease prevalence information was predominantly gathered through surveys. This study classified blood pressure based on clinical diagnostic criteria,

XSL•FO
RenderX

using clinical examination data from the KNHANES, and focused on a general population free from major chronic diseases.

The main analysis results indicated a significant association between low health literacy and prehypertension. Given the increased likelihood of developing hypertension during the prehypertensive stage, these findings were consistent with those of most previous studies investigating the correlation between health literacy and blood pressure [11-14,36]. In particular, Bonaccorsi and Modesti [36] reported that health literacy and lifestyle contributed to 21.7% of the variance in hypertension. They suggested that improving health literacy could be an effective strategy for managing hypertension through patient empowerment. The American College of Cardiology and the American Heart Association noted that low health literacy is associated with reduced medication adherence and higher readmission rates, underscoring its importance in the primary and secondary prevention of CVDs [13].

The association between health literacy and prehypertension was particularly significant in certain subgroups, suggesting that the protective impact of health literacy is not uniform but may be modulated by demographic and behavioral contexts. These observed patterns can be explained by three key mechanisms: (1) the motivation for and dependency on health information (eg, in women, middle-aged individuals, individuals with lower education, and unemployed individuals), (2) the synergy between health literacy and resources (eg, in high-income, urban, married, and employer-insured groups), and (3) the preventive efficacy of health literacy in low-risk populations.

First, the associations observed among women, middle-aged adults, individuals with lower educational levels, and unemployed individuals demonstrate how behavioral motivation and the relative importance of health literacy differ across groups. Regarding women and middle-aged people, the stronger association likely indicates a heightened motivation for health management. Studies have shown that women are more likely to use health services and comply with lifestyle advice than men [40-45], and middle-aged people often experience increased health concerns that drive active engagement [45-47]. Thus, in these groups, higher health literacy is more effectively converted into tangible actions. In contrast, for individuals with lower educational levels and those who are unemployed, the finding may be explained by a resource substitution effect. Unlike those with higher education or stable employment who benefit from various cognitive resources or workplace structures, vulnerable groups often lack these buffers [48,49]. Consequently, they must rely more heavily on their particular health literacy competencies as a primary resource for navigating disease prevention, making literacy a more critical determinant of health outcomes in these groups.

Second, the notable associations among individuals with a high household income, metropolitan or city residents, married individuals, and those with employee health insurance contrast with findings from prior studies that often highlighted correlations in socially disadvantaged groups [50,51]. This finding suggests a synergy between health literacy and available resources. High household income and urban living offer the infrastructure to engage with health information [52,53], while stable employment with health insurance ensures institutional access to regular health screenings and economic stability [49]. Furthermore, spousal support provides crucial social reinforcement [54,55]. In resource-rich environments, health literacy can be effectively enacted, amplifying its protective effect against prehypertension. It is noteworthy that although the point estimate for rural residents was considerable, it did not reach statistical significance, likely due to the smaller sample size and wider CIs.

Finally, the significant correlation observed among individuals with healthier lifestyles (eg, nonobese individuals [BMI <25 kg/m^2], nonsmokers, and moderate or nondrinkers) underscores the preventive potency of health literacy. In populations with a lower burden of risk factors, health literacy may serve as a critical primary defense, enabling early self-monitoring and the maintenance of good health behaviors before physiological dysregulation becomes irreversible [56]. This implies that health literacy is especially effective in the early stages of prevention, functioning as a key determinant for sustaining normal blood pressure in healthy individuals.

Within the domains of health literacy, disease prevention and health care–related literacy demonstrated notable associations with prehypertension. Drawing on the conceptual framework and measurement items established in our study [24,29], several behavioral mechanisms explain these associations. First, regarding disease prevention literacy, individuals with higher literacy are more likely to make informed decisions to adopt healthy behaviors than their counterparts [24]. They can better comprehend and apply recommendations such as quitting smoking, reducing salt intake, and exercising regularly, which are critical for maintaining normal blood pressure [57,58]. Conversely, those with lower literacy may struggle to grasp the risks associated with prehypertension and the necessity of these lifestyle modifications [59]. Second, regarding health care literacy, higher literacy correlates with improved adherence to medical advice and effective communication with health care providers [24]. This competence facilitates consistent blood pressure monitoring, symptom recognition, and timely medical care seeking [31-33]. Individuals with lower health literacy often encounter difficulties in performing these self-management activities, leading to poor blood pressure control. To address these deficits, multilevel intervention strategies are needed. For disease prevention, educational programs should prioritize practical methods (eg, interpreting nutritional labels) over passive information delivery [60]. For health care literacy, implementing “plain language” initiatives and training providers in the “teach-back” method are necessary to ensure that patients correctly understand and adhere to management protocols [61-64].

Our study, along with prior studies, highlights health literacy as a significant determinant of hypertension. Furthermore, enhancing health literacy via strengthened communication with health care providers and education on disease prevention may be effective in improving clinical outcomes. Therefore, enhancing health literacy is crucial for empowering patients

XSL•FO
RenderX

and for the prevention and management of hypertension. Improving individual health literacy involves providing tailored educational materials and intervention programs for the general population and vulnerable groups, as well as implementing health literacy-friendly frameworks within health care institutions [13,36]. Additionally, a cohort study by Halladay et al [57] suggested that the implementation of stepwise intervention tailored to varying health literacy levels can help bridge literacy disparities and improve long-term blood pressure control, even among individuals with differing degrees of health literacy.

Building on these general principles, our findings suggest specific strategies for implementation that align with the identified mechanisms of motivation, the synergy effect with health care resources, and preventive efficacy. First, for public health authorities, given the strong association between disease prevention literacy and prehypertension revealed in our study, integrating health literacy screening with the Health Literacy Index into National Health Screening Programs is recommended. This approach capitalizes on the preventive potency of literacy observed in our low-risk subgroups. Identifying literacy deficits allows for proactive interventions to address lifestyle risks crucial for preventing the progression to hypertension. The scale is highly feasible owing to its simplicity and established cutoff points. We further propose that screening should ideally be accompanied by a follow-up system, such as easy-to-understand reports or referrals for targeted education, where results prompt practical interventions to support lifestyle modification. Second, for clinicians, addressing health care literacy is essential for the effective management of prehypertension. Implementing screening is achievable through a “waiting room screening” approach, allowing patients to self-administer the assessment prior to their visit. This provides immediate information to adapt communication, a critical step for creating synergy with health care resources for vulnerable patients who rely heavily on provider guidance due to limited external resources. By tailoring explanations to the patient’s literacy level, clinicians can bridge the motivation gap and ensure patients accurately understand prehypertension management guidelines. Moreover, the future development of a validated short-form version would likely expedite its adoption in busy clinical settings.

This study has some limitations that should be acknowledged. The cross-sectional study design precludes causal inference regarding the relationship between low health literacy and prehypertension, as temporal sequencing cannot be determined. However, to minimize bias, individuals with chronic diseases were excluded, and adjustments were made for potential confounders to clarify the association under comparable conditions. Nevertheless, while this exclusion enhanced internal validity for prehypertension assessment, it may underestimate the broader impact of health literacy on cardiovascular health across varied groups. Even after adjusting for numerous sociodemographic and health-related confounders, the potential residual confounding from unmeasured or unadjusted variables remains a possibility. Specifically, complex factors such as psychosocial characteristics, dietary patterns, and health care access were not included in the model. Since these variables may function as potential mediators in the pathway between health literacy and blood pressure, adjusting for them could risk overadjustment and obscure the direct association of interest. Additionally, caution is needed when generalizing the findings. As this study was conducted in South Korea, the results may not fully capture the dynamics in settings with distinct social, cultural, and healthcare systems. Furthermore, the exclusion of participants with diabetes and dyslipidemia resulted in a healthier analytic sample; thus, our findings may not fully apply to broader populations with metabolic comorbidities. Finally, while the study used a context-specific Health Literacy Index, self-reported measures are inherently susceptible to response bias.

Despite these limitations, this study has several methodological strengths. It used a health literacy assessment tool specifically developed for the Korean population and health care context, thereby increasing its contextual relevance and strengthening the study’s internal validity. The use of clinically measured blood pressure data, rather than self-reported measures, enhanced the accuracy and reliability of the results. Thus, this methodological approach yielded a more robust dataset, facilitating a clearer interpretation of the relationship between health literacy and objectively measured health outcomes. Furthermore, the use of weighted survey data ensured a representative sample reflective of the South Korean population, thereby enhancing the external validity of the findings.

Conclusions

The findings of this study demonstrate a significant association between limited health literacy and the presence of prehypertension, underscoring the importance of domain-specific health literacy in the early stages of hypertension development. These findings highlight the role of health literacy as a compensatory resource and modifiable determinant for populations with limited health care access. Targeted interventions that address specific health literacy deficits, particularly in disease prevention and health care navigation, may offer a practical pathway to reduce the prehypertension burden and cardiovascular risk disparities.

To advance this field, future research should use longitudinal or prospective cohort designs to establish causal relationships between health literacy and the progression from prehypertension to hypertension. Incorporating behavioral and lifestyle factors, including diet quality, stress levels, and health care use patterns, would provide a more thorough understanding of the mechanisms linking health literacy to blood pressure management. Additionally, conducting qualitative evaluations or mixed methods strategies could help in understanding the complicated aspects of health literacy, such as communication barriers and cultural perceptions of disease prevention. Comparative analysis across diverse groups or regions would improve external validity and generalizability. Finally, evaluating the effectiveness of targeted interventions that enhance particular domains of health literacy would yield practical insights for public health and clinical practice.

XSL•FO
RenderX

Acknowledgments

The authors appreciate the Yonsei University Institute of Health Services Research for its administrative support and would like to thank Editage for English-language editing. The authors used Perplexity to assist with preliminary literature searches and the identification of relevant studies. All manuscript content, data analysis, and results interpretation are the authors' original work.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this manuscript.

Data Availability

The data that support the findings of this study are available from the Korea National Health and Nutrition Examination Survey (KNHANES).

Authors' Contributions

Conceptualization: JC, ECP
Formal analysis: JC
Supervision: ECP
Writing – original draft: JC
Writing – reviewing and editing: DBK, SYJ, ECP

Conflicts of Interest

None declared.

Checklist 1
STROBE checklist.
[PDF File, 175 KB - publichealth_v12i1e82684_app1.pdf]

References

1. Global report on hypertension: the race against a silent killer. : Inter-American Society of Hypertension, World Health Organization; 2023 URL: https://www.iashonline.org/pdf/21-9-23.pdf [accessed 2026-03-06]
2. Evaluation report [Report in Korean]. Health Insurance Review and Assessment Service. URL: https://www.hira.or.kr/ra/eval/asmWrptPopup.do?evlCd=50&pgmid=HIRAA030004000000#none [accessed 2025-08-20]
3. Dehmer SP, Maciosek MV, LaFrance AB, Flottemesch TJ. Health benefits and cost-effectiveness of asymptomatic screening for hypertension and high cholesterol and aspirin counseling for primary prevention. Ann Fam Med 2017 Jan;15(1):23-36. [doi: 10.1370/afm.2015] [Medline: 28376458]
4. Glasser SP, Basile JN, Lackland DT. Does prehypertension represent an increased risk for incident hypertension and adverse cardiovascular outcome? Hypertension 2009 Nov;54(5):954-955. [doi: 10.1161/HYPERTENSIONAHA.109.138545] [Medline: 19720951]
5. Huang Y, Wang S, Cai X, et al. Prehypertension and incidence of cardiovascular disease: a meta-analysis. BMC Med 2013 Aug 2;11(1):177. [doi: 10.1186/1741-7015-11-177] [Medline: 23915102]
6. Han M, Li Q, Liu L, et al. Prehypertension and risk of cardiovascular diseases: a meta-analysis of 47 cohort studies. J Hypertens 2019 Dec;37(12):2325-2332. [doi: 10.1097/HJH.0000000000002191] [Medline: 31335511]
7. Whelton PK, Carey RM, Aronow WS, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. J Am Coll Cardiol 2018 May 15;71(19):e127-e248. [doi: 10.1016/j.jacc.2017.11.006] [Medline: 29146535]
8. Jones DW, Whelton PK, Allen N, et al. Management of stage 1 hypertension in adults with a low 10-year risk for cardiovascular disease: filling a guidance gap: a scientific statement from the American Heart Association. Hypertension 2021 Jun;77(6):e58-e67. [doi: 10.1161/HYP.0000000000000195] [Medline: 33910363]
9. Health literacy. World Health Organization. URL: https://www.who.int/teams/health-promotion/enhanced-wellbeing/ninth-global-conference/health-literacy [accessed 2025-06-02]
10. Morrison AK, Gibson C, Higgins C, Gutzeit M. Health literacy-related safety events: a qualitative study of health literacy failures in patient safety events. Pediatr Qual Saf 2021;6(4):e425. [doi: 10.1097/pq9.0000000000000425] [Medline: 34235353]
11. McNaughton CD, Kripalani S, Cawthon C, Mion LC, Wallston KA, Roumie CL. Association of health literacy with elevated blood pressure: a cohort study of hospitalized patients. Med Care 2014 Apr;52(4):346-353. [doi: 10.1097/MLR.0000000000000101] [Medline: 24556896]

XSL•FO
RenderX

12. Sohrabi M, Karami M, Mirmoeini RS, Cheraghi Z. The relationship between health literacy and hypertension control: a cross-sectional study. J Tehran Heart Cent 2022 Oct;17(4):243-248. [doi: 10.18502/jthc.v17i4.11614] [Medline: 37143760]
13. Magnani JW, Mujahid MS, Aronow HD, et al. Health literacy and cardiovascular disease: fundamental relevance to primary and secondary prevention: a scientific statement from the American Heart Association. Circulation 2018 Jul 10;138(2):e48-e74. [doi: 10.1161/CIR.0000000000000579] [Medline: 29866648]
14. Rahimi T, Hashemi SS, Rezaei F, Aune D. Association between health literacy and Framingham risk score. Sci Rep 2024 Jun 4;14(1):12837. [doi: 10.1038/s41598-024-63607-6] [Medline: 38834663]
15. Outcome document of the high-level meeting of the general assembly on the comprehensive review and assessment of the progress achieved in the prevention and control of non-communicable diseases: draft resolution / submitted by the president of the general assembly. : United Nations Digital Library; 2014 URL: https://digitallibrary.un.org/record/774662?ln=en&v=pdf [accessed 2025-08-20]
16. Global patient safety report 2024. : World Health Organization; 2024 URL: https://iris.who.int/server/api/core/bitstreams/ae1caeae-bc04-4a71-a6b0-ab478f9a8c79/content [accessed 2026-03-06]
17. Urstad KH, Andersen MH, Larsen MH, Borge CR, Helseth S, Wahl AK. Definitions and measurement of health literacy in health and medicine research: a systematic review. BMJ Open 2022 Feb 14;12(2):e056294. [doi: 10.1136/bmjopen-2021-056294] [Medline: 35165112]
18. Coughlin SS, Vernon M, Hatzigeorgiou C, George V. Health literacy, social determinants of health, and disease prevention and control. J Environ Health Sci 2020;6(1):3061. [Medline: 33604453]
19. DeWalt DA, McNeill J. Integrating health literacy with health care performance measurement. : Institute of Medicine; 2013 URL: https://nam.edu/wp-content/uploads/2015/06/BPH-IntegratingHealthLiteracy.pdf [accessed 2026-03-06]
20. Gaffari-Fam S, Babazadeh T, Oliaei S, Behboodi L, Daemi A. Adherence to a health literacy and healthy lifestyle with improved blood pressure control in Iran. Patient Prefer Adherence 2020;14:499-506. [doi: 10.2147/PPA.S244820] [Medline: 32184576]
21. León-González R, García-Esquinas E, Paredes-Galán E, et al. Health literacy and health outcomes in very old patients with heart failure. Rev Esp Cardiol (Engl Ed) 2018 Mar;71(3):178-184. [doi: 10.1016/j.rec.2017.06.010] [Medline: 28697926]
22. Kanejima Y, Shimogai T, Kitamura M, Ishihara K, Izawa KP. Impact of health literacy in patients with cardiovascular diseases: a systematic review and meta-analysis. Patient Educ Couns 2022 Jul;105(7):1793-1800. [doi: 10.1016/j.pec.2021.11.021] [Medline: 34862114]
23. Choi SK. A study for improving health literacy [Report in Korean]. : Korea Institute for Health and Social Affairs; 2020 URL: https://repository.kihasa.re.kr/en/bitstream/201002/37301/1/%ec%97%b0%ea%b5%ac%eb%b3%b4%ea%b3%a0%ec%84%9c%202020-24.pdf [accessed 2026-03-06]
24. Yoon J, Kang SJ, Lee M, Cho J. Development and validation of the health literacy index for the community for the Korean National Health and Nutrition and Examination Survey. Epidemiol Health 2024;46(1):e2024061. [doi: 10.4178/epih.e2024061] [Medline: 39026432]
25. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. The Lancet 2007 Oct;370(9596):1453-1457. [doi: 10.1016/S0140-6736(07)61602-X]
26. Grundy SM, Cleeman JI, Daniels SR, et al. Diagnosis and management of the metabolic syndrome: an American Heart Association/National Heart, Lung, and Blood Institute scientific statement. Circulation 2005 Oct 25;112(17):2735-2752. [doi: 10.1161/CIRCULATIONAHA.105.169404] [Medline: 16157765]
27. Williams AR, Wilson-Genderson M, Thomson MD. A cross-sectional analysis of associations between lifestyle advice and behavior changes in patients with hypertension or diabetes: NHANES 2015-2018. Prev Med 2021 Apr;145:106426. [doi: 10.1016/j.ypmed.2021.106426] [Medline: 33450214]
28. Jackowski RM, Pogge EK, Early NK, Fairman KA, Sclar DA. Provision of lifestyle counseling and the prescribing of pharmacotherapy for hyperlipidemia among US ambulatory patients: a national assessment of office-based physician visits. Am J Cardiovasc Drugs 2018 Feb;18(1):65-71. [doi: 10.1007/s40256-017-0247-y] [Medline: 28849367]
29. Sørensen K, Van den Broucke S, Fullam J, et al. Health literacy and public health: a systematic review and integration of definitions and models. BMC Public Health 2012 Jan 25;12(1):80. [doi: 10.1186/1471-2458-12-80] [Medline: 22276600]
30. De la Cruz Olalde DJ, Vargas Hernandez F, Carro ES, et al. Lifestyle-associated health literacy in subjects with systemic arterial hypertension in a first-level unit. Int J Fam Community Med 2024;8(4):95-99. [doi: 10.15406/ijfcm.2024.08.00359]
31. Persell SD, Karmali KN, Lee JY, et al. Associations between health literacy and medication self-management among community health center patients with uncontrolled hypertension. Patient Prefer Adherence 2020;14:87-95. [doi: 10.2147/PPA.S226619] [Medline: 32021120]
32. Fu SN, Dao MC, Wong CKH, Cheung BMY. The association of health literacy with high‐quality home blood pressure monitoring for hypertensive patients in outpatient settings. Int J Hypertens 2020;2020(1):7502468. [doi: 10.1155/2020/7502468] [Medline: 32292599]
33. Tavakoly Sany SB, Behzhad F, Ferns G, Peyman N. Communication skills training for physicians improves health literacy and medical outcomes among patients with hypertension: a randomized controlled trial. BMC Health Serv Res 2020 Jan 23;20(1):60. [doi: 10.1186/s12913-020-4901-8] [Medline: 31973765]

XSL•FO
RenderX

34. Du S, Zhou Y, Fu C, Wang Y, Du X, Xie R. Health literacy and health outcomes in hypertension: an integrative review. Int J Nurs Sci 2018 Jul 10;5(3):301-309. [doi: 10.1016/j.ijnss.2018.06.001] [Medline: 31406840]
35. Santos RD. Better health literacy can make the difference when control of risk factors for cardiovascular disease and quality of life are concerned. Eur J Prev Cardiol 2017 Nov;24(17):1878-1879. [doi: 10.1177/2047487317735721] [Medline: 28969496]
36. Bonaccorsi G, Modesti PA. Health literacy, a new perspective for patient empowerment in the public health approach to hypertension. Intern Emerg Med 2017 Sep;12(6):737-739. [doi: 10.1007/s11739-017-1657-1] [Medline: 28378279]
37. Dennison Himmelfarb CR, Hughes S. Are you assessing the communication" vital sign"?: Improving communication with our low-health-literacy patients. J Cardiovasc Nurs 2011;26(3):177-179. [doi: 10.1097/JCN.0b013e318213aa2d] [Medline: 21483247]
38. Kim J, Kim S, Lim SC. The impact of health literacy on medication utilization behavior. Korean J Clin Pharm 2025 Mar 31;35(1):65-74. [doi: 10.24304/kjcp.2025.35.1.65]
39. Jae Yeon J, HeeJeung J. Identifying health literacy levels and related factors focusing on age. Korea J Hosp Manag 2024;29(1):64-75 [FREE Full text]
40. Bae H, Hong NS, Shin EK. Impact of health literacy on hypertension control of elderly women patients receiving home visiting health care services. J Health Inform Stat 2019 Aug;44(3):237-244. [doi: 10.21032/jhis.2019.44.3.237]
41. Hirooka N, Kusano T, Kinoshita S, Aoyagi R. Association of health literacy with the prevalence of cardiovascular diseases and their risk factors among older Japanese health management specialists. Gerontol Geriatr Med 2023;9:23337214231189059. [doi: 10.1177/23337214231189059] [Medline: 37492034]
42. Metlock FE, Kwapong YA, Evans C, et al. Design and rationale of the social determinants of the risk of hypertension in women of reproductive age (SAFE HEART) study: an American HEART Association research goes red initiative. Am Heart J 2024 Sep;275:151-162. [doi: 10.1016/j.ahj.2024.05.016] [Medline: 38862074]
43. Pahn J. Influence of health literacy and self-efficacy on compliance in hypertensive patients. J Korean Public Health Nurs 2024 Dec 31;38(3):331-343. [doi: 10.5932/JKPHN.2024.38.3.331]
44. Kim Y, Kong KA. Do hypertensive individuals who are aware of their disease follow lifestyle recommendations better than those who are not aware? PLoS ONE 2015;10(8):e0136858. [doi: 10.1371/journal.pone.0136858]
45. Deeks A, Lombard C, Michelmore J, Teede H. The effects of gender and age on health related behaviors. BMC Public Health 2009 Jun 30;9(1):213. [doi: 10.1186/1471-2458-9-213] [Medline: 19563685]
46. Mari FR, Alves GG, Aerts D, Camara S. The aging process and health: what middle-aged people think of the issue. Rev Bras Geriatr Gerontol 2016 Feb;19(1):35-44. [doi: 10.1590/1809-9823.2016.14122]
47. Halcomb E, Ashley C, Middleton R, et al. Understanding perceptions of health, lifestyle risks and chronic disease in middle age. J Clin Nurs 2021 Aug;30(15-16):2279-2286. [doi: 10.1111/jocn.15711] [Medline: 33594746]
48. Assari S. Life expectancy gain due to employment status depends on race, gender, education, and their intersections. J Racial and Ethnic Health Disparities 2018 Apr;5(2):375-386. [doi: 10.1007/s40615-017-0381-x]
49. van der Put AC, Mandemakers JJ, de Wit JBF, van der Lippe T. Worksite health promotion and social inequalities in health. SSM Popul Health 2020 Apr;10:100543. [doi: 10.1016/j.ssmph.2020.100543] [Medline: 32021901]
50. Moon S, Kim JH, Choi Y, Kim HN. Influencing factors on health literacy and the influence of health literacy on healthcare utilization. Health Soc Welf Rev 2024;44(4):130-150. [doi: 10.15709/hswr.2024.44.4.130]
51. Cabellos-García AC, Castro-Sánchez E, Martínez-Sabater A, et al. Relationship between determinants of health, equity, and dimensions of health literacy in patients with cardiovascular disease. Int J Environ Res Public Health 2020 Mar 20;17(6):2082. [doi: 10.3390/ijerph17062082] [Medline: 32245143]
52. Kjellstrom T, Mercado S. Towards action on social determinants for health equity in urban settings. Environ Urban 2008 Oct;20(2):551-574. [doi: 10.1177/0956247808096128]
53. Phelan JC, Link BG, Tehranifar P. Social conditions as fundamental causes of health inequalities: theory, evidence, and policy implications. J Health Soc Behav 2010;51:S28-S40. [doi: 10.1177/0022146510383498] [Medline: 20943581]
54. Han SH, Kim K, Burr JA. Social support and preventive healthcare behaviors among couples in later life. Gerontologist 2019 Nov 16;59(6):1162-1170. [doi: 10.1093/geront/gny144] [Medline: 30517629]
55. Franks MM, Stephens MAP, Rook KS, Franklin BA, Keteyian SJ, Artinian NT. Spouses' provision of health-related support and control to patients participating in cardiac rehabilitation. J Fam Psychol 2006 Jun;20(2):311-318. [doi: 10.1037/0893-3200.20.2.311] [Medline: 16756407]
56. Gonçalves-Fernández ML, Pino-Juste M. Health literacy in healthy adults: a systematic review of recent evidence. Aten Primaria 2025 Nov;57(11):103300. [doi: 10.1016/j.aprim.2025.103300] [Medline: 40680630]
57. Halladay JR, Donahue KE, Cené CW, et al. The association of health literacy and blood pressure reduction in a cohort of patients with hypertension: the Heart Healthy Lenoir trial. Patient Educ Couns 2017 Mar;100(3):542-549. [doi: 10.1016/j.pec.2016.10.015] [Medline: 27776790]
58. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. Ann Intern Med 2011 Jul 19;155(2):97-107. [doi: 10.7326/0003-4819-155-2-201107190-00005] [Medline: 21768583]

XSL•FO
RenderX

59. Williams MV, Baker DW, Parker RM, Nurss JR. Relationship of functional health literacy to patients' knowledge of their chronic disease. A study of patients with hypertension and diabetes. Arch Intern Med 1998 Jan 26;158(2):166-172. [doi: 10.1001/archinte.158.2.166] [Medline: 9448555]
60. Rothman RL, Housam R, Weiss H, et al. Patient understanding of food labels: the role of literacy and numeracy. Am J Prev Med 2006 Nov;31(5):391-398. [doi: 10.1016/j.amepre.2006.07.025] [Medline: 17046410]
61. National action plan to improve health literacy. : US Department of Health and Human Services, Office of Disease Prevention and Health Promotion; 2010 URL: https://odphp.health.gov/sites/default/files/2019-09/Health_Literacy_Action_Plan.pdf [accessed 2026-03-06]
62. Schillinger D, Piette J, Grumbach K, et al. Closing the loop: physician communication with diabetic patients who have low health literacy. Arch Intern Med 2003 Jan 13;163(1):83-90. [doi: 10.1001/archinte.163.1.83] [Medline: 12523921]
63. Stableford S, Mettger W. Plain language: a strategic response to the health literacy challenge. J Public Health Policy 2007;28(1):71-93. [doi: 10.1057/palgrave.jphp.3200102] [Medline: 17363939]
64. Yen PH, Leasure AR. Use and effectiveness of the teach-back method in patient education and health outcomes. Fed Pract 2019 Jun;36(6):284-289. [Medline: 31258322]

Abbreviations

CVD: cardiovascular disease
DBP: diastolic blood pressure
KDCA: Korea Disease Control and Prevention Agency
KNHANES: Korea National Health and Nutrition Examination Survey
OR: odds ratio
SBP: systolic blood pressure
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Mavragani, T Sanchez; submitted 20.Aug.2025; peer-reviewed by JH Lee, YJ Choi; accepted 24.Feb.2026; published 20.Mar.2026.

Please cite as:
Chun J, Kim DB, Jang SY, Park EC
Association Between Health Literacy and Prehypertension in South Korean Adults: Cross-Sectional Study Using the 2023 Korea National Health and Nutrition Examination Survey
JMIR Public Health Surveill 2026;12:e82684
URL: https://publichealth.jmir.org/2026/1/e82684
doi:10.2196/82684

© Jiyeon Chun, Dan Bi Kim, Suk-Yong Jang, Eun-Cheol Park. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 20.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Gendered Mediation Effects of Social Support on Fertility Intentions Among Childless Adults of Reproductive Age in China: National Cross-Sectional Study

Xinyu Xu[1,2], MPH; Peiyu Liu[2], MM; Peihua Ren[2], MM; Wenyan Shi[1,2], MM; Sha Lai[1], PhD

[1]Health Management and Policy Institute, School of Public Policy and Administration, Xi'an Jiaotong University, No. 74, Xianning West Road, Beilin District, Xi'an, Shaanxi, China

[2]Global Health Institute, Health Science Center, Xi'an Jiaotong University, Xi'an, Shaanxi, China

Corresponding Author:
Sha Lai, PhD
Health Management and Policy Institute, School of Public Policy and Administration, Xi'an Jiaotong University, No. 74, Xianning West Road, Beilin District, Xi'an, Shaanxi, China

Abstract

Background: China is currently facing a low fertility rate, making it crucial to explore the influence of psychosocial factors on fertility intentions to address demographic structural challenges. Social support, as a potentially significant influencing factor, is not yet fully understood in terms of its specific pathways and gender differences.

Objective: This study aimed to explore how social support impacts fertility intentions among Chinese adults aged 20 - 49 years, with an emphasis on gender-specific differences and the mediating roles of self-efficacy and conscientiousness.

Methods: Data were obtained from the Psychology and Behavior Investigation of Chinese Residents (PBICR). This study included 2653 childless adults of reproductive age. A decision tree model was used to identify key factors influencing fertility intentions. A mediation analysis was conducted to explore the mediating effects of self-efficacy and conscientiousness while controlling for demographic confounders.

Results: Among all 2653 participants, 71.3% (1892/2653) had fertility intentions. The proportion was significantly higher in men (weighted 79%, 95% CI 76.5%-81.3%) than in women (weighted 64.5%, 95% CI 61.8-67.1; $P<.001$). Participants with fertility intentions had a higher total social support score (mean 61.25, SD 14.02 vs mean 58.23, SD 13.01; $P=.001$). For women, family support significantly influenced fertility intentions, whereas support from friends was more relevant for men. Mediation analysis revealed that for men, self-efficacy significantly mediated the relationship between social support and fertility intention, with an indirect effect of 0.06 (95% CI 0.04-0.09; $P=.001$) and a mediation proportion of 52.54%. For women, conscientiousness played a significant mediating role, with an indirect effect of 0.011 (95% CI 0.002-0.018; $P=.001$) and a mediation proportion of 10.25%.

Conclusions: Enhancing targeted social support can increase fertility intentions, with implications for addressing demographic challenges. Tailored policies should prioritize providing family support and fostering conscientiousness for women, while boosting self-efficacy and friend-based social support for men.

(JMIR Public Health Surveill 2026;12:e77484) doi:10.2196/77484

KEYWORDS

fertility intention; social support; gender differences; self-efficacy; conscientiousness

Introduction

China is currently facing a profound demographic transition brought about by a low birth rate. According to the latest statistical data, the total number of births in China in 2023 was 9.02 million, with the birth rate dropping to 6.39‰ [1]. This trend has made China one of the countries with the lowest fertility rates in the world, with a total fertility rate of only 1.30 in 2020, significantly lower than the level required to maintain population replacement [2]. The low birth rate directly accelerates the process of social aging: the seventh national census shows that the proportion of people aged 65 years and older has risen from 8.92% in 2010 to 13.52%, with a total of over 190 million people [3]. To address this challenge, China has gradually relaxed birth restrictions since 2015 and implemented policies for 2 and 3 children [4]. However, policy adjustments have not brought about the expected rebound in fertility, and the trend of negative population growth continues [5]. In this context, it is urgent to delve into the influencing factors of reproductive behavior in the Chinese context.

XSL•FO
RenderX

Fertility intention refers to an individual's expectations and attitudes toward having children within a specific period [6]. It includes two aspects: rhythm intention and quantity intention. Rhythm intention involves the timing of childbirth, while quantity intention refers to the number of children expected to be born [7]. The willingness to have children is influenced by various factors, which can be divided into micro- and macrolevels, including personal characteristics, family environment, and social influence [8]. Previous studies have shown that the economic development of a society has a huge impact on fertility intentions and that sociocultural trends also influence people's perceptions of fertility [9]. A study of 6680 students nationwide in China showed that child health services or support were significantly associated with higher fertility intentions [10]. Thus, social support, as an important influencing factor, is gradually receiving attention.

Previous studies have shown that social support is an important factor affecting fertility intentions. It comes from emotional, informational, and instrumental assistance from family, friends, and partners and plays a crucial role in individuals' fertility decisions [11-13]. Specifically, social support can alleviate parenting-related pressures and enhance individuals' confidence in their parenting abilities [14]. Positive family communication and good subjective well-being have also been found to help increase the fertility intention of childless women of childbearing age [15]. However, the role of social support extends beyond its potential direct effect on fertility intentions to its mediating mechanisms. For instance, it may act as a mediator in the relationship between other psychological traits, such as self-efficacy or conscientiousness, and fertility intentions. In addition, due to the different roles and responsibilities assigned to men and women by social expectations and norms, there may be significant gender differences in the impact of social support on fertility intentions [16]. Therefore, the specific psychological pathways through which social support affects fertility intentions, and how these pathways differ between men and women, still need to be clarified through empirical research.

Therefore, this national cross-sectional study is designed to specifically investigate the psychological pathways through which social support influences fertility intentions among childless adults of reproductive age in China, with a central focus on delineating gender-specific mechanisms. It aims to (1) quantify the direct associations between multidimensional social support and fertility intentions, (2) empirically test the mediating roles of key factors within these associations, and (3) explicitly compare the strength and significance of these direct and indirect pathways between man and woman respondents.

Methods

Study Design and Population

Our data were derived from the large-scale cross-sectional Psychology and Behavior Investigation of Chinese Residents (PBICR). The survey used a multistage sampling design to ensure the representativeness and generalizability of the collected data. A total of 23 provinces, 5 autonomous regions, and 4 directly administered municipalities (Beijing, Tianjin, Shanghai, and Chongqing) were directly incorporated into the first stage of sampling in this study. Additionally, 2 to 6 cities in each nonprovincial capital prefecture-level administrative region of each province and autonomous region were randomly selected using a random number table, amounting to a total of 120 cities. In the second stage of sampling, a quota sampling method was used for the residents in each selected community, using quotas based on gender, age, and urban-rural distribution. The gender ratio was stipulated at 1:1, and the age distribution was similar to the age distribution in the Seventh National Population Census of China (2020).

The survey data were acquired through one-on-one interviews, using an electronic questionnaire administered via a networking questionnaire tool. The eligibility criteria for the overarching PBICR survey are shown in Textbox 1. The initial survey involved 11,031 participants.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria:

- Age 18 years and older.
- Chinese citizenship.
- Permanent residency in China (with an annual absence of ≤1 month).
- Voluntary participation with signed informed consent.
- The ability to comprehend questionnaire items and the capacity to complete the survey independently or with noninterferential assistance.

Exclusion criteria:

- Impaired consciousness or psychiatric abnormalities.
- Cognitive impairment.
- Concurrent participation in comparable research.
- Unwillingness to participate.

For this analysis, we applied additional filters to this dataset to define our target subpopulation. We restricted our analysis to childless participants of childbearing age, specifically those aged 20 to 49 years. Consequently, the final analytical sample

XSL•FO
RenderX

for this study comprised 2653 participants. Figure 1 shows the complete sample selection process.

Figure 2 shows the comprehensive experimental design framework for analyzing social support and fertility intentions.

Figure 1. Flowchart of participant selection for the analysis of social support and fertility intentions in a national cross-sectional study.

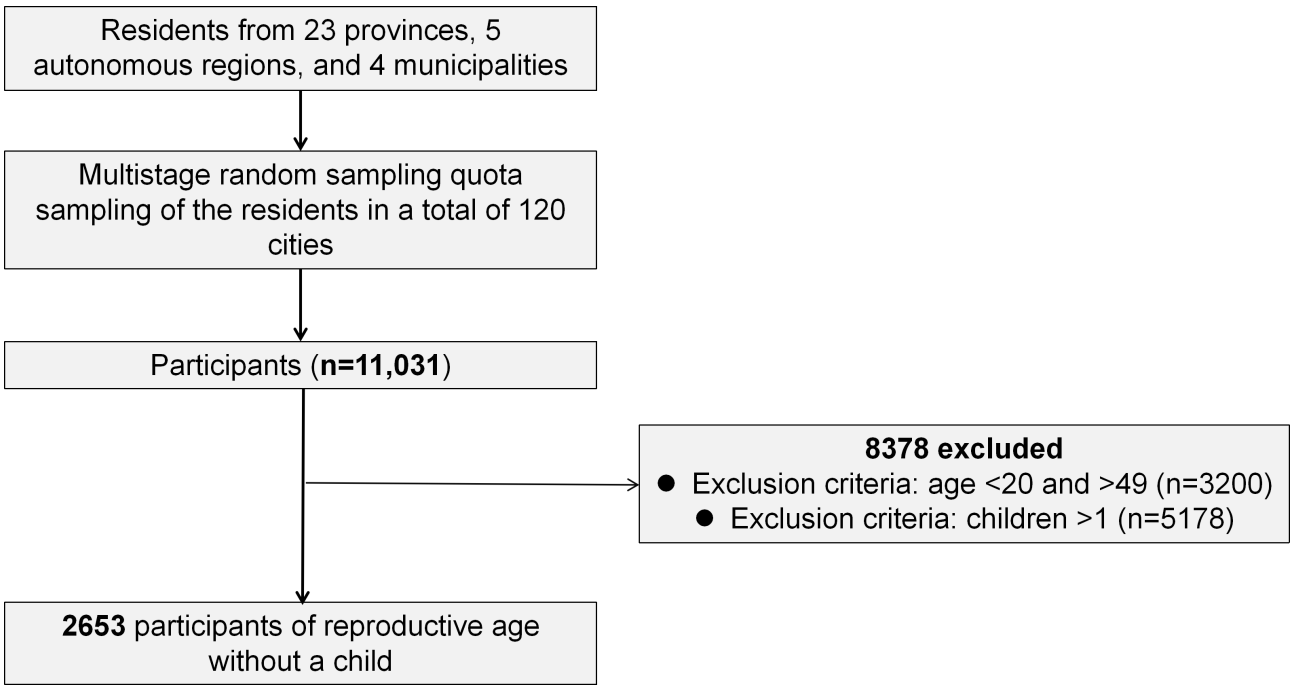
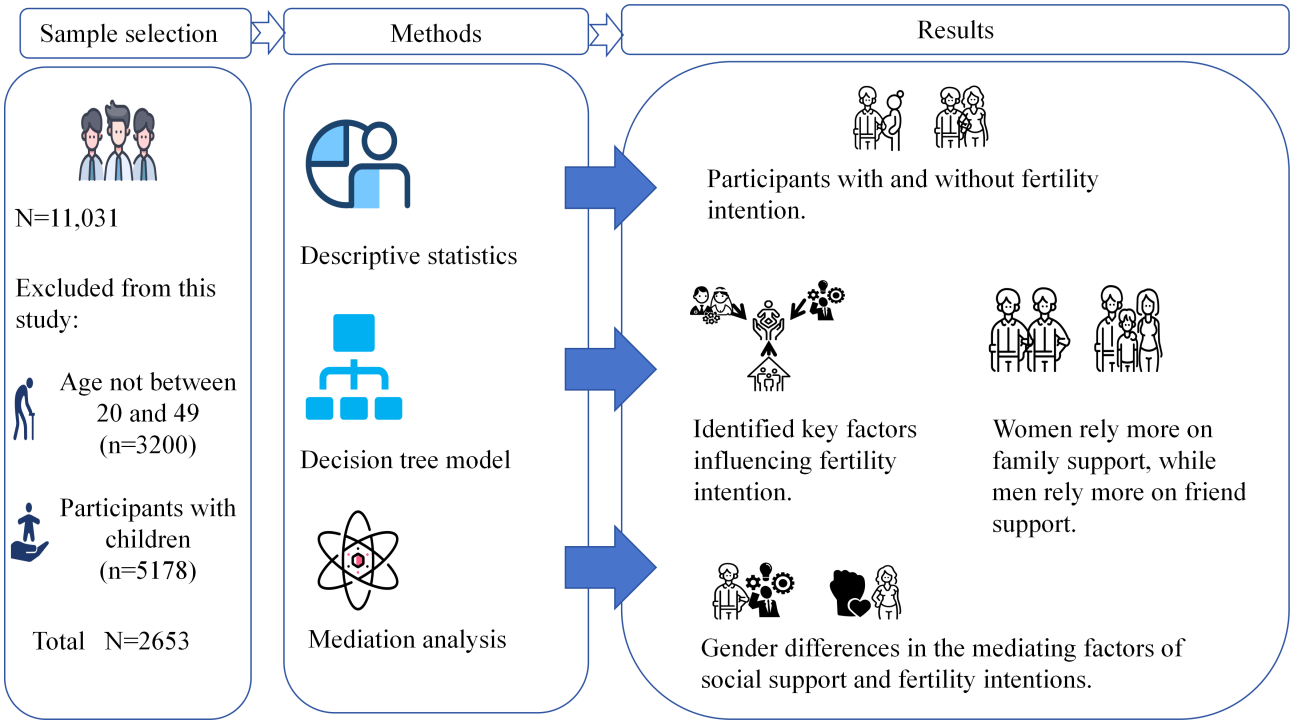


Figure 2. Conceptual framework of the study analyzing the gendered pathways from social support to fertility intention.



Ethical Considerations

This study was reviewed and approved by the institutional review committee of Jinan University (approval number: JNUKY-2021-018). All procedures involving human participants were conducted in accordance with the ethical standards of this committee and with the principles of the WMA Declaration of Helsinki. Prior to participation, informed consent was obtained from all individual participants involved in the study. The consent process ensured that participants were fully informed of the study's nature, purpose, and procedures, and they were advised that their participation was voluntary. To protect participant privacy, all collected data were anonymized at the point of entry using unique identification codes. No personally identifiable information was stored in the research databases. All electronic data were stored securely on a password-protected server with access strictly limited to authorized members of the research team.

XSL•FO
RenderX

Outcome Variable: Fertility Intention

Fertility intention was measured using a single question, which has been widely accepted and used in demographic surveys and fertility studies [17]. The participants were asked: "How strong is your willingness to have your first child?" The response scale consisted of five options, with 1 indicating "completely unwilling," 2 indicating "not willing," 3 indicating "average," 4 indicating "willing," and 5 indicating "strongly willing." For analysis, the fertility intention variable was used as a dichotomized variable. Participants who selected 1 or 2 were categorized as "having no intention to have children," whereas those who selected 3, 4, or 5 were categorized as "having the intention to have children." Unlike the descriptive analysis where the outcome was dichotomized, for the mediation analysis, fertility intention was utilized in its original 1-5 Likert scale form. It was treated as a continuous variable to facilitate linear regression modeling.

Explanatory Variable: Social Support

The participants completed the Perceived Social Support Scale (PSSS) developed by Zimet et al [18] to assess social support. This scale comprises 12 items divided into three dimensions: (1) family support, (2) friend support, and (3) significant other support. Each item is rated on a 7-point Likert scale (1=strongly disagree to 7=strongly agree). Total scores from 12 to 36 indicate a low level of social support, scores from 37 to 60 indicate a moderate level, and scores from 61 to 84 indicate a high level of social support. We conducted separate analyses of each dimension, with each dimension consisting of 4 questions. The average score was calculated to determine the level of support for each dimension, with higher scores indicating higher levels of support. The reliability coefficient for the total social support dimension was 0.96, 0.93 for the friend support dimension, 0.90 for the significant other support dimension, and 0.90 for the family support dimension.

Mediating Variables

The mediating variables of interest in this study were self-efficacy level, personality traits, depression symptoms, and anxiety symptoms of the respondents.

The New General Self-Efficacy Scale was used to measure the respondents' self-efficacy level. It consists of 8 items, and each item was scored on a 5-point Likert scale (1=completely disagree, 2=disagree, 3=neutral, 4=agree, and 5=strongly agree) [19]. All items were scored positively, and the total score on the scale was calculated by summing all item scores and ranged from 8 to 40 points. The higher the score, the higher the self-efficacy level of the respondents. The Cronbach α of the scale in this study was 0.94, indicating that the scale had good reliability in this study.

The Big Five Inventory-10 (BFI-10) was used to assess the personality traits of individuals, including extraversion, agreeableness, conscientiousness, neuroticism, and openness [20,21]. This assessment uses a 5-point Likert scale, with responses ranging from 1 (completely disagree) to 5 (completely agree). A higher score in a particular personality trait indicates a stronger presence of that trait in the respondent. Numerous studies have demonstrated that the BFI-10 possesses strong reliability and validity [22,23]. In our study, the reliability of the BFI-10 was confirmed through Cronbach α analysis, yielding satisfactory results: extraversion (Cronbach α=.80), agreeableness (Cronbach α=.81), conscientiousness (Cronbach α=.86), neuroticism (Cronbach α=.79), and openness to experience (Cronbach α=.90).

The Patient Health Questionnaire-9 (PHQ-9) is a 9-item instrument used to measure the severity of depressive symptoms experienced over the past 2 weeks. Each item is rated on a scale from 0 (not at all) to 3 (nearly every day), with total scores ranging from 0 to 27 [24]. Higher scores indicate more severe depression. The PHQ-9 has been validated and shown to be reliable in various populations [25,26]. In our research, the reliability was 0.94.

The Generalized Anxiety Disorder-7 (GAD-7) tool is a 7-item scale designed to assess the severity of generalized anxiety symptoms over the previous 2 weeks. Items are scored from 0 (not at all) to 3 (nearly every day), with total scores ranging between 0 and 21. Higher scores suggest more severe anxiety [27]. The GAD-7 has demonstrated strong reliability across different groups [28,29]. In our research, the reliability was 0.96.

Control Variables

The control variables included in this study were gender, age, BMI, marital status (married or single), rural or urban residence, number of siblings, education level (primary school or below or middle school or high school or vocational school or junior college or undergraduate or graduate or above), employment status (full-time work or part-time work or no fixed job or retirement), religious beliefs (yes or no), ethnic groups (Han or non-Han), recent medication use (none, 1, 2, 3, 4, or 5 types or more), smoking status (never or current or past), alcohol consumption (never or current or past), and annual per capita household income.

Statistical Analysis

We constructed a decision tree model to identify the key factors influencing fertility intentions. In constructing the decision tree, the splitting criteria were based on minimizing Gini impurity, a standard measure used to quantify node homogeneity. Parameters, such as the minimum number of observations per leaf node (minsplit=20) and the maximum depth of the tree (maxdepth=5), were prespecified to prevent overfitting. These parameters were determined through cross-validation to optimize the trade-off between model complexity and predictive performance. The performance and accuracy of the decision tree model were evaluated using receiver operating characteristic (ROC) curves, which plotted the relationship between the true positive rate and the false positive rate across various classification thresholds. The area under the curve (AUC) of ROC was calculated using the *pROC* package to quantify the model's classification performance, with values closer to 1 indicating better accuracy. We used the *rpart* package in R statistical software (version 4.4.1; R Core Team), which implements the classification and regression tree (CART) algorithm, enabling recursive partitioning of the dataset to generate interpretable decision tree models. We used mediation

XSL•FO
RenderX

analysis using the *mediation* package in R to investigate the mechanisms through which independent variables influence fertility intentions through mediating variables. The parameters for the mediation model were carefully designed. Independent variables and mediators were selected based on theoretical relevance and previous empirical evidence. Continuous variables were standardized to ensure comparability across scales. The significance of the mediating effects was tested using the bootstrap method, providing robust estimates of the indirect effects. These analyses helped to elucidate the pathways through which key variables impact fertility intentions. To account for the complex sampling design and potential deviations from the population benchmarks, all analyses in this study used survey weights. The weights were constructed using a poststratification raking procedure to calibrate the sample to the national population distributions of gender, age group, and urban-rural residence among adults aged 20‑49 years, based on the 2020 Chinese National Population Census. All descriptive statistics (reported as weighted percentages or means with 95% CIs) and inferential analyses (including χ^2 tests and regression models) were performed using these weights to obtain estimates that are representative of the target national subpopulation. All statistical analyses and graphing were performed using R statistical software. Numerical data were subjected to normality tests. Continuous data are presented as mean ± standard deviation, while categorical data are presented as frequency.

Results

Description of the Study Population

Table 1 presents the sociodemographic characteristics of the study participants based on weighted analyses to ensure national representativeness. Among the 2653 childless adults of reproductive age, an estimated 49.2% (95% CI 46.9%-51.5%) were women, and 41.2% (95% CI 38.9%-43.5%) were aged 20‑25 years. Overall, 71.3% (1892/2653) of respondents reported having fertility intentions. The fertility intention was significantly higher among man participants (weighted 79%, 95% CI 76.5%-81.3%) than among women participants (weighted 64.5%, 95% CI 61.8-67.1; $P<.001$). Marital status was also strongly associated with fertility intention ($P<.001$), with weighted 77.2% of those intending to have children being married. Other characteristics that showed significant differences between the groups with and without fertility intentions included age group ($P=.045$), monthly income ($P=.003$), number of siblings ($P=.02$), BMI ($P<.001$), smoking status ($P=.01$), and alcohol use ($P=.04$). Variables including religion, ethnicity, education, employment status, and place of residence showed no significant associations.

XSL•FO
RenderX

Table . Demographic and sociological characteristics of childless adults of reproductive age in China (N=2653).

Variable	Total (N=2653)		No fertility intention (n=761)	Having fertility intention (n=1892)	*P* value
	Unweighted n	Weighted n (95% CI)	Unweighted n (weighted %)	Unweighted n (weighted %)	
Sex					<.001
Man	1192	50.8 (48.5-53.1)	245 (33.5)	947 (49.5)	
Woman	1461	49.2 (46.9-51.5)	516 (66.5)	945 (50.5)	
Age group (years)					.045
20- 25	1301	41.2 (38.9-43.5)	388 (45.1)	913 (39.8)	
26 - 30	898	35.7 (33.5-38.0)	249 (34.2)	649 (36.3)	
31 - 35	286	14.1 (12.5-15.9)	63 (11.2)	223 (15.0)	
36 - 40	86	5.3 (4.2-6.7)	29 (6.1)	57 (5.0)	
41 - 45	53	2.5 (1.8-3.5)	23 (3.0)	30 (2.3)	
46 - 49	29	1.2 (0.8-1.8)	9 (0.4)	20 (1.6)	
Marital status					<.001
Married	1930	78.5 (76.4-80.5)	420 (53.1)	1510 (77.2)	
Single	723	21.5 (19.5-23.6)	341 (46.9)	382 (22.8)	
Religion					.72
No	2585	97.5 (96.7-98.1)	744 (97.7)	1841 (97.4)	
Yes	68	2.5 (1.9-3.3)	17 (2.3)	51 (2.6)	
Ethnicity					.89
Han	2498	94.3 (93.2-95.3)	715 (94.1)	1783 (94.4)	
Other	155	5.7 (4.7-6.8)	46 (5.9)	109 (5.6)	
Education					.85
Primary school or low	35	1.5 (1.0-2.2)	13 (1.8)	22 (1.4)	
Secondary school	36	1.5 (1.1-2.1)	9 (1.3)	27 (1.6)	
High school or technical secondary	226	8.7 (7.6-10.0)	62 (8.2)	164 (8.9)	
Junior college or undergraduate	2055	77.0 (75.0-78.9)	588 (76.8)	1467 (77.1)	
Graduate or above	301	11.3 (10.0-12.7)	89 (11.9)	212 (11.0)	
Employment status					.22
Student	1441	52.1 (49.8-54.4)	439 (55.2)	1002 (51.0)	
Retired	4	0.2 (0.1-0.5)	2 (0.3)	2 (0.1)	
Employed	945	36.8 (34.7-39.0)	248 (33.1)	697 (38.1)	
Unemployed	263	10.9 (9.6-12.3)	72 (11.4)	191 (10.8)	
Monthly income, yuan					.003
≤3000	755	29.1 (27.0-31.3)	221 (30.1)	534 (28.7)	
3001 - 6000	967	36.0 (33.9-38.2)	288 (37.1)	679 (35.6)	
6001 - 9000	466	17.3 (15.7-19.0)	121 (15.2)	325 (18.0)	
>9000	465	17.6 (15.9-19.4)	131 (17.6)	354 (17.7)	
Residence					.22
Rural area	661	31.6 (29.4-33.9)	172 (29.8)	489 (32.2)	

XSL•FO
RenderX

Variable	Total (N=2653)		No fertility intention (n=761)	Having fertility intention (n=1892)	*P* value
	Unweighted n	Weighted n (95% CI)	Unweighted n (weighted %)	Unweighted n (weighted %)	
Urban area	1992	68.4 (66.1-70.6)	589 (70.2)	1403 (67.8)	
Number of siblings					.02
0	1105	40.2 (38.0-42.4)	339 (42.8)	766 (39.2)	
1	1118	42.5 (40.3-44.7)	315 (41.0)	803 (43.0)	
2	284	11.3 (10.0-12.8)	75 (10.1)	209 (11.8)	
≥3	146	6.0 (5.1-7.1)	32 (6.1)	114 (6.0)	
BMI (kg/m²), mean (95% CI)	2653	21.5 (21.3-21.7)	20.9 (20.6-21.2)	21.7 (21.5-21.9)	<.001
Smoking status					.01
Never	2316	86.5 (84.8-88.0)	685 (89.1)	1631 (85.6)	
Current	242	10.1 (8.9-11.5)	48 (6.7)	194 (11.3)	
Past	95	3.4 (2.8-4.2)	28 (4.2)	67 (3.1)	
Alcohol use					.04
Never	1462	54.3 (52.0-56.6)	438 (56.8)	1024 (53.4)	
Current	773	30.5 (28.5-32.6)	193 (26.1)	580 (32.1)	
Past	418	15.2 (13.7-16.8)	130 (17.1)	288 (14.5)	

All percentages, means, and *P* values are calculated based on survey weights to ensure that the sample is consistent with the distribution of the 20‑49 age group in the 2020 Chinese National Population Census in terms of gender, age group, and urban-rural distribution. Continuous variables are represented by weighted means and their 95% CIs. The *P* value of categorical variables is based on weighted Rao Scott χ^2 test, while the *P* value of continuous variables is based on weighted linear regression.

Analysis accounting for the complex survey design indicated significant associations between fertility intention and several psychosocial characteristics. Weighted scores for self-efficacy (28.85, SD 5.51 vs 27.65, SD 5.73; *P*=.001), conscientiousness (6.38, SD 1.50 vs 6.21, SD 1.53; *P*=.002), total social support (61.25, SD 14.02 vs 58.23, SD 13.01; *P*=.001), family support (4.99, SD 1.25 vs 4.85, SD 1.30; *P*=.001), significant other support (4.98, SD 1.20 vs 4.86, SD 1.27; *P*=.002), and friend support (5.09, SD 1.20 vs 5.02, SD 1.28; *P*=.049) were significantly higher in the group with fertility intention. Conversely, weighted scores for life stress (3.13, SD 1.49 vs 3.37, SD 1.46; *P*=.048) and neuroticism (5.78, SD 1.51 vs 6.12, SD 1.59; *P*=.001) were significantly lower in this group. No statistically significant differences were observed for ability to handle stress, depression, anxiety, extraversion, agreeableness, or openness. Complete data are presented in Table 2.

XSL•FO
RenderX

Table . Distribution of psychological factors, personality traits, and social support scores by fertility intention status among childless adults in China (N=2653).

Characteristic	Overall	Fertility intention		*P* value
	N=2653	Yes (n=1892)	No (n=761)	
Psychological factors				
Life stress	3.20 (1.48)	3.13 (1.49)	3.37 (1.46)	.048
Ability to handle stress	3.12 (1.65)	3.08 (1.67)	3.31 (1.61)	.08
Depression	7.43 (6.35)	7.35 (6.50)	7.72 (6.01)	.44
Anxiety	5.34 (5.26)	5.29 (5.35)	5.55 (5.04)	.60
Self-efficacy	28.48 (5.62)	28.85 (5.51)	27.65 (5.73)	.001
Personality traits				
Extraversion	6.76 (1.55)	6.71 (1.52)	6.91 (1.60)	.10
Conscientiousness	6.31 (1.51)	6.38 (1.50)	6.21 (1.53)	.002
Neuroticism	5.86 (1.54)	5.78 (1.51)	6.12 (1.59)	.001
Agreeableness	6.84 (1.52)	6.88 (1.54)	6.75 (1.48)	.06
Openness	6.19 (1.72)	6.22 (1.69)	6.16 (1.78)	.09
Social supports				
Family dimension score	4.93 (1.27)	4.99 (1.25)	4.85 (1.30)	.001
Friend dimension score	5.05 (1.23)	5.09 (1.20)	5.02 (1.28)	.049
Significant others dimension score	4.92 (1.22)	4.98 (1.20)	4.86 (1.27)	.002
Total social support score	60.38 (13.85)	61.25 (14.02)	58.23 (13.01)	.001

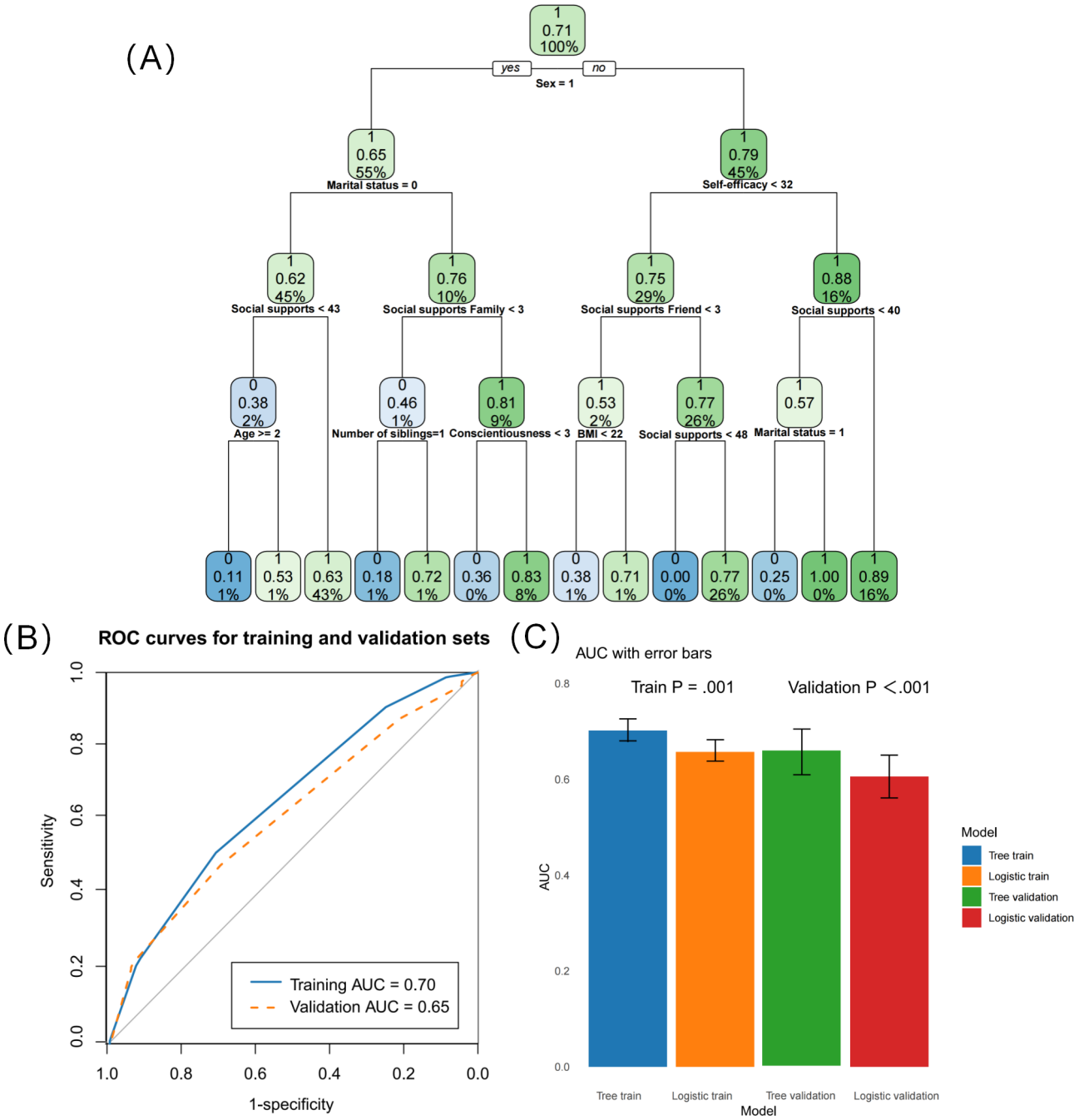
Data are presented as weighted mean (weighted SD). All estimates (means, SDs) and *P* values were derived from analyses that accounted for the complex, multistage sampling design by applying survey weights. The weights were constructed using poststratification raking to calibrate the sample to the national population distributions of gender, age, and urban-rural residence for adults aged 20‑49 years, based on the 2020 Chinese National Population Census. *P* values for comparisons between groups (Yes vs No fertility intention) were obtained from weighted linear regression models. The sample sizes (n) represent the unweighted number of participants in each group.

Association Between Fertility Intentions and Social Support in the Decision Tree

The fertility intention rate of women was lower than that of men (65% vs 79%, Figure 3A). The results of our study showed that marital status was the most significant variable associated with fertility intentions among women. For women, being married and having strong family support (family support >3) was associated with a fertility intention rate of 81%. Additionally, respondents whose marital status=0 (married) and social supports of family >3 and reported Conscientiousness >3 had a higher proportion of fertility intention than married individuals with high family support but low conscientiousness (83% vs 36%). For those whose marital status was 0 but had the support of <4 family members, the proportion of fertility intention increased when the number of siblings was 1 (72% vs 18%). Our study found self-efficacy to be the strongest predictor of fertility intention among men. When self-efficacy >32 and social support >40, there was a high proportion of fertility intention (89%). Respondents with self-efficacy scores of <32, social support of friends >3, and social support >48 also presented high fertility intention rates, at 77%.

XSL•FO
RenderX

Figure 3. Decision tree model for predicting fertility intention and its performance evaluation among childless adults in China. (A) The DT model for predicting fertility intention. (B) ROC curves for training and validation sets. (C) AUC with error bars for training and validation sets. AUC: area under the curve; DT: decision tree; ROC: receiver operating characteristic.



The decision tree model demonstrated strong predictive accuracy for fertility intention, with an AUC of 0.70, indicating good overall model performance. The model showed excellent diagnostic ability in both the training set (AUC=0.70) and test set (AUC=0.65), confirming its ability to generalize across different datasets (Figures 3B and 3C).

Figure 4 indicates that increased social support correlated with higher fertility intentions for both genders, but women generally had lower intentions than men at the same support levels (Figure 4A). For women, family was the most crucial form of social support, while for men, friends held the highest importance (Figures 4B and 4C). Therefore, we explored different mediation pathways for men and women.

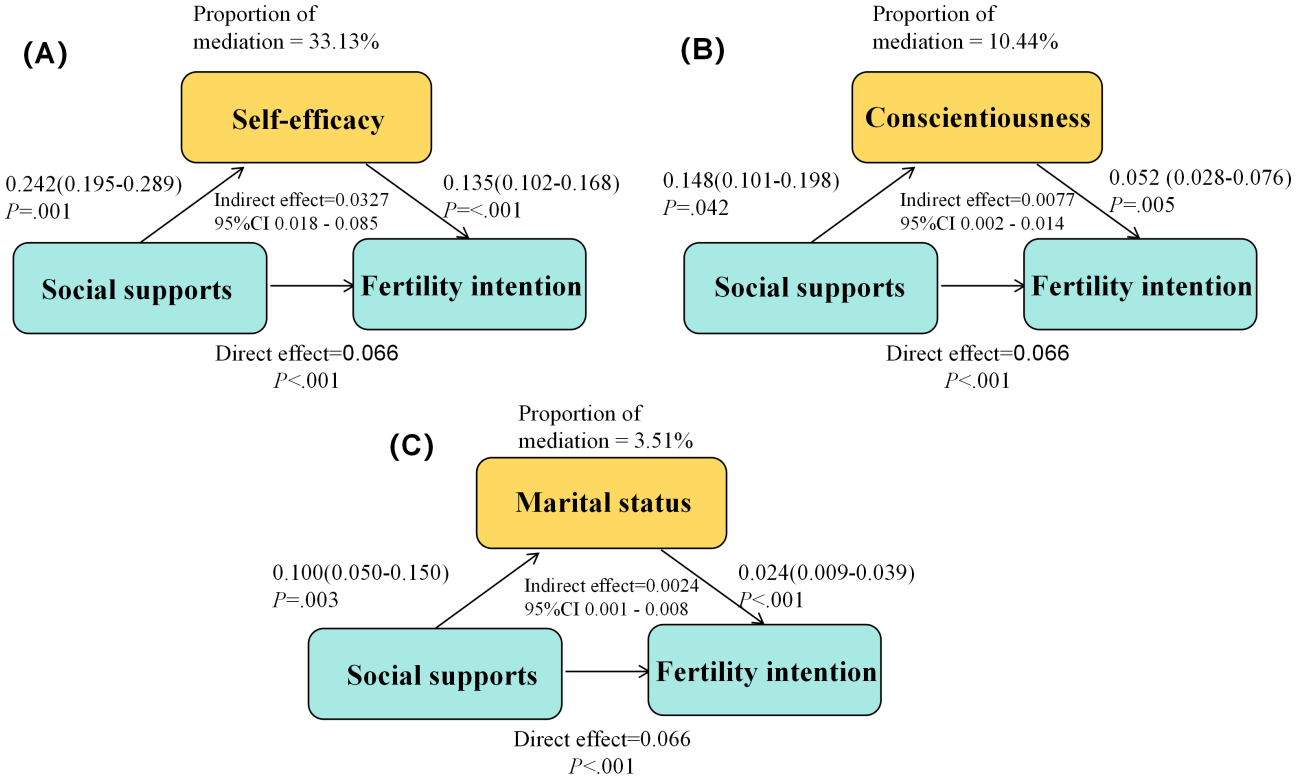
XSL•FO
RenderX

Figure 4. Gender differences in fertility intentions and the relative importance of different sources of social support. (A) Line chart of gender differences in fertility intentions. The horizontal axis represents the quartiles of social support, and the vertical axis represents the fertility intention percentage in this population. (B) The importance ranking of the 3 dimensions of fertility intention among men. (C) The importance ranking of the 3 dimensions of fertility intention in women.

XSL•FO
RenderX

The results showed significant indirect effects of conscientiousness (mediation proportion: 10.44%), self-efficacy (mediation proportion: 33.13%), and marital status (mediation proportion: 3.51%) on the association between social support and fertility intention (Figure 5).

Figure 5. Mediation analyses examining the indirect effects of conscientiousness, self-efficacy, and marital status on the relationship between social support and fertility intention. Statistical results from mediation analyses for (A) self-efficacy (B), conscientiousness, and (C) marital status as mediators. Each panel reports the Average Causal Mediation Effect (ACME; ie, the indirect effect), the average direct effect (ADE), and the proportion of the total effect mediated, quantifying the contribution of each factor in the mediation pathway. Mediation analysis adjusted for demographic confounders



The mediation analysis revealed gender differences in the mediating effects of self-efficacy and conscientiousness on fertility intention (Table 3). Specifically, self-efficacy played a significant mediating role in the relationship between social support and fertility intention for men, with a mediation effect of 0.06 (95% CI 0.04 to 0.09; *P*<.001) and a mediation proportion of 52.54%. In women, self-efficacy did not demonstrate a significant mediation effect (95% CI −0.02 to 0.04; *P*=.85).

Table . Gender-stratified mediation analysis of the effect of social support on fertility intention through self-efficacy and conscientiousness among childless Chinese adults.

Exposure[a]	Mediator[b]	Outcome	Man					Woman				
			Total effect	Direct effect	Mediation effect (95% CI)	*P* value	Mediation proportion, %	Total effect	Direct effect	Mediation effect (95% CI)	*P* value	Mediation proportion, %
Social supports	Self-efficacy	Fertility intention	0.118	0.055	0.062 (0.040 to 0.087)	<.001	52.54	0.100	0.097	0.003 (−0.023 to 0.043)	.85	0
Social supports	Conscientiousness	Fertility intention	0.119	0.056	0.001 (−0.014 to 0.011)	.95	0	.100	0.089	0.011 (0.002 to 0.018)	<.001	10.25

[a]The Bootstrap method is used to test the mediating effect, and the 95% confidence interval (95% CI) of the mediating effect is estimated through repeated sampling (with 5000 repetitions).

[b]Mediation analysis adjusted for demographic confounders.

For conscientiousness, the mediation effect was significant in women, with a mediation effect of 0.011 (95% CI 0.002 to 0.018; *P*<.001) and a mediation proportion of 10.25%. In men, conscientiousness did not significantly mediate the relationship (95% CI –0.014 to 0.011, *P*=.95).

XSL•FO
RenderX

Adjustments were made for some confounding factors, including age, ethnicity, the amount of medication taken (excluding health supplements), smoking, BMI, educational level, alcohol consumption, annual income, and marital status.

Discussion

Principal Findings

This study examined the influence of social support on fertility intentions in Chinese adults, focusing on gender differences and the mediating effects of self-efficacy and conscientiousness, and revealed 3 key findings. First, social support was positively linked to fertility intention, suggesting that individuals with higher levels of support are more likely to express a desire to have children. Second, self-efficacy, conscientiousness, and marital status were significant mediators of the relationship between social support and fertility intention. Self-efficacy was especially relevant in men, while conscientiousness was more impactful in women. Third, a gender difference was seen in fertility intention. At the same level of social support, women had lower fertility intentions than men. However, the type of support mattered; family support was more critical for women, while men benefited significantly from friend-based support. This relationship was mediated by self-efficacy in men and conscientiousness in women. The study findings highlight important pathways for enhancing fertility intentions among men and women by enhancing social support, especially in strategy development for gender differences.

Our study revealed a positive relationship between social support and fertility intention, highlighting that individuals with stronger social networks are more likely to express a desire to have children. This finding aligns with previous research indicating that emotional and instrumental support can foster the intention to start or expand a family. Higher levels of social support, including support from family and friends, may provide the psychological and practical resources needed to manage the demands of child-rearing, thus promoting fertility intentions [30-32].

Our study identified self-efficacy, conscientiousness, and marital status as key mediators in the relationship between social support and fertility intention. Self-efficacy, which refers to one's belief in one's ability to achieve goals, was particularly influential for men, whereas conscientiousness, the tendency to be organized and responsible, had a stronger mediating effect in women. This supports earlier findings that personality traits and psychological factors can significantly influence fertility decisions [33]. In addition, marital status moderated this effect, with married individuals showing stronger fertility intentions than single people [30]. Together, these mediators emphasize the importance of both psychological resources and relationship status in shaping fertility intentions, providing critical pathways for interventions aimed at enhancing fertility decisions.

The study findings also highlight important gender differences in fertility intention. At equivalent levels of social support, women demonstrated lower fertility intentions than men. This phenomenon may reflect the potential influence of sociocultural and gender roles on woman fertility decisions, consistent with other studies [34-36]. Okun et al [35] suggest that Confucian culture profoundly influences China, and women take on traditional family service roles in their households. Therefore, Chinese women shoulder the dual responsibility of social work and family care and face a career fertility dilemma, an essential factor that reduces the fertility intentions of women. Furthermore, the study found that the type of social support mattered. For women, family support emerged as the most influential, while men benefited more from support from friends. These gender-specific differences in the role of support were mediated by self-efficacy in men and conscientiousness in women. These gender differences may stem from the traditional division of gender roles, where women take more responsibility and pressure in the family and, thus, are more dependent on family support when making reproductive decisions. In contrast, men are less involved in the family domain [37-39]. In many cultures, men are often expected to play the role of breadwinner and decision-maker in the family [40]. Thus, their self-efficacy is closely related to their confidence in their abilities. Social support enhances men' sense of self-efficacy and makes them more confident in their fatherly roles and responsibilities. For women, the mediating role of responsibility between social support and fertility intentions may be related to traditional societal role expectations. Women are usually expected to take more responsibility for caring for and educating the children in the family, and therefore, their fertility intentions are closely related to their sense of responsibility [40]. Social support can help women feel more resourceful and emotionally supported, thus enhancing their sense of responsibility and making them more willing to take on the role of motherhood. Although gender roles have diversified and become more flexible with time, traditional gender role expectations still play a role.

In summary, our study findings underscore the critical role of social support in shaping fertility intentions, with self-efficacy, conscientiousness, and marital status identified as significant mediators. Importantly, gender differences were evident in how support influences reproductive decisions, with distinct pathways for men and women. This suggests that future strategies to promote fertility intentions should consider gender-specific needs and personality traits, particularly by enhancing the types of support most relevant to each gender. Policymakers and health professionals could focus on promoting family and friend support in ways that align with individual gender dynamics and personality characteristics, potentially leading to more effective fertility interventions. For example, the fertility intentions of women can be enhanced by strengthening family support and developing a sense of responsibility, whereas for men, fertility intentions can be enhanced by boosting self-efficacy and friend support.

Limitations

However, the study has some limitations. This study has several limitations that should be considered. First, the observational cross-sectional design hinders causal inference regarding the relationships between social support, psychological mechanisms, and fertility intentions. Future longitudinal studies are needed to verify the directionality and temporal dynamics of these associations. Second, while statistically significant, some of the observed effect sizes were modest. Their practical significance

XSL•FO
RenderX

should be interpreted with caution.While social support, self-efficacy, and conscientiousness are significant predictors, they account for limited variance in fertility intentions. This indicates that fertility decisions are heavily influenced by unmeasured factors beyond the psychosocial domain, such as economic constraints (e.g., housing, childcare costs) and public policies. Future research should integrate these macro-level variables to build more comprehensive models. Third, the generalizability of our findings may be limited as the study specifically focused on childless individuals of reproductive age. The applicability of the proposed model to other populations requires further examination. Finally, a key methodological limitation is the use of a single-item measure for fertility intention. Although practical in large-scale surveys, such a measure may not fully capture the complexity and multifaceted nature of reproductive planning (eg, timing and desired number of children), which could influence the depth and nuance of our conclusions.

Conclusion

The study revealed that individuals with greater support are more likely to want children, with self-efficacy, conscientiousness, and marital status significantly influencing the link between social support and fertility intentions. Additionally, women were generally found to have lower intentions than men at the same level of social support. Family support was more crucial for women, whereas men benefited more from friend-based support. For men, this relationship was influenced by self-efficacy, whereas for women, it was affected by conscientiousness. This study enhances our understanding of gender differences in fertility intentions and offers practical guidelines for policymakers to improve these intentions by increasing social support and psychological factors.

Acknowledgments

The authors sincerely thank the participants of the Psychology and Behavior Investigation of Chinese Residents for their contributions. We also acknowledge the research assistants and collaborators for their support in data collection and analysis. No generative artificial intelligence was used to generate, analyze, or write the study's content.

Funding

This study was funded by the National Natural Science Foundation of China (72004180).

Data Availability

The datasets analyzed during this study are available in the Psychology and Behavior Investigation of Chinese Residents repository [41].

Authors' Contributions

Conceptualization: XX (lead), PL (equal)
Methodology: XX (lead), PR (equal)
Resources: SL (lead), PL (supporting), WS (supporting)
Formal analysis: XX (lead), PL (equal), PR (supporting), WS (supporting)
Investigation: PR (lead), WS (equal)
Visualization: XX (lead), PR (equal), PL (supporting)
Writing – original draft: XX (lead), PR (equal), WS (supporting)
Writing – review and editing: SL (lead), XX (supporting), PL (supporting), WS (supporting)

Conflicts of Interest

None declared.

References

1. Statistical communiqué of the People's Republic of China on the 2023 national economic and social development [Website in Chinese]. National Bureau of Statistics of China. URL: https://www.stats.gov.cn/sj/zxfb/202402/t20240228_1947915.html [accessed 2026-02-04]
2. Bulletin of the seventh national population census [Website in Chinese]. National Bureau of Statistics of China. 2021. URL: https://www.stats.gov.cn/sj/pcsj/rkpc/d7c/ [accessed 2026-02-04]
3. Hanmo Y. Dynamic trend of China's population ageing and new characteristics of the elderly. Popul Res 2022 Sep 29;46(5):104 [FREE Full text]
4. Wang Y, Kong F, Qiao J. Gender equity, caregiving, and the 1-2-3-child policy in China - authors' reply. Lancet 2021 Sep 11;398(10304):953-954. [doi: 10.1016/S0140-6736(21)01749-9] [Medline: 34509225]
5. Wang Y, Kong F, Fu Y, Qiao J. How can China tackle its declining fertility rate? BMJ 2024 Aug 30;386:e078635. [doi: 10.1136/bmj-2023-078635] [Medline: 39214553]

XSL•FO
RenderX

6. Jing W, Liu J, Ma Q, Zhang S, Li Y, Liu M. Fertility intentions to have a second or third child under China's three-child policy: a national cross-sectional study. Hum Reprod 2022 Jul 30;37(8):1907-1918. [doi: 10.1093/humrep/deac101] [Medline: 35554542]
7. Balbo N, Billari FC, Mills M. Fertility in advanced societies: a review of research [La fécondité dans les sociétés avancées: un examen des recherches]. Eur J Popul 2013 Feb;29(1):1-38. [doi: 10.1007/s10680-012-9277-y] [Medline: 23440941]
8. Preis H, Tovim S, Mor P, Grisaru-Granovsky S, Samueloff A, Benyamini Y. Fertility intentions and the way they change following birth- a prospective longitudinal study. BMC Pregnancy Childbirth 2020 Apr 17;20(1):228. [doi: 10.1186/s12884-020-02922-y] [Medline: 32303192]
9. Morgan SP, Taylor MG. Low fertility at the turn of the twenty-first century. Annu Rev Sociol 2006 Aug 1;32:375-399. [doi: 10.1146/annurev.soc.31.041304.122220] [Medline: 20376287]
10. Zhang C, Wei L, Zhu Y, et al. Fertility intentions among young people in the era of China's three-child policy: a national survey of university students. BMC Pregnancy Childbirth 2022 Aug 12;22(1):637. [doi: 10.1186/s12884-022-04873-y] [Medline: 35962340]
11. Shenk MK, Kaplan HS, Hooper PL. Status competition, inequality, and fertility: implications for the demographic transition. Philos Trans R Soc Lond B Biol Sci 2016 Apr 19;371(1692):20150150. [doi: 10.1098/rstb.2015.0150] [Medline: 27022077]
12. Innocenti N, Vignoli D, Lazzeretti L. Economic complexity and fertility: insights from a low fertility country. Reg Stud 2021;55(8):1388-1402. [doi: 10.1080/00343404.2021.1896695] [Medline: 34381283]
13. Yu X, Liang J. Social norms and fertility intentions: evidence from China. Front Psychol 2022;13:947134. [doi: 10.3389/fpsyg.2022.947134] [Medline: 36467244]
14. Simons RL, Johnson C. The impact of marital and social network support on quality of parenting. In: Pierce GR, Sarason BR, Sarason IG, editors. Handbook of Social Support and the Family: Springer US; 1996:269-287. [doi: 10.1007/978-1-4899-1388-3_12]
15. Zhao J, Qi W, Cheng Y, et al. Influence of perceived stress on fertility intention among women of childbearing age without children: multiple mediating effect of anxiety, family communication and subjective well-being. Reprod Health 2024 Sep 18;21(1):135. [doi: 10.1186/s12978-024-01855-5] [Medline: 39294666]
16. Ahinkorah BO, Seidu AA, Budu E, et al. Which factors predict fertility intentions of married men and women? Results from the 2012 Niger Demographic and Health Survey. PLoS One 2021;16(6):e0252281. [doi: 10.1371/journal.pone.0252281] [Medline: 34106940]
17. Zeng T, Li B, Zhang K, et al. The association between childbirth-related fear, childbirth readiness, and fertility intentions, and childbirth readiness as the mediator. Reprod Health 2023 Apr 21;20(1):62. [doi: 10.1186/s12978-023-01607-x] [Medline: 37085786]
18. Zimet GD, Powell SS, Farley GK, Werkman S, Berkoff KA. Psychometric characteristics of the Multidimensional Scale of Perceived Social Support. J Pers Assess 1990;55(3-4):610-617. [doi: 10.1080/00223891.1990.9674095] [Medline: 2280326]
19. Chen G, Gully SM, Eden D. Validation of a new general self-efficacy scale. Organ Res Methods 2001 Jan;4(1):62-83. [doi: 10.1177/109442810141004]
20. Rammstedt B, John OP. Measuring personality in one minute or less: a 10-item short version of the Big Five Inventory in English and German. J Res Pers 2007 Feb;41(1):203-212. [doi: 10.1016/j.jrp.2006.02.001]
21. Rammstedt B. The 10-Item Big Five Inventory. Eur J Psychol Assess 2007 Jan;23(3):193-201. [doi: 10.1027/1015-5759.23.3.193]
22. Carciofo R, Yang J, Song N, Du F, Zhang K. Psychometric evaluation of Chinese-language 44-item and 10-item big five personality inventories, including correlations with chronotype, mindfulness and mind wandering. PLoS One 2016;11(2):e0149963. [doi: 10.1371/journal.pone.0149963] [Medline: 26918618]
23. Kunnel John R, Xavier B, Waldmeier A, Meyer A, Gaab J. Psychometric evaluation of the BFI-10 and the NEO-FFI-3 in Indian adolescents. Front Psychol 2019;10:1057. [doi: 10.3389/fpsyg.2019.01057] [Medline: 31143151]
24. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9. J Gen Intern Med 2001 Sep;16(9):606-613. [doi: 10.1046/j.1525-1497.2001.016009606.x]
25. Levis B, Benedetti A, Thombs BD, DEPRESsion Screening Data (DEPRESSD) Collaboration. Accuracy of Patient Health Questionnaire-9 (PHQ-9) for screening to detect major depression: individual participant data meta-analysis. BMJ 2019 Apr 9;365:l1476. [doi: 10.1136/bmj.l1476] [Medline: 30967483]
26. Negeri ZF, Levis B, Sun Y, et al. Accuracy of the Patient Health Questionnaire-9 for screening to detect major depression: updated systematic review and individual participant data meta-analysis. BMJ 2021 Oct 5;375:n2183. [doi: 10.1136/bmj.n2183] [Medline: 34610915]
27. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
28. Shih YC, Chou CC, Lu YJ, Yu HY. Reliability and validity of the traditional Chinese version of the GAD-7 in Taiwanese patients with epilepsy. J Formos Med Assoc 2022 Nov;121(11):2324-2330. [doi: 10.1016/j.jfma.2022.04.018] [Medline: 35584970]

XSL•FO
RenderX

29. Martínez-Vázquez S, Martínez-Galiano JM, Peinado-Molina RA, Gutiérrez-Sánchez B, Hernández-Martínez A. Validation of General Anxiety Disorder (GAD-7) questionnaire in Spanish nursing students. PeerJ 2022;10:e14296. [doi: 10.7717/peerj.14296] [Medline: 36340193]
30. Hashemzadeh M, Shariati M, Mohammad Nazari A, Keramat A. Childbearing intention and its associated factors: a systematic review. Nurs Open 2021 Sep;8(5):2354-2368. [doi: 10.1002/nop2.849] [Medline: 33705606]
31. Park SM. Social networks and second-childbirth intentions of Korean married women. J Reprod Infant Psychol 2012 Sep;30(4):398-412. [doi: 10.1080/02646838.2012.725129]
32. Yang Y, Yang Y, Xia R, et al. “Hit a bottleneck”: a convergent mixed-methods approach to the Influencing factors and experience of fertility intention among mothers of childbearing age with two children in mainland China. Reprod Health 2024 Nov 5;21(1):159. [doi: 10.1186/s12978-024-01888-w] [Medline: 39501305]
33. Peters S. The prospective power of personality for childbearing: a longitudinal study based on data from Germany. Genus 2023 Feb;79(1):6. [doi: 10.1186/s41118-023-00184-y]
34. Goldscheider F, Bernhardt E, Lappegård T. The gender revolution: a framework for understanding changing family and demographic behavior. Population & Development Rev 2015 Jun;41(2):207-239. [doi: 10.1111/j.1728-4457.2015.00045.x]
35. Okun BS, Raz‐Yurovich L. Housework, gender role attitudes, and couples’ fertility intentions: reconsidering men’s roles in gender theories of family change. Population & Development Rev 2019 Mar;45(1):169-196. [doi: 10.1111/padr.12207]
36. Raybould A, Sear R. Children of the (gender) revolution: a theoretical and empirical synthesis of how gendered division of labour influences fertility. Popul Stud (Camb) 2021 Jul;75(2):169-190. [doi: 10.1080/00324728.2020.1851748] [Medline: 33321056]
37. Kim HW, Kim SY. Gender differences in willingness for childbirth, fertility knowledge, and value of motherhood or fatherhood and their associations among college students in South Korea, 2021. Arch Public Health 2023 Jun 16;81(1):110. [doi: 10.1186/s13690-023-01127-x] [Medline: 37328899]
38. Li X. Fathers’ involvement in Chinese societies: increasing presence, uneven progress. Child Dev Perspect 2020 Sep 1;14(3):150-156. [doi: 10.1111/cdep.12375]
39. Shen Y, Jiao L. “Widow-style parenting”: the discursive formation of gendered parenting roles and popular feminism on social media in China. Chinese Journal of Communication 2024 Oct;17(4):399-415. [doi: 10.1080/17544750.2024.2329063]
40. Farre L. The role of men in the economic and social development of women: implications for gender equality. World Bank Res Obs 2013 Feb 1;28(1):22-51. [doi: 10.1093/wbro/lks010]
41. Psychology and behavior investigation of Chinese residents (PBICR) [Website in Chinese]. URL: https://www.x-mol.com/groups/pbicr [accessed 2025-01-16]

Abbreviations

AUC: area under the curve
BFI-10: Big Five Inventory-10
CART: classification and regression tree
GAD-7: Generalized Anxiety Disorder-7
PBICR: Psychology and Behavior Investigation of Chinese Residents
PHQ-9: Patient Health Questionnaire-9
PSSS: Perceived Social Support Scale
ROC: receiver operating characteristic

Edited by A Mavragani, T Sanchez; submitted 14.May.2025; peer-reviewed by MS Shafi, Y Wu; revised version received 29.Dec.2025; accepted 30.Dec.2025; published 11.Feb.2026.

Please cite as:
Xu X, Liu P, Ren P, Shi W, Lai S
The Gendered Mediation Effects of Social Support on Fertility Intentions Among Childless Adults of Reproductive Age in China: National Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e77484
URL: https://publichealth.jmir.org/2026/1/e77484
doi:10.2196/77484

© Xinyu Xu, Peiyu Liu, Peihua Ren, Wenyan Shi, Sha Lai. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 11.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete

XSL•FO
RenderX

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

XSL•FO
RenderX

Survey Modalities and COVID-19 Vaccine Uptake in Vietnamese Americans: Cross-Sectional Study

Celine Nguyen[1], BA; Alexander Le[2], BA, MD; Paul Yeh[3], MPH, MD; Ben King[3], MPH, PhD; Angelica Nguyen[4], BA; Jannette Diep[4], BS; Bich-May Nguyen[3], MPH, MD

[1]The University of Texas Southwestern Medical Center, Dallas, TX, United States

[2]College of Medicine, Texas A&M University, College Station, TX, United States

[3]Tilman J Fertitta Family College of Medicine, University of Houston, 5055 Medical Circle, Houston, TX, United States

[4]Boat People SOS, Houston, TX, United States

Corresponding Author:
Bich-May Nguyen, MPH, MD
Tilman J Fertitta Family College of Medicine, University of Houston, 5055 Medical Circle, Houston, TX, United States

Abstract

Background: Diverse survey methodologies are essential to ensure equitable representation in public health research, particularly among minority populations. This study evaluates demographic differences among Vietnamese Americans who completed paper versus electronic surveys while administering the National Institutes of Health Community Engagement Alliance Common Survey 2, which focused on COVID-19–related topics.

Objective: The study aimed to (1) describe the sociodemographic characteristics of survey respondents; (2) compare paper versus electronic survey modalities and their associations with respondent demographics; and (3) highlight the role of community-based organizations (CBOs) and community-engaged research in improving representativeness and inclusivity.

Methods: Vietnamese adults in Texas were recruited in two phases. In phase 1 (September 2021-March 2022), surveys were administered online. In response to CBOs' feedback, phase 2 (December 2022-April 2023) added paper surveys administered by bilingual recruiters; surveys were available in English and Vietnamese. Descriptive analyses were conducted for both phases. Multivariate logistic regression, limited to phase 2, assessed factors associated with survey method preferences, including language, sex, education, COVID-19 history, and willingness to participate in COVID-19 trials.

Results: Phase 1 included 224 electronic surveys. Phase 2 included 359 surveys (electronic: n=124, 34.5% and paper: n=235, 65.5%). Vietnamese speakers were significantly more likely to complete paper surveys (adjusted odds ratio [AOR] 100.9, 95% CI 24.3 - 418.9; $P<.001$), as were female participants (AOR 5.09, 95% CI 1.43 - 18.1; $P=.01$). Conversely, those with a history of COVID-19 (AOR 0.16, 95% CI 0.05 - 0.52; $P=.002$), a college or higher education level (AOR 0.18, 95% CI 0.05 - 0.67; $P=.01$), and high willingness to participate in COVID-19 trials (AOR 0.21, 95% CI 0.06 - 0.81; $P=.02$) were less likely to complete paper rather than electronic surveys.

Conclusions: Incorporating paper surveys and engaging CBOs improved participation among Vietnamese speakers and those without postsecondary education, addressing the underrepresentation observed in phase 1. These findings highlight the importance of tailored survey methodologies to achieve demographic inclusivity in public health research.

(JMIR Public Health Surveill 2026;12:e77520) doi:10.2196/77520

KEYWORDS

health disparities; immigrant health; population health; public health; paper survey; methodology

Introduction

Historically marginalized populations are often underrepresented in public health research, which can exacerbate existing health disparities by failing to capture the unique health challenges these groups face adequately. Among these populations, the Vietnamese American community presents distinctive challenges and opportunities for researchers. Despite significant progress in medical advancements, Asian Americans, including Vietnamese Americans, continue to experience disparities in health outcomes, largely driven by nonmedical drivers of health such as language accessibility, cultural differences, and varying levels of trust in public institutions [1]. These disparities are further compounded by the limited representation of Asian Americans in public health research, which often relies on online surveys that may not fully engage this population [2-5].

Survey methodology is crucial in public health research, especially when engaging economically disadvantaged populations. While online surveys are increasingly used for convenience and broad reach, they may not be as effective for

XSL•FO
RenderX

populations with limited digital access or health literacy, such as older Vietnamese Americans [6-8]. For example, one study conducted in Chicago found that disparities in digital access were associated with race, ethnicity, and income [9]. In contrast, paper surveys provide an accessible alternative that may enhance participation among these groups and potentially lead to more representative data collection [10]. However, the comparative effectiveness of paper surveys versus electronic ones in accurately reaching and representing Vietnamese Americans remains underexplored, indicating a need for further research in this area [11,12].

Vietnamese Americans face several specific health disparities, including higher rates of chronic conditions such as hepatitis B, cancer [13-17], and non–small cell lung cancer [18] compared to other racial and ethnic groups. The historical context of migration, coupled with socioeconomic challenges, has also led to significant mental health issues within this community, with elevated rates of posttraumatic stress disorder and depression among refugees and their descendants [19-21]. However, these health issues are frequently underreported and understudied, partly due to the barriers hindering their research participation [22,23]. Vietnamese Americans, particularly older adults and recent immigrants, have a high level of non-English language preference, which restricts their ability to engage with research materials that are not available in their native language [24-26]. Additionally, there is often a deep-seated mistrust in public institutions among Vietnamese Americans, stemming from historical experiences of political oppression and displacement, which can further deter participation in research conducted by these institutions [27-30].

Community-based organizations (CBOs) have long been recognized as vital partners in addressing these challenges and promoting equity in public health research [31,32]. Research has shown that CBOs can increase participation rates and data accuracy by leveraging their established relationships within these communities [33]. Community-engaged research (CEnR) emphasizes actively involving community organizations, stakeholders, and members in the research process to ensure that studies are culturally relevant and inclusive [34-36]. CBO involvement in CEnR is particularly important in studies involving populations with specific cultural and linguistic needs, as they can enhance the reliability and representativeness of the collected data [33].

In response to the pressing need to address research gaps among communities of color during the COVID-19 pandemic, the National Institutes of Health (NIH) launched the Community Engagement Alliance (CEAL) Against COVID-19 in Disproportionately Affected Communities Consortium. This initiative aimed to conduct CEnR and outreach in collaboration with local entities. As part of the Texas CEAL initiative, our team fielded the Common Survey 2 instrument twice to examine the impact of the pandemic on Vietnamese American adults in Texas. During the first year, data were collected exclusively through electronic surveys [37,38], while in the second year, a mixed approach incorporating both electronic and paper surveys was used in response to CBO feedback. This study compares participant demographics across these two survey modalities to evaluate how different data collection strategies influence engagement and representation. Additionally, it examines COVID-19–related challenges and explores facilitators and barriers to vaccine uptake, providing valuable insights into the role of social determinants of health in shaping health behaviors and outcomes in this population.

This study had three objectives: (1) to describe the sociodemographic characteristics of Vietnamese American respondents; (2) to compare paper versus electronic survey modalities and evaluate their associations with respondent demographics; and (3) to highlight the role of CBOs and CEnR in enhancing representativeness and inclusivity in minority population health research. These findings aim to inform the development of equitable interventions and policies that address nonmedical drivers of health, ultimately contributing to the reduction of health disparities among Vietnamese Americans and other marginalized communities.

Methods

Study Design and Population

As part of the NIH CEAL initiative, we used the Common Survey 2 instrument in this repeated cross-sectional survey approved by the Institutional Review Board at the University of Houston (June 4, 2021; STUDY00003046).

The CEAL Common Survey 2 is a 23-item questionnaire designed to comprehensively assess COVID-19–related attitudes, behaviors, and experiences. This instrument containing questions pertaining to vaccine hesitancy, adherence to preventive measures, and willingness to participate in clinical trials has been described previously [27]. The CEAL Common Survey 2 has been widely implemented across CEAL-affiliated projects but has not undergone formal psychometric validation, such as reliability or construct validity testing [39]. This reflects the program's rapid launch during the COVID-19 public health emergency in September 2020, when urgent deployment precluded full validation procedures. Despite this limitation, the survey has been applied across diverse communities and study contexts, supporting its use as a standardized, though not formally validated, instrument. The English and Vietnamese versions are provided in Multimedia Appendices 1 and 2.

A translator from a collaborating CBO translated the survey from English into Vietnamese. Subsequently, a different translator from the same CBO performed a back-translation of the survey from Vietnamese to English. To ensure clarity and cultural relevance for the target population, native Vietnamese speakers at another CBO reviewed and reconciled the translations. The survey and consent form were accessible online in English and Vietnamese. On average, participants required 15 minutes to complete the survey in English and 20 minutes in Vietnamese [27].

The study design was divided into two phases. Phase 1, conducted between September 2021 and March 2022, exclusively used an online survey format. However, recognizing the limitations of an online-only approach in response to the CBO feedback and preliminary analyses from phase 1 and phase 2 (December 2022-April 2023) incorporated a paper-based survey option in addition to online recruitment.

XSL•FO
RenderX

Participants were eligible for inclusion in the study if they met the following criteria: they were at least 18 years of age, self-identified as being of Vietnamese heritage, resided in the state of Texas, and were proficient in reading and writing either English or Vietnamese.

Recruitment

A convenience sampling method was used to recruit Vietnamese Americans for the study. Data collection in phase 1 was primarily facilitated through digital channels, capitalizing on the existing networks of local and national Vietnamese American civic groups, religious organizations, and networks on social media platforms.

The survey remained open for responses from September 20, 2021, to March 4, 2022. Recruitment efforts were conducted virtually, in both English and Vietnamese, through email distribution lists through the Vietnamese American Medical Association, Vietnamese American Nurses Association, Vietnamese American Pharmacists Association, the Progressive Vietnamese American Organization, and the Vietnamese Culture and Science Association; social media platforms such as Facebook groups; an online webinar hosted by Boat People SOS—Houston (BPSOS); and an internet advertisement placed on the Google Marketing Platform that ran from January 6 to 23, 2022. Additional recruitment was conducted in person with recruitment flyers in both English and Vietnamese with QR codes at BPSOS literacy classes and two health fairs organized by Vietnamese American Nurses Association, Vietnamese American Medical Association, and Vietnamese American Pharmacists Association. Furthermore, the recruitment flyers were displayed at two health clinics in southwest Houston serving predominantly Vietnamese populations. Participants who completed the survey and provided a valid email address or phone number were entered into a raffle to win one of five US $50 gift cards [27].

Phase 2 used the same electronic recruitment methods from phase 1 and expanded survey administration to include simultaneous paper survey distribution. These surveys remained open for responses from December 26, 2022 to April 30, 2023. Paper surveys were strategically distributed at community events and health fairs led by researchers and local CBO-trained bilingual community members who recruited participants at a local clinic in southwest Houston with a large Vietnamese American patient population and CBO events sponsored by the organizations above. Trained community members used the paper surveys and did not offer electronic participation. Research team members performed double data entry of paper survey responses into a separate online survey. All entries were compared for accuracy and quality assurance. Unclear or inconsistent responses underwent a structured review process, with discrepancies addressed through a consensus-driven decision-making process.

Measures

Dependent Variable

The primary outcome of interest in this study was the mode of survey completion, categorized as either electronic or paper-based. This reflects both a participant's opportunity to encounter the survey in a given format and their decision to complete it. Differences in where and how each modality was used, such as online recruitment versus CBO events, are therefore embedded in this outcome and influence interpretation of participation patterns. Participants who completed the survey online were classified under the "electronic survey" group, while those who completed a paper version were assigned to the "paper survey" group.

Independent Variables

Sociodemographic characteristics were assessed to examine differences among respondents based on the mode of survey completion. Gender was categorized as female, male, other, or prefer not to answer. Age was recorded as a continuous variable and further categorized into the following age groups: 18 - 24, 25 - 34, 35 - 44, 45 - 54, 55 - 64, 65 - 74, and ≥75 years. Sexual orientation was classified as heterosexual, homosexual, other, or prefer not to answer. Language preference was determined based on participants' self-reported preferred language for survey completion (English or Vietnamese). Educational attainment was recorded as a college degree or higher versus less than a college degree. Insurance status was classified as insured or uninsured.

In addition to sociodemographic characteristics, health and COVID-19–related variables were included in the analysis. Participants reported their history of prior COVID-19 infection (yes, no, or unsure diagnosis). COVID-19 vaccination status was categorized as fully vaccinated, first dose of a 2-dose vaccine, not vaccinated, or prefer not to answer.

To assess trust in health care and research institutions, participants rated their level of trust in various sources, including the NIH, health care providers, university hospitals, community clinics, pharmaceutical companies, and family or friends. Response options included a great deal, a fair amount, not very much, not at all, or no opinion. Trust was dichotomized into trust (a great deal or a fair amount) versus no trust (not very much or not at all). Willingness to participate in COVID-19 clinical trials was measured using a 7-point Likert scale. For analysis, responses were categorized into high willingness (scores of 6 or 7 on the scale) versus less than high willingness (≤5 on the scale).

To evaluate health literacy and access to health care information, participants responded to the question, "How often do you need someone to help you read written information from your doctor or drug store?" Response options included always, often, sometimes, rarely, never, or prefer not to respond. For analysis, responses were categorized as requiring assistance if participants selected always, often, or sometimes.

The independent variables included in regression models, including language, education, trust, COVID-19 history, and willingness to participate in trials, were selected based on prior CEAL studies [27] and conceptual frameworks identifying social determinants of health and community trust as key drivers of health engagement during the COVID-19 pandemic. These constructs were prioritized given their centrality to CEAL program aims and their relevance in understanding disparities in research participation among Vietnamese Americans.

XSL•FO
RenderX

Statistical Analysis

We analyzed data using STATA (version 17; StataCorp LLC). Multivariate logistic regression models were used to explore the factors associated with the choice of survey mode in phase 2. Phase 1 data were not included in these models because all respondents in that phase completed the survey electronically, resulting in no variability with respect to survey modality. Although combining electronic data from both phases could have increased statistical power, this approach was not pursued due to meaningful contextual differences between phases, including the period of data collection, the recruitment strategies used, and the introduction of CBO-led in-person events during phase 2. Restricting analyses to phase 2, therefore, ensured analytic consistency and reduced the potential for bias introduced by structural differences between phases. These models adjusted for key demographic and socioeconomic variables, including age, gender, language preference (English or Vietnamese), and education level. In addition, the models adjusted for key constructs about COVID-19 included in the survey, such as history of COVID-19 infection, COVID-19 vaccination history, trusted sources of health information, and willingness to participate in COVID-19 clinical trials. The logistic regression models generated adjusted odds ratios (AOR) and 95% CIs for each predictor variable, with statistical significance determined by a 2-sided P value threshold of $P<.05$.

Ethical Considerations

The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the institutional review board of the University of Houston (STUDY00003046; June 4, 2021).

Informed consent was obtained from all subjects involved in the study. The consent was obtained via the cover letter of both the electronic and paper survey, which participants reviewed prior to starting the survey. The consent form clarified the voluntary nature of participation, the length of time required, and the purpose of the study. The form did not explicitly mention future data analyses. All responses were deidentified prior to analysis. Survey data were stored on a secure institutional cloud drive accessible only to members of the research team. Incentives for participation included the option to enter a raffle to win one of five US $50 gift cards. No identification of individual participants has been included in this manuscript or supplementary materials.

Results

The study identified differences in respondents' demographic characteristics based on the survey completion mode. In phase 1, 224 electronic surveys were completed. In phase 1, 359 surveys were completed, of which 124 were electronic and 235 were paper surveys. Table 1 presents the sociodemographic characteristics of participants from phase 1 and phase 2.

XSL•FO
RenderX

Table . Sociodemographic characteristics of survey respondents (phase 1 and phase 2)[a].

Characteristic	Phase 1[b] electronic (n=224), n (%)	Phase 2[c] electronic (n=124), n (%)	Phase 2 paper (n=235), n (%)
Gender			
Woman	125 (55.8)	53 (42.7)	149 (63.4)
Man	94 (42)	62 (50)	81 (34.5)
Other	2 (0.9)	2 (1.6)	2 (0.9)
Prefer not to answer	3 (1.3)	7 (5.65)	3 (1.28)
Sexual orientation			
Heterosexual	142 (67)	104 (83.9)	150 (63.8)
Homosexual	39 (18.4)	8 (6.5)	11 (4.7)
Other	31 (14.6)	3 (2.4)	21 (8.9)
Prefer not to answer	12 (5.4)	9 (7.3)	52 (22.1)
Age (y)			
18 - 24	23 (10.3)	20 (16.1)	4 (1.7)
25 - 34	40 (17.9)	43 (34.7)	9 (3.8)
35 - 44	36 (16)	15 (12.1)	22 (9.4)
45 - 54	36 (16)	10 (8.1)	30 (12.8)
55 - 64	35 (15.6)	19 (15.3)	35 (14.9)
65 - 74	34 (15.2)	8 (6.5)	65 (27.7)
≥75	8 (3.6)	2 (1.6)	23 (9.8)
Prefer not to answer	12 (5.4)	0 (0)	0 (0)
Preferred language for survey			
English	167 (74.6)	100 (80.7)	12 (5.1)
Vietnamese	57 (25.4)	24 (19.4)	223 (94.9)
Insurance status			
Insured	194 (86.6)	106 (85.5)	205 (87.2)
Uninsured	23 (10.3)	9 (7.3)	21 (8.9)
Prefer not to answer	7 (3.1)	9 (7.3)	9 (3.8)
Education			
Less than high school	25 (11.2)	12 (9.7)	69 (29.4)
High school or GED[d] equivalent	47 (21.0)	21 (16.9)	76 (32.3)
Some college	37 (16.5)	25 (20.2)	20 (8.5)
Associate or college degree	87 (38.8)	46 (37.1)	48 (20.4)
Professional or doctorate degree	21 (9.4)	18 (14.5)	4 (1.7)
Prefer not to answer	7 (3.1)	8 (6.5)	18 (7.7)
COVID-19 vaccination status			
Fully vaccinated	181 (80.8)	102 (82.3)	206 (87.7)
First dose of a 2-dose vaccine	20 (8.9)	14 (11.3)	17 (7.2)
Not vaccinated	9 (4.0)	3 (2.4)	8 (3.4)
Prefer not to answer	14 (6.3)	5 (4)	4 (1.70)
Need assistance reading written health information	34 (15.18)	39 (31.5)	165 (70.2)

[a]Descriptive characteristics of Vietnamese American adults in Texas who participated in a cross-sectional COVID-19 study conducted between 2021 and 2023.

XSL•FO
RenderX

[b]Phase 1 (September 20, 2021-March 4, 2022) involved online-only surveys.

[c]Phase 2 (December 26, 2022-April 30, 2023) involved both online and paper surveys.

[d]GED: General Education Development.

Of the 224 phase 1 respondents, 125 (55.8%) identified as female, 94 (42%) as male, and 2 (0.9%) as other genders. Regarding language preference, 167 (74.6%) participants completed the survey in English, while 57 (25.4%) completed it in Vietnamese. In phase 1, 194 (86.6%) participants were insured, while 23 (10.3%) were uninsured. Educational attainment was nearly evenly distributed, with 108 (48.2%) participants reporting a college degree or higher and 109 (48.7%) reporting less than a college degree. In this phase, 181 (80.8%) participants were fully vaccinated, 20 (8.9%) received the first dose of a 2-dose vaccine, and 9 (4%) were unvaccinated.

In phase 2, which incorporated both electronic (n=124, 34.5%) and paper (n=235, 65.5%) surveys, 359 participants completed the survey. Of these, 202 (56.3%) identified as female participants, 143 (39.8%) as male participants, and 4 (1.1%) as other genders. Seven (1.9%) preferred not to answer. In terms of language preference, 112 (31.2%) participants completed the survey in English, while 247 (68.8%) completed it in Vietnamese. Among the respondents, 311 (86.6%) were insured, 30 (8.4%) were uninsured, and 18 (5%) preferred not to answer. Educational attainment was reported as 98 (27.3%) with a college degree or higher, 235 (65.5%) with less than a college degree, and 26 (7.2%) preferred not to answer. In terms of vaccination status, 308 (85.8%) were fully vaccinated, 31 (8.6%) had received the first dose of a 2-dose vaccine, 11 (3.1%) were unvaccinated, and 9 (2.5%) preferred not to answer. Further details on the breakdown between electronic and paper survey responses are presented in Table 1.

Table 2 highlights the multivariate logistic regression analysis using data gathered solely within phase 2. Being a Vietnamese speaker was the strongest predictor of completing a paper survey (AOR 100.9, 95% CI 24.3 - 418.9; $P<.001$). Additionally, female participants were significantly more likely to complete a paper survey than male participants (AOR 5.09, 95% CI 1.43 - 18.1; $P=.01$). Conversely, respondents with a history of COVID-19 infection (AOR 0.16, 95% CI 0.05 - 0.52; $P=.002$), those with higher education levels (AOR 0.18, 95% CI 0.05 - 0.67; $P=.01$), and those who expressed a strong willingness to participate in COVID-19 clinical trials (AOR 0.21, 95% CI 0.06 - 0.81; $P=.02$) were more likely to complete the survey electronically.

XSL•FO
RenderX

Table . Multivariate logistic regression examining differences in characteristics between participants who completed paper versus electronic surveys in phase 2. Results from a multivariate logistic regression analysis examining associations between participant sociodemographic characteristics and survey modality among Vietnamese American adults in Texas during phase 2 (December 26, 2022-April 30, 2023) of a cross-sectional COVID-19 study. Measures the odds of a respondent having completed a paper survey in the survey mode of administration versus an electronic survey.

Sociodemographic variables	AOR[g] (95% CI)	*P* value
Language use (reference: English)		
Vietnamese	100.91 (24.31 - 418.95)	<.001
Age		
Per increasing year	1.03 (0.99 - 1.07)	.11
Sex (reference: biological male)		
Biological female	5.09 (1.43 - 18.10)	.01
Health insurance status (reference: uninsured)		
Insured	0.33 (0.04 - 3.06)	.33
Educational attainment (reference: less than high school)		
Some college	0.23 (0.05 - 1.12)	.07
College or higher	0.18 (0.05 - 0.67)	.01
History of prior COVID-19 infection (reference: no)		
Yes	0.16 (0.05 - 0.52)	.002
Completed COVID-19 vaccination[b] (reference: no)		
Yes	1.45 (0.23 - 9.10)	.69
Great deal of trust in the following sources[c] (reference: less than a great deal of trust)		
NIH[d]	0.57 (0.12 - 2.75)	.48
Health care provider	3.63 (0.54 - 24.59)	.19
Local community clinic	1.28 (0.19 - 8.78)	.80
University hospital	1.68 (0.24 - 11.73)	.60
Pharmaceutical company	1.47 (0.20 - 10.92)	.71
Researchers	0.81 (0.19 - 3.49)	.78
Friend/family	0.35 (0.04 - 3.11)	.35
FDA[e]	1.11 (0.33 - 3.79)	.86
Level of willingness to participate in COVID-19 clinical trial (reference: less than high willingness[f])		
High willingness[f]	0.21 (0.06 - 0.81)	.02

[a]AOR: adjusted odds ratio.

[b]Completed COVID-19 vaccination series (either 1 vaccination of a 1-vaccination series or both vaccinations of a 2-vaccination series).

[c]Defined as the proportion of those who trust each row's source a great deal versus those who endorse less than a great deal of trust.

[d]NIH: National Institutes of Health.

[e]FDA: Food and Drug Administration.

[f]Defined as a high level of willingness to participate in a COVID-19 clinical trial (6 or 7 on the Likert scale) versus those with less than a high level of willingness (≤5 on the 7-point Likert scale).

Discussion

Principal Findings

This study provides insights into the challenges of conducting public health research among Vietnamese Americans, a historically underrepresented population in public health data. The initial reliance on electronic surveys in phase 1 predominantly engaged younger, more educated, and technologically proficient individuals, leading to the unintended exclusion of key population segments, including older adults, non-English speakers, and those with lower educational attainment [38]. These social determinants are well-recognized factors that contribute to health disparities, not only in terms of access to health care but also in the context of research participation and representation [1,8,40,41]. Importantly, survey completion in a given modality reflects both access to that modality and willingness to engage with it. Digital participation

XSL•FO
RenderX

requires digital access, just as paper-based participation presupposes in-person contact. Therefore, differences in survey completion across modalities should be interpreted not only as participant preference, but as an outcome shaped by where and how participants encountered the survey.

Demographic differences between respondents who completed paper versus electronic surveys highlight the limitations of digital-only data collection. Studies conducted among ethnic subpopulations have examined response rates from web or mail surveys [42]. However, this study is one of the first to highlight this in a population where digital literacy is a barrier. For instance, research involving women with cancer found that in ethnic subpopulation questionnaires, data quality and responses significantly varied by mode of survey administration [43]. These findings align with the results of our study, reinforcing the value of accessible survey methodologies that account for the diverse needs of marginalized populations.

Nonmedical Drivers of Health

Language and Digital Access Barriers

Language accessibility is an important nonmedical driver of health that not only influences access to health care information and services but also impacts participation in health research, contributing to health disparities [44,45]. Research studies conducted exclusively in English may systematically exclude non-English speakers, limiting the representation and generalizability of findings. A systematic review assessing health outcomes in Asian American subgroups in the United States found that all 76 studies only considered English-speaking individuals, underscoring barriers faced by linguistically diverse populations in research participation [1]. In response, this study partnered with CBOs in a CEnR approach to translate surveys into Vietnamese, improving accessibility for a historically underrepresented subgroup and facilitating engagement in a culturally relevant manner [46]. This study's findings indicate that Vietnamese language preference was significantly associated with completing paper surveys compared to electronic surveys. While online surveys are commonly used in large-scale health studies due to their convenience, cost-effectiveness, and onsite data error prevention, they may unintentionally exclude individuals with limited internet access or digital literacy skills [47]. In phase 2 of our study, 90.3% (223/247) of Vietnamese-speaking participants completed paper surveys compared to 10.7% (12/112) of English-speaking participants, suggesting that an electronic-only approach may underrepresent certain populations. Paper surveys are often more accessible to individuals who lack reliable internet access or familiarity with digital platforms, which may be particularly relevant for older Vietnamese Americans, who may encounter linguistic barriers and have lower levels of digital literacy than younger individuals [6,48,49]. Additionally, a potential contributing factor may be the recruitment context. Given that in phase 2, paper surveys were distributed in part at community events via CBOs (eg, BPSOS), whose mission was to support recent Vietnamese immigrants, refugees, and low-income families, participants may have felt more comfortable participating in research through organizations they trust, which could partly explain the higher rates of paper survey completion among Vietnamese-speaking respondents. Overall, a multipronged recruitment strategy, including both translated paper and electronic surveys, may include a broader range of participants and mitigate research participation disparities among linguistically diverse communities.

Education and Digital Disparities

One limitation of phase 1 was the underrepresentation of participants with less than a high school education. The 2022 Asian Americans, Native Hawaiians, and Pacific Islanders (AANHPI) report used data from the 2020 American Community Survey, applying a rolling multiyear sampling methodology to generate weighted socioeconomic and demographic characteristics estimates [38,50,51]. While the 2022 AANHPI report found that 25% of Vietnamese Americans had not completed high school, only 11.2% (25/224) of our phase 1 respondents fell into this category [38,49,50]. However, in phase 2, this percentage increased to 22.6% (69/305), suggesting that paper-based survey administration may improve participation among individuals with lower educational attainment. This finding aligns with prior research highlighting that education level is a determinant of digital literacy and online survey participation [9,51]. Consequently, digital-only survey methods may inadvertently exclude this population.

Health Insurance

Comparing our sample demographics to the 2022 AANHPI report, we found similar rates of uninsured participants in both study phases (phase 1: 10.3%, 23/224 vs phase 2: 8.4%, 30/359) [50,51]. Insurance status often correlates with socioeconomic factors influencing digital access, but our findings suggest that insurance coverage alone did not significantly predict survey modality preference. Uninsured participants were neither disproportionately excluded from digital surveys nor more likely to complete paper surveys, indicating that other social determinants, such as language, education, or gender, played a larger role in survey participation.

Role of Community-Based Organizations in Community-Engaged Research

Given the digital and language barriers identified in our study, CBOs played a critical role in bridging accessibility gaps, improving representation, recruiting non-English speakers, distributing paper surveys, and increasing research trust. Often, CBO's deep-rooted connections within the community and understanding of cultural and linguistic nuances uniquely position them to enhance research participation [32,34,36].

CBO involvement can ensure that research is methodologically sound, culturally appropriate, and respectful of the community's unique needs [33-36]. Prior studies have demonstrated that CBOs can effectively bridge the gap between researchers and participants, fostering trust and facilitating engagement in a way that traditional research institutions often cannot [35,52]. This study adds to the growing body of evidence supporting the involvement of CBOs, demonstrating that their contributions are critical for recruitment and ensuring that research methodologies align with the target population's cultural and linguistic needs [35].

XSL•FO
RenderX

Limitations

This study contains some limitations. Using a cross-sectional survey limits the ability to establish causality, preventing generalizability beyond the sample population. In addition, this research was conducted within a specific geographic and demographic context, focusing exclusively on Vietnamese Americans in Texas. As such, the findings may not fully reflect the experiences of Vietnamese Americans in other regions or contexts. Additionally, this survey instrument did not assess participants' years of immigration to the United States, which may influence acculturation, language proficiency, and technology use [8]. Indeed, neither sample in this study (electronic or paper survey respondents) can be assumed to reflect or represent the broader population of Vietnamese Americans, even in the region studied here. Because recruitment was based on convenience sampling through CBOs, events, and listservs, the sample is subject to selection bias. In-person recruitment using paper surveys took place, in part, at Houston-area events offering free health care screenings. This likely attracts a different population than who receives a recruitment email from a Vietnamese American health professional listserv. In addition, participants who are more health-seeking or community-engaged may be overrepresented across respondents. The differences between the two groups of respondents even suggest that there is heterogeneity in how such a representative sample might be constructed but do not suggest that one or the other is more generalizable. In addition, overall, the sample size collected for either group is small, and the multivariate approach risks overfitting the models reported here, resulting in large (occasionally very large) CIs. As a result, point estimates of effect sizes should be treated as exploratory findings, and the risk of type II error (falsely null associations) is higher than the usual likelihood.

Although our sample size in our exploratory cross-sectional study design limits our external validity of our exploratory observations to the broader Vietnamese American population, statistical methods such as poststratification or propensity score adjustment would enable future researchers to mitigate selection bias by reweighting the analytic sample to reflect the distribution of observed covariates in the general population.

Conclusion

In conclusion, this study underscores the importance of using flexible and adaptive research methodologies when engaging with marginalized populations. The combination of digital and paper survey methods, informed by community insights and tailored to the specific needs of the Vietnamese American community, proved to be an effective strategy for enhancing participation and representativeness. These findings have broader implications for public health research, particularly in addressing health disparities among historically marginalized populations to achieve health equity. As the field moves toward more inclusive and equitable research practices, researchers must continue to draw on the expertise of community partners and pay close attention to the social determinants of health that shape participation and outcomes. By doing so, public health research can better serve the needs of all populations, ultimately contributing to reducing health disparities and promoting health equity.

In summary, this study underscores the role of CBOs and CEnR in enhancing the inclusivity and representativeness of public health research. The study achieved a more comprehensive representation of the Vietnamese American community by integrating paper surveys and bilingual recruitment strategies. Future public health research should continue prioritizing the involvement of CBOs and CEnR to ensure that different population segments are adequately represented. This approach is fundamental to addressing health disparities and developing public health interventions responsive to different communities' needs.

Acknowledgments

The authors would like to thank medical librarian Stefanie Lapka for assisting with the literature search and Saba Siddiqi and Shielene Vargas for entering data. We also acknowledge the staff and members of Boat People SOS—Greater Houston, Progressive Vietnamese American Organization, and the Vietnamese Culture and Science Association for assistance with translations and recruitment.

Funding

This research was, in part, funded by the National Institutes of Health Agreement OT2HL158287.

Disclaimer

The views and conclusions in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the National Institutes of Health.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: JD, BMN
Data curation: CN, AL, AN, JD
Formal analysis: CN, AL, BK, PY

XSL•FO
RenderX

Funding acquisition: BMN
Investigation: CN, AL, JD, BMN
Methodology: BMN
Project administration: BMN
Resources: JD, AN
Supervision: BMN
Visualization: CN, AL
Writing – original draft: CN, AL
Writing – review & editing: PY, BK, AN, JD, BMN

Conflicts of Interest

While none of the authors have any financial conflicts of interest to disclose, many of us were involved in COVID-19 vaccine outreach efforts during the pandemic. CN and AL are affiliated with the Vietnamese Culture and Science Association. AN and JD are affiliated with Boat People SOS–Houston. BMN is affiliated with the Progressive Vietnamese American Organization and serves on the board of Doctors for America. These community organizations were involved in COVID-19 vaccine campaigns during the pandemic.

Multimedia Appendix 1
Common survey (English).
[PDF File, 316 KB - publichealth_v12i1e77520_app1.pdf]

Multimedia Appendix 2
Common survey (Vietnamese).
[PDF File, 377 KB - publichealth_v12i1e77520_app2.pdf]

References

1. Yom S, Lor M. Advancing health disparities research: the need to include Asian American subgroup populations. J Racial Ethn Health Disparities 2022 Dec;9(6):2248-2282. [doi: 10.1007/s40615-021-01164-8] [Medline: 34791615]
2. Quach T, Đoàn LN, Liou J, Ponce NA. A rapid assessment of the impact of COVID-19 on Asian Americans: cross-sectional survey study. JMIR Public Health Surveill 2021 Jun 11;7(6):e23976. [doi: 10.2196/23976] [Medline: 34019478]
3. McGarity-Palmer R, Saw A, Tsoh JY, Yellow Horse AJ. Trends in racial discrimination experiences for Asian Americans during the COVID-19 pandemic. J Racial Ethn Health Disparities 2024 Feb;11(1):168-183. [doi: 10.1007/s40615-022-01508-y] [Medline: 36602751]
4. McGarity-Palmer R, Saw A, Horse AJY, Yi SS, Tsoh J, Takeuchi D. Profiles of a COVID-19 syndemic: anti-Asian racism, economic challenges, and mental and physical health. J Racial Ethn Health Disparities 2024 Feb;11(1):300-312. [doi: 10.1007/s40615-023-01519-3] [Medline: 36692660]
5. Pahl K, Wang J, Sanichar N, et al. Anti-Asian attitudes in the context of the COVID-19 pandemic: an exploratory study. J Racial Ethn Health Disparities 2023 Aug;10(4):1947-1954. [doi: 10.1007/s40615-022-01376-6] [Medline: 35913546]
6. Van Hoa H, Giang HT, Vu PT, Van Tuyen D, Khue PM. Factors associated with health literacy among the elderly people in Vietnam. Biomed Res Int 2020;2020:3490635. [doi: 10.1155/2020/3490635] [Medline: 32309429]
7. Lee HY, Rhee TG, Kim NK, Ahluwalia JS. Health literacy as a social determinant of health in Asian American immigrants: findings from a population-based survey in California. J Gen Intern Med 2015 Aug;30(8):1118-1124. [doi: 10.1007/s11606-015-3217-6] [Medline: 25715993]
8. Meyer OL, Park VT, Kanaya AM, et al. Inclusion of Vietnamese Americans: opportunities to understand dementia disparities. Alzheimers Dement (N Y) 2023;9(2):e12392. [doi: 10.1002/trc2.12392] [Medline: 37251911]
9. Kan K, Heard-Garris N, Bendelow A, et al. Examining access to digital technology by race and ethnicity and child health status among Chicago families. JAMA Netw Open 2022 Aug 1;5(8):e2228992. [doi: 10.1001/jamanetworkopen.2022.28992] [Medline: 36018593]
10. Ebert JF, Huibers L, Christensen B, Christensen MB. Paper- or web-based questionnaire invitations as a method for data collection: cross-sectional comparative study of differences in response rate, completeness of data, and financial cost. J Med Internet Res 2018 Jan 23;20(1):e24. [doi: 10.2196/jmir.8353] [Medline: 29362206]
11. Singh S, Sagar R. A critical look at online survey or questionnaire-based research studies during COVID-19. Asian J Psychiatr 2021 Nov;65:102850. [doi: 10.1016/j.ajp.2021.102850] [Medline: 34534919]
12. Jäckle A, Roberts C, Lynn P. Assessing the effect of data collection mode on measurement. Int Stat Rev 2010 Apr;78(1):3-20. [doi: 10.1111/j.1751-5823.2010.00102.x]
13. Pham C, Fong TL, Zhang J, Liu L. Striking racial/ethnic disparities in liver cancer incidence rates and temporal trends in California, 1988–2012. J Natl Cancer Inst 2018 Nov 1;110(11):1259-1269. [doi: 10.1093/jnci/djy051] [Medline: 29617913]

XSL•FO
RenderX

14. Kwong SL, Chen MS, Snipes KP, Bal DG, Wright WE. Asian subgroups and cancer incidence and mortality rates in California. Cancer 2005 Dec 15;104(12 Suppl):2975-2981. [doi: 10.1002/cncr.21511] [Medline: 16247792]
15. McCracken M, Olsen M, Chen MS Jr, et al. Cancer incidence, mortality, and associated risk factors among Asian Americans of Chinese, Filipino, Vietnamese, Korean, and Japanese ethnicities. CA Cancer J Clin 2007;57(4):190-205. [doi: 10.3322/canjclin.57.4.190] [Medline: 17626117]
16. Miller BA, Chu KC, Hankey BF, Ries LAG. Cancer incidence and mortality patterns among specific Asian and Pacific Islander populations in the U.S. Cancer Causes Control 2008 Apr;19(3):227-256. [doi: 10.1007/s10552-007-9088-3] [Medline: 18066673]
17. Lauderdale DS, Huo D. Cancer death rates for older Asian-Americans: classification by race versus ethnicity. Cancer Causes Control 2008 Mar;19(2):135-146. [doi: 10.1007/s10552-007-9079-4] [Medline: 17992577]
18. Raz DJ, Gomez SL, Chang ET, et al. Epidemiology of non-small cell lung cancer in Asian Americans: incidence patterns among six subgroups by nativity. J Thorac Oncol 2008 Dec;3(12):1391-1397. [doi: 10.1097/JTO.0b013e31818ddff7] [Medline: 19057262]
19. Norris FH, Vanlandingham MJ, Vu L. PTSD in Vietnamese Americans following Hurricane Katrina: prevalence, patterns, and predictors. J Trauma Stress 2009 Apr;22(2):91-101. [doi: 10.1002/jts.20389] [Medline: 19235888]
20. Boscarino JA, Boscarino JJ. Conceptualization of PTSD from the Vietnam war to current conflicts and beyond. Int J Emerg Ment Health 2015;17(3):661-663. [doi: 10.4172/1522-4821.1000278] [Medline: 26568701]
21. Vaage AB, Thomsen PH, Rousseau C, Wentzel-Larsen T, Ta TV, Hauff E. Paternal predictors of the mental health of children of Vietnamese refugees. Child Adolesc Psychiatry Ment Health 2011 Jan 10;5:2. [doi: 10.1186/1753-2000-5-2] [Medline: 21219651]
22. Lee S, Martinez G, Ma GX, et al. Barriers to health care access in 13 Asian American communities. Am J Health Behav 2010;34(1):21-30. [doi: 10.5993/ajhb.34.1.3] [Medline: 19663748]
23. Liu Y, Elliott A, Strelnick H, Aguilar-Gaxiola S, Cottler LB. Asian Americans are less willing than other racial groups to participate in health research. J Clin Transl Sci 2019 Jun;3(2-3):90-96. [doi: 10.1017/cts.2019.372] [Medline: 31660231]
24. Jang Y, Kim MT. Limited English proficiency and health service use in Asian Americans. J Immigr Minor Health 2019 Apr;21(2):264-270. [doi: 10.1007/s10903-018-0763-0] [Medline: 29797103]
25. Pandey M, Maina RG, Amoyaw J, et al. Impacts of English language proficiency on healthcare access, use, and outcomes among immigrants: a qualitative study. BMC Health Serv Res 2021 Jul 26;21(1):741. [doi: 10.1186/s12913-021-06750-4] [Medline: 34311712]
26. Nguyen C, Gilbert L, Diep J, Nguyen BM. Identifying facilitators and barriers to increasing COVID-19 vaccination and trial participation in vaccinated Vietnamese Americans. Health Equity 2022;6(1):485-493. [doi: 10.1089/heq.2022.0032] [Medline: 35801146]
27. Vargas S, Siddiqi S, King B, et al. Vietnamese Americans' level of trust in sources of information and willingness to participate in COVID-19 clinical trials. J Clin Transl Sci 2024;8(1):e88. [doi: 10.1017/cts.2024.495] [Medline: 38784109]
28. Lemke AA, Choi SH, Dang V, Dang TQ, Yu JH. Assessing Vietnamese American patient perspectives on population genetic testing in primary care: a community-engaged approach. HGG Adv 2022 Oct 13;3(4):100134. [doi: 10.1016/j.xhgg.2022.100134] [Medline: 36039118]
29. Kim PY, Kendall DL, Cheon HS. Racial microaggressions, cultural mistrust, and mental health outcomes among Asian American college students. Am J Orthopsychiatry 2017;87(6):663-670. [doi: 10.1037/ort0000203] [Medline: 27732011]
30. Sung CL. Asian Patients' distrust of western medical care: one perspective. Mt Sinai J Med 1999 Sep;66(4):259-261. [Medline: 10477478]
31. Israel BA, Schulz AJ, Parker EA, Becker AB. Review of community-based research: assessing partnership approaches to improve public health. Annu Rev Public Health 1998;19:173-202. [doi: 10.1146/annurev.publhealth.19.1.173] [Medline: 9611617]
32. Agonafer EP, Carson SL, Nunez V, et al. Community-based organizations' perspectives on improving health and social service integration. BMC Public Health 2021 Mar 6;21(1):452. [doi: 10.1186/s12889-021-10449-w] [Medline: 33676470]
33. Haldane V, Chuah FLH, Srivastava A, et al. Community participation in health services development, implementation, and evaluation: a systematic review of empowerment, health, community, and process outcomes. PLoS ONE 2019;14(5):e0216112. [doi: 10.1371/journal.pone.0216112] [Medline: 31075120]
34. Irby MB, Moore KR, Mann-Jackson L, et al. Community-engaged research: common themes and needs identified by investigators and research teams at an emerging academic learning health system. Int J Environ Res Public Health 2021 Apr 8;18(8):3893. [doi: 10.3390/ijerph18083893] [Medline: 33917675]
35. Balls-Berry JE, Acosta-Pérez E. The use of community engaged research principles to improve health: community academic partnerships for research. P R Health Sci J 2017 Jun;36(2):84-85. [Medline: 28622404]
36. Key KD, Furr-Holden D, Lewis EY, et al. The continuum of community engagement in research: a roadmap for understanding and assessing progress. Prog Community Health Partnersh 2019;13(4):427-434. [doi: 10.1353/cpr.2019.0064] [Medline: 31866597]

XSL•FO
RenderX

37. Le A, Siddiqi S, Nguyen C, et al. Examining health insurance and non-medical challenges among Vietnamese Americans in Texas during the COVID-19 pandemic. Int J Environ Res Public Health 2025 Jan 29;22(2):189. [doi: 10.3390/ijerph22020189] [Medline: 40003415]
38. Nguyen C, King B, Diep J, Gilbert L, Nguyen BM. Perspectives of Vietnamese Americans regarding COVID-19 vaccine acceptance, trusted sources of information, and pandemic-related challenges. J Racial Ethn Health Disparities 2025 Mar 8. [doi: 10.1007/s40615-025-02327-7] [Medline: 40057648]
39. Walker RJ, Eisenhauer E, Thompson EL, et al. COVID-19 information, trust, and risk perception across diverse communities in the United States: Initial findings from a multistate Community Engagement Alliance (CEAL). Am J Public Health 2024 Jan;114(S1):S112-S123. [doi: 10.2105/AJPH.2023.307504] [Medline: 38207271]
40. Do MP, Hutchinson PL, Mai KV, Vanlandingham MJ. Disparities in health care among Vietnamese New Orleanians and the impacts of Hurricane Katrina. Res Sociol Health Care 2009;27:301-319. [doi: 10.1108/S0275-4959(2009)0000027016] [Medline: 20725605]
41. Obra JK, Lin B, Đoàn LN, Palaniappan L, Srinivasan M. Achieving equity in Asian American healthcare: critical issues and solutions. J Asian Health 2021 Jul;1(1):e202103. [doi: 10.59448/jah.v1i1.3] [Medline: 37872960]
42. West BT, Zhang S, Wagner J, Gatward R, Saw HW, Axinn WG. Methods for improving participation rates in national self-administered web/mail surveys: evidence from the United States. PLoS ONE 2023;18(8):e0289695. [doi: 10.1371/journal.pone.0289695] [Medline: 37540678]
43. Hagan TL, Belcher SM, Donovan HS. Mind the mode: differences in paper vs. web-based survey modes among women with cancer. J Pain Symptom Manage 2017 Sep;54(3):368-375. [doi: 10.1016/j.jpainsymman.2017.07.005] [Medline: 28711752]
44. Chen WT, Sun W, Huang F, et al. Lost in translation: impact of language barriers and facilitators on the health care of Asian Americans living with HIV. J Racial Ethn Health Disparities 2024 Aug;11(4):2064-2072. [doi: 10.1007/s40615-023-01674-7] [Medline: 37306920]
45. Ngo-Metzger Q, Massagli MP, Clarridge BR, et al. Linguistic and cultural barriers to care. J Gen Intern Med 2003 Jan;18(1):44-52. [doi: 10.1046/j.1525-1497.2003.20205.x] [Medline: 12534763]
46. Andrade C. The limitations of online surveys. Indian J Psychol Med 2020 Nov;42(6):575-576. [doi: 10.1177/0253717620957496] [Medline: 33354086]
47. Zeleke AA, Naziyok T, Fritz F, Christianson L, Röhrig R. Data quality and cost-effectiveness analyses of electronic and paper-based interviewer-administered public health surveys: systematic review. J Med Internet Res 2021 Jan 22;23(1):e21382. [doi: 10.2196/21382] [Medline: 33480859]
48. Lee G, Chang A, Pal A, Tran TA, Cui X, Quach T. Understanding and addressing the digital health literacy needs of low-income limited English proficient Asian American patients. Health Equity 2022;6(1):494-499. [doi: 10.1089/heq.2022.0045] [Medline: 36186613]
49. Campanozzi LL, Gibelli F, Bailo P, Nittari G, Sirignano A, Ricci G. The role of digital literacy in achieving health equity in the third millennium society: a literature review. Front Public Health 2023;11:1109323. [doi: 10.3389/fpubh.2023.1109323] [Medline: 36891330]
50. State of Asian Americans, Native Hawaiians, and Pacific Islanders in the United States. : AAPI Data and National Council of Asian Pacific Americas; 2022 URL: https://aapidata.com/wp-content/uploads/2024/02/State-AANHPIs-National-June2022.pdf [accessed 2026-01-28]
51. Sabado-Liwag MD, Kwan PP, Okihiro M, et al. Asian American and Native Hawaiian and Pacific Islander visibility and data equity: insights from NIH RADx-UP. Am J Public Health 2024 May;114(S5):S354-S359. [doi: 10.2105/AJPH.2024.307590] [Medline: 38547462]
52. Kraft SA, Duenas DM, Lewis H, Shah SK. Bridging the researcher-participant gap: a research agenda to build effective research relationships. Am J Bioeth 2020 Jun;20(5):31-33. [doi: 10.1080/15265161.2020.1745936] [Medline: 32364474]

Abbreviations

AANHPI: Asian Americans, Native Hawaiians, and Pacific Islanders
AOR: adjusted odds ratio
BPSOS: Boat People SOS—Houston
CBO: community-based organization
CEAL: Community Engagement Alliance
CEnR: community-engaged research
NIH: National Institutes of Health

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 28.May.2025; peer-reviewed by B Foroutan, C Obianyo, F Guo, Y Deng; accepted 17.Dec.2025; published 25.Feb.2026.

Please cite as:
Nguyen C, Le A, Yeh P, King B, Nguyen A, Diep J, Nguyen BM
Survey Modalities and COVID-19 Vaccine Uptake in Vietnamese Americans: Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e77520
URL: https://publichealth.jmir.org/2026/1/e77520
doi:10.2196/77520

© Celine Nguyen, Alexander Le, Paul Yeh, Ben King, Angelica Nguyen, Jannette Diep, Bich-May Nguyen. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 25.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Knowledge, Attitudes, Practices, and Vaccination Willingness Toward Mpox (Monkeypox) Among Chinese Medical Students: Cross-Sectional Study

Yang Liu[1], MPH; Yuehui Jia[2], PhD; Honglong Li[3], MD; Jie Ge[2], PhD; Yunfeng Han[2], MPH; Zhiping Xie[2], MD; Jiaxin Chen[2], MD

[1]Scientific Research Department, Qiqihar Medical University, Qiqihar, Heilongjiang, China
[2]School of Public Health, Qiqihar Medical University, 333 Bukui Street, Jianhua District, Qiqihar, Heilongjiang, China
[3]School of Mental Health, Qiqihar Medical University, Qiqihar, Heilongjiang, China

Corresponding Author:
Jiaxin Chen, MD
School of Public Health, Qiqihar Medical University, 333 Bukui Street, Jianhua District, Qiqihar, Heilongjiang, China

Abstract

Background: Mpox (monkeypox) remains a global public health threat. However, data on mpox-related knowledge, attitudes, and practices (KAP) and vaccination willingness among Chinese medical students, who are key future health care practitioners, remain lacking.

Objective: This study aimed to investigate systematically the KAP and mpox vaccination willingness of Chinese medical students and identify the factors influencing their vaccination decisions.

Methods: A nationwide cross-sectional survey was conducted from November 2023 to March 2024. An anonymous self-designed questionnaire was used to assess basic information, KAP toward mpox, vaccination-related behaviors, and willingness. Categorical data were presented as frequency (constituent ratio). The normality of continuous variables was assessed using the Kolmogorov-Smirnov test. Continuous variables that did not conform to a normal distribution were presented as median (IQR). Data were analyzed using the chi-square test, 2-tailed *t* test, ANOVA, Kruskal-Wallis *H* test, and multinomial logistic regression.

Results: Among the 4098 participants, 84.63% (n=3468) accepted mpox vaccination. The median scores of KAP toward mpox were 43 (IQR 33-50), 33 (IQR 32-36), and 20 (IQR 19-24), respectively, with a median score of 73 (IQR 68-79) for vaccination-related practices. Multinomial logistic regression showed that factors associated with vaccination hesitancy (vs acceptance) included male individuals (odds ratio [OR] 1.416, 95% CI 1.158 - 1.732), being an only child (OR 1.340, 95% CI 1.098 - 1.635), no history of COVID-19 in family or friends (OR 1.520, 95% CI 1.161 - 1.991), no influenza vaccination (OR 1.429, 95% CI 1.146 - 1.783), and low mpox knowledge (OR 0.948, 95% CI 0.941 - 0.955). Factors associated with vaccination rejection (vs acceptance) included male sex (OR 1.641, 95% CI 1.003 - 2.686), high academic grade (OR 1.442, 95% CI 1.154 - 1.802), family or friends working on COVID-19 frontlines (OR 2.243, 95% CI 1.337 - 3.764), no internship experience (OR 2.049, 95% CI 1.076 - 3.901), presence of organic diseases (OR 3.733, 95% CI 1.778 - 7.838), and low mpox knowledge (OR 0.954, 95% CI 0.938 - 0.971). Good self-reported health status was a protective factor against refusal (OR 0.748, 95% CI 0.580 - 0.965).

Conclusions: The high willingness to receive mpox vaccination among Chinese medical students and its determinants, as identified in this study, carry clear implications for both education and policy. These findings inform the design of targeted health education programs for students and guide the development of evidence-based prevention strategies on campuses during public health emergencies.

(JMIR Public Health Surveill 2026;12:e86981) doi:10.2196/86981

KEYWORDS

monkeypox; mpox; mpox vaccination; vaccine hesitancy; determinants; knowledge; attitude; practice

Introduction

Monkeypox (mpox) is a zoonotic infectious disease caused by the mpox virus (MPXV) [1], with universal susceptibility in the human population. Its clinical manifestations are similar to those of smallpox, including fever, rash, and malaise. A key difference from smallpox is the higher prevalence of lymphadenopathy in mpox cases [2]. The first human case of mpox was identified in the Democratic Republic of the Congo (then known as Zaire) in 1970, after which the virus was primarily endemic in west

XSL•FO
RenderX

and central Africa [3]. Since May 2022, a large-scale mpox outbreak has re-emerged globally [4]. In July 2024, a more virulent new strain of MPXV (clade Ib) began spreading rapidly in the Democratic Republic of the Congo. Initially prevalent among sex workers, this strain has now spread to other populations [5,6]. As of now, a total of 162,785 mpox cases and 424 deaths have been reported across 140 countries worldwide [7]. In China, 767 infections have been documented in 2025 [8]. The mpox epidemic may have catastrophic consequences for public health, socioeconomic factors, and the entire health care system. On July 23, 2022, the World Health Organization declared the mpox epidemic a "Public Health Emergency of International Concern" [9], and reaffirmed this declaration on August 14, 2024 [10]. On September 15, 2023, the National Health Commission of the People's Republic of China classified mpox as a category B infectious disease for management, which took effect on September 20, 2023 [11]. In accordance with the Law of the People's Republic of China on the Prevention and Control of Infectious Diseases, category B infectious diseases are subject to strict management [12].

Public health measures to prevent mpox transmission include enhancing public awareness of the disease; however, vaccination remains the core measure. Previous studies have demonstrated that smallpox vaccination provides at least 85% effectiveness in preventing MPXV infection [13]. The World Health Organization recommends vaccination for priority populations, but vaccine hesitancy is widespread globally [14-17]. Previous experience with COVID-19 vaccine hesitancy highlights the need for targeted research on vaccination willingness toward mpox [18,19].

Medical students represent a crucial demographic in epidemic preparedness. As future frontline health care professionals and influential health communicators, their knowledge, attitudes, practices (KAP), and willingness to be vaccinated directly will impact clinical response capacity, the effectiveness of community health education, and public health messaging during outbreaks. While prior research indicates a generally high willingness to receive the mpox vaccine among Chinese health care workers [17], data regarding medical students' KAP related to mpox and their willingness to receive the vaccine are lacking. This study used a cross-sectional online survey to investigate the mpox-related KAP of Chinese medical students and their willingness to be vaccinated against mpox, as well as the factors influencing their vaccination willingness. These factors may serve as key targets for controlling the current mpox outbreak and rapidly responding to future epidemic outbreaks. Therefore, this study provides a basis for policy decisions regarding targeted educational interventions and vaccination strategies, thereby enhancing preparedness for potential mpox outbreaks.

Methods

Study Design

A nationwide population-based cross-sectional survey was conducted from November 2023 to March 2024, and the snowball sampling method was adopted to recruit participants. Data were collected through an anonymous, self-administered questionnaire designed to assess Chinese medical students' demographic characteristics and KAP related to mpox, as well as vaccination-related behaviors and willingness. The study follows the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines (Checklist 1).

Study Objects

The study population met the following inclusion criteria: (1) full-time medical students in mainland China, (2) currently enrolled in a degree program, and (3) willing to complete a questionnaire. A total of 4320 students participated in the survey, among whom 4098 valid questionnaires were collected, with an effective response rate of 94.86%.

Survey Methods

An anonymous online survey was conducted via the Wenjuanxing platform, a professional online survey platform in China, and snowball sampling was used to recruit participants. All the questionnaires were edited, and an electronic 2D code for the survey was generated through the Wenjuanxing online platform. The process was initiated by the research team contacting academic advisors and student leaders from various medical schools. These initial contacts were provided with a standardized recruitment package, which included a brief study description, eligibility criteria, and the unique 2D code link to the questionnaire. The initial contacts distributed the recruitment package to potential participants via QQ (Tencent), a popular instant messaging software in China, and WeChat (Tencent), a multifunctional social media and messaging app. Specifically, dissemination occurred through a combination of private messages, academic group chats, and institutional social media channels. Additionally, participants were encouraged to assist in recruiting other study individuals by sharing the electronic 2D code of the questionnaire.

Students scanned the 2D code of the questionnaire using electronic devices, participated in the survey voluntarily, and completed the questionnaire. All items in the questionnaire were set as mandatory to ensure data integrity. Each IP address was restricted to submitting the questionnaire only once to prevent duplicate submissions. Furthermore, the time spent completing the questionnaire was automatically monitored by the Wenjuanxing platform; responses were considered invalid if the completion time was less than 180 seconds.

Survey Instruments

A self-designed questionnaire was used to investigate medical students' KAP regarding mpox, their willingness to receive mpox vaccination, and the influencing factors thereof. The questionnaire was developed on the basis of the *Technical Guidelines for Mpox Prevention and Control* (2022 version) [20]. To ensure its content validity, the initial draft was reviewed by 4 experts in public health and health education, after which it was revised. Subsequently, a pilot survey was conducted to test the clarity and feasibility of the instrument, leading to the final version used in the study.

The questionnaire consisted of 6 sections with a total of 81 items, including basic information (20 items), knowledge (23 items), attitude (11 items), practice (5 items), vaccination-related

XSL•FO
RenderX

behaviors (21 items), and vaccination willingness (1 item). For the knowledge section, 2 points were awarded for each correct answer to single-choice questions, and no points for incorrect answers; for multiple-choice questions, 1 point was given for each correctly selected option. The attitude, practice, vaccination-related behaviors, and vaccination willingness sections were scored using a 5-point Likert scale, with response options of "Strongly disagree," "Disagree," "Neutral," "Agree," and "Strongly agree," corresponding to scores of 1, 2, 3, 4, and 5, respectively. Respondents were required to rate each question. The total score of the questionnaire was 245 points. The maximum scores for the 4 sections, knowledge, attitude, practice, and vaccination-related behaviors, were 60, 55, 25, and 105 points, respectively. The Cronbach α coefficient of the questionnaire was 0.915, indicating good reliability of the questionnaire.

Statistical Methods

SPSS (Statistical Package for the Social Sciences) 24.0 software (IBM Corp) was used for data analysis. Categorical data were presented as frequency (constituent ratio), and comparisons between groups were performed using the chi-square test or the Fisher exact test. The Kolmogorov-Smirnov test was used to test the normality of continuous variables. Continuous variables that conformed to a normal distribution were expressed as mean (SD), and comparisons between groups were conducted using the 2-tailed *t* test or ANOVA. Continuous variables that did not conform to a normal distribution were presented as median [IQR], and comparisons between groups were performed using the Kruskal-Wallis *H* test. Multinomial logistic regression was performed to identify the factors correlated with the willingness to receive the mpox vaccine. The willingness to receive the mpox vaccine was taken as the dependent variable, which included 3 categories: acceptance, hesitancy, and rejection. In the multinomial logistic regression model, the status of the acceptance group served as the reference category for the dependent variable. Variables with $P<.10$ in univariable analyses were considered as independent variables in the multinomial logistic regression model. A stepwise selection method (entry and removal criteria: $P=.05$) was used for predictor selection. The risks of the factors were displayed as odds ratios (OR) and the corresponding 95% CIs. The *P* value <.05 (2-sided) was considered statistically significant.

Ethical Considerations

This study was approved by the ethics committee of Qiqihar Medical University (approval: 202371). This study adhered to the Declaration of Helsinki. Participants in this research participated voluntarily and were conducted anonymously through an online platform. The questionnaire did not include any identifying or sensitive content. All participants provided online informed consent. Participants did not receive any financial or nonfinancial compensation for their participation.

Results

Demographic Characteristics of Medical Students and Their Willingness to Receive Mpox Vaccination

A total of 4098 participants were included in the study, which included 1739 male participants and 2359 female participants. Among them, 3468 individuals accepted mpox vaccination, accounting for 84.63%; 550 individuals expressed hesitancy, representing 13.42%; and 80 individuals refused, making up 1.95%.

The willingness of medical students to receive mpox vaccination showed significant differences between sex ($\chi^2_2=42.5$; $P<.001$), grade ($\chi^2_8=23.8$; $P=.003$), being an only child ($\chi^2_2=8.2$; $P=.02$), the reason for choosing medicine ($\chi^2_2=11.1$; $P=.004$), having parents or relatives who worked in the frontline during the COVID-19 outbreak ($\chi^2_2=10.0$; $P=.007$), self-health status ($\chi^2_8=59.4$; $P<.001$), chronic disease ($\chi^2_2=14.1$; $P<.001$), allergic constitution ($\chi^2_2=8.6$; $P=.01$), organic disease ($\chi^2_2=26.0$; $P<.001$), having family members or friends who had experienced COVID-19 ($\chi^2_2=40.6$; $P<.001$), having received the influenza vaccination ($\chi^2_2=24.2$; $P<.001$), having received the COVID-19 vaccination ($\chi^2_2=47.6$; $P<.001$), and having considered the reason for receiving COVID-19 vaccination ($\chi^2_2=35.4$; $P<.001$). The details are presented in Multimedia Appendix 1.

KAP Status Regarding Mpox Among Medical Students and Their Willingness to Receive Mpox Vaccination

The maximum scores for the 4 sections—knowledge, attitude, practice, and vaccination-related behaviors—were 60, 55, 25, and 105 points, respectively. The median scores for KAP status regarding mpox among medical students were 43 (IQR 33-50), 33 (IQR 32-36), and 20 (IQR 19-24), respectively. The median for practice status regarding receiving the mpox vaccine among medical students was 73 (IQR 68-79). A significant difference was found between willingness to receive mpox vaccination and knowledge level regarding mpox ($H=236.2$; $P<.001$), practice status regarding mpox ($H=945.2$; $P<.001$), and practice status regarding receiving mpox vaccine ($H=1174.2$; $P<.001$). Table 1 outlines the details, and Figure 1 illustrates the results.

XSL•FO
RenderX

Table . Knowledge, attitude, and practice (KAP) status regarding monkeypox (mpox) among Chinese medical students and its association with their willingness to receive mpox vaccination.

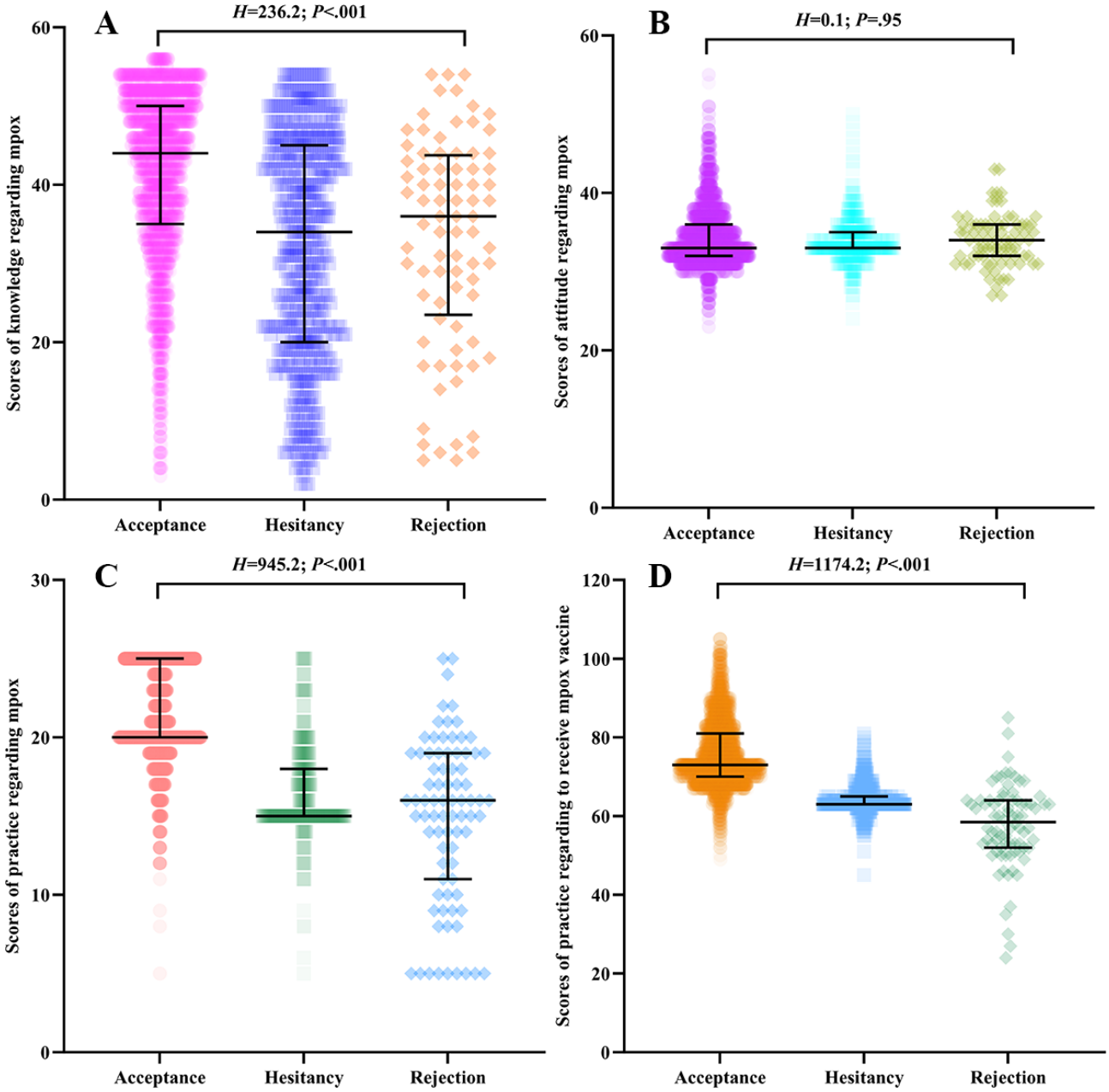
Variables	All participants, median (IQR)	Willingness to receive mpox vaccine, median (IQR)			*H* value[a]	*P* value
		Acceptance	Hesitancy	Rejection		
Knowledge status regarding mpox	43 (33-50)	44 (35-50)	34 (20-45)	36 (24-43.5)	236.2	<.001[b]
Attitude status regarding mpox	33 (32-36)	33 (32-36)	33 (33-35)	34 (32-36)	0.1	.95
Practice status regarding mpox	20 (19-24)	20 (20-25)	15 (15-18)	16 (11-19)	945.2	<.001[b]
Practice status regarding receiving mpox vaccine	73 (68-79)	73 (70-81)	63 (63-65)	58.5 (52-64)	1174.2	<.001[b]

[a]Kruskal-Wallis *H* test.

[b]$P<.05$ and the difference was statistically significant.

XSL•FO
RenderX

Figure 1. KAP status regarding monkeypox (mpox) and practice status regarding receiving mpox vaccination among Chinese medical students: (A) scores of knowledge regarding mpox; (B) scores of attitude regarding mpox; (C) scores of practice regarding mpox; and (D) scores of practice regarding receiving mpox vaccine. The numbers are represented by median (IQR). KAP: knowledge, attitude, and practice.



Multinomial Logistic Regression for Identification of the Factors Correlated With the Willingness to Receive Mpox Vaccination

Compared with the acceptance group, factors such as sex, only-child status, history of COVID-19 infection among students or their relatives and friends, prior influenza vaccination, and knowledge level regarding mpox were significantly correlated with hesitancy to receive the mpox vaccine. Male sex was a risk factor for vaccine hesitancy, with a 1.416-fold risk than female sex (OR 1.416, 95% CI 1.158-1.732). Compared with their counterpart groups, participants who were the only child in their family (OR 1.340, 95% CI 1.098 - 1.635), those who had never contracted COVID-19 with their relatives or friends (OR 1.520, 95% CI 1.161-1.991), those who had not received the influenza vaccine (OR 1.429, 95% CI 1.146 - 1.783), and those with low knowledge level regarding mpox (OR 0.948, 95% CI 0.941 - 0.955) displayed an increased hesitancy to receive mpox vaccination. Table 2 outlines the details, and Figure 2 illustrates the results.

XSL•FO
RenderX

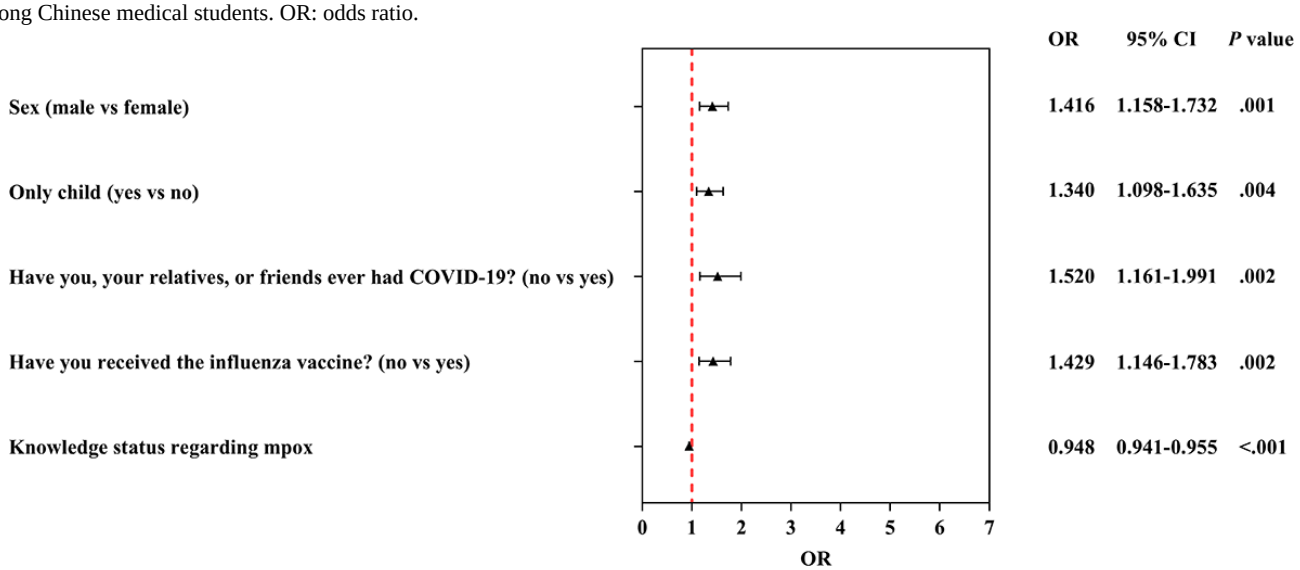
Table . Multivariable analysis of factors correlated with monkeypox (mpox) vaccination willingness among medical students in China[a].

Variables	β		Wald chi-square (*df*)	*P* value	OR[b] (95% CI)
Hesitancy					
Sex (reference: female)					
Male	0.3481	0.1028	11.5 (1)	.001	1.416 (1.158 - 1.732)
Only child (reference: no)					
Yes	0.2928	0.1016	8.3 (1)	.004	1.340 (1.098 - 1.635)
Have you, your relatives, or friends ever had COVID-19? (reference: yes)					
No	0.4190	0.1376	9.3 (1)	.002	1.520 (1.161 - 1.991)
Have you received the influenza vaccine? (reference: yes)					
No	0.3573	0.1127	10.1 (1)	.002	1.429 (1.146 - 1.783)
Knowledge status regarding mpox	−0.0537	0.0039	194.5 (1)	<.001	0.948 (0.941 - 0.955)
Rejection					
Sex (reference: female)					
Male	0.4955	0.2513	3.9 (1)	.049	1.641 (1.003 - 2.686)
Grade	0.3662	0.1137	10.4 (1)	.001	1.442 (1.154 - 1.802)
Do your parents or relatives engage in frontline work during the COVID-19 outbreak? (reference: no)					
Yes	0.8079	0.2640	9.4 (1)	.002	2.243 (1.337 - 3.764)
Do you have internship experience? (reference: yes)					
No	0.7174	0.3285	4.8 (1)	.03	2.049 (1.076 - 3.901)
Self-health status	−0.2898	0.1298	5.0 (1)	.03	0.748 (0.580 - 0.965)
Organic disease (reference: no)					
Yes	1.3171	0.3785	12.1 (1)	.001	3.733 (1.778 - 7.838)
Knowledge status regarding mpox	−0.0467	0.0090	27.3 (1)	<.001	0.954 (0.938 - 0.971)

[a]The willingness to receive the mpox vaccine was used as the dependent variable, which included 3 categories: acceptance, hesitancy, and rejection. In the multinomial logistic regression model, the status of the acceptance group served as the reference category for the dependent variable.

[b]OR: odds ratio.

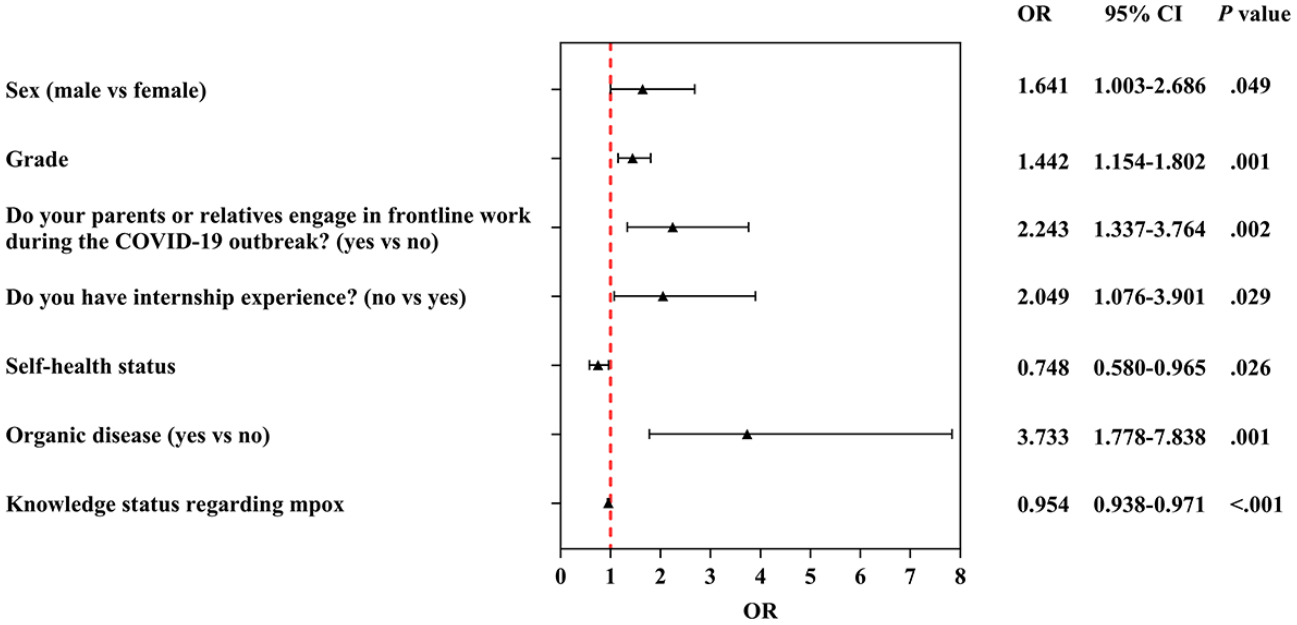
Figure 2. Multinomial logistic regression analysis of the determinants significantly associated with hesitancy to receive monkeypox (mpox) vaccination among Chinese medical students. OR: odds ratio.



XSL•FO
RenderX

Compared with the acceptance group, factors such as sex, grade, parents' or relatives' engagement in frontline work during the COVID-19 outbreak, internship experience, self-reported health status, presence of organic disease, and knowledge level regarding mpox were significantly correlated with rejection of mpox vaccination. Male sex was the risk factor for rejection status, with a 1.641-fold risk compared to the female sex (OR 1.641, 95% CI 1.003 - 2.686). Grade was the risk factor for rejection status with a 1.442-fold risk as grade increased (OR 1.442, 95% CI 1.154 - 1.802). Self-health status was the protective factor for rejection status as the health scores increased (OR 0.748, 95% CI 0.580 - 0.965). Compared with their counterpart groups, participants whose parents or relatives engaged in frontline work during the COVID-19 outbreak (OR 2.243, 95% CI 1.337 - 3.764), those who had no internship experience (OR 2.049, 95% CI=1.076 - 3.901), those who had organic disease (OR 3.733, 95% CI 1.778 - 7.838), and those with low knowledge level regarding mpox (OR 0.954, 95% CI 0.938 - 0.971) displayed an increased rejection to receive mpox vaccination. Table 2 outlines the details, and Figure 3 visualizes the results.

Figure 3. Multinomial logistic regression analysis of the determinants significantly associated with rejection to receive monkeypox (mpox) vaccination among Chinese medical students. OR: odds ratio.



Discussion

Principal Findings

This study conducted a nationwide cross-sectional online survey, targeting 4098 full-time medical students in mainland China to investigate their KAP regarding mpox and their willingness to receive the mpox vaccination systematically. As the first study to focus on Chinese medical students, a group poised to become core public health practitioners, this research fills a critical gap in domestic data on mpox-related health literacy and vaccine acceptance. The findings reveal high overall vaccine acceptance among medical students, alongside clear demographic, health-related, and KAP-associated factors influencing vaccine hesitancy and rejection. These results are not only timely, given the ongoing global mpox epidemic (with 162,785 cases and 424 deaths across 140 countries) and China's classification of mpox as a class B notifiable disease, but also provide actionable evidence for formulating targeted educational interventions and vaccine distribution strategies. By identifying modifiable determinants of vaccine willingness, this study supports enhanced preparedness for potential future mpox outbreaks and strengthens the role of medical students as frontline communicators of public health knowledge.

The results of this study showed that 84.63% of Chinese medical students were willing to accept mpox vaccination, with 13.42% expressing hesitancy and only 1.95% refusing. This acceptance rate is remarkably higher than that reported in most studies on health care–related populations or college students. For example, a study in the Democratic Republic of the Congo involving health care workers found that the willingness rate to receive vaccines was 61% [21]; in a Saudi Arabian study, 52.7% of health care workers were willing to receive vaccines [22]; in Pakistan, the vaccine acceptance rate among college students was 67.7% [23]; while a meta-analysis conducted by Yappalparvi et al [24] showed that only 58.6% of college students were willing to receive the mpox vaccine. Another cross-sectional multinational study targeting health care providers found that the acceptability of the mpox vaccine was just over half of the participants (54.5%) [25]. Notably, the acceptance rate in our study is slightly lower than the 90.1% reported for Chinese health care workers in a 2023 systematic review by Lounis and Riad [17] but remains among the highest globally. The high vaccine acceptance among Chinese medical students can be attributed to three key factors. First, their professional medical background enables them to gain a deeper understanding of the pathogenesis of mpox, transmission risks, and the 85% protective efficacy of smallpox vaccines against mpox [13]. Furthermore, they are more capable of comprehending the high-efficiency protective mechanism of mRNA vaccines developed by scholars, which is achieved through the synergistic activation of "humoral immunity +

XSL•FO
RenderX

cellular immunity" [26,27]. This knowledge reserve mitigates doubts arising from information asymmetry. Second, differences in the policy environment are equally critical. China's strict category B management of mpox, coupled with targeted health education conducted via official guidelines, such as *Monkeypox Prevention and Control Technology* (2022 Edition) [20], has established a clear public health narrative regarding the necessity of vaccines. In addition, the experience accumulated from large-scale COVID-19 vaccination has led to a remarkably higher recognition of vaccines as public health intervention tools among Chinese medical students [28,29]. The marginal gap between medical students and practicing health care workers likely reflects differences in direct clinical exposure risk; health care workers face immediate occupational exposure, while medical students' risk perception is more knowledge-based than experiential.

Multivariate logistic regression identified five independent factors associated with vaccine hesitancy: male sex, being an only child, no personal or familial COVID-19 history, no influenza vaccine uptake, and low mpox knowledge levels. Male participants were 1.416-fold more likely to be hesitant than female participants (OR 1.416, 95% CI 1.158 - 1.732). This result is consistent with the findings of a study on COVID-19 vaccines among Saudi university students, where male individuals often exhibit low health-seeking behavior and a great tendency to downplay infectious disease risks [30]. The observed association in our study may reflect a similar pattern of risk underestimation, which could be a particularly relevant factor in decision-making regarding nonhighly lethal infectious diseases such as mpox. Only children had a 34.0% higher hesitancy risk (OR 1.340, 95% CI 1.098 - 1.635). This unique finding is closely associated with the characteristics of family structures shaped by the long-term impact of China's family planning policy. Only children often face greater parental caution regarding medical interventions, and family decision-making may prioritize avoiding potential, even rare, vaccine risks over mpox prevention. Furthermore, only children may have less exposure to collective health narratives compared with those with siblings, reducing perceived personal responsibility for vaccination. Participants without a history of COVID-19 in themselves or relatives were 1.520-fold more hesitant (OR 1.520, 95% CI 1.161 - 1.991). A study in Arab countries also found that experiencing COVID-19 firsthand or through family members likely reinforces awareness of infectious disease severity and the value of preventive measures [31]. By contrast, those without such experience may perceive mpox as a "distant threat," weakening motivation to accept vaccination. Nonrecipients of influenza vaccines had a 42.9% higher hesitancy risk (OR 1.429, 95% CI 1.146 - 1.783). Influenza vaccine uptake reflects established trust in routine immunization. Nonrecipients may hold persistent doubts about vaccine safety or effectiveness, attitudes that generalize to novel vaccines such as mpox. This result suggests that routine vaccine adherence is a strong predictor of willingness to accept emerging vaccines. The level of correct knowledge about mpox among medical students was associated with vaccine willingness: each unit decrease in knowledge score increased hesitancy (OR 0.948, 95% CI 0.941 - 0.955). This finding echoes the conclusion from a Nigerian study that "knowledge gaps lead to barriers in prevention and control" [32]. Medical students with incomplete knowledge may misunderstand transmission routes or overestimate vaccine side effects, resulting in hesitation. Within the KAP framework, these factors do not exist in isolation. Knowledge of emerging infectious diseases, as an identified influencing factor, may affect individuals' vaccination attitudes, and such attitudes may further impact vaccination willingness. Past vaccination practices may consolidate vaccination attitudes, which is supported by our single-factor analysis results showing that previous influenza and COVID-19 vaccination history strengthened mpox vaccination willingness. These findings indicate that vaccine hesitancy may arise from the interaction of knowledge, attitudes, and practices.

Vaccine refusal was associated with 6 factors: male sex, high academic year, having relatives or friends who worked on the COVID-19 frontline, no internship experience, presence of organic diseases, and low mpox knowledge. Conversely, good self-reported health status was a protective factor. Male participants were 1.641-fold more likely to refuse (OR 1.641, 95% CI 1.003 - 2.686), and each increase in academic year raised refusal risk by 44.2% (OR 1.442, 95% CI 1.154 - 1.802). The sex effect is more pronounced for refusal than hesitancy, suggesting male individuals may hold more entrenched skepticism, possibly due to overconfidence in clinical knowledge or exposure to niche medical debates about vaccine use. Higher-year students, while more knowledgeable about clinical practice, may also be more exposed to anecdotal reports of vaccine adverse events, which could foster increased skepticism toward public health guidelines. This skepticism may manifest as deliberate rejection rather than indecision. Studies on American medical students also found that 25.1% of the students believed that vaccine education was insufficient, and 8.6% refused to encourage patients to receive the COVID-19 vaccine [33]. Such findings suggest that clinical training should strengthen education on "the dialectical relationship between individual cases and population data." Participants with relatives or friends on the COVID-19 frontline had a 2.233-fold higher refusal risk (OR 2.233, 95% CI 1.337 - 3.764). This counterintuitive finding may stem from unique knowledge sources and consequent attitude shifts. Their proximity to frontline workers provides them with first-hand, often stressful, knowledge regarding systemic challenges, health care worker overwork, and rare adverse events. This direct or relayed knowledge can critically shape attitudes, eroding trust in official vaccine recommendations, which is consistent with the research findings on high vaccine hesitancy toward the COVID-19 vaccine in health care workers and health care students worldwide [34]. This finding also aligns with the conclusion that pandemic fatigue among the Chinese public in the postpandemic era may hinder the conversion of vaccine uptake willingness into actual behavior [35], suggesting that authorities must rebuild the trust of professional groups by enhancing transparency in epidemic prevention and control processes. Students without internships were 2.049-fold more likely to refuse (OR 2.049, 95% CI 1.076 - 3.901). Clinical internships expose students to real-world cases and reinforce vaccine use through direct observation of patient outcomes. Without this experiential learning, vaccine refusal may stem from abstract, theoretical skepticism rather than evidence-based reasoning.

XSL•FO
RenderX

Individuals with organic diseases had the highest refusal risk (OR 3.733, 95% CI 1.780 - 7.838). This result is clinically rational: those with chronic conditions may fear vaccine-disease interactions or adverse effects exacerbating their condition. Even with medical training, concerns about individual health vulnerability likely override general preventive benefits, highlighting the need for personalized vaccine counseling [36,37]. Concerns about vaccine safety are also reflected in surveys conducted among college students in southwest China [38]. Low knowledge increased refusal (OR 0.954, 95% CI 0.938 - 0.971), while good health reduced it (OR 0.748, 95% CI 0.580 - 0.965). Healthy students may perceive fewer barriers to vaccination and more benefits. By contrast, knowledge deficits amplify misconceptions [39]; for example, conflating mpox with low-prevalence diseases, leading to active rejection.

Limitations

The following limitations should be acknowledged. First, the snowball sampling in this study was a nonprobability sample, as all participants enrolled on a voluntary basis. Therefore, the recruitment networks may have over-represented students with stronger social connections or a pre-existing interest in public health, which could introduce selection bias. Second, 4 experts in public health and health education reviewed the questionnaire for content validity, while factor analysis results were lacking. Third, self-reported vaccination willingness may be subject to social desirability bias, which may have been heightened by mandatory response items, regulated minimum response time, and the participants' identity as medical students. Fourth, the study sample was confined to medical students, which limits the generalizability of the findings to nonmedical students or other populations. Finally, the use of a self-designed questionnaire may limit comparability with other studies.

Conclusions

This study demonstrates high mpox vaccination acceptance among Chinese medical students. Factors associated with vaccination hesitancy included male sex, being an only child, no history of COVID-19 in family or friends, no influenza vaccination, and low mpox knowledge. Factors associated with vaccination rejection included male sex, high academic grade, family or friends working on COVID-19 frontlines, no internship experience, and presence of organic diseases. Good self-reported health status was a protective factor against refusal. The identified influencing factors provide actionable insights for policymakers, offering a reference for future risk assessment, emergency response, and management during similar public health emergencies. For medical education, the findings point to specific, operational measures: integrating mpox education into medical curricula to address knowledge gaps; designing targeted messaging for higher-risk subgroups such as male individuals and senior students; leveraging medical students as trusted community health advocates; and rebuilding trust through transparent communication, particularly among those affected by pandemic fatigue. These measures not only contribute to controlling current mpox outbreaks but also strengthen long-term public health resilience against future zoonotic disease threats.

Acknowledgments

The authors are deeply grateful to the medical students who took part in the survey. The authors also appreciate ShineWrite for providing professional language editing support.

Funding

This work was supported by the Department of Education, Heilongjiang Province (grant 2024-KYYWF-0365).

Data Availability

The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

Authors' Contributions

JC and YL contributed to conceptualization, project management, resources, supervision, writing, and review. YJ and JG analyzed the data and reviewed the manuscript. HL, YH, and ZX were responsible for data curation, investigation, supervision, and software. JC acquired funding. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Demographic characteristics of Chinese medical students and their willingness to receive monkeypox (mpox) vaccination.
[DOC File, 114 KB - publichealth_v12i1e86981_app1.doc]

Checklist 1
STROBE checklist.
[DOCX File, 45 KB - publichealth_v12i1e86981_app2.docx]

XSL•FO
RenderX

References

1. Gong Q, Wang C, Chuai X, Chiu S. Monkeypox virus: a re-emergent threat to humans. Virol Sin 2022 Aug;37(4):477-482. [doi: 10.1016/j.virs.2022.07.006] [Medline: 35820590]
2. Rahmani E, Bayat Z, Farrokhi M, et al. Monkeypox: A comprehensive review of virology, epidemiology, transmission, diagnosis, prevention, treatment, and artificial intelligence applications. Arch Acad Emerg Med 2024;12(1):e70. [doi: 10.22037/aaem.v12i1.2491] [Medline: 39296520]
3. Velavan TP, Meyer CG. Monkeypox 2022 outbreak: an update. Tropical Med Int Health 2022 Jul;27(7):604-605. [doi: 10.1111/tmi.13785]
4. Asaad A, El-Sokkary R. The current monkeypox outbreak: knowledge gaps and research priorities. J Infect Public Health 2023 Jan;16(1):78-79. [doi: 10.1016/j.jiph.2022.12.002] [Medline: 36495815]
5. Marziano V, Guzzetta G, Longini I, Merler S. Epidemiologic quantities for monkeypox virus clade I from historical data with implications for current outbreaks, Democratic Republic of the Congo. Emerg Infect Dis 2024 Oct;30(10):2042-2046. [doi: 10.3201/eid3010.240665] [Medline: 39255234]
6. Wawina-Bokalanga T, Akil-Bandali P, Kinganda-Lusamaki E, et al. Co-circulation of monkeypox virus subclades Ia and Ib in Kinshasa Province, Democratic Republic of the Congo, July to August 2024. Euro Surveill 2024 Sep;29(38):2400592. [doi: 10.2807/1560-7917.ES.2024.29.38.2400592] [Medline: 39301745]
7. Global mpox trends. World Health Organization. 2025. URL: https://worldhealthorg.shinyapps.io/mpx_global/#global-data [accessed 2025-10-09]
8. National situation of notifiable infectious diseases [Article in Chinese]. Chinese Center for Disease Control and Prevention. 2025. URL: https://www.chinacdc.cn/jksj/jksj01/ [accessed 2026-01-16]
9. Second meeting of the International Health Regulations (2005) (IHR) Emergency Committee regarding the multi-country outbreak of monkeypox. World Health Organization. 2022. URL: https://www.who.int/news/item/23-07-2022-second-meeting-of-the-international-health-regulations-(2005)-(ihr)-emergency-committee-regarding-the-multi-country-outbreak-of-monkeypox [accessed 2025-09-03]
10. WHO Director-General declares mpox outbreak a public health emergency of international concern. World Health Organization. 2024. URL: https://www.who.int/news/item/14-08-2024-who-director-general-declares-mpox-outbreak-a-public-health-emergency-of-international-concern [accessed 2024-09-23]
11. Announcement of the National Health ommission of the People's Republic of China [Article in Chinese]. National Health Commission of the People' s Republic of China. 2023. URL: https://www.nhc.gov.cn/ylyjs/gzdt/202309/a631afc87ad34172b2e96607a43545a2.shtml [accessed 2023-09-20]
12. Law of the People's Republic of China on the prevention and control of infectious diseases [Article in Chinese]. The Central People's Government of the People's Republic of China. 2025. URL: https://www.gov.cn/yaowen/liebiao/202504/content_7021966.htm [accessed 2025-09-20]
13. Sallam M, Al-Mahzoum K, Al-Tammemi AB, et al. Assessing healthcare workers' knowledge and their confidence in the diagnosis and management of human monkeypox: a cross-sectional study in a Middle Eastern country. Healthcare (Basel) 2022 Sep 8;10(9):1722. [doi: 10.3390/healthcare10091722] [Medline: 36141334]
14. Monkeypox. World Health Organization. 2024. URL: https://www.who.int/news-room/questions-and-answers/item/mpox [accessed 2024-10-20]
15. Riad A, Rybakova N, Dubatouka N, et al. Belarusian healthcare professionals' views on monkeypox and vaccine hesitancy. Vaccines (Basel) 2023 Aug 15;11(8):1368. [doi: 10.3390/vaccines11081368] [Medline: 37631936]
16. Mahameed H, Al-Mahzoum K, AlRaie LA, et al. Previous vaccination history and psychological factors as significant predictors of willingness to receive mpox vaccination and a favorable attitude towards compulsory vaccination. Vaccines (Basel) 2023 Apr 25;11(5):897. [doi: 10.3390/vaccines11050897] [Medline: 37243001]
17. Lounis M, Riad A. Monkeypox (MPOX)-related knowledge and vaccination hesitancy in non-endemic countries: concise literature review. Vaccines (Basel) 2023 Jan 19;11(2):229. [doi: 10.3390/vaccines11020229] [Medline: 36851107]
18. Zhao T, Cai X, Zhang S, et al. COVID-19 vaccine hesitancy in Chinese residents: a national cross-sectional survey in the community setting. Hum Vaccin Immunother 2025 Dec;21(1):2481003. [doi: 10.1080/21645515.2025.2481003] [Medline: 40156199]
19. Pereira da Silva A, Castaneda LR, Cavalcante de Oliveira AP, et al. COVID-19 vaccination acceptance and hesitancy in healthcare workers and the general population: a systematic review and policy recommendations. Int J Environ Res Public Health 2024 Aug 28;21(9):1134. [doi: 10.3390/ijerph21091134] [Medline: 39338017]
20. Technical guidelines for mpox prevention and control [Article in Chinese]. National Health Commission of the People's Republic of China. 2022. URL: https://www.nhc.gov.cn/wjw/c100378/202207/b4d0736aebbe438a8d049dd0cb92df8c.shtml [accessed 2023-09-13]
21. Petrichko S, Kindrachuk J, Nkamba D, et al. Mpox vaccine acceptance, Democratic Republic of the Congo. Emerg Infect Dis 2024 Dec;30(12):2614-2619. [doi: 10.3201/eid3012.241226] [Medline: 39486156]

XSL•FO
RenderX

22. Alarifi AM, Alshahrani NZ, Sah R. Are Saudi healthcare workers willing to receive the monkeypox virus vaccine? Evidence from a descriptive-baseline survey. Trop Med Infect Dis 2023 Aug 2;8(8):396. [doi: 10.3390/tropicalmed8080396] [Medline: 37624334]
23. Kumar N, Ahmed F, Raza MS, et al. Monkeypox cross-sectional survey of knowledge, attitudes, practices, and willingness to vaccinate among university students in Pakistan. Vaccines (Basel) 2022 Dec 31;11(1):97. [doi: 10.3390/vaccines11010097] [Medline: 36679942]
24. Yappalparvi A, Gaidhane S, Padmapriya G, et al. Prevalence of mpox vaccine acceptance among students: a systematic review and meta-analysis. Vaccines (Basel) 2025 Feb 13;13(2):183. [doi: 10.3390/vaccines13020183] [Medline: 40006730]
25. Swed S, Alibrahim H, Bohsas H, et al. A multinational cross-sectional study on the awareness and concerns of healthcare providers toward monkeypox and the promotion of the monkeypox vaccination. Front Public Health 2023;11:1153136. [doi: 10.3389/fpubh.2023.1153136] [Medline: 37228716]
26. Tai W, Tian C, Shi H, et al. An mRNA vaccine against monkeypox virus inhibits infection by co-activation of humoral and cellular immune responses. Nat Commun 2025 Mar 26;16(1):2971. [doi: 10.1038/s41467-025-58328-x] [Medline: 40140411]
27. Mucker EM, Freyn AW, Bixler SL, et al. Comparison of protection against mpox following mRNA or modified vaccinia Ankara vaccination in nonhuman primates. Cell 2024 Oct 3;187(20):5540-5553. [doi: 10.1016/j.cell.2024.08.043] [Medline: 39236707]
28. Gao L, Su S, Du N, et al. Medical and non-medical students' knowledge, attitude and willingness towards the COVID-19 vaccine in China: a cross-sectional online survey. Hum Vaccin Immunother 2022 Nov 30;18(5):2073757. [doi: 10.1080/21645515.2022.2073757] [Medline: 35612817]
29. Lucia VC, Kelekar A, Afonso NM. COVID-19 vaccine hesitancy among medical students. J Public Health (Oxf) 2021 Sep 22;43(3):445-449. [doi: 10.1093/pubmed/fdaa230] [Medline: 33367857]
30. Daher S, Fakhoury HMA, Tamim H, et al. Attitudes toward COVID-19 vaccination and clinical trials among Saudi university students. J Epidemiol Glob Health 2025 Apr 16;15(1):61. [doi: 10.1007/s44197-025-00393-y] [Medline: 40237969]
31. Swed S, Bohsas H, Alibrahim H, et al. Monkeypox post-COVID-19: knowledge, worrying, and vaccine adoption in the Arabic general population. Vaccines (Basel) 2023 Mar 29;11(4):759. [doi: 10.3390/vaccines11040759] [Medline: 37112671]
32. Iliyasu Z, Kwaku AA, Nass NS, et al. Risk perception and mpox vaccine acceptability among people living with HIV in northern Nigeria. Trans R Soc Trop Med Hyg 2025 May 1;119(5):487-497. [doi: 10.1093/trstmh/trae135] [Medline: 39731207]
33. Vangsness KL, Eklof L, Moore S, Coffey T, Evans J, Sloan R. Medical students' attitudes toward the COVID-19 vaccine and medical school vaccine education: a survey study. Cureus 2024 Dec;16(12):e75168. [doi: 10.7759/cureus.75168] [Medline: 39759758]
34. McCready JL, Nichol B, Steen M, Unsworth J, Comparcini D, Tomietto M. Understanding the barriers and facilitators of vaccine hesitancy towards the COVID-19 vaccine in healthcare workers and healthcare students worldwide: an umbrella review. PLoS ONE 2023;18(4):e0280439. [doi: 10.1371/journal.pone.0280439] [Medline: 37043505]
35. Qin C, Deng J, Du M, et al. Pandemic fatigue and vaccine hesitancy among people who have recovered from COVID-19 infection in the post-pandemic era: cross-sectional study in China. Vaccines (Basel) 2023 Oct 5;11(10):1570. [doi: 10.3390/vaccines11101570] [Medline: 37896973]
36. Gagneur A, Gutnick D, Berthiaume P, Diana A, Rollnick S, Saha P. From vaccine hesitancy to vaccine motivation: a motivational interviewing based approach to vaccine counselling. Hum Vaccin Immunother 2024 Dec 31;20(1):2391625. [doi: 10.1080/21645515.2024.2391625] [Medline: 39187772]
37. Ding J, Liu XC, Hong J, et al. Knowledge about, attitudes toward and acceptance and predictors of intention to receive the mpox vaccine among cancer patients in China: a cross-sectional survey. Hum Vaccin Immunother 2024 Dec 31;20(1):2337157. [doi: 10.1080/21645515.2024.2337157] [Medline: 38644633]
38. Yang X, Yang X, Jiang W, et al. A cross-sectional investigation of factors influencing mpox vaccine hesitancy for students in southwest China. Hum Vaccin Immunother 2024 Dec 31;20(1):2309704. [doi: 10.1080/21645515.2024.2309704] [Medline: 38300140]
39. Suleiman AK. Knowledge, attitudes, and practices toward Mpox and vaccination: a cross-sectional study in Saudi Arabia. Libyan J Med 2025 Dec;20(1):2528299. [doi: 10.1080/19932820.2025.2528299] [Medline: 40591822]

Abbreviations

KAP: knowledge, attitudes, and practices
mpox: monkeypox
MPXV: monkeypox virus
OR: odds ratio
SPSS: Statistical Package for the Social Sciences
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 03.Nov.2025; peer-reviewed by A Girma, O Allela; revised version received 29.Dec.2025; accepted 30.Dec.2025; published 06.Feb.2026.

Please cite as:
Liu Y, Jia Y, Li H, Ge J, Han Y, Xie Z, Chen J
Knowledge, Attitudes, Practices, and Vaccination Willingness Toward Mpox (Monkeypox) Among Chinese Medical Students: Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e86981
URL: https://publichealth.jmir.org/2026/1/e86981
doi:10.2196/86981

© Yang Liu, Yuehui Jia, Honglong Li, Jie Ge, Yunfeng Han, Zhiping Xie, Jiaxin Chen. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 6.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Digital Exclusion Among People Experiencing Homelessness and Residents of Urban Communities in Brazil: Cross-Sectional Study

Ariela Fehr Tártaro[1], BS; Dulce Gomes[2], Prof Dr; Thaís Zamboni Berra[1], PhD; Reginaldo Bazon Vaz Tavares[1], BS; Yan Mathias Alves[1], MSc; Letícia Perticarrara Ferezin[1], PhD; Antônio Carlos Vieira Ramos[3], Prof Dr; Nathalia Zini[1], PhD; Maria Eduarda Pagano Pelodan[1], BS; Marcela Antunes Paschoal Popolin[1], Prof Dr; Ricardo Alexandre Arcêncio[1], Prof Dr

[1]Escola de Enfermagem de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, Brazil

[2]Departamento de Matemática, Escola de Ciências e Tecnologia, Universidade de Évora, Évora, Portugal

[3]Unidade Acadêmica de Passos, Universidade do Estado de Minas Gerais, Passos, Brazil

Corresponding Author:
Ariela Fehr Tártaro, BS
Escola de Enfermagem de Ribeirão Preto
Universidade de São Paulo
Avenida dos Bandeirantes, 3900 Campus Universitário - Bairro Monte Alegre
Ribeirão Preto, 14040-902
Brazil
Phone: 55 16 33154408
Email: ariela.fehr@gmail.com

Abstract

Background: The COVID-19 pandemic amplified digital divides in Brazil, restricting vulnerable groups' online access to health information and preventive guidance, with limited intersectional analyses of these inequities.

Objective: This study aimed to investigate inequalities in digital exclusion and access to online COVID-19 information among people experiencing homelessness and residents of urban communities in Brazil by using an intersectional multilevel analysis.

Methods: A cross-sectional study (2021-2023) involving 2652 participants (n=1353, 51% experiencing homelessness and n=1299, 49% from urban communities across 26 state capitals) was conducted using the adapted COVID-19 Social Thermometer questionnaire administered via face-to-face interviews. Multilevel analysis of individual heterogeneity and discriminatory accuracy examined 115 intersectional strata (gender, race and ethnicity, schooling, income, and Brazilian Unified Health System use) with online COVID-19 information seeking as the binary outcome; multilevel logistic models estimated additive effects and between-strata variance.

Results: Most participants were men (1600/2652, 60.3%), self-identified as Black or Brown individuals (1942/2652, 73.2%), and were Unified Health System users (2433/2652, 91.7%) without private insurance (2469/2652, 93.1%). Over one-third (905/2652, 34.1%) had no formal schooling; 62.4% (1655/2652) reported low income. A total of 39.2% (1040/2652) sought online COVID-19 information. Being a woman (odds ratio [OR] 1.49, 95% CI 1.13-1.97), higher schooling (OR 1.78-5.59, 95% CI 3.52-8.88), and higher income (OR 2.37-4.54, 95% CI 2.59-7.93) showed a stronger association with online COVID-19 information seeking; public health system use was not associated with the outcome (OR 0.92, 95% CI 0.64-1.33). Predicted probabilities ranged between 14% and 85% across 115 strata, with the lowest among Black or Brown men (no schooling or low income) and the highest among women and higher schooling or income. The intersectional analysis (n=2405) null model showed 24% between-strata variance; the full additive model reduced it to 1% (proportional change in variance=97%).

Conclusions: Intersectional analysis reveals structural informational exclusion driven by additive disadvantages in schooling, income, and gender among participants, calling for digital inclusion policies, critical health literacy programs, and equitable multichannel communication strategies to address persistent COVID-19 information seeking disparities.

(JMIR Public Health Surveill 2026;12:e77124) doi:10.2196/77124

KEYWORDS

digital health; intersectionality; COVID-19; information-seeking behavior; internet

XSL•FO
RenderX

Introduction

The COVID-19 pandemic exacerbated structural inequalities in health and information systems, highlighting barriers to accessing reliable content [1,2]. In Brazil, the unequal distribution of digital infrastructure, combined with low health literacy, a phenomenon known as the digital divide, undermined the ability of vulnerable social groups to access and apply relevant information for adopting preventive measures [3,4].

Digital exclusion refers to the inability of individuals or groups to fully participate in digital society due to barriers in access, skills, or use, resulting in limited benefits from online services such as health information [5].

In the COVID-19 context, it manifested as restricted access to online pandemic guidance, exacerbated by 3 layers: physical or material access (device or internet costs), skills (digital literacy), and outcomes (effective information seeking) [5]. During lockdowns, services shifted online, excluding people in vulnerable situations without smartphones (eg, 22% of UK adults lacked basic digital skills before the pandemic), creating "informational exclusion" in which low-income, low-literacy groups missed preventive advice, worsening health inequities [6].

Despite growing evidence of the digital divide in Brazil, most studies focus on the general population or on socioeconomically disadvantaged households and rarely include people experiencing homelessness or residents of urban communities as specific populations of interest [4-9]. Furthermore, previous work has typically analyzed social dimensions such as gender, race, educational level, income, and use of health services separately rather than examining how these characteristics intersect to shape digital exclusion and access to online COVID-19 information.

By applying an intersectional multilevel analysis to people experiencing homelessness and residents of urban communities in Brazil, this study addresses this gap and provides new evidence on how overlapping social positions jointly influence inequalities in digital exclusion.

From a methodological perspective, few studies have used multilevel analysis of individual heterogeneity and discriminatory accuracy (MAIHDA), a statistical approach that estimates the effects of specific social combinations on outcomes and can overcome several limitations of traditional additive models [10,11]. Grounded in Black feminist scholarship, this approach provides robust tools for measuring structural inequality [12-14].

In this context, the COVID-19 pandemic presents a critical opportunity to identify patterns of digital and informational exclusion that, if not addressed, are likely to persist and deepen in future public health emergencies. Therefore, this study aimed to investigate inequalities in digital exclusion and access to online COVID-19 information among people experiencing homelessness and residents of urban communities in Brazil using an intersectional multilevel analysis to examine how overlapping social characteristics jointly shape these disparities.

Methods

Study Design

This was a cross-sectional study guided by the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines [15,16].

Study Setting

This study was conducted in all 26 Brazilian state capitals and the Federal District, covering the 5 major regions of the country: north, northeast, central-west, southeast, and south. This geographic coverage was defined to encompass diverse urban and regional contexts considering Brazil's socioterritorial diversity and the deep structural inequalities that characterize the country [17,18].

Brazil has a total area of approximately 8.5 million km^2 and an estimated population of 203 million [19]. The target population of this study included people experiencing homelessness and residents of urban communities to capture their experiences and vulnerabilities during the COVID-19 pandemic.

The selection of these groups is justified by historical inequalities in access to basic rights, including health care, housing, and information, which were exacerbated during the pandemic [20,21]. It is estimated that the country has approximately 281,400 individuals experiencing homelessness and approximately 16.4 million residents of urban communities, representing 8.1% of the national population [22,23].

Study Population and Sample

Inclusion criteria were Brazilian nationality, age of ≥18 years, and either residence in an urban community or living in a situation of homelessness. For the intersectional analysis, participants who did not provide responses to key variables, namely, gender, race and ethnicity, income, place of residence, use of the Unified Health System (SUS; *Sistema Único de Saúde* in Portuguese), and information-seeking behavior related to COVID-19, were excluded. This represented 9.3% (247/2652) of the total sample.

Urban communities refer to densely populated territories characterized by unregulated growth and inadequate infrastructure. In contrast, people experiencing homelessness are those without conventional housing, living in public spaces such as sidewalks, squares, and abandoned buildings and exposed to multiple forms of vulnerability [24].

Sampling was sequential, where participants were invited to take part as they were approached in public spaces, shelters, hostels, boarding houses, and urban communities provided they were willing to take part and gave informed consent [25].

Due to the unique characteristics of each of Brazil's macroregions, the sample did not have a fixed size for each subgroup. Therefore, we applied the standard formula for simple random sampling with finite populations following classic references for opinion polls, epidemiological research, and surveys [26]. Specifically, the Cochran formula for infinite populations is as follows:

XSL•FO
RenderX

The adjusted Cochran formula for finite populations is as follows:

In these formulas, z is the percentile of the standard normal distribution, ε is the margin of error, N is the population size, and is the estimated population proportion. On the basis of the calculation for finite populations, a minimum sample size of 385 individuals was adopted for the vulnerable population considering a random margin of error of 5%, a confidence level of 95%, a statistical power of 80%, and a variance of 50%. Additionally, a 10% increase was calculated to account for possible sample losses [26,27].

Data Collection Instrument

The instrument used for data collection was the COVID-19 Social Thermometer: Social Opinion questionnaire, originally developed by the National School of Public Health at NOVA University Lisbon and later adapted and validated for the Brazilian context using the Delphi method. The questionnaire comprises 141 variables, including structured questions formatted as checklists, multiple-choice items, and Likert-scale responses [28].

To ensure data integrity and security, the questionnaire was hosted on REDCap (Research Electronic Data Capture; Vanderbilt University), a secure, web-based platform widely used by academic institutions worldwide for the collection and management of clinical and epidemiological research data. The REDCap version used in this study was installed at the University of São Paulo (USP) Ribeirão Preto campus and includes features that ensure response traceability and internal consistency [29].

Data Collection

Data collection was conducted between 2021 and 2023 across multiple urban settings characterized by high social vulnerability. Participant mobilization took place through a coordinated network of professionals from research institutions, universities, civil society organizations, and social movements, which facilitated access to different territories and ensured broad geographic coverage.

Participants were recruited as they were encountered in the field following predefined routes and schedules established by the research team in collaboration with local partners. Recruitment took place in public spaces (such as parks, squares, and streets), shelters, social assistance centers, hostels, boarding houses, and urban communities.

Among people experiencing homelessness, interviewers circulated through designated areas previously identified by local teams familiar with the territory and the population. Recruitment followed a standardized protocol: individuals were approached individually, informed about the study, and invited to participate if they met the eligibility criteria and demonstrated willingness to participate. This procedure also ensured participant safety and adherence to ethical standards.

In urban communities, recruitment occurred with the support of local partners in each city, who acted as intermediaries. These groups identified the most appropriate locations and times for conducting interviews, facilitated safe entry into the territory, and guided the research team toward areas with greater resident flow. This collaboration contributed to a systematic recruitment approach and enhanced both the feasibility and safety of fieldwork.

All interviews were conducted face-to-face using mobile devices (tablets or smartphones) by interviewers who were previously trained and exclusively assigned to the project. Training emphasized consistent administration of the questionnaire, standardized participant approach techniques, and the reduction of potential measurement bias. Each interview lasted approximately 20 to 30 minutes and was conducted only once, a strategy necessary due to the high mobility of people experiencing homelessness.

Study Variables

The outcome variable in this study was the search for information about COVID-19 on the internet, operationalized as a dichotomous variable (“yes” or “no”). In the questionnaire, the question used was as follows: “Which sources of information do you use to stay informed about covid-19?” Among the 8 available response options, the category “internet” was isolated as the main outcome indicator. Participants who selected this option were classified as “yes,” and those who did not select it were classified as “no.”

The internet was selected as the outcome variable due to its central role as a mediator of access to information during public health crises. During the pandemic, the web became the primary channel for disseminating health-related content. However, its use depends on physical access, connectivity, digital literacy, and information seeking skills. Thus, internet use constitutes not only an informational practice but also an indirect marker of structural inequality. By analyzing this variable in isolation, the aim was to identify social profiles with a lower propensity to rely on this resource, particularly among vulnerable populations in territories historically marked by digital exclusion.

To examine differences in information-seeking behavior across dimensions, five socioeconomic characteristics were used to construct intersectional strata: (1) gender (man, woman, or other); (2) race or ethnicity (White; Black or Brown; or other, comprising East Asian and Indigenous identities); (3) household income (low income [no income and less than 1 minimum wage per month], 1 to 2 minimum wages per month, 2 to 3 minimum wages per month, and more than 3 minimum wages per month); (4) schooling (no schooling, basic education, secondary education, and higher education); and (5) use of the SUS (yes and no), the Brazilian public and universal health system, constitutionally guaranteed and responsible for comprehensive health care delivery to the population [30]. Race/ethnicity was self-reported and categorized according to the Brazilian Institute of Geography and Statistics (IBGE) as White, Black, Brown (parda), Asian, and Indigenous.

XSL•FO
RenderX

For race and ethnicity, the categories "other," "East Asian," and "Indigenous" were grouped into a single category due to their very low frequency in the sample. This procedure resulted in 3 analytically robust categories: White, Black or Brown, and "other." A similar limitation occurred for the gender variable, which was collected through self-identification with the response options "woman," "man," and "other," the latter allowing participants to specify another gender identity, including transgender identities. However, no participant selected the "other" option in the final dataset. Consequently, although the study design allowed for the identification of gender diversity, the empirical distribution did not include respondents outside the "woman" or "man" categories.

Furthermore, it is important to highlight that the minimum wage in Brazil varied between R$1100 (US $197.68) and R$1320 (US $237.22) over the course of the study. These variables (sex, race/ethnicity, schooling, household income, and use of the public health system) were selected based on the literature on inequalities in access to information with the goal of analyzing differences in information-seeking behavior across distinct social groups.

Data Analysis

We used the MAIHDA approach within an intersectional framework [11]. Following the 2-model intersectional MAIHDA approach described by Evans et al [10], the analysis was designed to estimate inequalities across social strata, quantify the role of additive main effects, and detect potential intersectional interaction effects.

An intersectional strata matrix was constructed based on the combination of 5 sociodemographic and health-related dimensions: gender (2 categories), race and ethnicity (3 categories), schooling (4 categories), family income (4 categories), and use of the public health system (2 categories).

This resulted in 192 possible intersectional strata (2 × 3 × 4 × 4 × 2) [10,11,31-33]. Among these, 115 intersectional strata were represented in the study sample and constituted the analytical strata. The selection of these combinations was constrained by data availability but achieved the maximum feasible level of intersectional detail.

Variable selection was guided by the premise that inequalities may exist in access to information relevant to coping with COVID-19. On the basis of the matrix, an intersectional MAIHDA model was conducted, with individual-level data nested within intersectional strata. To estimate the probability of seeking COVID-19–related information, we applied a sequence of 2 multilevel logistic regression models [10,11].

The null model (model 1) included only a random intercept for each intersectional stratum with no fixed covariates. From this model, the between-stratum variance (σ^2) was estimated, and the variance partition coefficient (VPC) was calculated on the latent scale assuming a fixed residual variance (approximately σ^2=3.29) according to the latent response approach for logistic models. The VPC represents the proportion of the total variance attributable to differences between strata. We also calculated the area under the receiver operating characteristic curve (AUC) to assess the discriminatory accuracy of the strata in predicting internet use [10,11,31-33].

The second model (model 2) included all variables defining the intersectional strata (race and ethnicity, schooling, family income, and use of the public health system) as fixed effects at level 2 while retaining the stratum random intercept. This model allowed us to break down the between-stratum variance into a component explained by additive main effects and a residual component attributable to intersectional interaction effects [10,11,31-33].

The proportional change in variance (PCV) was calculated to quantify the reduction in between-stratum variance relative to the null model. Its complement (1−PCV) represents the portion of residual variance that may be attributed to intersectional interactions. Additionally, predicted probabilities were estimated for each stratum and broken down into an additive component (based on fixed effects) and an interaction component (the difference between the total predicted probability and the additive component). This allowed for the identification of strata with meaningful deviations from additivity [10,11,31-33].

The AUC was interpreted following the work by Axelsson Fisk et al [11]. In MAIHDA models, the AUC is treated as a comparative measure of the strata's ability to discriminate the outcome, avoiding conventional absolute thresholds for predictive classification. The AUC ranges from 50% (no discriminatory accuracy) to 100% (perfect discriminatory accuracy), reflecting how well intersectional strata distinguish between individuals with and without the outcome [10,11,31-33].

All models were estimated using the maximum likelihood estimation method via the *lme4* package (*lmer* and *glmer* functions) in the R software (version 4.4.0; R Foundation for Statistical Computing) [34].

Ethical Considerations

This study was approved by the Research Ethics Committee of the Ribeirão Preto School of Nursing at USP, with certificate of submission for ethical appraisal (57933622.4.1001.5393). The entire investigation was conducted in accordance with resolution 466 of December 12, 2012, of the National Health Council considering the relevant ethical and scientific foundations. Before participation, all individuals were informed about the purpose of the study, their rights, and the voluntary nature of participation. Participants received no compensation for participation. The informed consent form was read and explained to participants, and only those who agreed and provided written informed consent were interviewed. For participants who were unable to read or write, consent was obtained verbally and documented through a fingerprint signature. Participant confidentiality was ensured throughout the study. Personal information obtained during data collection was stored securely in encrypted REDCap servers hosted at USP, with access limited to authorized members of the research team. Before any analytical procedures were performed, all datasets were anonymized, and any variables that could identify individuals were removed. Statistical analyses were conducted exclusively with deidentified data, and no results are presented

XSL•FO
RenderX

in a manner that could allow for the reidentification of participants.

Results

Descriptive Analysis

A total of 2652 individuals participated in the study (Table 1), of whom 1353 (51%) self-identified as people experiencing homelessness and 1299 (49%) were residents of urban communities. Most participants were men (1600/2652, 60.3%), self-identified as Black or Brown individuals (1942/2652, 73.2%), were uninsured (2469/2652, 93.1%), and used the Brazilian SUS (2433/2652, 91.7%). A considerable proportion of participants had no formal education (905/2652, 34.1%) and reported low income (1655/2652, 62.4%). Regarding the use of the internet to seek information about COVID-19, most (1612/2652, 60.8%) stated that they did not use this source.

Table 1. Socioeconomic characteristics of vulnerable populations seeking online information about COVID-19 (Brazil, 2021-2023; N=2652).

Variable	Values, n (%)
Participant category	
People experiencing homelessness	1353 (51)
People residing in urban communities	1299 (49)
Gender	
Man	1600 (60.3)
Woman	1001 (37.7)
No answer	51 (1.9)
Race or ethnicity	
Black or Brown	1942 (73.2)
White	601 (22.7)
Other	109 (4.1)
Schooling	
No schooling	905 (34.1)
Basic education	726 (27.4)
Secondary education	871 (32.8)
Higher education	148 (5.6)
Family income	
Low income	1655 (62.4)
1 to 2 minimum wages	570 (21.5)
2 to 3 minimum wages	157 (5.9)
More than 3 minimum wages	94 (3.5)
Did not know	176 (6.6)
Use of the public health system	
Yes	2433 (91.7)
No	219 (8.3)
Health insurance	
Yes	172 (6.5)
No	2469 (93.1)
No answer	11 (0.4)
Seeking COVID-19 information online	
Yes	1040 (39.2)
No	1612 (60.8)

XSL•FO
RenderX

MAIHDA Analysis

For the intersectional analysis using the MAIHDA approach, 2405 participants were included after applying eligibility criteria. Table 2 presents the results from the hierarchical logistic models used to estimate the likelihood of seeking information about COVID-19 on the internet based on intersectional strata defined by sociodemographic characteristics.

Table 2. Intersectional multilevel analysis of individual heterogeneity and discriminatory accuracy results with parameter estimates from logistic models predicting seeking online information about COVID-19 among vulnerable populations in Brazil (2021-2023).

	Model 1	Model 2
Fixed effects: regression coefficients, OR[a] (95% CI)		
Intercept	1.02[b] (0.78-1.31)	0.19[b] (0.14-0.25)
Gender		
Man (reference)	—[c]	—
Woman	—	1.49[b] (1.13-1.97)
Race or ethnicity		
Black or Brown	—	—
White (reference)	—	1.05 (0.82-1.34)
Other	—	1.35 (0.82-2.24)
Schooling		
No schooling (reference)	—	—
Basic education	—	1.78[b] (1.32-2.41)
Secondary education	—	3.37[b] (2.53-4.49)
Higher education	—	5.59[b] (3.52-8.88)
Family income		
Low income (reference)	—	—
1-2 minimum wages	—	2.37[b] (1.84-3.06)
2-3 minimum wages	—	2.85[b] (1.87-4.34)
More than 3 minimum wages	—	4.54[b] (2.59-7.93)
Use of the public health system		
Yes (reference)	—	—
No	—	0.92 (0.64-1.33)
Summary statistics		
Variance partition coefficient (%)	24	1
Proportional change in variance (%)	—	97
Area under the receiver operating characteristic curve	0.74	0.73

[a]OR: odds ratio.

[b]Statistically significant at $P<.05$.

[c]Not applicable.

Model 1 (null) showed a VPC of 24%, indicating that nearly a quarter of the total variation in information-seeking behavior lay between social strata; that is, between-group inequality was substantial. The AUC was 0.74.

Model 2, which included the individual components of the strata as additive effects, reduced the VPC to 1%, with a PCV of 97%. This means that almost all the inequality observed between strata can be explained by additive effects of gender, race, educational level, income, and age. The AUC remained practically unchanged (0.73), suggesting that incorporating additive effects did not significantly improve the predictive capacity of the model.

Women were more likely to seek information online than men (odds ratio [OR] 1.49, 95% CI 1.13-1.97), representing a 49% increase in odds relative to the reference group.

Schooling was positively associated with the outcome. Compared to individuals with no formal schooling, those with primary schooling had a 78% higher likelihood of seeking

XSL•FO
RenderX

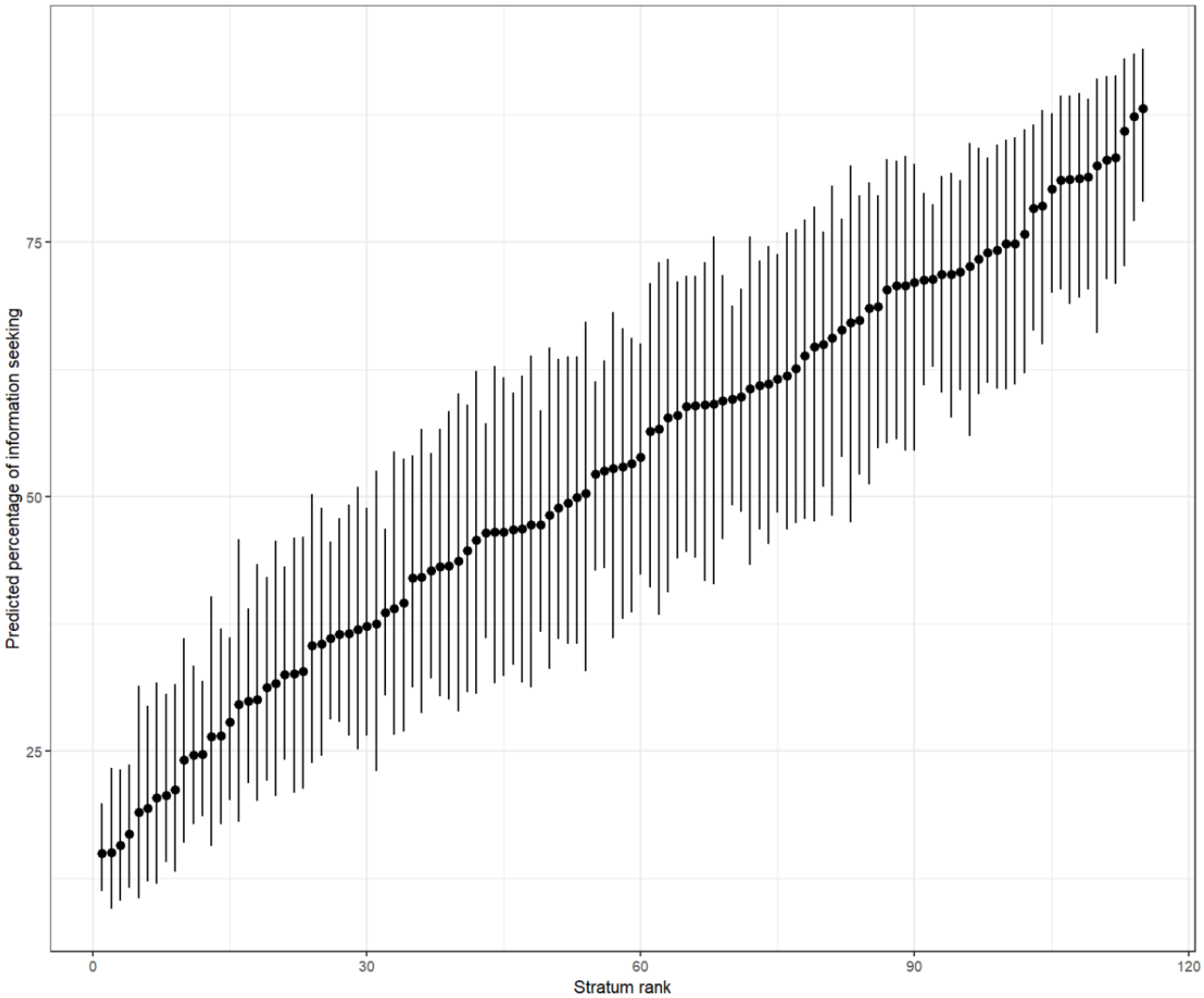
information (OR 1.78, 95% CI 1.32-2.41), those with secondary schooling had a 237% increase (OR 3.37, 95% CI 2.53-4.49), and those with higher education had a 459% increase (OR 5.59, 95% CI 3.52-8.88).

Household income was also positively associated with information seeking. Compared to participants in the lowest income category, those earning between 1 and 2 minimum wages had an OR of 2.37 (95% CI 1.84-3.06), those earning between 2 and 3 minimum wages had an OR of 2.85 (95% CI 1.87-4.34), and those who earned more than 3 minimum wages had an OR of 4.54 (95% CI 2.59-7.93). Use of the public health system was not significantly associated with the outcome (OR 0.92, 95% CI 0.64-1.33).

Comparison between models revealed a substantial reduction in between-strata variance: the VPC dropped from 25% in the null model to 0.01% in the full model, resulting in a PCV of 97%. The VPC of 25% indicates that one-quarter of the variance in information-seeking behavior was attributable to intersectional groupings. This finding suggests that most of the explained variance was due to the additive effects of individual variables rather than intersectional interactions. The AUC for model 2 was 0.73, maintaining a similar discriminatory performance to that of the null model.

Figure 1 shows the predicted probabilities of seeking COVID-19 information across intersectional strata (null model). Strata are ranked from the lowest to the highest predicted probability. Predicted values range from approximately 15% to nearly 90%, revealing a wide gradient of intersectional inequality.

Figure 1. Predicted probabilities of seeking COVID-19 information across intersectional strata (null model) in Brazil in 2021-2023.



The continuous distribution—without discrete clusters—suggests cumulative and overlapping disadvantage. Wider CIs for some strata reflect smaller sample sizes and greater estimation uncertainty. Because the null model captures both additive and interactive components, this distribution represents the overall pattern of intersectional heterogeneity in information-seeking behavior.

Table 3 presents the 5 intersectional strata with the lowest and highest predicted probabilities of seeking information about COVID-19 online based on estimates from the null model (model 1). The strata are organized in ascending order of predicted probability.

XSL•FO
RenderX

Table 3. Classification of strata by predicted probability levels (highest and lowest) of seeking online information about COVID-19 among vulnerable populations (null model; Brazil, 2021-2023).

Rank	Stratum ID	Gender	Race or ethnicity	Schooling	Family income	Use of the public health system	n (participants per stratum)	Predicted probability of seeking online information about COVID-19 (95% CI)
Five strata with the lowest information seeking								
1	12111	Man	Black or Brown	No schooling	Low income	Yes	335	14.26 (9.20-23.13)
2	13111	Man	Other	No schooling	Low income	Yes	19	18.31 (10.80-31.78)
3	23111	Woman	Other	No schooling	Low income	Yes	8	18.40 (10.80-31.78)
4	21111	Woman	White	No schooling	Low income	Yes	38	19.48 (11.71-23.84)
5	12112	Man	Black or Brown	No schooling	Low income	No	53	19.61 (9.91-23.86)
Five strata with the highest information seeking								
110	21321	Woman	White	High schooling	1-2 minimum wages	Yes	35	79.17 (69.54-90.65)
111	22421	Woman	Black or Brown	Higher education	1-2 minimum wages	Yes	11	79.17 (69.54-90.65)
113	12421	Man	Black or Brown	Higher education	1-2 minimum wages	Yes	12	80.32 (73.28-93.25)
114	11441	Man	White	Higher education	More than 3 minimum wages	Yes	8	81.80 (76.58-92.86)
115	11331	Man	White	Higher education	2-3 minimum wages	Yes	11	84.82 (73.28-93.25)

Among the 5 strata with the lowest predicted probabilities, there was a predominance of men; Black, Brown, or other race or ethnicity; no formal education; low-income households; and use of the public health system. The lowest predicted probability was 14.26% (95% CI 9.2%-23.13%).

In contrast, the 5 strata with the highest predicted probabilities showed an opposite profile: most were women or men with higher education, middle to high income (1-3 minimum wages), and users of the public health system. The predicted probabilities of information seeking in these strata ranged from 79.17% (95% CI 69.54%-90.65%) to 84.82% (95% CI 73.28%-93.25%).

These findings highlight the coexistence of significant intersectional disparities in health information seeking during the pandemic associated with structural markers of vulnerability such as race or ethnicity, educational attainment, income, and access to public health services

Discussion

Principal Findings

This study aimed to investigate inequalities in digital exclusion and access to online COVID-19 information among people experiencing homelessness and urban community residents in Brazil using intersectional multilevel analysis. Key findings show that only 39.2% (1040/2652) sought online COVID-19 information, with being a woman, higher schooling, and income strongly linked to higher access across 115 intersectional strata. The findings also revealed predicted probabilities ranging from 14% (Black or Brown men, no schooling, and low income) to 85% (higher educated, higher income groups), with 97% of the between-strata variance (VPC=24% to 1%) explained by additive effects rather than interactions. The findings confirm structural digital exclusion driven by overlapping gender, schooling, income, race or ethnicity, and SUS use disadvantages, fulfilling the study's goal of mapping how social positions jointly shaped informational inequities during the pandemic [10,11,32].

The association between higher levels of education and greater use of the internet as a source of health information reflects not only technical access but also, more importantly, the possession of cognitive and informational competencies that support autonomy in health care [35]. Income, in turn, functions as a proxy for time availability, connectivity, access to digital devices, and overall stability, factors that are all essential for sustained online engagement. This underscores the structural nature of digital exclusion [36,37].

The absence of an association between the use of the SUS and online health information seeking suggests that the public health care system still lacks effective digital strategies for informational mediation targeting vulnerable populations. This finding offers a critical lesson for future public health emergencies: institutional digital channels alone are not sufficient. Health communication requires multichannel, territory-based approaches supported by community-based mediation [38].

XSL•FO
RenderX

Notably, all 5 of the strata with the lowest predicted probability of seeking COVID-19 information online were users of the SUS, whereas among the strata with the highest predicted probability, SUS users also predominated. This suggests that the inequality does not lie between those who do or do not use the public health system but rather in how intersectional social characteristics moderate digital engagement even within the same health care system.

By identifying that men with low levels of education and income exhibit the lowest rates of health information seeking online, this study highlights a social group that remains systematically overlooked in communication strategies. This population profile embodies multiple layers of exclusion (informational, economic, and institutional), placing them at constant risk in the face of misinformation and informational disengagement [20,21].

The relationship between gender and health information seeking observed in this study contrasts with international findings that identified greater male engagement in COVID-19–related communicative uses of the internet. While the international study assessed online communication practices such as content sharing and interaction on social networks, our focus was on the active search for health information.

This distinction suggests that gender differences in the digital sphere are not homogeneous but vary according to the type of activity analyzed. These findings also point to an expanded understanding of digital exclusion, which is not limited to physical access but involves symbolic and functional disconnections from the networks through which knowledge circulates. The greater tendency among women to seek health information may be associated with their central role in care networks, indicating the potential for future strategies anchored in female and community leadership [39-41].

This pattern is not exclusive to the Brazilian context. Both national and international studies have shown that groups with lower educational levels and those residing in marginalized or racialized territories consistently exhibit reduced access to and critical use of digital health information [6,7,37,42,43]. In the United States, Suh et al [37] found that communities facing greater socioeconomic vulnerability showed limited increases in online information seeking even during a global crisis. This suggests that exclusion patterns are not only resistant to exceptional circumstances but may, in fact, be exacerbated by them [37].

Moreover, research has shown that overcoming digital exclusion requires more than infrastructure expansion. Public policies must address informational competencies and invest in effective symbolic mediation channels, especially in vulnerable contexts [31,41]. In Brazil, qualitative studies indicate that community leaders and public health agents remain the primary bridges between the population and health information, suggesting that trust and accessibility are as crucial as connectivity [44].

These results engage with international debates on digital health inequities, which highlight how structurally marginalized groups face not only limited physical access but also a drastically reduced capacity to translate information into health action [4]. Such processes intensify during public health emergencies, when the infodemic rather than democratizing access often amplifies existing asymmetries by demanding complex digital literacies and continuous connectivity, 2 factors profoundly shaped by class, gender, and race [37].

Corroborating this picture, studies show that older adults, people with lower educational attainment, and those with low traditional literacy are precisely those who were least able to obtain positive outcomes from internet use during the COVID-19 pandemic [4].

Such identification is strategically valuable for the formulation of proactive public policies that can mitigate disparities before they become irreversible in future crises. Thus, the trends observed during the pandemic should be interpreted as early warnings of chronic vulnerabilities. The adoption of exclusively digital solutions without territorial grounding or community mediation tends to reproduce and even amplify the very asymmetries this study reveals.

Therefore, promoting informational equity is not a peripheral goal but a central tenet of health justice. It is essential to ensure that all social groups not only have access to information but are also able to understand, contextualize, and use it to make informed care decisions. Tackling the infodemic demands integrated approaches that combine digital inclusion, critical literacy, trust networks, and communication strategies sensitive to territorial diversity [2,4].

These findings underscore the urgency of territorialized strategies to combat infodemics, as evidenced by studies in Brazilian favelas in which digital divides during COVID-19 amplified misinformation uptake among low-income residents lacking community mediation [45-47]. Literature on health communication highlights how ungrounded digital interventions exacerbate asymmetries in rural and indigenous groups, reducing trust and engagement due to cultural mismatches [45-47]. Taken together, these findings suggest that future research and interventions should further explore how territorially grounded and socially mediated communication strategies can mitigate informational inequalities during public health emergencies.

This study has limitations. Its cross-sectional design does not allow for causal inferences, and using a single question to measure the outcome may not fully capture the complexity of information-seeking behaviors. In addition, self-reported data are vulnerable to recall and social desirability bias, although the use of a validated instrument, face-to-face data collection, and trained interviewers strengthens the reliability of the findings [28].

There were also important constraints related to measuring gender and race and ethnicity. Although the questionnaire allowed for self-identification beyond binary gender and included detailed racial and ethnic categories, no participant identified as gender diverse or transgender, and very few identified as Indigenous, East Asian, or “other.” This underrepresentation is consistent with evidence that gender and racial and ethnic minority groups are often less visible in surveys involving socially vulnerable populations because of stigma, safety concerns, and structural barriers to participation [48].

XSL•FO
RenderX

From a methodological perspective, the absence of participants outside the "woman" and "man" categories prevented exploration of gender diversity within the intersectional MAIHDA framework. Very small cell sizes in some strata also required grouping categories to preserve statistical stability and confidentiality, which reduced the granularity of intersectional comparisons [10].

Finally, the analysis was limited to participants with complete data for all intersectional variables. While this complete-case strategy was necessary to construct comparable strata, it may have disproportionately excluded individuals in more vulnerable situations, introducing potential selection bias.

We highlight the need for future studies using mixed methods approaches to investigate the meanings attributed to information, the relationship between formal and informal sources, and the impact of community mediation on the development of informational attitudes. Research that integrates qualitative and quantitative methods may offer new insights into the information behavior of vulnerable populations and help guide more inclusive, effective, and sustainable health communication policies.

Conclusions

Beyond mapping intersectional patterns, this analysis reveals how overlapping social positions (gender, race and ethnicity, schooling, and income) drive structural disparities in online COVID-19 information seeking among people experiencing homelessness and urban community residents in Brazil. These patterns underscore additive disadvantages as the primary mechanism of digital exclusion rather than synergistic interactions, highlighting the cumulative impact of vulnerability markers even within the public health system. Identifying these extreme social profiles enables policymakers to prioritize territory-based digital inclusion strategies, advancing informational equity and health justice in future public health crises. These intersectional disparities reveal that, even within Brazil's constitutionally guaranteed SUS, a fundamental right under Article 196 of the 1988 Federal Constitution, no vulnerable group should be left behind in accessing essential health information.

Funding

This work was supported by (1) Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (funding code 001 and process 88887.657730/2021-00), (2) Doctoral Scholarship Conselho Nacional de Desenvolvimento Científico e Tecnológico (process 403872/2022-7), and (3) Fundação de Amparo à Pesquisa do Estado de São Paulo under grants 2021/08263-7 and 2022/08510-7.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: AFT and RAA
Data curation: AFT and RAA
Formal analysis: AFT, DG, and RAA
Funding acquisition: AFT and RAA
Investigation: TZB, RBVT, MAPP, AFT, RAA, and LPF
Methodology: AFT, RAA, RBVT, DG, MAPP, YMA, NZ, and ACVR
Supervision: RAA
Visualization: AFT and RAA
Writing—original draft: AFT, RAA, DG, MEPP, and TZB
Writing—review and editing: AFT, RAA, YMA, LPF, ACVR, MEPP, NZ, and MAPP

Conflicts of Interest

None declared.

References

1. Sachs JD, Karim SS, Aknin L, Allen J, Brosbøl K, Colombo F, et al. The Lancet Commission on lessons for the future from the COVID-19 pandemic. Lancet 2022 Oct 08;400(10359):1224-1280 [FREE Full text] [doi: 10.1016/S0140-6736(22)01585-9] [Medline: 36115368]
2. Cooks EJ, Vilaro MJ, Dyal BW, Wang S, Mertens G, Raisa A, et al. What did the pandemic teach us about effective health communication? Unpacking the COVID-19 infodemic. BMC Public Health 2022 Dec 13;22(1):2339 [FREE Full text] [doi: 10.1186/s12889-022-14707-3] [Medline: 36514047]
3. Understanding infodemic and misinformation in the fight against COVID-19. Pan American Health Organization. URL: https://iris.paho.org/handle/10665.2/52054 [accessed 2025-04-20]

XSL•FO
RenderX

4. van Deursen AJ. Digital inequality during a pandemic: quantitative study of differences in COVID-19-related internet uses and outcomes among the general population. J Med Internet Res 2020 Aug 20;22(8):e20073 [FREE Full text] [doi: 10.2196/20073] [Medline: 32750005]
5. Wang J, Lu X, Ngai SB, Xie L, Liu X, Yao Y, et al. Digital exclusion and depressive symptoms among older people: findings from five aging cohort studies across 24 countries. Health Data Sci 2025;5:0218 [FREE Full text] [doi: 10.34133/hds.0218] [Medline: 39790099]
6. Holmes H, Burgess G. "Pay the wi-fi or feed the children": coronavirus has intensified the UK's digital divide. University of Cambridge. URL: https://www.cam.ac.uk/stories/digitaldivide [accessed 2025-04-20]
7. van Deursen AJ, van Dijk JA. The first-level digital divide shifts from inequalities in physical access to inequalities in material access. New Media Soc 2019 Mar 07;21(2):354-375 [FREE Full text] [doi: 10.1177/1461444818797082] [Medline: 30886536]
8. Kontos E, Blake KD, Chou WY, Prestin A. Predictors of eHealth usage: insights on the digital divide from the Health Information National Trends Survey 2012. J Med Internet Res 2014 Jul 16;16(7):e172 [FREE Full text] [doi: 10.2196/jmir.3117] [Medline: 25048379]
9. Van Deursen AJ, Helsper EJ. The third-level digital divide: who benefits most from being online? In: Robinson L, Cotten SR, Schulz J, Hale TM, Williams A, editors. Communication and Information Technologies Annual. Berlin, Germany: Bingleymerald Group Publishing Limited; 2015:29-52.
10. Evans CR, Leckie G, Subramanian SV, Bell A, Merlo J. A tutorial for conducting intersectional multilevel analysis of individual heterogeneity and discriminatory accuracy (MAIHDA). SSM Popul Health 2024 Jun;26:101664 [FREE Full text] [doi: 10.1016/j.ssmph.2024.101664] [Medline: 38690117]
11. Axelsson Fisk S, Mulinari S, Wemrell M, Leckie G, Perez Vicente R, Merlo J. Chronic obstructive pulmonary disease in Sweden: an intersectional multilevel analysis of individual heterogeneity and discriminatory accuracy. SSM Popul Health 2018 Apr;4:334-346 [FREE Full text] [doi: 10.1016/j.ssmph.2018.03.005] [Medline: 29854918]
12. Crenshaw K. Demarginalizing the intersection of race and sex: a Black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. In: Crenshaw K, editor. Feminist Legal Theories. Abingdon, UK: Routledge; 2013:139-167.
13. Cho S, Crenshaw KW, McCall L. Toward a field of intersectionality studies: theory, applications, and praxis. Signs 2013 Jun;38(4):785-810. [doi: 10.1086/669608]
14. Collins PH. Black feminist thought in the matrix of domination. In: Collins PH, editor. Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment. Boston, MA: Unwin Hyman; 1990:221-238.
15. Malta M, Cardoso LO, Bastos FI, Magnanini MM, Silva CM. Iniciativa STROBE: subsídios para a comunicação de estudos observacionais. Rev Saude Publica 2010 Jun;44(3):559-565 [FREE Full text] [doi: 10.1590/s0034-89102010000300021] [Medline: 20549022]
16. Amadeu CV, Silva JL, Manochio-Pina MG. Inclusão digital e suas relações com o empoderamento, a qualidade de vida e o bem-estar. Aletheia 2022;55(1):207-223. [doi: 10.29327/226091.55.1-11]
17. Barbosa IR, Galvão MH, Souza TA, Gomes SM, Medeiros AD, Lima KC. Incidence of and mortality from COVID-19 in the older Brazilian population and its relationship with contextual indicators: an ecological study. Rev Bras Geriatr Gerontol 2020;23(1):e200171. [doi: 10.1590/1981-22562020023.200171]
18. Barbosa TP, Costa FB, Ramos AC, Berra TZ, Arroyo LH, Alves YA, et al. Morbimortalidade por COVID-19 associada a condições crônicas, serviços de saúde e iniquidadesvidências de sindemia. Rev Panam Salud Publica 2022:e6. [doi: 10.26633/RPSP.2022]
19. Censo Demográfico 2022: identificação étnico-racial da população, por sexo e idade. Instituto Brasileiro de Geografia e Estatística. URL: https://biblioteca.ibge.gov.br/visualizacao/periodicos/3105/cd_2022_etnico_racial.pdf [accessed 2025-04-23]
20. Natalino MA. Publication: estimativa da população em situação de rua no Brasil. Instituto de Pesquisa Econômica Aplicada Ipea. URL: http://repositorio.ipea.gov.br/handle/11058/11604 [accessed 2025-04-20]
21. A população em situação de rua nos números do Cadastro Único. Instituto de Pesquisa Econômica Aplicada (IPEA). 2024. URL: https://www.econstor.eu/bitstream/10419/290130/1/TD2944.pdf [accessed 2026-01-02]
22. New census and housing program for the homeless population. Revista Pesquisa FAPESP. URL: https://revistapesquisa.fapesp.br/en/new-census-and-housing-program-for-the-homeless-population/ [accessed 2025-04-20]
23. Censo 2022: Brazil had 16.4 million people living in favelas and urban communities. Instituto Brasileiro de Geografia e Estatística. URL: https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias/noticias/41797-censo-2022-brasil-tinha-16-4-milhoes-de-pessoas-morando-em-favelas-e-comunidades-urbanas [accessed 2025-04-20]
24. Morais MD, Krause CH, Lima Neto VC. Caracterização e tipologia de assentamentos precários: estudos de caso brasileiros. Rev Bras Estud Urbanos Reg 2016 Apr 30;18(1):61-78. [doi: 10.22296/2317-1529.2016v18n1]
25. Etikan I. Comparison of convenience sampling and purposive sampling. Am J Theor Appl Stat 2016;5(1):1. [doi: 10.11648/j.ajtas.20160501.11]
26. Cochran WG. Sampling Techniques. 3rd edition. New York, NY: John Wiley & Sons; 1977.
27. Bolfarine H, Bussab WO. Elementos de Amostragem. São Paulo, Brazil: Blucher; 2005.

XSL•FO
RenderX

28. Moura HS, Arcêncio RA, Berra TZ, Rosa RJ, Campos MC, Canatto MC, et al. Data collection instrument in Brazilian Portuguese. figshare. 2024. URL: https://figshare.com/articles/online_resource/Data_collection_instrument_in_Brazilian_Portuguese/25917571?file=46630732 [accessed 2026-01-02]
29. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
30. Gomes I, Marli M. IBGE shows the colors of inequality. Brazilian Institute of Geography and Statistics. 2018. URL: https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias/noticias/21206-ibge-mostra-as-cores-da-desigualdade [accessed 2025-05-07]
31. Bauer GR. Incorporating intersectionality theory into population health research methodology: challenges and the potential to advance health equity. Soc Sci Med 2014 Jun;110:10-17 [FREE Full text] [doi: 10.1016/j.socscimed.2014.03.022] [Medline: 24704889]
32. Green MA, Evans CR, Subramanian SV. Can intersectionality theory enrich population health research? Soc Sci Med 2017 Apr;178:214-216. [doi: 10.1016/j.socscimed.2017.02.029] [Medline: 28238539]
33. Persmark A, Wemrell M, Zettermark S, Leckie G, Subramanian SV, Merlo J. Precision public health: mapping socioeconomic disparities in opioid dispensations at Swedish pharmacies by Multilevel Analysis of Individual Heterogeneity and Discriminatory Accuracy (MAIHDA). PLoS One 2019 Aug 27;14(8):e0220322 [FREE Full text] [doi: 10.1371/journal.pone.0220322] [Medline: 31454361]
34. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. URL: https://www.R-project.org/ [accessed 2025-05-08]
35. Lorini C, Caini S, Ierardi F, Bachini L, Gemmi F, Bonaccorsi G. Health literacy as a shared capacity: does the health literacy of a country influence the health disparities among immigrants? Int J Environ Res Public Health 2020 Mar 12;17(4):1149 [FREE Full text] [doi: 10.3390/ijerph17041149] [Medline: 32059496]
36. Souza NF. Literacia digital em saúde: convergências, divergências e possíveis caminhos para um campo em evolução. Rev Eletron Comun Inf Inov Saúde 2024 Sep 30;18(3):695-715. [doi: 10.29397/reciis.v18i3.3746]
37. Suh J, Horvitz E, White RW, Althoff T. Disparate impacts on online information access during the Covid-19 pandemic. Nat Commun 2022 Nov 19;13(1):7094-7015 [FREE Full text] [doi: 10.1038/s41467-022-34592-z] [Medline: 36402817]
38. Borges VC. Impactos da covid-19: desafios da pandemia na perspectiva da comunicação e saúde. Reciis 2023;17(4):976-983 [FREE Full text]
39. Fox S, Duggan M. Health online 2013. Pew Research Center. 2013. URL: https://www.pewresearch.org/internet/2013/01/15/health-online-2013/ [accessed 2025-04-16]
40. Fox S, Jones S. The social life of health information. Pew Research Center. 2009. URL: https://www.pewresearch.org/internet/2009/06/11/the-social-life-of-health-information/ [accessed 2025-04-02]
41. Kontos E, Bennett G, Viswanath K. Barriers and facilitators to home computer and internet use among urban novice computer users of low socioeconomic position. J Med Internet Res 2007 Oct 22;9(4):e31 [FREE Full text] [doi: 10.2196/jmir.9.4.e31] [Medline: 17951215]
42. Ribeiro KB, Moreira MR, Andrade MV. Desigualdade digital e acesso à internet durante a pandemia da COVID-19 no Brasil. Rev Adm Publica 2021;55(4):704-726. [doi: 10.1590/0034-761220200652]
43. Nóbrega MP, Silva RC, Souza NF, Lima KM, Oliveira JS. Acesso à informação sobre COVID-19 e fatores associados em moradores de comunidade urbana vulnerável. Cad Saude Publica 2021;37(7):e00151920. [doi: 10.1590/0102-311X00151920]
44. Santos SL, Barbosa CA, Carvalho MC. Informação em saúde e populações vulneráveis: estratégias comunitárias frente à COVID-19. Saude Soc 2021;30(4):e210092. [doi: 10.1590/S0104-12902021210092]
45. Choukou MA, Sanchez-Ramirez DC, Pol M, Uddin M, Monnin C, Syed-Abdul S. COVID-19 infodemic and digital health literacy in vulnerable populations: a scoping review. Digit Health 2022;8:20552076221076927 [FREE Full text] [doi: 10.1177/20552076221076927] [Medline: 35223076]
46. Basile P. Vulnerability, neglect, and collectivity in Brazilian favelas: surviving the threats of the COVID-19 pandemic and the state's necropolitics. Urban Stud 2023 Jul 12;60(9):1690-1706 [FREE Full text] [doi: 10.1177/00420980221103342] [Medline: 37416833]
47. Asaad C, Khaouja I, Ghogho M, Baïna K. When infodemic meets epidemic: systematic literature review. JMIR Public Health Surveill 2025 Mar 03;11:e55642 [FREE Full text] [doi: 10.2196/55642] [Medline: 39899850]
48. Wali N, Huda MN, Gill T, Green J, Renzaho AM. A systematic review of recruitment and retention of ethnic minorities and migrants in obesity prevention randomised controlled trials. Int J Obes (Lond) 2024 Aug 04;48(8):1065-1079. [doi: 10.1038/s41366-024-01545-z] [Medline: 38834795]

Abbreviations

AUC: area under the receiver operating characteristic curve
MAIHDA: multilevel analysis of individual heterogeneity and discriminatory accuracy
OR: odds ratio

XSL•FO
RenderX

PCV: proportional change in variance
REDCap: Research Electronic Data Capture
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
SUS: Unified Health System (Sistema Único de Saúde in Portuguese)
USP: University of São Paulo
VPC: variance partition coefficient

Edited by A Mavragani, T Sanchez; submitted 07.May.2025; peer-reviewed by E Chavarria, L Grilli; comments to author 26.Nov.2025; accepted 22.Dec.2025; published 12.Jan.2026.

Please cite as:
Tártaro AF, Gomes D, Berra TZ, Tavares RBV, Alves YM, Ferezin LP, Ramos ACV, Zini N, Pelodan MEP, Popolin MAP, Arcêncio RA
Digital Exclusion Among People Experiencing Homelessness and Residents of Urban Communities in Brazil: Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e77124
URL: https://publichealth.jmir.org/2026/1/e77124
doi:10.2196/77124
PMID:41525683

©Ariela Fehr Tártaro, Dulce Gomes, Thaís Zamboni Berra, Reginaldo Bazon Vaz Tavares, Yan Mathias Alves, Letícia Perticarrara Ferezin, Antônio Carlos Vieira Ramos, Nathalia Zini, Maria Eduarda Pagano Pelodan, Marcela Antunes Paschoal Popolin, Ricardo Alexandre Arcêncio. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 12.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Drug-Facilitated Sexual Assault Pornography and Sexual Violence While Partying: Cross-Sectional Study

Pablo Prego-Meleiro[1,2,3,4], PhD; Guadalupe Pastor-Moreno[2,5,6], PhD; Irantzu Recalde-Esnoz[4,7], PhD; Luis Sordo[1,2,4,8], PhD

[1]Department of Public Health and Maternal and Child Health, Faculty of Medicine, Universidad Complutense de Madrid, Madrid, Spain
[2]Centro de Investigación Biomédica en Red de Epidemiología y Salud Pública, Madrid, Spain
[3]Instituto Universitario de Investigación en Ciencias Policiales, Universidad de Alcalá, Alcalá de Henares, Spain
[4]Observatorio Universitario de Violencia Sexual Facilitada por Drogas, Universidad de Alcalá, Alcalá de Henares, Spain
[5]Instituto de Investigación Biosanitaria de Granada, Granada, Spain
[6]Andalusian School of Public Health, Granada, Andalusia, Spain
[7]Department of Sociology and Social Work, Universidad Publica de Navarra, Pamplona, Navarre, Spain
[8]Fundación para la Investigación Biomédica del Hospital Clínico San Carlos (IDISCC), Madrid, Spain

Corresponding Author:
Guadalupe Pastor-Moreno, PhD
Andalusian School of Public Health
Cuesta del Observatorio, 4
Granada, Andalusia, 18011
Spain
Phone: 1 958027400
Email: guadalupe.pastor.easp@juntadeandalucia.es

Abstract

Background: Drug-facilitated sexual assaults (DFSAs) in youth partying contexts represent a growing public health concern, affecting approximately half of women and 1 in 4 men. These assaults often occur in environments where alcohol and other psychoactive substances are consumed, leading to impaired consent and increased vulnerability. At the same time, young people are increasingly exposed to pornography, often using it as a primary source of sexual information. However, pornography can disseminate misleading or harmful messages about sexuality and consent. Of particular concern is a subtype of pornographic material (hereafter referred to as DFSA pornography) that depicts nonconsensual sexual acts involving individuals who are asleep, unconscious, or under the influence of psychoactive substances, including alcohol and other drugs.

Objective: This study aimed to examine the prevalence of DFSA pornography consumption among young adults in Spain and analyze its association with self-reported DFSA perpetration and victimization in party settings.

Methods: A cross-sectional online survey (computer-assisted web interviewing) was conducted among individuals aged 18 to 35 years residing in Spain. Participants (N=1601; n=1534, 95.8% valid responses) were recruited from a certified online panel using quota sampling stratified by sex, age group, and region to ensure national representativeness. The questionnaire was adapted from the Sexual Experiences Survey–Short Form Victimization and the Spanish Macro-Survey on Violence Against Women. It assessed DFSA perpetration and victimization in partying contexts under the influence of alcohol or drugs. A specific variable, DFSA pornography, was created to measure intentional viewing of explicit sexual content depicting unconscious or intoxicated individuals. Additional sociodemographic variables included sex, age, educational level, sexual orientation, political ideology, nationality, and socioeconomic level. Descriptive, bivariate, and binary logistic regression analyses were performed, estimating associations between DFSA experiences and both general and DFSA pornography consumption.

Results: Among respondents (800/1593, 50.2% female participants; mean age 27.0, SD 5.1 years), 78.4% (1233/1572) identified as heterosexual, and 52% (825/1587) held a university degree. Overall, 66.6% (1013/1521) reported consuming pornography in the previous year, with higher prevalence among male participants (638/753, 84.7%) than among female participants (370/762, 48.6%). DFSA pornography consumption was reported by 22.2% (167/753) of male participants and 11.3% (86/762) of female participants, and increased with overall pornography use frequency. Multivariate logistic regression indicated that DFSA perpetration (adjusted odds ratio 3.78, 95% CI 1.72-8.28; $P<.001$) and victimization (adjusted odds ratio 1.86, 95% CI 1.24-2.78; $P=.003$) were associated with DFSA pornography consumption.

XSL•FO
RenderX

Conclusions: The findings reveal an association between exposure to DFSA pornography and both DFSA perpetration and victimization among young people in Spain. These results underscore the need for comprehensive sexual education that critically addresses pornography as a source of misinformation, emphasizing accurate understanding of consent and substance-impaired sexual activity. Public health strategies should integrate media literacy and consent education to mitigate the normalization of sexual violence depicted in pornography.

(JMIR Public Health Surveill 2026;12:e80110) doi:10.2196/80110

KEYWORDS

drug-facilitated sexual assault; drugs; sexual assault; pornography consumption; sexual violence

Introduction

Sexual violence is a pervasive public health concern that demands immediate attention [1]. According to the World Health Organization, sexual violence encompasses unwanted sexual acts, comments, or insinuations, from verbal harassment to forced penetration [1,2]. Drug-facilitated sexual assault (DFSA) is a specific form of sexual violence in which perpetrators take advantage of individuals who are incapacitated by psychoactive substances, including alcohol, illicit drugs, or medications, consumed either voluntarily or involuntarily [1,3]. DFSA episodes particularly affect partying, dating, and hookup contexts, where alcohol and other drug use is often combined with high expectations for sexual interaction [3,4]. This type of violence occurs across diverse settings worldwide, including Europe, within which Spain has shown particular concerns [5]. A recent study showed that 1 in 2 women and 1 in 4 men among young people in Spain have experienced DFSA at least once while partying [6]. The relationship between nightlife partying and DFSA has been a focus of intense research in Spain in recent years [7-11]. Multiple studies have highlighted how rape culture leads young people of both sexes to normalize sexually violent practices such as nonconsensual sex, thereby rendering sexual violence invisible and minimizing its impact [12-16].

Pornography consumption is widespread among the young population, especially men, with prevalence of consumption in the last year ranging between 60% and 80% in Spain and other high-income countries [17-21]. In addition, this consumption has significantly increased in recent years, driven by its growing accessibility online [19,21,22]. One of the main problems related to the consumption of pornography is that many young people perceive it as a source of sexual education, turning to it for information in the absence of reliable and quality guidance; in Spain, just over 10% of young people express satisfaction with the sexual education they have received, whereas approximately half report that pornography helps them better understand and learn about sex and acknowledge using it as a source of inspiration [19]. Similar figures have been observed in other countries [23], highlighting how particularly young men incorporate practices depicted in pornography into their sexual lives [24,25].

Pornography can be a source of misinformation about sex [26,27]. Violent content in pornography can encompass a wide range of physical, verbal, and sexual violence. Numerous studies have highlighted that mainstream pornography often portrays violent scenes featuring practices characterized by aggression that demean and sexually objectify people [23,28,29]. This type of pornographic content also includes scenes of nonconsensual sexual activity while someone is asleep, unconscious, or under the influence of alcohol or other drugs [29,30]. This pornographic content resembles DFSA situations, where individuals are rendered incapacitated by the psychoactive effects of licit or illicit substances, including alcohol, other drugs of abuse, and medications. There are no data available on the consumption of this type of pornography with DFSA content. Nevertheless, as a proxy, in Spain, at least 18.2% of men and 14.5% of women among young people admit to consuming pornography classified as highly violent, particularly degrading, or humiliating [18]. These are particularly concerning figures considering that numerous studies indicate that some young people do not distinguish between pornography and real-life sexual behaviors, potentially leading to the normalization of degrading acts or extreme violence [31]; 31% of young people in Spain agree that pornography shapes sexual fantasies involving either the perpetration or reception of violence, which they perceive as a harmful effect of pornography on its consumers [18]. Therefore, given the high prevalence of DFSA in party contexts, combined with the widespread use of pornography as a sexual education source and the accessibility of violent content, this study aimed to examine the prevalence of DFSA pornography consumption and its relationship with having either experienced or perpetrated DFSA while partying.

Methods

Design and Participants

This was a cross-sectional study using an anonymized, self-administered online survey for sexual victimization and perpetration (computer-assisted web interviewing). The study population consisted of individuals aged 18 to 35 years residing in Spain, totaling 9,250,779 people in 2022, with 50.85% being men [32]. A minimum sample size of 1537 individuals was required, assuming 50% prevalence, 95% confidence, and a 2.5% margin of error, with a prevalence of 0.5 due to the lack of prior data as recommended by statistical guidelines [33]. Participants were recruited from an online panel managed by a research company certified under ISO 20252, which provides verified samples for social and health research. The panel includes more than 100,000 registered members residing in Spain. Panelists are recruited through paid social strategies and undergo a validation process, with sociodemographic profiles periodically updated to ensure data accuracy.

Quota sampling was applied to ensure proportional representation by sex, age group (18-24 and 25-35 years), and region (17 autonomous communities) according to the

XSL•FO
RenderX

demographic distribution of the Spanish population. Study participants were selected through a routing system designed to minimize self-selection bias and received personalized invitations via email and mobile notifications. To maintain quota compliance, invitations were dynamically adjusted based on response and dropout rates.

Survey access included duplicate checks, captcha verification, and validation of self-reported gender and age. Standard quality controls, such as attention-check questions, a minimum completion time threshold, and mandatory responses, were applied to ensure data integrity.

Of the 1707 invited individuals, 106 (6.2%) declined participation or did not complete the questionnaire, resulting in a final sample of 1601 respondents.

The questionnaire was specifically tailored by an interdisciplinary team, drawing on sources such as the Sexual Experiences Survey–Short Form Victimization [34] and the 2019 Macro-Survey on Violence Against Women [35]. At the time this study was conducted, to the authors' knowledge, no fully validated instruments existed to specifically assess DFSA perpetration and victimization or the use of pornography depicting DFSA content. A pretest assessed clarity, reliability, and validity, including cognitive debriefing with over 100 participants to address any understanding issues and ensure clarity. Questions were designed to avoid identifiable information and social desirability biases. IP addresses and cookies were not collected.

Variables

Two main variables provided information about DFSA perpetration or victimization experiences while partying. Sexual violence included unwanted sexual touching, kissing, masturbation by another person, oral sex, or penetration. DFSA perpetration was defined as "having engaged in any of these sexual behaviors non-consensually with someone who was under the influence of alcohol or other drugs and, because of that, was unable to communicate their sexual consent, while partying or immediately afterwards" ("no" or "yes"). DFSA victimization was defined as "having experienced any of these sexual acts non-consensually when you were under the influence of alcohol or other drugs, while partying or immediately afterwards" ("no" or "yes").

Variables concerning pornography use were pornography consumption frequency, defined as "how often one intentionally seeks out and views videos or photographs containing explicit sexual content" ("no consumption," "Yes, between every day and two to three times per week," "Yes, between once per week and two to three times per month," and "Yes, less than once per month") and DFSA pornography consumption, defined as "seeking out and intentionally viewing videos or photographs containing explicit sexual content in which a person is asleep, unconscious, or under the influence of alcohol or other drugs (including sedation or other effects)" ("no" or "yes"). The item assessing DFSA pornography consumption was developed and reviewed by the research team to ensure conceptual clarity. During the pilot questionnaire, all participants completed the item, and only 1 reported discomfort, indicating good comprehension and acceptability. Other independent variables were sex; age; educational level (university or nonuniversity); sexual orientation (heterosexual or nonheterosexual); political ideology (left or right), originally measured on a scale from 1 to 10 and dichotomized into 1 to 5 for "left" and 6 to 10 for "right"; nationality (Spanish or Spanish and/or other); and socioeconomic level (family income [net per month in €]; low and low to medium: ≤€2000 [US $2357.44]; medium to high or high: >€2000 [US $2357.44]). Details on variable measurement and the full list of items are provided in the questionnaire in Multimedia Appendix 1.

Statistical Analysis

A descriptive analysis of pornography use focusing on the overall and gender-specific prevalence rates of pornography consumption frequency and DFSA pornography consumption was conducted. Relationships between the 2 main variables (DFSA perpetration and DFSA victimization) and covariates were examined using the chi-square and Fisher exact tests for categorical variables and ANOVA for continuous variables. Binary logistic regression models were performed to explore the relationship between DFSA perpetration and DFSA victimization and pornography use including all sociodemographic variables, with $P<.05$ in the bivariate analysis. Odds ratios and their corresponding 95% CIs were provided. Data analyses were conducted using Stata (version 19; StataCorp).

Ethical Considerations

Ethics approval was granted by the Research Ethics Committee of the University of Alcalá (approval code CEIP2022/2/040) on March 31, 2022. Informed consent was obtained from all participants before taking part. Participation was voluntary, and participants were free to decline or withdraw at any time. Privacy and confidentiality were ensured through the use of an anonymous survey. No personal identifiers were collected, and all data were analyzed in aggregate form. Consequently, individual participants cannot be identified from the dataset.

As a result of their participation in the surveys, panelists receive incentive points that can be redeemed through a catalog of gifts, prize draws, and even donations to nongovernmental organizations. These incentives are flexible and tailored to the specific conditions of each survey, taking into account factors such as survey duration, and are designed to reward panelists according to their level of commitment and effort. The incentive system allows for variations to be defined according to the survey process, ensuring that panelists receive appropriate compensation for their time and contributions. This approach not only motivates panelists to participate actively in surveys but also encourages them to provide thoughtful and comprehensive responses, resulting in higher-quality data by optimizing participant engagement and maximizing response rates.

Results

Among the 1601 respondents, (800/1593, 50.2%) were female participants, the mean age was 27.0 (SD 5.1) years, 78.4% (1233/1572) were heterosexual, 51.9% (825/1587) had a

XSL•FO
RenderX

university degree, 54.6% (689/1262) had a low to medium or low socioeconomic status, 91.4% (1446/1583) were Spanish, and 75.8% (1213/1601) had a left-wing political ideology. Ages ranged from 18 to 35 years for both male and female participants (Table 1).

The prevalence of pornography consumption in the last year was 66.6% (1013/1521), with 24.5% (373/1521) consuming it daily or 2 to 3 times a week. Among male participants, 84.7% (638/753) had consumed pornography, with 44.2% (333/753) consuming it daily or 2 to 3 times a week and 28.2% (212/753) consuming it between once a week and 2 to 3 times a month. Among female participants, 27.2% (207/762) reported consuming pornography less than once a month, whereas 16.3% (124/762) consumed it between once a week and 2 to 3 times a month. Regarding the type of pornography consumed, 22.2% (167/753) of male participants and 11.3% (86/762) of female participants reported consuming pornography with DFSA content (Table 2). The higher the frequency of pornography use, the greater the consumption of DFSA pornography. A total of 30.8% (115/373) of those who consumed pornography more frequently acknowledged having seen DFSA pornography (Figure 1).

The results of the bivariate analysis showed that the odds of having perpetrated DFSA at least once while partying were 6.27 times higher among those who consumed DFSA pornography (P<.001), whereas they were 3.01 times higher among those who had more frequent pornography consumption (daily or 2-3 times a week; P<.001). However, the results of the multivariate analysis model, adjusted for both variables and other sociodemographic characteristics, revealed that perpetrating DFSA while partying appeared only in relation to the type of pornography consumed, not its frequency. The odds of having perpetrated DFSA while partying were 3.78 times higher among people who reported consuming DFSA pornography than among those who did not (P<.001). Regarding other factors, in the multivariate analysis, DFSA perpetration was higher among male participants (adjusted odds ratio [aOR] 2.03, 95% CI 1.17-3.51; P=.01), nonheterosexual individuals (aOR 2.10, 95% CI 1.33-3.32; P<.001), and individuals of foreign origin (aOR 2.55, 95% CI 1.40-4.62; P=.002). The logistic model (χ^2_8=88.8; P<.001; McFadden pseudo-R^2=0.123) fit the data well (Hosmer-Lemeshow χ^2_8=7.5; P=.48; Table 3).

Table 1. Sociodemographic characteristics of the studied sample, overall and stratified by sex (N=1601).[a]

	Global	Female (n=800)	Male (n=793)
Age (y), mean (SD; 95% CI)	27.0 (5.1; 26.8-27.3)	26.1 (4.9; 25.7-26.4)	28.0 (5.2; 27.7-28.4)
Sexual orientation, n (%; 95% CI)			
Heterosexual	1233 (78.4; 76.3-80.4)	623 (79.6; 76.6-82.2)	608 (77.8; 74.7-80.5)
Nonheterosexual	339 (21.6; 19.6-23.7)	160 (20.4; 17.8-23.4)	174 (22.3; 19.5-25.3)
Political ideology, n (%; 95% CI)			
Left	1231 (75.8; 73.6-77.8)	635 (79.4; 76.4-82.0)	571 (72.0; 68.8-75.0)
Right	388 (24.2; 22.2-26.4)	165 (20.6; 18.0-23.6)	222 (28.0; 25.0-32.1)
Nationality, n (%; 95% CI)			
Spanish	1446 (91.4; 89.9-92.6)	719 (90.9; 88.7-92.7)	723 (92.2; 90.1-93.9)
Spanish and/or other	137 (8.7; 7.4-10.1)	72 (9.1; 7.3-11-3)	61 (7.8; 6.1-9.9)
Socioeconomic level, n (%; 95% CI)			
Low to medium low	689 (54.6; 51.8-57.3)	358 (57.9; 54.0-61.8)	329 (51.3; 47.4-55.2)
Medium high to high	573 (45.4; 42.7-48.2)	260 (42.1; 38.2-46.0)	312 (48.7; 44.8-52.6)
Educational level, n (%; 95% CI)			
University	825 (52.0; 49.5-54.4)	406 (51.2; 47.7-54.7)	414 (52.7; 49.2-56.1)
Nonuniversity	762 (48.0; 45.6-50.5)	387 (48.8; 45.3-52.3)	372 (47.3; 43.9-50.8)

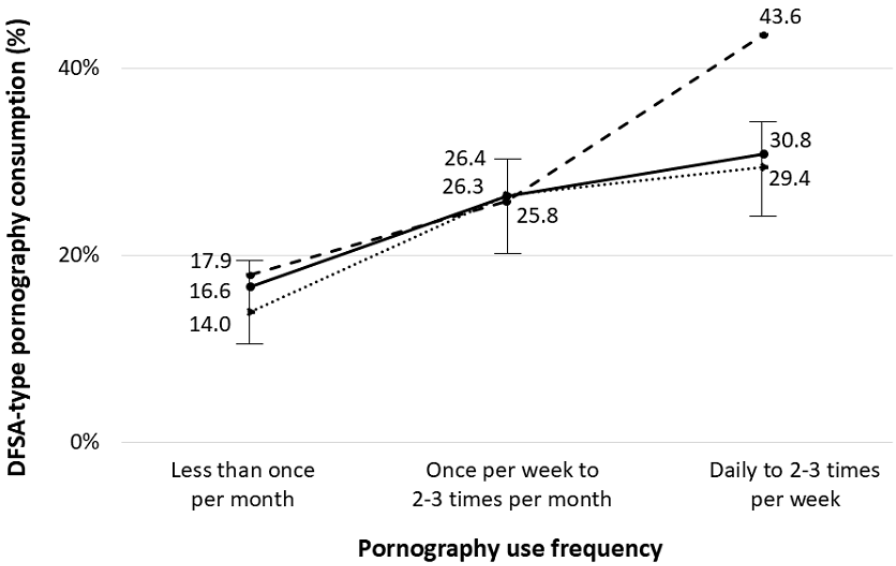
[a]Totals do not equal 1601 because 8 (0.5%) participants preferred not to specify their sex.

XSL•FO
RenderX

Table 2. Prevalence of pornography consumption categorized by use frequency and type of pornography consumed based on whether drug-facilitated sexual assault (DFSA) pornography was used (N=1521).

	Global, n (%; 95% CI)	Female, n (%; 95% CI)	Male, n (%; 95% CI)
Total responding "yes"	1013 (66.6; 64.2-68.9)	370 (48.6; 45.0-52.1)	638 (84.7; 82.0-87.1)
Daily to 2 to 3 times per week	373 (24.5; 22.4-26.8)	39 (5.1; 3.8-6.9)	333 (44.2; 40.7-47.8)
Once per week to 2 to 3 times per month	339 (22.3; 20.3-24.5)	124 (16.3; 13.8-19.1)	212 (28.2; 25.1-31.5)
Less than once per month	301 (19.8; 17.9-21.9)	207 (27.2; 47.9-55.0)	93 (12.4; 10.2-14.9)
Pornography without DFSA	759 (49.9; 47.4-52.4)	284 (37.3; 33.9-40.8)	471 (62.6; 59.0-65.9)
Pornography with DFSA	254 (16.7; 14.9-18.7)	86 (11.3; 9.2-13.7)	167 (22.2; 19.3-25.3)
No pornography use	508 (33.4; 31.1-35.8)	392 (51.4; 47.8-55.0)	115 (15.3; 12.9-18.0)

Figure 1. Relationship between drug-facilitated sexual assault (DFSA)–type pornography consumption and pornography use frequency (low, medium, and high), overall and by gender. The solid line represents the overall group, the dotted line represents men, and the dashed line represents women. The 95% CIs are shown for the overall group (*P*<.001).



XSL•FO
RenderX

Table 3. The relationship between lifetime drug-facilitated sexual assault (DFSA) perpetration while partying and the type of pornography consumed and frequency of consumption (N=1482).[a]

Sociodemographic and behavioral variables	Perpetrated DFSA at any point in their life while partying		*P* value	Crude OR[b] (95% CI)	Adjusted OR (95% CI)
	Yes	No			
Type of pornography consumed, n (%)			<.001		
No consumption	18 (3.6)	485 (96.4)		Reference	Reference
Without DFSA	35 (4.6)	719 (95.4)		1.31 (0.73-2.34)	0.91 (0.43-1.95)
With DFSA	47 (18.9)	202 (81.1)		6.27 (3.55-11.06)[c]	3.78 (1.72-8.28)[d]
Frequency of pornography consumption, n (%)			<.001		
Never to less than once per month	33 (4.1)	770 (95.9)		Reference	Reference
Once per week to 2 to 3 times per month	25 (7.5)	310 (92.5)		1.88 (1.10-3.22)[d]	1.01 (0.49-2.06)
Daily to 2 to 3 times per week	42 (11.4)	326 (88.6)		3.01 (1.87-4.83)[c]	1.30 (0.65-2.59)
Sex, n (%)			<.001		
Female	32 (4.1)	757 (95.9)		Reference	Reference
Male	72 (9.2)	711 (90.8)		2.40 (1.56-3.68)[c]	2.03 (1.17-3.51)[d]
Age (y), mean (SD)	27 (5)	27 (5)	.79	0.99 (0.96-1.03)	1.00 (0.98-1.05)
Sexual orientation, n (%)			<.001		
Heterosexual	61 (5.0)	1162 (95.0)		Reference	Reference
Nonheterosexual	42 (12.6)	292 (87.4)		2.74 (1.81-4.14)[c]	2.10 (1.33-3.32)[d]
Nationality, n (%)			<.001		
Spanish	81 (5.7)	1351 (94.3)		Reference	Reference
Spanish and/or other	20 (14.8)	115 (85.2)		2.90 (1.72-4.90)[c]	2.55 (1.40-4.62)[d]
Educational level, n (%)			.68		
University	56 (6.9)	759 (93.1)		Reference	—[e]
Nonuniversity	48 (6.4)	707 (93.6)		0.92 (0.62-1.37)	—

[a]The multivariate model included as covariates those variables that showed a *P* value of <.05 in the bivariate analysis.

[b]OR: odds ratio.

[c]*P*<.001.

[d]*P*<.05.

[e]Variables not included in the multivariate model according to the results of the bivariate analysis.

Regarding DFSA victimization while partying, the results of the bivariate analysis were largely consistent with those of the multivariate analysis. In this model, the probability was 1.86 times higher among DFSA pornography users (95% CI 1.24-2.78, *P*=.003). This association was also 3.36 (95% CI 2.53-4.46) times higher among female participants (*P*<.001), 1.67 (95% CI 1.28-2.20) times higher among nonheterosexual individuals (*P*<.001), 1.65 (95% CI 1.10-2.47) times higher among individuals of foreign origin (*P*=.01), and 1.36 (95% CI 1.09-1.71) times higher among those with lower educational levels (*P*=.008). The overall model (χ^2_9=128.4; *P*<.001) explained 6.6% of the variance (McFadden pseudo-R^2=0.066) and showed good fit to the data (Hosmer-Lemeshow χ^2_8=4.5; *P*=.81; Table 4).

XSL•FO
RenderX

Table 4. The relationship between lifetime drug-facilitated sexual assault (DFSA) victimization while partying and the type of pornography consumed and frequency of consumption (N=1470).[a]

Sociodemographic and behavioral variables	Experienced DFSA at any point in their life while partying		*P* value	Crude OR[b] (95% CI)	Adjusted OR (95% CI)
	Yes	No			
Type of pornography consumed, n (%)			.07		
No consumption	190 (37.8)	313 (62.2)		Reference	Reference
Without DFSA	275 (36.5)	478 (63.5)		0.95 (0.75-1.20)	1.26 (0.92-1.72)
With DFSA	109 (44.7)	135 (55.3)		1.33 (0.98-1.81)	1.86 (1.24-2.78)[c]
Frequency of pornography consumption, n (%)			.02		
Never to less than once per month	319 (39.9)	481 (60.1)		Reference	Reference
Once per week to 2 to 3 times per month	137 (41.4)	194 (58.6)		1.06 (0.82-1.38)	1.28 (0.90-1.82)
Daily to 2 to 3 times per week	118 (32.0)	251 (68.0)		0.71 (0.55-0.92)	1.12 (0.77-1.64)
Sex, n (%)			<.001		
Male	210 (27.0)	567 (73.0)		Reference	Reference
Female	381 (48.4)	406 (51.6)		2.53 (2.05-3.13)[d]	3.36 (2.53-4.46)[d]
Age (y), median (SD)	27 (5)	27 (5)	.88	1.00 (0.98-1.02)	1.03 (1.00-1.05)[c]
Sexual orientation, n (%)			<.001		
Heterosexual	430 (35.3)	790 (64.8)		Reference	Reference
Nonheterosexual	156 (47.7)	171 (52.3)		1.68 (1.31-2.15)[d]	1.67 (1.28-2.20)[d]
Nationality, n (%)			.002		
Spanish	521 (36.5)	905 (63.5)		Reference	Reference
Spanish and/or other	66 (50.0)	66 (50.0)		1.74 (1.21-2.48)[c]	1.65 (1.10-2.47)[c]
Educational level, n (%)			.001		
University	274 (34.0)	532 (66.0)		Reference	Reference
Nonuniversity	319 (42.2)	437 (57.8)		1.42 (1.15-1.74)[c]	1.36 (1.09-1.71)[c]

[a]The multivariate model included as covariates those variables that showed a *P* value of <.05 in the bivariate analysis.

[b]OR: odds ratio.

[c]*P*<.05.

[d]*P*<.001.

Discussion

This study confirms the widespread use of pornography among young people, indicating that those who use it more frequently are also more likely to consume sexually violent pornography. Additionally, evidence was provided regarding the relationship between the use of sexually violent pornography and both DFSA perpetration and victimization.

In this study, 2 out of 3 young people (1013/1521, 66.6%) admitted to consuming pornography, with consumption being almost twice as high in male participants compared to female participants. These results align with a recent nationwide study in Spain, according to which 6 out of 10 young people aged 16 to 29 years admitted to using pornography, with 72.1% being men and 59.3% being women [19]. Similarly, other studies have found that approximately 70% of men and 30% of women use pornography in high-income countries [17,18]. Regarding frequency, in our study, almost half (333/753, 44.2%) of all male participants consumed pornography daily or 2 to 3 times a week. Among female participants, consumption was much less frequent. These figures are slightly higher than those reported in similar studies conducted in Spain [19] and Europe [36]. The data from our study are more up-to-date and are based on more representative samples. Therefore, they may reflect current consumption, especially considering that pornography consumption has been growing in recent years [22,37].

In terms of type of pornography, more than 20% of male participants (167/753, 22.2%) and 10% of female participants (86/762, 11.3%) in our study reported consumption of pornography featuring scenes in which a person is asleep, unconscious, or under the influence of alcohol or other drugs

XSL•FO
RenderX

(referred to as DFSA pornography). These figures cannot be directly compared as there is currently no scientific literature providing prevalence data on the consumption of violent pornographic content specifically depicting DFSA situations. As such, comparisons could be made with broader data on the consumption of violent pornography. Some studies indicate that 10% of adolescents have been exposed to violent pornography, with a gradual increase in violent themes as age progresses [38]. In this regard, a study in New Zealand highlighted how easily young people can be exposed to nonconsensual sexual behavior in online pornography, including scenes of sexual activity while someone is sleeping [39], findings similar to those described by other authors [30,31]. In Spain, 40.2% of individuals have viewed pornographic content classified as high in violence, particularly degrading, or humiliating, and 16.6% acknowledge doing so with high or moderate frequency (18.2% of men and 14.5% of women) [19]. Additionally, 5.2% of men and 6.9% of women indicate that the presence of violence is the factor that influences them the most when selecting pornographic material [19]. These results show the magnitude of pornography consumption of this type. There are many related considerations, but these are beyond the scope of this study. What we must emphasize again is that many young people acquire their sexual education through pornography (in the absence of other forms of education) [19,40], with a significant portion of this pornography being of a violent nature.

According to our study, using DFSA pornography is related to approximately fourfold higher odds of perpetrated DFSA while partying. Similarly, other researchers have observed a connection between sexual assault perpetration and the use of violent pornography [23,41]. Models explaining this relationship suggest that the risk of perpetration correlates with male pornography users who have high levels of hostility and sexual promiscuity [42]. Exposure to violent pornography also shapes sexual behavior by reinforcing scripts that are perceived as normative, acceptable, and gratifying, which are then activated and applied in dating and sexual relationships [43]. This correlation was obtained in analyses adjusting for other factors. In addition to the type of pornography consumption, it was noted in this study that there were more perpetrators among male participants, foreigners, and nonheterosexual individuals. The presence of men in these studies is nothing new [23,44], hence the importance of adjusting the results by gender. Regarding country of origin and sexual orientation, these categories encompass highly diverse populations. These groups are heterogeneous and shaped by intersecting social determinants that may influence risk and experiences in complex ways [45]. Further research is needed stratifying by country of origin and other relevant factors [46,47].

Concerning victimization, experiencing DFSA while partying was twice as frequent among users of DFSA pornography content. Although various studies have indicated a link between sexual victimization and exposure to pornography [44,48,49], there is currently no known research specifically addressing DFSA. On the one hand, many studies on violence against women suggest that conventional pornography often depicts violent scenes that sexually degrade and objectify women [23,28,29,40]. It is noted that female sexual objectification in pornography is linked to a higher likelihood of sexual victimization in young women due to the normalization of or desensitization to violent sexual behaviors [50]. On the other hand, studies linking violence using psychoactive substances and pornography primarily focus on alcohol, consistently finding an increased risk of sexual victimization among women who consume pornography [48]. While it is difficult to infer a cause, it has been suggested that increased exposure to violent pornography and its normalization, combined with the effects of alcohol, may impair the initial ability to detect an aggression [51]. At the same time, considering the absence of temporality in cross-sectional designs, the observed relationship between having experienced DFSA and DFSA pornography consumption leads us to hypothesize that, in the absence of adequate social support, some DFSA survivors may resort to viewing such content to understand the episode even at the risk of revictimization. However, this hypothesis is not supported by empirical data, suggesting directions for future research. Regardless, this highlights the imperative to strengthen support systems for survivors of DFSA. Consequently, it is essential to conduct longitudinal studies that establish causal relationships and qualitative research that explores DFSA survivors' experiences in depth. In the same way, related adjustment variables indicate that this is an issue in which global messages must be accompanied by other messages particularized according to gender, educational level, or country of origin.

Considering the high prevalence of DFSA experiences in youth party settings [6] and the fact that many young people turn to pornography as a source of sexual education [19], the influence of DFSA-themed pornography in normalizing this form of sexual violence among young people is deeply concerning. While we cannot establish a causal relationship, there is a clear correlation between DFSA—whether perpetration or victimization—and the consumption of this specific type of pornography. This correlation does not emerge when considering the overall quantity of pornography consumed. Although higher levels of consumption are related to the consumption of violent porn, it is specifically this violent content that shows a significant correlation with DFSA. Consequently, we should not wait for causal evidence to implement the necessary sexual health promotion interventions. Preventive action is essential, particularly to ensure that the normalization of sexual behaviors is not shaped primarily by the content of the pornography that young people consume.

This study's findings contribute to addressing some global challenges highlighted by the United Nations Sustainable Development Goals, particularly those aimed at ending violence and fostering just societies. As DFSA is a paradigmatic example of how violence affects the most vulnerable individuals, this study reinforces the commitment to "leave no one behind" in achieving sustainability.

This study has several limitations. First, the cross-sectional design precludes causal inference. Second, although quota sampling approximated the demographic distribution of the Spanish population in terms of age, gender, and region, participation was voluntary and based on an online panel, which may have introduced self-selection bias and limited generalizability. Social desirability bias cannot be entirely ruled

XSL•FO
RenderX

out despite assurances of anonymity. In addition, the representation of sexually diverse and gender-diverse groups was limited, and the sample size did not allow for analyses by DFSA subtype or level of invasiveness. Third, although the questionnaire was piloted, some constructs—such as the use of DFSA-related pornographic content—were assessed using nonvalidated items, which may affect measurement precision. Finally, no psychological support was available for participants who might have experienced discomfort. Despite these limitations, this study provides novel and valuable evidence on DFSA-related pornographic content and its association with sexual violence perpetration and victimization in party contexts.

The use of pornography depicting nonconsensual sex involving individuals who are asleep, unconscious, or under the influence of alcohol or other drugs correlates with experiences of DFSA perpetration and victimization in youth party settings. This association emphasizes the importance of addressing the impact of pornography as a primary source of sexual knowledge for young people, particularly in a context in which the widespread use of pornography intersects with insufficient sexual education and a lack of effective mechanisms to control consumption among minors. Future public health strategies may include ongoing research regarding the implications of pornography for sexual violence. From a public health policy perspective, it is recommended to provide training programs for educators and clinicians to enhance their skills in sex education considering the realities of pornography and its influence on young people. Accurate information on these realities should also be integrated into comprehensive sexual education initiatives and educational materials aimed directly at youth. Finally, legal responsibilities should be expanded, and the obligations of online pornography distribution platforms should be strengthened to improve the monitoring, filtering, and removal of violent content.

Acknowledgments

The authors thank the study participants for their contribution and The Cocktail Analysis for fieldwork support. The authors acknowledge that a generative artificial intelligence tool (ChatGPT; OpenAI) was used exclusively to assist with language editing and phrasing improvements. The authors verified all content for accuracy and take full responsibility for the manuscript's content.

Funding

This work was funded by the Margarita Salas postdoctoral grant financed by the University of Alcalá, NextGenerationEU, and the Spanish Ministry of Universities, as well as by the University Institute for Research in Police Sciences (IUICP-2023/07). The funding source had no involvement in the study.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

PP-M took the lead in writing the manuscript and performing the analytic calculations with support from LS. Both authors designed and planned the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questionnaire used in the study.
[DOCX File , 24 KB - publichealth_v12i1e80110_app1.docx]

References

1. Sardinha L, Maheu-Giroux M, Stöckl H, Meyer SR, García-Moreno C. Global, regional, and national prevalence estimates of physical or sexual, or both, intimate partner violence against women in 2018. Lancet 2022 Feb 26;399(10327):803-813 [FREE Full text] [doi: 10.1016/S0140-6736(21)02664-7] [Medline: 35182472]
2. Krug EG, Mercy JA, Dahlberg LL, Zwi AB. The world report on violence and health. Lancet 2002 Oct 05;360(9339):1083-1088. [doi: 10.1016/S0140-6736(02)11133-0] [Medline: 12384003]
3. Recalde-Esnoz I, Prego-Meleiro P, Montalvo G, Del Castillo H. Drug-facilitated sexual assault: a systematic review. Trauma Violence Abuse 2024 Jul;25(3):1814-1825. [doi: 10.1177/15248380231195877] [Medline: 37650508]
4. Prego-Meleiro P, Montalvo G, Quintela-Jorge Ó, García-Ruiz C. Increasing awareness of the severity of female victimization by opportunistic drug-facilitated sexual assault: a new viewpoint. Forensic Sci Int 2020 Oct;315:110460 [FREE Full text] [doi: 10.1016/j.forsciint.2020.110460] [Medline: 32858463]

XSL•FO
RenderX

5. Cruz-Landeira A, Quintela-Jorge Ó, López-Rivadulla M. [Chemical submission, epidemiology and some clues for the diagnosis]. Med Clin (Barc) 2008 Dec 06;131(20):783-789 [FREE Full text] [doi: 10.1016/s0025-7753(08)75505-2] [Medline: 19094881]
6. Prego-Meleiro P, Recalde-Esnoz I, Sordo L, Del Castillo H, García-Ruiz C, Montalvo G, et al. Drug-facilitated sexual assault in youth nightlife contexts in Spain. Public Health 2024 Aug;233:157-163 [FREE Full text] [doi: 10.1016/j.puhe.2024.05.022] [Medline: 38901300]
7. Diagnóstico sobre informaciones mediáticas en torno a las violencias sexuales, el uso de drogas y la sumisión química. Drogas y Género. URL: https://www.drogasgenero.info/diagnostico-medioscomunicacion/ [accessed 2025-05-29]
8. Rial A, García-Couceiro N, Gómez P, Braña T, Isorna M. Menores y agresiones sexuales facilitadas por drogas: entre la sumisión y la vulnerabilidad química. Adicciones 2024 Sep 25;36(3):317-328 [FREE Full text] [doi: 10.20882/adicciones.1871] [Medline: 39033520]
9. Recalde-Esnoz I, Prego-Meleiro P, Montalvo G, Del Castillo H. Drug-facilitated sexual assault acceptance scale: construction and validation of a new scale in Spanish context. J Public Health (Berl.) 2023 Sep 20;33(5):1157-1166 [FREE Full text] [doi: 10.1007/S10389-023-02084-z]
10. Panyella-Carbó MN, Agustina JR, Martin-Fumadó C. Sumisión química versus vulnerabilidad química: análisis criminológico de los delitos sexuales facilitados mediante el uso de sustancias psicoactivas a partir de una muestra de sentencias. Rev Esp Invest Criminol 2019;17:1-13 [FREE Full text] [doi: 10.46381/reic.v17i0.247]
11. Espada JP, Morales A, Orgilés M, Piqueras JA, Carballo JL. Comportamiento sexual bajo la influencia del alcohol en adolescentes españoles. Adicciones 2013;25(1):55-62 [FREE Full text] [doi: 10.20882/adicciones.72]
12. Madureira AS. El asesinato de mujeres por ser mujeres. Methaodos Rev Cienc Soc 2015;3(2):85. [doi: 10.17502/m.rcs.v3i2.94]
13. José S, Briceño H. Cultura de la violación, un análisis del continuo en la violencia sexual que viven las mujeres. Pacha Rev Est Contemp Sur Global 2020;1(3):89-103 [FREE Full text] [doi: 10.46652/pacha.v1i3.44]
14. Recio BT. Todo es mentira: cultura de la violación, mitos y falsas creencias sobre violencia sexual hacia las mujeres. Política Soc 2022;59(1):e78892 [FREE Full text] [doi: 10.5209/poso.78892]
15. Meneses AC, Barroso JG, Martín-Palomino ET. Los procesos de pornificación social y su relación con el incremento de la violencia sexual contra mujeres y niñas. Feminismo/s 2025;1(46):387 [FREE Full text] [doi: 10.14198/fem.2025.46.15]
16. Gutiérrez Barroso J, Romero MY, González RA, Alario GM, Ranea TB, Robles SA. Sexualidad y consumo de pornografía en adolescentes y jóvenes de 16 a 29 años. Universidad de La Laguna. URL: https://doi.org/10.25145/b.SexAdolesct.2021 [accessed 2025-05-29]
17. Peter J, Valkenburg PM. Adolescents and pornography: a review of 20 years of research. J Sex Res 2016;53(4-5):509-531 [FREE Full text] [doi: 10.1080/00224499.2016.1143441] [Medline: 27105446]
18. Rissel C, Richters J, de Visser RO, McKee A, Yeung A, Caruana T. A profile of pornography users in Australia: findings from the second Australian study of health and relationships. J Sex Res 2017 Feb;54(2):227-240 [FREE Full text] [doi: 10.1080/00224499.2016.1191597] [Medline: 27419739]
19. Juventud y pornografía en la era digital: consumo, percepción y efectos. Centro Reina Sofía de Fad Juventud. URL: https://www.centroreinasofia.org/publicacion/juventud-y-pornografia-en-la-era-digital-consumo-percepcion-y-efectos/ [accessed 2025-05-29]
20. Informe (Des)información sexual: pornografía y adolescencia. Save the Children. URL: https://www.savethechildren.es/informe-desinformacion-sexual-pornografia-y-adolescencia [accessed 2025-05-29]
21. Díaz Hernández C, Gutiérrez Barroso J, Torrado Martín-Palomino E. Pornography consumption and sexual behaviors in Spanish adolescents and young adults: findings from a sample of girls and boys aged 15 to 29 years. Géneros 2024 Sep 18;13(3):1-19. [doi: 10.17583/generos.14080]
22. Ashton S, McDonald K, Kirkman M. What does ‘pornography’ mean in the digital age? Revisiting a definition for social science researchers. Porn Studies 2019 Feb 07;6(2):144-168 [FREE Full text] [doi: 10.1080/23268743.2018.1544096]
23. Rostad WL, Gittins-Stone D, Huntington C, Rizzo CJ, Pearlman D, Orchowski L. The association between exposure to violent pornography and teen dating violence in grade 10 high school students. Arch Sex Behav 2019 Oct;48(7):2137-2147 [FREE Full text] [doi: 10.1007/s10508-019-1435-4] [Medline: 31309432]
24. Rothman EF, Kaczmarsky C, Burke N, Jansen E, Baughman A. "Without porn … I wouldn't know half the things i know now": a qualitative study of pornography use among a sample of urban, low-income, black and Hispanic youth. J Sex Res 2015;52(7):736-746. [doi: 10.1080/00224499.2014.960908] [Medline: 25350847]
25. Svedin CG, Akerman I, Priebe G. Frequent users of pornography: a population based epidemiological study of Swedish male adolescents. J Adolesc 2011 Aug;34(4):779-788 [FREE Full text] [doi: 10.1016/j.adolescence.2010.04.010] [Medline: 20888038]
26. Aznar-Martínez B, Lorente-De-Sanz J, López-i-Martín X, Castillo-Garayoa J. Pornography and gender-based violence: two neglected topics in sexuality education. A systematic review. Sex Education 2024 Feb 19;25(2):256-272 [FREE Full text] [doi: 10.1080/14681811.2024.2316154]
27. Litsou K, Byron P, McKee A, Ingham R. Learning from pornography: results of a mixed methods systematic review. Sex Educa 2020 Jul 09;21(2):236-252. [doi: 10.1080/14681811.2020.1786362]

XSL•FO
RenderX

28. Carrotte ER, Davis AC, Lim MS. Sexual behaviors and violence in pornography: systematic review and narrative synthesis of video content analyses. J Med Internet Res 2020 May 14;22(5):e16702 [FREE Full text] [doi: 10.2196/16702] [Medline: 32406863]
29. Mestre-Bach G, Villena-Moya A, Chiclana-Actis C. Pornography use and violence: a systematic review of the last 20 years. Trauma Violence Abuse 2024 Apr;25(2):1088-1112 [FREE Full text] [doi: 10.1177/15248380231173619] [Medline: 37309642]
30. Paul P. Pornified: How Pornography Is Transforming Our Lives, Our Relationships, and Our Families. Internet Archive. 2005. URL: https://archive.org/details/pornifiedhowporn0000paul [accessed 2025-05-29]
31. La reproducción de la violencia sexual en las sociedades formalmente igualitarias: un análisis filosófico de la cultura de la violación actual a través de los discursos y el imaginario de la pornografía. Ministerio de Igualdad. URL: https://violenciagenero.igualdad.gob.es/violenciaencifras/estudios/tesis/tesis-10-reproduccion-violencia-pornografia/ [accessed 2025-05-29]
32. Demografía y población / Padrón / Estadística del Padrón continuo / Últimos datos. Instituto Nacional de Estadística (INE). URL: https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177012&menu=ultiDatos&idp=1254734710990 [accessed 2025-05-29]
33. Lwanga SK, Lemeshow S. Sample size determination in health studies: a practical manual. World Health Organization. 1991. URL: https://iris.who.int/handle/10665/40062 [accessed 2025-05-29]
34. Koss MP, Abbey A, Campbell R, Cook S, Norris J, Testa M, et al. Revising the SES: a collaborative process to improve assessment of sexual aggression and victimization. Psychol Women Q 2007 Dec 01;31(4):357-370. [doi: 10.1111/j.1471-6402.2007.00385.x]
35. Macroencuesta de violencia contra la mujer. Ministerio de Igualdad. 2019. URL: https://violenciagenero.igualdad.gob.es/wp-content/uploads/Macroencuesta_2019_estudio_investigacion.pdf [accessed 2025-05-29]
36. Andrie EK, Sakou II, Tzavela EC, Richardson C, Tsitsika A. Adolescents' online pornography exposure and its relationship to sociodemographic and psychopathological correlates: a cross-sectional study in six European countries. Children (Basel) 2021 Oct 16;8(10):925 [FREE Full text] [doi: 10.3390/children8100925] [Medline: 34682190]
37. Giordano AL, Cashwell CS. Cybersex addiction among college students: a prevalence study. Sex Addict Compulsivity 2017 Mar 28;24(1-2):47-57 [FREE Full text] [doi: 10.1080/10720162.2017.1287612]
38. Vandenbosch L. Antecedents of adolescents' exposure to different types of sexually explicit internet material: a longitudinal study. Comput Human Behav 2015 Sep;50:439-448. [doi: 10.1016/j.chb.2015.04.032]
39. Breaking down porn. Classification Office. 2019 Dec. URL: https://www.classificationoffice.govt.nz/resources/research/breaking-down-porn/ [accessed 2025-05-29]
40. Miller DJ, McBain KA. The content of contemporary, mainstream pornography: a literature review of content analytic studies. Am J Sex Educ 2022;17(2):219-256 [FREE Full text] [doi: 10.1080/15546128.2021.2019648]
41. Antevska A, Gavey N. "Out of sight and out of mind". Men Masc 2015 Apr 07;18(5):605-629. [doi: 10.1177/1097184x15574339]
42. Malamuth NM, Addison T, Koss M. Pornography and sexual aggression: are there reliable effects and can we understand them? Annu Rev Sex Res 2000;11:26-91. [Medline: 11351835]
43. Wright PJ. Mass media effects on youth sexual behavior: assessing the claim for causality. Ann Int Commun Assoc 2016 May 18;35(1):343-385 [FREE Full text] [doi: 10.1080/23808985.2011.11679121]
44. Rodenhizer KA, Edwards KM. The impacts of sexual media exposure on adolescent and emerging adults' dating and sexual violence attitudes and behaviors: a critical review of the literature. Trauma Violence Abuse 2019 Oct 13;20(4):439-452. [doi: 10.1177/1524838017717745] [Medline: 29333966]
45. Mayo D, Lozano A, Williams R, Safren S, Harkness A. The relationship between intersectional oppression and affirmation and latino sexual minority men's mental health. LGBT Health 2023;10(8):629-638 [FREE Full text] [doi: 10.1089/lgbt.2022.0212] [Medline: 37466482]
46. Bauer GR, Churchill SM, Mahendran M, Walwyn C, Lizotte D, Villa-Rueda A. Intersectionality in quantitative research: a systematic review of its emergence and applications of theory and methods. SSM Popul Health 2021 Jun;14:100798 [FREE Full text] [doi: 10.1016/j.ssmph.2021.100798] [Medline: 33997247]
47. Ghasemi E, Majdzadeh R, Rajabi F, Vedadhir A, Negarandeh R, Jamshidi E, et al. "Applying Intersectionality in designing and implementing health interventions: a scoping review". BMC Public Health 2021 Jul 16;21(1):1407 [FREE Full text] [doi: 10.1186/s12889-021-11449-6] [Medline: 34271905]
48. de Heer B, Prior S, Fejervary J. Women's pornography consumption, alcohol use, and sexual victimization. Violence Against Women 2021 Aug;27(10):1678-1695. [doi: 10.1177/1077801220945035] [Medline: 32791027]
49. Maas MK, Bray BC, Noll JG. Online sexual experiences predict subsequent sexual health and victimization outcomes among female adolescents: a latent class analysis. J Youth Adolesc 2019 May 18;48(5):837-849 [FREE Full text] [doi: 10.1007/s10964-019-00995-3] [Medline: 30778831]
50. Franz MR, DiLillo D, Gervais SJ. Sexual objectification and sexual assault: do self-objectification and sexual assertiveness account for the link? Psychol Violence 2016 Apr;6(2):262-270. [doi: 10.1037/vio0000015]

XSL•FO
RenderX

51. Davis KC, Norris J, George WH, Martell J, Heiman JR. Rape-myth congruent beliefs in women resulting from exposure to violent pornography: effects of alcohol and sexual arousal. J Interpers Violence 2006 Sep 01;21(9):1208-1223. [doi: 10.1177/0886260506290428] [Medline: 16893966]

Abbreviations

aOR: adjusted odds ratio
DFSA: drug-facilitated sexual assault

Edited by A Mavragani, T Sanchez; submitted 04.Jul.2025; peer-reviewed by MDM Sanchez-Fuentes, JG Barroso, T Church; comments to author 24.Sep.2025; revised version received 13.Nov.2025; accepted 17.Nov.2025; published 22.Jan.2026.

Please cite as:
Prego-Meleiro P, Pastor-Moreno G, Recalde-Esnoz I, Sordo L
Drug-Facilitated Sexual Assault Pornography and Sexual Violence While Partying: Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e80110
URL: https://publichealth.jmir.org/2026/1/e80110
doi:10.2196/80110
PMID:41571270

©Pablo Prego-Meleiro, Guadalupe Pastor-Moreno, Irantzu Recalde-Esnoz, Luis Sordo. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 22.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Problematic Alcohol Use Among Adolescents in Germany: Representative Cross-Sectional Study

Rebekka Schröder[1], Dr phil; Tim Hamer[1], MSc; Ralf Suhr[1], PD, Dr med; Lars König[1], Dr rer nat

Stiftung Gesundheitswissen, Berlin, Germany

Corresponding Author:
Lars König, Dr rer nat
Stiftung Gesundheitswissen
Friedrichstr. 134
Berlin, 10117
Germany
Phone: 49 30419549262
Email: lars.koenig@stiftung-gesundheitswissen.de

Abstract

Background: Alcohol is a widely used psychoactive substance, and its use constitutes a major public health challenge due to its immediate and long-term adverse effects on various health-related outcomes. Adolescence has been identified as a particularly vulnerable phase regarding alcohol use. Although consumption rates in this age group have declined in Germany over the past decades, a plateau has been reached, and there is a continued need for interventions to further reduce consumption rates.

Objective: This study aimed to assess problematic alcohol use among adolescents in Germany and explore associations with sociodemographic and psychosocial characteristics, particularly with health literacy, to inform future interventions tailored to the specific needs of this target group.

Methods: In a cross-sectional quota-based survey, 2006 adolescents (aged 12-17 years) completed an online survey (n=1406) or face-to-face interview (n=600) assessing the frequency of weekly alcohol use, the presence of problematic alcohol use (German version of the Car-, Relax-, Alone-, Forget-, Friends-, Trouble- questionnaire [CRAFFT-d]), sociodemographic information, and health literacy (European Health Literacy Survey instrument [HLS-EU-Q16]). Based on their CRAFFT-d and HLS-EU-Q16 scores, participants were identified as exhibiting problematic alcohol use (vs no problematic alcohol use) and inadequate or problematic health literacy levels (vs adequate health literacy levels), respectively. Chi-square tests were computed to analyze differences between different groups (as defined by the sociodemographic factors, weekly alcohol consumption frequency, and health literacy) in terms of problematic alcohol use (binary CRAFFT-d outcome).

Results: Approximately 20% (390/2006) of the participants reported consuming alcohol on at least 1 day per week, and 12.7% (255/2006) of the sample met the CRAFFT-d screening criterion for problematic alcohol use. Problematic alcohol use was significantly associated with gender (χ^2_1=20.96, V=0.10; P<.001), age (χ^2_2=85.88, V=0.21; P<.001), subjective social status (χ^2_2=8.23, V=0.06; P=.02), and migration background (χ^2_1=5.60, V=0.05; P=.02), but there were no significant associations with level of education (χ^2_1=3.43, V=0.04; P=.06), and health literacy (χ^2_1=1.54, V=0.03; P=.21). In addition, participants who reported more frequent alcohol consumption per week, also met the screening criterion for problematic alcohol consumption more frequently (χ^2_7=698.65, V=0.59; P<.001).

Conclusions: The findings demonstrate that problematic alcohol use is more common in boys than girls, in older vs younger adolescents, in those with high or low (vs intermediate) social status, in individuals with (vs without) a migration background, and in those who drink alcohol more frequently. These results emphasize the necessity of implementing targeted prevention strategies that address the specific risk profiles of adolescents concerning alcohol consumption.

(JMIR Public Health Surveill 2026;12:e87800) doi:10.2196/87800

KEYWORDS

adolescents; alcohol, Germany; health literacy; problematic alcohol use; representative survey

XSL•FO
RenderX

Introduction

Background

Alcohol use poses a major public health challenge [1,2], and adolescence has emerged as a particularly vulnerable period for its detrimental short- and long-term effects [3-5]. This study addresses the problematic use of alcohol among adolescents in Germany and how this problematic use relates to psychosocial and sociodemographic factors with cross-sectional data.

Public Health Relevance of Alcohol Use

Alcohol is one of the most widely used psychoactive substances worldwide [6,7]. At the same time, more than 3 million deaths worldwide are attributable to alcohol, corresponding to 5.3% of all deaths [2]. This effect is exacerbated in younger age, with 18.4% of all deaths in men aged 25-29 years attributed to alcohol. Women are less affected than men but also experience a substantial health burden due to alcohol [2]. While there is some evidence that alcohol might have beneficial effects when consumed in smaller quantities [8-10], this position has been challenged in recent years and might be attributed to methodological artifacts [11-13]. Importantly, it is undisputed that health effects are detrimental when alcohol is consumed in large quantities [10,14-16]. These harmful effects on health can be observed in the short and long term [11]. Short-term effects refer to the immediate consequences of alcohol consumption, including acute alcohol intoxication as well as decreased visuomotor coordination, increased reaction times, and increased impulsivity [11,17,18]. These may lead to more severe consequences, such as injuries or deaths, for example, from falls, traffic accidents, or drowning [11,19]. The arguably most apparent long-term effects of excessive alcohol consumption are clinically relevant alcohol abuse and dependence, which are highly prevalent worldwide, and also in many European countries, including Germany [1,20]. These effects are associated with a significant decline in physical and mental health and a loss of quality of life [21,22]. Other long-term effects of alcohol consumption on health-related outcomes include, but are not limited to, cancer, cardiovascular, gastrointestinal, and neuropsychiatric conditions, which may result in premature mortality [11,16]. In addition to these individual consequences, alcohol consumption has also been associated with substantial direct and indirect economic costs at a societal level [23].

Alcohol Use Among Adolescents

Given the detrimental effects of alcohol consumption on various health outcomes, more recent nutritional policies in Germany recommend abstaining from alcohol entirely [11]. However, alcohol consumption remains at a high level in Germany compared to other Western-European countries [24,25]. Critically, alcohol is consumed from a young age [26]. Recent evidence from representative samples estimates the prevalence of regular alcohol use, defined as consumption of alcoholic beverages at least weekly in the past 12 months, at about 9.7% in those aged 12-17 years, with substantially higher rates in boys (12.4%) than girls (6.9%). In addition, 63% of all adolescents in that age group report lifetime alcohol use, and about 4% show risky drinking behaviors, exceeding established thresholds for risky alcohol consumption in adults (equivalent to drinking more than 12 grams of pure alcohol for women and 24 grams of pure alcohol for men daily). However, a decrease in general and risky alcohol consumption over the past decades has also been observed in adolescents [26]. Crucially, this decline appears to have plateaued in the last few years, showing only minimal variation [26].

Problematic Alcohol Use Among Adolescents and Its Correlates

The stagnation in alcohol use in recent years underscores the need for renewed public health strategies to address the persistent harm associated with alcohol consumption from a young age. A particular focus should be placed on problematic alcohol use in adolescents as this has been associated with severe consequences, including functional and structural brain damage and increased risks for developing alcohol use disorders [3-5,27]. A recent investigation with data from a representative sample from Germany found that 11.3% of the participants aged 12-17 years met the criteria for problematic alcohol use on a standardized screening tool. Moreover, it was shown that problematic alcohol use was associated with greater psychopathology, lower mindfulness, and lower quality of life, even when potentially confounding factors were controlled for [28]. In order to develop interventions that are specifically tailored to societal needs, particularly for adolescents, it is necessary to collect more detailed data on the associations between problematic alcohol use and other factors.

In particular, there is currently insufficient data on the relationship between problematic alcohol consumption and specific sociodemographic and psychosocial factors. For example, the association with health literacy has not yet been adequately explored. Health literacy encompasses the knowledge, motivation, and competencies in the process of accessing, understanding, appraising, and applying health information [29]. Health-literate individuals are more likely to successfully take care of their own health and navigate the health care system [29]. As a consequence, individuals with greater health literacy often exhibit healthier lifestyles (eg, adequate medication intake, less sedentary behavior), have a better overall health status, and lower mortality than individuals with lower health literacy levels [29-32]. This might also give them an advantage when making decisions regarding their alcohol consumption [33,34]. For example, greater levels of health literacy may help individuals to understand and appraise content, units, strengths, and harms of alcohol and make healthier choices regarding their alcohol intake [35,36].

Furthermore, while some relevant sociodemographic correlates (eg, gender) of (problematic) alcohol use among adolescents have already been obtained [26], there is still insufficient data on other possible associations and a strong need for replication of prior findings. Particularly, some associations have not been found consistently, for example, concerning levels of education and social status [37-40].

Study Aims

Therefore, the aim of this study was to assess the frequency of alcohol use and the prevalence of individuals with problematic

XSL•FO
RenderX

alcohol use among adolescents in Germany, and to explore associations between problematic alcohol use and various sociodemographic and psychosocial characteristics, including health literacy.

Methods

Participants, Recruitment, and Survey Methodology

The target population for this study was adolescents aged 12-17 years living in Germany with sufficient knowledge of the German language to participate in the study. We aimed to recruit 2000 individuals in order to allow precise estimates of prevalences of problematic drinking, to test for differences between the sociodemographic and psychosocial groups, and to allow representativeness of the sample for the characteristics described below (more information is given in the Data Weighting section). For example, an a priori power analysis for a chi-square test of independence for comparing the prevalences of problematic drinking across different age groups yielded a minimum sample size of 1283 participants (input parameters: 80% power, α=.05, *df*=5, effect size ω=0.1 [41]). Recruitment of the participants, study administration, and data acquisition were carried out by the German market research institute Gesellschaft für Innovative Marktforschung mbH and took place from November 2024 to January 2025. Only complete datasets without missing data were provided by the market research institute. A total of 2006 individuals participated in the study. Data were collected both online via web-based surveys (n=1406 participants, 70.1% of all participants) and via in-person face-to-face interviews (n=600 participants, 29.9% of all participants). Online participants were recruited from 3 online-access panels via personalized e-mail invitation links. Face-to-face participants were recruited from validated and regularly updated address pools of private households in Germany with known household composition. All participants aged 15 and younger were recruited via their parents. A quota-based sampling strategy was used, with quotas for 3 age categories (12-13 years, 14-15 years, and 16-17 years), gender, type of school, and federal state. Quotas were based on current data from the MA Audio (agma) study and the Federal Statistical Office (Statistisches Bundesamt). The MA Audio study is a large nation-wide and representative study with more than 66,000 yearly interviews. Detailed descriptive statistics for the final sample are in Table 1.

Data Weighting

Despite careful quota-based recruitment, some disparities between the sample and the reference population regarding relevant sociodemographic characteristics remain, for example, because some sociodemographic characteristics, such as level of education, negatively correlate with the willingness to participate in online surveys. Therefore, the data were weighted by the market research institute to match the target distribution of relevant reference studies (again using the latest MA Audio study and specifications of the Federal Statistical Office). The reference population for the weighting procedure is the German-speaking population living in Germany aged 12 years and older. The following weighting variables and combinations were used: age × gender; education (operationalized as type of school); and federal state. The calculation method involved using an iterative process in which all weighting variables of the sample are used simultaneously to inform the weighting with the aim of achieving minimal weights. This procedure resulted in a single weighting factor per individual.

Measures

Sociodemographic Information

We collected sociodemographic information concerning gender (boys, girls, and diverse), age, level of education, subjective social status, and migration background. For further analyses, participants were categorized into 3 age groups (12-13 years, 14-15 years, and 16-17 years).

To assess levels of education, participants were asked to indicate the type of school they currently attend. They were grouped into 2 categories: those participants attending the most advanced type of secondary school (Gymnasium, equivalent to grammar schools where a university entrance qualification can be obtained) were assigned a high level of education, and all other participants were assigned a low level of education.

Subjective social status was assessed using the German version of the MacArthur scale [42,43]. The scale uses the analogy of a ladder to represent social status, with the top rung (10) symbolizing the highest social status and the bottom rung (1) denoting the lowest social status. Participants are asked to identify the rung that best represents their individual position relative to other members of society. In total, 3 categories of subjective social status were determined: low subjective social status (scores 1-4), intermediate subjective social status (scores 5-7), and high subjective social status (scores 8-10).

Finally, participants were asked to specify whether they had a migration background, which was defined as having at least 1 parent who was not born in Germany or having been born outside of Germany themselves.

Health Literacy

The German translation of the short version of the European Health Literacy Survey instrument (HLS-EU-Q16) was administered to assess health literacy [44-46]. Participants were asked to indicate perceived difficulty in accessing, understanding, appraising, and applying information in the domains of health care, disease prevention, and health promotion. A total of 16 items was presented with a 4-point Likert scale (“very easy,” “fairly easy,” “fairly difficult,” and “very difficult”). Individual data points were preprocessed by dichotomizing item responses, that is, by assigning one point to any “fairly easy” and “very easy” responses and zero points to any “fairly difficult” and “very difficult” responses. Then, a total score was calculated by summing up points across these dichotomized items (possible range 0-16). Finally, participants were grouped according to their overall score, for example, those with inadequate or problematic health literacy (scores 0-12) and those with adequate health literacy (scores 13-16) [47]. The HLS-EU-Q16 questionnaire was originally developed from a more comprehensive questionnaire with 47 items. It has been thoroughly validated and analyzed regarding its reliability in samples of the (adult) general population [45,46]. In addition,

XSL•FO
RenderX

the 16-item version of the questionnaire was found to be both valid and reliable in studies of adolescents [48,49].

Frequency of Alcohol Use and Problematic Alcohol Use

Two relevant statistics were collected regarding the alcohol use of the participants. First, to assess the frequency of alcohol consumption, participants were asked to indicate in a closed-response format with 8 response options (0-7 days) on how many days they consume alcohol in an average week. Second, to assess problematic alcohol use, the German version of the Car-, Relax-, Alone-, Forget-, Friends-, Trouble- (CRAFFT-d) questionnaire [50,51] was administered. The questionnaire's name is an acronym for its 6 items used to assess problematic alcohol use. The original questionnaire was developed as a screening instrument for both alcohol and drug misuse in adolescents [51], but the German CRAFFT-d only assesses alcohol use [50]. In 6 items the participants are asked to indicate whether they have ever ridden in a car with a driver (including themselves) who had consumed alcohol, whether they ever drink to relax, feel better about themselves or fit in, whether they have ever drunk by themselves, whether they have ever forgotten things while using alcohol, whether they were ever told by friends or family to cut down drinking, and whether they had ever gotten into trouble while drinking [51]. Each item is a simple yes-no question, and an overall score can be calculated as the number of positive answers to these questions. The outcome of the questionnaire is binary: if the overall score of an individual is equivalent to or exceeds 2, these individuals are classified as having potentially problematic alcohol consumption; a score lower than 2 indicates no problematic alcohol consumption [50-53]. The German CRAFFT-d has been validated with a sensitivity of 88.8% and a specificity of 66.2% [53].

Statistical Analysis

Data preprocessing and all statistical analyses were conducted with the statistical software SPSS (version 29.0.2.0; IBM Corp). All inferential statistical analyses are reported for the weighted data (Data Weighting for details on the weighting procedures). Internal consistencies of the CRAFFT-d and HLS-EU-Q16 were calculated as Cronbach α. To test for significant associations between the sociodemographic and psychosocial factors and problematic alcohol consumption according to the CRAFFT-d screening criterion, chi-square tests of independence were calculated with Cramer *V* as a measure of effect size. Cramer *V* assesses the strength of the association between 2 nominal variables, for example, when analyzing contingency tables. It can be interpreted according to the following rule of thumb: 0.10 small, 0.30 medium, and 0.50 large [54].

Ethical Considerations

The ethics committee of the Berlin Medical Association did not raise any ethical or professional objections to the study protocol (reference number Eth-SB-24-047). Informed consent was obtained from all participants before data collection was initiated. For all participants aged 15 and younger, informed consent was also provided by a parent or legal guardian. The market research institute provided only anonymized data to the Stiftung Gesundheitswissen. Confidentiality was maintained throughout the study. Participants were not directly compensated by the Stiftung Gesundheitswissen. They received panel-specific compensation or credits that can be redeemed for bank transfers, vouchers, or raffle entries. In face-to-face settings, compensation is handled individually by the interviewers and typically involves direct cash payments or small gifts. This study was part of a larger study that assessed several different health-related constructs and behaviors (eg, anxiety and eating habits) with a specific focus on health literacy in 2 independent samples of the adult general population and adolescents in a cross-sectional study design. Further analyses focusing on other thematic areas are ongoing and are expected to result in additional publications.

Results

Sample Characteristics and Descriptive Statistics

A total of 2006 individuals aged 12-17 years (mean 14.47, SD 1.70 years) participated in the study. Detailed descriptive statistics of the sample across the relevant sociodemographic and psychosocial categories before and after the weighting process are presented in Table 1.

XSL•FO
RenderX

Table 1. Sample characteristics of weighted and unweighted data from a cross-sectional survey (2024-2025) of N=2006 adolescents (aged 12-17 years) in Germany.

Variables[a]	Unweighted sample, n (%)	Weighted sample, n (%)
Gender		
Boys	1025 (51.1)	1035 (51.6)
Girls	975 (48.6)	964 (48.1)
Diverse	6 (0.3)	6 (0.3)
Age groups (years)		
12-13	657 (32.8)	676 (33.7)
14-15	688 (34.3)	681 (34)
16-17	661 (33)	649 (32.4)
Level of education		
Low	1167 (58.2)	1370 (68.3)
High	839 (41.8)	636 (31.7)
Social status		
Low	248 (12.4)	258 (12.9)
Intermediate	1324 (66)	1329 (66.3)
High	434 (21.6)	419 (20.9)
Migration background		
Yes	292 (14.6)	295 (14.7)
No	1714 (85.4)	1711 (85.3)
Health literacy		
Inadequate or problematic	1180 (58.8)	1203 (60)
Adequate	826 (41.2)	803 (40)

[a]Cumulative percentages and absolute numbers may exceed or fall below 100% due to weighting and rounding.

Reliability of the Measures

Internal consistency of the CRAFFT-d overall score to assess problematic alcohol use was α=.75, and internal consistency of the HLS-EU-Q16 to assess health literacy was α=.86.

Descriptive Findings on Frequency of Alcohol Consumption

On average, participants drank alcohol on a mean of 0.35 (SD 0.89) days per week, with the majority of the participants (1615/2006, 80.5%) reporting not drinking at all across an average week. Details of the frequency of alcohol consumption are presented in Table 2.

Table 2. Frequency of weekly alcohol consumption across the sample of 2006 adolescents (aged 12-17 years) from a cross-sectional survey (2024-2025) in Germany.

How many days per week, on average, do you consume alcohol?[a]	Values, n (%)
0	1615 (80.5)
1	221 (11)
2	91 (4.6)
3	46 (2.3)
4	15 (0.8)
5	7 (0.4)
6	6 (0.3)
7	4 (0.2)

[a]Note: Cumulative percentages and absolute numbers may exceed or fall below 100% due to weighting and rounding.

XSL•FO
RenderX

Across the entire sample, n=255 (12.7%; 95% CI 11.2%-14.2%) individuals reported problematic alcohol consumption according to the CRAFFT-d screening questionnaire, and n=1751 (87.3%) individuals did not. Response frequencies for each CRAFFT-d item are reported in Table 3. The items referring to drinking to relax, forgetting things when drinking, and getting into trouble while drinking were affirmed more frequently (ie, by more than 10% of the sample) than the items on drunk driving, drinking alone, and being told to cut down drinking, which were affirmed by less than 10% of the sample. On average, participants reached a sum score of mean 0.50 (SD 1.11) in the CRAFFT-d questionnaire.

Table 3. Individual item responses for the German version of the Car-, Relax-, Alone-, Forget-, Friends-, Trouble- questionnaire to assess problematic alcohol use across the sample of N=2006 adolescents (aged 12-17 years) from a cross-sectional survey (2024-2025) in Germany.

Item number	Items[a]	Yes, n (%)	No, n (%)
1	Have you ever ridden in a car driven by someone (including yourself) who had been using alcohol?	113 (5.6)	1893 (94.4)
2	Do you ever use alcohol to relax, feel better about yourself, or fit in?	209 (10.4)	1797 (89.6)
3	Do you ever use alcohol when you are by yourself (alone)?	104 (5.2)	1902 (94.8)
4	Do you ever forget things you did while using alcohol?	209 (10.4)	1797 (89.6)
5	Do your family or friends ever tell you that you should cut down on your drinking?	166 (8.3)	1840 (91.7)
6	Have you ever gotten into trouble while you were using alcohol?	205 (10.2)	1801 (89.8)

[a]In contrast to the original English version, the German CRAFFT-d only assesses alcohol consumption and does not refer to drug consumption [50,51]. Therefore, the items are reported with adjusted wording to match the German version, which was presented to the participants in this study. Cumulative percentages and absolute numbers may exceed or fall below 100% due to weighting and rounding.

Prevalences of Problematic Alcohol Use According to the CRAFFT-d Questionnaire and Associations With Sociodemographic and Psychosocial Variables

Table 4 presents the descriptive results regarding rates of problematic and nonproblematic alcohol use according to the CRAFFT-d questionnaire across the sociodemographic and psychosocial categories.

Boys reported significantly more problematic alcohol consumption than girls (χ^2_1=20.96, *V*=0.10; *P*<.001). Participants with diverse genders were not included in this analysis because cell frequencies were too low to allow testing.

Problematic alcohol consumption significantly increased with increasing age (χ^2_2=85.88, *V*=0.21; *P*<.001). Participants aged 12-13 years had the lowest frequency of problematic alcohol consumption (34/676, 5%), and those aged 16-17 years had the highest frequency of problematic alcohol consumption (142/649, 21.9%), with participants aged 14-15 years falling in between these 2 groups (79/681, 11.6%).

Problematic alcohol consumption was not significantly associated with level of education (χ^2_1=3.43, *V*=0.04; *P*=.06), with a descriptively marginally higher frequency in those with a low (187/1370, 13.6%) compared to a higher level of education (68/636, 10.7%).

Participants who reported low subjective social status had the highest frequency of problematic alcohol consumption (46/258, 17.8%), followed by participants with high subjective social status (56/418, 13.4%) and intermediate subjective social status (152/1329, 11.4%). This association was significant (χ^2_2=8.23, *V*=0.06; *P*=.02).

Problematic alcohol consumption was also significantly associated with migration background, with a higher rate of problematic alcohol consumption in individuals with a migration background (50/295, 16.9%) compared to those without a migration background (205/1711, 12%; χ^2_1=5.60, *V*=0.05; *P*=.02).

Problematic alcohol consumption was not significantly associated with levels of health literacy (χ^2_1=1.54, *V*=0.03; *P*=.21). There was a descriptively higher rate of problematic alcohol consumption in individuals with inadequate or problematic health literacy (162/1203, 13.5%) than in those with adequate health literacy (93/803, 11.6%).

Participants who reported more frequent alcohol consumption per week also showed more frequent problematic alcohol consumption (χ^2_7=698.65, *V*=0.59; *P*<.001). Rates of problematic alcohol consumption were lowest in those who reported drinking on average on zero days per week (57/1614, 3.5%), and highest in those who drank every day (4/4, 100%), with those drinking on 2 to 6 days per week falling in between these 2 extremes (Table 4). However, in some cells, total frequencies were very low, which should be kept in mind when interpreting the results.

A summary of the key significant findings is presented in Figure 1.

XSL•FO
RenderX

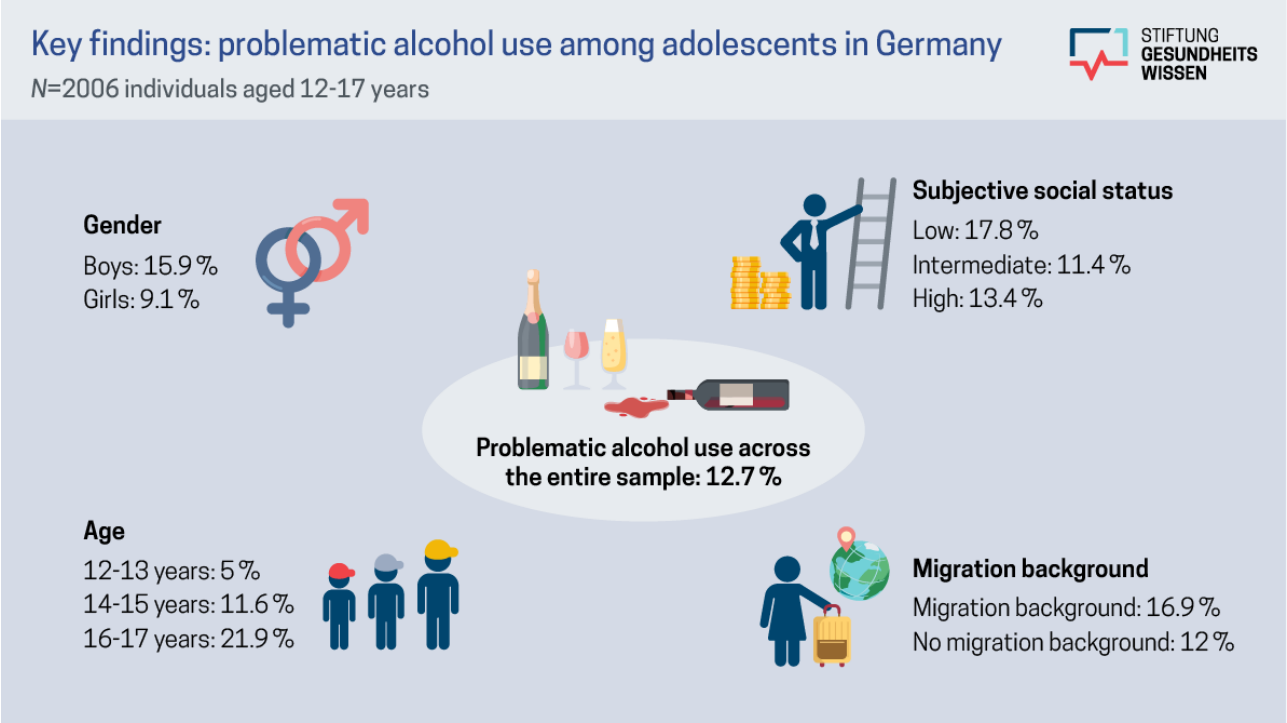
Table 4. Proportions of positive and negative screens obtained with the German version of the Car-, Relax-, Alone-, Forget-, Friends-, Trouble-questionnaire indicating problematic alcohol use across the sociodemographic and psychosocial categories in a sample of N=2006 adolescents (aged 12-17 years) from a cross-sectional survey (2024-2025) in Germany.

Categories[a]	Problematic alcohol use, n (%)	No problematic alcohol use, n (%)
Overall	255 (12.7)	1751 (87.3)
Gender		
Boys (n=1035)	165 (15.9)	870 (84.1)
Girls (n=964)	88 (9.1)	876 (90.9)
Age groups (years)		
12-13 (n=676)	34 (5)	642 (95)
14-15 (n=681)	79 (11.6)	602 (88.4)
16-17 (n=649)	142 (21.9)	507 (78.1)
Level of education		
Low (n=1370)	187 (13.6)	1183 (86.4)
High (n=636)	68 (10.7)	568 (89.3)
Social status		
Low (n=258)	46 (17.8)	212 (82.2)
Intermediate (n=1329)	152 (11.4)	1177 (88.6)
High (n=418)	56 (13.4)	362 (86.6)
Migration background		
Yes (n=295)	50 (16.9)	245 (83.1)
No (n=1711)	205 (12)	1506 (88)
Health literacy		
Inadequate or problematic (n=1203)	162 (13.5)	1041 (86.5)
Adequate (n=803)	93 (11.6)	710 (88.4)
Frequency of weekly alcohol consumption		
Zero days (n=1614)	57 (3.5)	1557 (96.5)
One day (n=221)	88 (39.8)	133 (60.2)
Two days (n=91)	48 (52.7)	43 (47.3)
Three days (n=46)	34 (73.9)	12 (26.1)
Four days (n=15)	13 (86.7)	2 (13.3)
Five days (n=7)	5 (71.4)	2 (28.6)
Six days (n=6)	4 (66.7)	2 (33.3)
Seven days (n=4)	4 (100)	0 (0)

[a]Cumulative percentages and absolute numbers may exceed or fall below 100% due to weighting and rounding.

XSL•FO
RenderX

Figure 1. Visual summary of the key findings regarding problematic alcohol use in a sample of N=2006 adolescents (aged 12-17 years) from a cross-sectional survey (2024-2025) in Germany.



Discussion

Principal Findings

The aim of this study was to assess current data on the rates of problematic alcohol use and the frequency of alcohol use among adolescents aged 12-17 years in Germany, and to explore how problematic alcohol use relates to sociodemographic and psychosocial factors as well as frequency of alcohol use. Approximately 20% (390/2006) of all participants reported drinking alcohol on at least 1 day per week, and 12.7% (255/2006) of the sample met the CRAFFT-d screening criterion for problematic alcohol use. Problematic alcohol use was significantly associated with gender, age, subjective social status, migration background, and frequency of weekly alcohol consumption. However, there were no significant associations between problematic alcohol consumption and level of education or health literacy.

The fact that about 20% (390/2006) of the sample reported alcohol consumption on at least 1 day in an average week is concerning, as dose-response relationships between levels of alcohol use and alcohol use disorder incidence and mortality have been found in previous studies [55]. Consistent with this line of evidence, the frequency of alcohol consumption was associated with problematic alcohol use in our data, with higher rates of problematic use observed in those individuals who consumed alcohol more frequently across an average week. As a recent meta-analysis using longitudinal data suggests that even small doses of alcohol are associated with an increased risk of developing an alcohol use disorder and increased mortality [55], there is a clear public health interest in further reducing alcohol consumption. Regarding the prevalence of problematic alcohol use, our findings obtained with the CRAFFT-d questionnaire (255/2006, 12.7%) are comparable to recent findings that found a rate of 13.6% (using a liberal threshold) or 11.3% (using a more conservative threshold) of problematic alcohol use with a different screening instrument in a sample of more than 4000 adolescents from Germany [28], highlighting the validity of our approach. Moreover, when compared to data from 2015 from slightly older samples (CRAFFT-d baseline means of 0.95, SD 1.19 and 1.10, SD 1.29 in 2 groups), we obtained lower CRAFFT-d mean scores (mean 0.50, SD 1.11). This pattern of results converges with our finding of higher scores in older individuals and with a decline in the frequency of adolescent drinking in recent years [26]. Overall, our findings demonstrate that there is still a substantial proportion of adolescents (255/2006, 12.7%) who show problematic drinking patterns. Examining the CRAFFT-d items individually reveals that approximately 1 in 10 adolescents have experienced trouble because of drinking, have forgotten things due to alcohol use, and have drunk to relax, feel better, or fit in. The other items concerning drunk driving, having been told to cut down on drinking, and drinking alone were affirmed less frequently, but still had prevalence rates exceeding 5% in the sample. These results emphasize the need for continued and intensified evidence-based efforts to reduce alcohol consumption and problematic drinking among adolescents in Germany.

When analyzing the reasons why adolescents drink alcohol, 3 major motives emerge in the literature: social enhancement (eg, to obtain social rewards), coping (eg, to attenuate negative affect), and dominance (eg, to obtain respect) [56]. In addition, on a neurobiological level, adolescents might be particularly sensitive to the rewarding effects of alcohol and less sensitive to its sedative effects [57]. Crucially, adolescence is characterized as a period of rapid changes in various areas of life, including physical, social, hormonal, and mental transitions, during which individuals have to cope with multiple

XSL•FO
RenderX

developmental tasks [58,59]. These transitions also manifest in concrete behaviors, such as tendencies to take more risks, which might be explained by immature self-regulation competencies and changes in personality factors such as sensation seeking and impulsivity. Of note, these factors have been found to influence alcohol use in longitudinal studies [60-62]. Moreover, alcohol expectancies, that is, the individual beliefs about alcohol effects, have been shown to change in young adolescence and to predict alcohol use in that age group. Environmental factors (eg, influences from parents and peers) play an important role in the development of these expectancies [63], and there is first evidence that programs that include peer groups might be effective when tackling alcohol use in adolescents [64]. Taken together, the findings on predictors of alcohol use among adolescents suggest that psycho-educative and other efforts to reduce alcohol use in that age group might be more effective when they involve the social environment of the targeted population and when they particularly target alcohol expectancies and individual motives to drink [56,63,64], but more research is needed to make causal claims on these relationships.

Problematic alcohol use was associated with some—but not all—sociodemographic factors assessed in our study. For example, problematic alcohol use was found to be more prevalent among boys than girls, consistent with the finding that boys generally consume alcohol more frequently than girls [26]. This finding suggests that gender-specific prevention strategies may be a viable option that should be further investigated in the future [65].

Problematic alcohol consumption was more frequent in the older (16-17 years) compared to the intermediate (14-15 years) and the younger age groups (12-13 years). In the oldest adolescents, problematic alcohol use was observed in more than 1 in 5 participants. Although problematic drinking rates were lower in the youngest age group (5%), the reported frequency in these individuals is still concerning, as alcohol consumption is associated with detrimental developmental consequences [3,4]. Moreover, early adolescence is a particularly vulnerable period as drinking patterns in early adolescence predict problematic alcohol consumption later in life [66,67]. Therefore, prevention strategies specifically targeting both young and older adolescents should be further developed and evaluated [68].

Concerning subjective social status, frequency of problematic alcohol use followed a V-shaped pattern, with the most frequent problematic use in those with the lowest and highest subjective social status and the least frequent problematic consumption in those with intermediate subjective social status. However, problematic drinking was not associated with the level of education in our sample. This pattern of results is only partly in line with evidence that demonstrates that general alcohol consumption is typically higher in adolescents with lower compared to higher socioeconomic status [37] and that it is more often harmful (eg, leading to more frequent hospitalizations) in these individuals [38]. However, other studies point to opposite associations and higher alcohol consumption in individuals with higher socioeconomic status [39] or, in the case of binge drinking, do not find any clear association [40]. Increased problematic consumption among adolescents from higher socioeconomic levels might be explained by cultural differences, financial resources to buy alcohol, and the availability of alcohol at home [39], whereas more harmful alcohol consumption in individuals with a lower socioeconomic background might be attributed to a lack of parental support or monitoring alongside increased distress among the adolescents [38,69]. When interpreting these results, it should be kept in mind that the evidence comes from studies in different samples, from different cultural backgrounds, in different age groups, and using different modes to assess social or socioeconomic status, which makes it difficult to compare these findings. Our results add to the literature by demonstrating that among adolescents aged 12-17 years in Germany, both low and high subjective social status are associated with higher problematic alcohol use. Furthermore, the level of education was not associated with problematic alcohol use.

More frequent problematic alcohol use was observed in individuals with a migration background compared to those without a migration background. These findings can be further elucidated with the help of evidence from a recent study assessing both risky alcohol consumption (exceeding hazardous dose thresholds concerning total alcohol content) and binge drinking (exceeding thresholds for the number of drinks consumed on a single occasion) in a sample similar to ours. The study found that there are interindividual differences between individuals from various cultural backgrounds concerning their drinking patterns [26]. Risky use rates were similar in individuals with no migration background and those from Western Europe, Eastern Europe, and Turkey or Asia, but significantly lower in individuals from other regions. In contrast, binge drinking was more prevalent in individuals from Western Europe compared to those with no migration background, but less prevalent (compared to those with no migration background) in individuals from Eastern Europe and other countries, while there was no difference between those with no migration background and individuals from Turkey or Asia. Future studies need to investigate more comprehensively how problematic drinking is associated with specific cultural backgrounds and how this information can be used to develop culturally sensitive prevention strategies.

Interestingly, problematic alcohol use was not associated with general health literacy. This finding is surprising, given that health literacy predicts many other health-related behaviors [31,70] and has been shown to be associated with alcohol-related behaviors in adult populations [33,34]. It is possible that the administered instrument to measure general health literacy was too broad to capture its specific facets that more closely apply to the use of alcohol. In this context, recently, the term “alcohol health literacy,” or “alcohol literacy,” has been introduced and conceptualized. While its precise definition is still subject to ongoing debate, it appears to closely relate to general health literacy and more specifically focuses on the “capacity to obtain, process, and understand knowledge about alcohol content, units, strengths, and harms” [35,36]. First attempts to measure alcohol health literacy have been made in adults, but to date, there is no data on levels in adolescents, which should be addressed in future research [71]. Recently, new recommendations have been published to increase alcohol health literacy and reduce alcohol

XSL•FO
RenderX

consumption in Germany [35]. These recommendations include action on the levels of education and information, in the health care system, and concerning alcohol control policy. One important pillar concerning education and information is the implementation of effective alcohol prevention programs in schools and the provision of easily accessible information about alcohol, especially for adolescents and young adults [35,72].

Limitations

When interpreting the results of this study, it is important to bear in mind the following limitations. First, as noted above, we used a measure of general health literacy that might not be specific enough to measure relevant aspects of alcohol-related health literacy in the present sample [35,36]. Second, this was a cross-sectional study, which means that the observed associations cannot be interpreted as causal relationships [73]. Third, problematic alcohol use was assessed with a short screening questionnaire, which does not replace a formal clinical diagnosis for alcohol misuse or dependence. However, importantly, the questionnaire has been validated thoroughly with high sensitivity and specificity scores and therefore provides good estimates for problematic drinking [51-53,74]. Fourth, problematic alcohol use was assessed with a self-report questionnaire. As there is some level of stigma concerning alcohol use, especially among minors who are not legally allowed to drink, there is always a risk of bias due to social desirability and impression management even in anonymous surveys [75]. This might have led to an underestimation of the actual prevalence of problematic drinking in our sample [75]. Fifth, we only collected data on the frequency of drinking and on potential problematic drinking, but not on drinking quantity, which would have allowed us to conduct more in-depth analyses. Finally, the proportion of individuals with a migration background was lower than the most recent census data from Germany indicates. This discrepancy might be attributed to insufficient knowledge of the German language and the more challenging recruitment of this population segment. Therefore, our results concerning migration background should be interpreted cautiously.

Conclusions

This study presents data on problematic alcohol consumption among adolescents in Germany in addition to examining its associations with various psychosocial and sociodemographic factors. The findings show that problematic alcohol use is more prevalent among boys than girls, among older age groups than younger ones, among individuals with a higher or lower subjective social status than those with an intermediate subjective social status, among participants with a migration background than those without, and among those who consume alcohol more frequently on a weekly basis. However, there were no significant associations between problematic alcohol use and levels of education or health literacy. These results underscore the importance of targeted prevention strategies that address the specific risk profiles of adolescents. By tailoring interventions to individuals with a higher risk, policymakers might be able to more successfully mitigate problematic alcohol use and promote healthier life choices.

Acknowledgments

The German market research institute Gesellschaft für Innovative Marktforschung mbH was responsible for the recruitment of study participants, study administration, and data acquisition.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to copyright restrictions, but are available from the corresponding author on reasonable request.

Funding

This research was funded by the independent, nonprofit foundation Stiftung Gesundheitswissen.

Authors' Contributions

Conceptualization, formal analysis, methodology, writing – original draft: R Schröder
Conceptualization, validation, writing, review, and editing: TH
Writing – review and editing: R Suhr
Conceptualization, supervision, methodology, writing – review and editing: LK

Conflicts of Interest

All authors are employees of the independent, nonprofit foundation Stiftung Gesundheitswissen.

References

1. Global status report on alcohol and health and treatment of substance use disorders. World Health Organization. 2024. URL: https://www.who.int/publications/i/item/9789240096745 [accessed 2026-01-21]
2. Shield K, Manthey J, Rylett M, Probst C, Wettlaufer A, Parry CDH, et al. National, regional, and global burdens of disease from 2000 to 2016 attributable to alcohol use: a comparative risk assessment study. Lancet Public Health 2020;5(1):e51-e61 [FREE Full text] [doi: 10.1016/S2468-2667(19)30231-2] [Medline: 31910980]

XSL•FO
RenderX

3. Zeigler DW, Wang CC, Yoast RA, Dickinson BD, McCaffree MA, Robinowitz CB, Council on Scientific Affairs, American Medical Association. The neurocognitive effects of alcohol on adolescents and college students. Prev Med 2005 Jan;40(1):23-32. [doi: 10.1016/j.ypmed.2004.04.044] [Medline: 15530577]
4. Jacobus J, Tapert SF. Neurotoxic effects of alcohol in adolescence. Annu Rev Clin Psychol 2013;9:703-721 [FREE Full text] [doi: 10.1146/annurev-clinpsy-050212-185610] [Medline: 23245341]
5. Marshall EJ. Adolescent alcohol use: risks and consequences. Alcohol Alcohol 2014;49(2):160-164. [doi: 10.1093/alcalc/agt180] [Medline: 24402246]
6. Manthey J, Shield KD, Rylett M, Hasan OSM, Probst C, Rehm J. Global alcohol exposure between 1990 and 2017 and forecasts until 2030: a modelling study. Lancet 2019;393(10190):2493-2502. [doi: 10.1016/S0140-6736(18)32744-2] [Medline: 31076174]
7. Peacock A, Leung J, Larney S, Colledge S, Hickman M, Rehm J, et al. Global statistics on alcohol, tobacco and illicit drug use: 2017 status report. Addiction 2018;113(10):1905-1926. [doi: 10.1111/add.14234] [Medline: 29749059]
8. Ronksley PE, Brien SE, Turner BJ, Mukamal KJ, Ghali WA. Association of alcohol consumption with selected cardiovascular disease outcomes: a systematic review and meta-analysis. BMJ 2011;342:d671 [FREE Full text] [doi: 10.1136/bmj.d671] [Medline: 21343207]
9. Zhong L, Chen W, Wang T, Zeng Q, Lai L, Lai J, et al. Alcohol and health outcomes: an umbrella review of meta-analyses base on prospective cohort studies. Front Public Health 2022;10:859947 [FREE Full text] [doi: 10.3389/fpubh.2022.859947] [Medline: 35602135]
10. Wood AM, Kaptoge S, Butterworth AS, Willeit P, Warnakula S, Bolton T, et al. Risk thresholds for alcohol consumption: combined analysis of individual-participant data for 599 912 current drinkers in 83 prospective studies. Lancet 2018;391(10129):1513-1523 [FREE Full text] [doi: 10.1016/S0140-6736(18)30134-X] [Medline: 29676281]
11. Richter M, Tauer J, Conrad J, Heil E, Kroke A, Virmani K, et al. Alcohol consumption in Germany, health and social consequences and derivation of recommendations for action - position statement of the German Nutrition Society (DGE). Ernährungs Umschau 2024;71(10):125-139. [doi: 10.4455/eu.2024.033]
12. Stockwell T, Zhao J, Panwar S, Roemer A, Naimi T, Chikritzhs T. Do "Moderate" drinkers have reduced mortality risk? A systematic review and meta-analysis of alcohol consumption and all-cause mortality. J Stud Alcohol Drugs 2016;77(2):185-198 [FREE Full text] [doi: 10.15288/jsad.2016.77.185] [Medline: 26997174]
13. Zhao J, Stockwell T, Naimi T, Churchill S, Clay J, Sherk A. Association between daily alcohol intake and risk of all-cause mortality: a systematic review and meta-analyses. JAMA Netw Open 2023;6(3):e236185 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.6185] [Medline: 37000449]
14. Hendriks HFJ. Alcohol and human health: what is the evidence? Annu Rev Food Sci Technol 2020;11:1-21. [doi: 10.1146/annurev-food-032519-051827] [Medline: 32209032]
15. Anderson BO, Berdzuli N, Ilbawi A, Kestel D, Kluge HP, Krech R, et al. Health and cancer risks associated with low levels of alcohol consumption. Lancet Public Health 2023;8(1):e6-e7 [FREE Full text] [doi: 10.1016/S2468-2667(22)00317-6] [Medline: 36603913]
16. Rehm J, Gmel GE, Gmel G, Hasan OSM, Imtiaz S, Popova S, et al. The relationship between different dimensions of alcohol use and the burden of disease-an update. Addiction 2017;112(6):968-1001 [FREE Full text] [doi: 10.1111/add.13757] [Medline: 28220587]
17. Zoethout RWM, Delgado WL, Ippel AE, Dahan A, van Gerven JMA. Functional biomarkers for the acute effects of alcohol on the central nervous system in healthy volunteers. Br J Clin Pharmacol 2011;71(3):331-350 [FREE Full text] [doi: 10.1111/j.1365-2125.2010.03846.x] [Medline: 21284693]
18. Dougherty DM, Marsh-Richard DM, Hatzis ES, Nouvion SO, Mathias CW. A test of alcohol dose effects on multiple behavioral measures of impulsivity. Drug Alcohol Depend 2008;96(1-2):111-120 [FREE Full text] [doi: 10.1016/j.drugalcdep.2008.02.002] [Medline: 18378098]
19. Hingson RW, Zha W, Weitzman ER. Magnitude of and trends in alcohol-related mortality and morbidity among U.S. college students ages 18-24, 1998-2005. J Stud Alcohol Drugs Suppl 2009(16):12-20 [FREE Full text] [doi: 10.15288/jsads.2009.s16.12] [Medline: 19538908]
20. Atzendorf J, Rauschert C, Seitz N, Lochbühler K, Kraus L. The use of alcohol, tobacco, illegal drugs and medicines: an estimate of consumption and substance-related disorders in Germany. Dtsch Arztebl Int 2019;116(35-36):577-584 [FREE Full text] [doi: 10.3238/arztebl.2019.0577] [Medline: 31587705]
21. Odlaug BL, Gual A, DeCourcy J, Perry R, Pike J, Heron L, et al. Alcohol dependence, co-occurring conditions and attributable burden. Alcohol Alcohol 2016;51(2):201-209 [FREE Full text] [doi: 10.1093/alcalc/agv088] [Medline: 26246514]
22. Levola J, Aalto M, Holopainen A, Cieza A, Pitkänen T. Health-related quality of life in alcohol dependence: a systematic literature review with a specific focus on the role of depression and other psychopathology. Nord J Psychiatry 2014;68(6):369-384. [doi: 10.3109/08039488.2013.852242] [Medline: 24228776]
23. Manthey J, Hassan SA, Carr S, Kilian C, Kuitunen-Paul S, Rehm J. What are the economic costs to society attributable to alcohol use? A systematic review and modelling study. Pharmacoeconomics 2021;39(7):809-822 [FREE Full text] [doi: 10.1007/s40273-021-01031-8] [Medline: 33970445]

XSL•FO
RenderX

24. Making the WHO European region SAFER: developments in alcohol control policies, 2010-2019. World Health Organization. 2021. URL: https://www.who.int/europe/publications/i/item/9789289055048 [accessed 2026-01-21]
25. Seitz N, Lochbühler K, Atzendorf J, Rauschert C, Pfeiffer-Gerschel T, Kraus L. Trends in substance use and related disorders: analysis of the epidemiological survey of substance abuse 1995 to 2018. Dtsch Arztebl Int 2019;116(35-36):585-591 [FREE Full text] [doi: 10.3238/arztebl.2019.0585] [Medline: 31587706]
26. Orth B, Spille-Merkel C, Nitzsche A. Die Drogenaffinität Jugendlicher in der Bundesrepublik Deutschland 2023: Substanzkonsum und Internetnutzung im Jahr 2023 und Trends [Report in German]. BZgA-Forschungsbericht. 2025. URL: https://www.bioeg.de/fileadmin/user_upload/Studien/PDF/DAS_2023_Forschungsbericht_final.pdf [accessed 2026-01-21]
27. Silins E, Horwood LJ, Najman JM, Patton GC, Toumbourou JW, Olsson CA, et al. Adverse adult consequences of different alcohol use patterns in adolescence: an integrative analysis of data to age 30 years from four Australasian cohorts. Addiction 2018;113(10):1811-1825. [doi: 10.1111/add.14263] [Medline: 29749666]
28. Wartberg L, Belau M, Arnaud N, Thomasius R, IMAC Mind Consortium. Problematic consumption of alcohol, cannabis and cigarettes—a German nationwide survey on psychopathology, stress, mindfulness, and quality of life. Dtsch Arztebl Int 2024;121(13):422-427 [FREE Full text] [doi: 10.3238/arztebl.m2024.0061] [Medline: 38657177]
29. Sørensen K, Van den Broucke S, Fullam J, Doyle G, Pelikan J, Slonska Z, et al. Health literacy and public health: a systematic review and integration of definitions and models. BMC Public Health 2012;12:80 [FREE Full text] [doi: 10.1186/1471-2458-12-80] [Medline: 22276600]
30. Buja A, Rabensteiner A, Sperotto M, Grotto G, Bertoncello C, Cocchio S, et al. Health literacy and physical activity: a systematic review. J Phys Act Health 2020;17(12):1259-1274 [FREE Full text] [doi: 10.1123/jpah.2020-0161] [Medline: 33129198]
31. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. Ann Intern Med 2011;155(2):97-107. [doi: 10.7326/0003-4819-155-2-201107190-00005] [Medline: 21768583]
32. Paasche-Orlow MK, Wolf MS. The causal pathways linking health literacy to health outcomes. Am J Health Behav 2007;31(1):19-26. [doi: 10.5993/ajhb.31.s1.4]
33. Zanobini P, Parsa N, Bonaccorsi G, Minardi V, Possenti V, Masocco M, et al. Exploring the relationship between health literacy and alcohol consumption patterns in Tuscany. Eur J Public Health 2024;34(Supplement_3). [doi: 10.1093/eurpub/ckae144.1711]
34. Gould DA, Lubin RE, McGrew SJ, Smit T, Vujanovic AA, Otto MW, et al. The role of functional health literacy in terms of hazardous alcohol use in adults with probable posttraumatic stress disorder and alcohol use disorder. J Dual Diagn 2025;21(1):49-55. [doi: 10.1080/15504263.2024.2433775] [Medline: 39635875]
35. Manthey J, Kokole D, Riedel-Heller S, Rowlands G, Schäfer I, Schomerus G, et al. Improving alcohol health literacy and reducing alcohol consumption: recommendations for Germany. Addict Sci Clin Pract 2023;18(1):28 [FREE Full text] [doi: 10.1186/s13722-023-00383-0] [Medline: 37161561]
36. Okan O, Rowlands G, Sykes S, Wills J. Shaping alcohol health literacy: a systematic concept analysis and review. Health Lit Res Pract 2020;4(1):e3-e20 [FREE Full text] [doi: 10.3928/24748307-20191104-01] [Medline: 31935296]
37. Lemstra M, Bennett NR, Neudorf C, Kunst A, Nannapaneni U, Warren LM, et al. A meta-analysis of marijuana and alcohol use by socio-economic status in adolescents aged 10-15 years. Can J Public Health 2008;99(3):172-177 [FREE Full text] [doi: 10.1007/BF03405467] [Medline: 18615935]
38. Tolstrup JS, Kruckow S, Becker U, Andersen O, Sawyer SM, Katikireddi SV, et al. Socioeconomic inequalities in alcohol-related harm in adolescents: a prospective cohort study of 68,299 Danish 15-19-year-olds. EClinicalMedicine 2023;62:102129 [FREE Full text] [doi: 10.1016/j.eclinm.2023.102129] [Medline: 37576460]
39. Gautam N, Dessie G, Rahman MM, Khanam R. Socioeconomic status and health behavior in children and adolescents: a systematic literature review. Front Public Health 2023;11:1228632 [FREE Full text] [doi: 10.3389/fpubh.2023.1228632] [Medline: 37915814]
40. Kwok KHR, Yuan SNV. Parental socioeconomic status and binge drinking in adolescents: a systematic review. Am J Addict 2016;25(8):610-619. [doi: 10.1111/ajad.12461] [Medline: 27749983]
41. Faul F, Erdfelder E, Lang A, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007;39(2):175-191. [doi: 10.3758/bf03193146] [Medline: 17695343]
42. Hoebel J, Müters S, Kuntz B, Lange C, Lampert T. [Measuring subjective social status in health research with a German version of the MacArthur scale]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2015;58(7):749-757. [doi: 10.1007/s00103-015-2166-x] [Medline: 25986532]
43. Adler NE, Epel ES, Castellazzo G, Ickovics JR. Relationship of subjective and objective social status with psychological and physiological functioning: preliminary data in healthy white women. Health Psychol 2000;19(6):586-592. [doi: 10.1037//0278-6133.19.6.586] [Medline: 11129362]
44. Jordan S, Hoebel J. [Health literacy of adults in Germany: findings from the German health update (GEDA) study]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2015;58(9):942-950. [doi: 10.1007/s00103-015-2200-z] [Medline: 26227894]

XSL•FO
RenderX

45. Pelikan JM, Röthlin F, Ganahl K. Measuring comprehensive health literacy in general populations: validation of instrument, indices and scales of the HLS-EU study. 2014 Presented at: Proceedings of the 6th Annual Health Literacy Research Conference; November 3-4, 2014; Bethesda, Maryland, USA URL: https://www.bumc.bu.edu/healthliteracyconference/files/2014/06/Pelikan-et-al-HARC-2014-fin.pdf
46. Sørensen K, Van den Broucke S, Pelikan JM, Fullam J, Doyle G, Slonska Z, et al. Measuring health literacy in populations: illuminating the design and development process of the European health literacy survey questionnaire (HLS-EU-Q). BMC Public Health 2013;13:948 [FREE Full text] [doi: 10.1186/1471-2458-13-948] [Medline: 24112855]
47. Pelikan JM, Ganahl K, van DBS, Sørensen K. Measuring health literacy in Europe: introducing the European health literacy survey questionnaire (HLS-EU-Q). In: Bauer U, Levin-Zamir D, Okan O, Pinheiro P, Sørensen K, editors. International Handbook of Health Literacy: Research, Practice and Policy Across the Life-Span. Bristol, UK: Policy Press; 2019:115-138.
48. König L, Schröder R, Hamer T, Suhr R. Depression and health literacy among adolescents and adults in Germany: findings from two representative samples. Front Psychol 2024;15:1494333 [FREE Full text] [doi: 10.3389/fpsyg.2024.1494333] [Medline: 39726619]
49. Röthlin F, Pelikan J, Ganahl K. Die Gesundheitskompetenz von 15-jährigen Jugendlichen in Österreich: Abschlussbericht der Österreichischen Gesundheitskompetenz Jugendstudie im Auftrag des Hauptverbands der Österreichischen Sozialversicherungsträger (HVSV) [Report in German]. Ludwig Boltzmann Institut Health Promotion Research. 2013. URL: https://www.sozialversicherung.at/cdscontent/load?contentid=10008.715507 [accessed 2026-01-21]
50. Tossmann P, Kasten L, Lang P, Strüber E. CRAFFT-d: CRAFFT – deutsche Fassung [Web page in German]. Leibniz-Institut für Psychologie (ZPID). 2011. URL: https://www.testarchiv.eu/en/test/9006232 [accessed 2026-01-21]
51. Knight JR, Shrier LA, Bravender TD, Farrell M, Vander Bilt J, Shaffer HJ. A new brief screen for adolescent substance abuse. Arch Pediatr Adolesc Med 1999;153(6):591-596. [doi: 10.1001/archpedi.153.6.591] [Medline: 10357299]
52. Mitchell SG, Kelly SM, Gryczynski J, Myers CP, O'Grady KE, Kirk AS, et al. The CRAFFT cut-points and DSM-5 criteria for alcohol and other drugs: a reevaluation and reexamination. Subst Abus 2014;35(4):376-380 [FREE Full text] [doi: 10.1080/08897077.2014.936992] [Medline: 25036144]
53. Tossmann P, Kasten L, Lang P, Strüber E. [Determination of the concurrent validity of the CRAFT-d - a screening instrument for problematic alcohol consumption]. Z Kinder Jugendpsychiatr Psychother 2009;37(5):451-459. [doi: 10.1024/1422-4917.37.5.451] [Medline: 19739063]
54. Ellis PD. The Essential Guide to Effect Sizes : Statistical Power, Meta-Analysis, and the Interpretation of Research Results. England: Cambridge University Press; 2010.
55. Carr T, Kilian C, Llamosas-Falcón L, Zhu Y, Lasserre AM, Puka K, et al. The risk relationships between alcohol consumption, alcohol use disorder and alcohol use disorder mortality: a systematic review and meta-analysis. Addiction 2024;119(7):1174-1187 [FREE Full text] [doi: 10.1111/add.16456] [Medline: 38450868]
56. Comasco E, Berglund K, Oreland L, Nilsson KW. Why do adolescents drink? Motivational patterns related to alcohol consumption and alcohol-related problems. Subst Use Misuse 2010;45(10):1589-1604. [doi: 10.3109/10826081003690159] [Medline: 20590377]
57. Chung T, Creswell KG, Bachrach R, Clark DB, Martin CS. Adolescent binge drinking. Alcohol Res 2018;39(1):5-15 [FREE Full text] [doi: 10.35946/arcr.v39.1.02] [Medline: 30557142]
58. Sawyer SM, Azzopardi PS, Wickremarathne D, Patton GC. The age of adolescence. Lancet Child Adolesc Health 2018;2(3):223-228. [doi: 10.1016/S2352-4642(18)30022-1] [Medline: 30169257]
59. Blakemore S. Adolescence and mental health. Lancet 2019;393(10185):2030-2031. [doi: 10.1016/S0140-6736(19)31013-X] [Medline: 31106741]
60. Steinberg L. Risk taking in adolescence: what changes, and why? Ann N Y Acad Sci 2004;1021:51-58. [doi: 10.1196/annals.1308.005] [Medline: 15251873]
61. Quinn PD, Harden KP. Differential changes in impulsivity and sensation seeking and the escalation of substance use from adolescence to early adulthood. Dev Psychopathol 2013;25(1):223-239 [FREE Full text] [doi: 10.1017/S0954579412000284] [Medline: 22824055]
62. Harden KP, Tucker-Drob EM. Individual differences in the development of sensation seeking and impulsivity during adolescence: further evidence for a dual systems model. Dev Psychol 2011;47(3):739-746. [doi: 10.1037/a0023279] [Medline: 21534657]
63. Smit K, Voogt C, Hiemstra M, Kleinjan M, Otten R, Kuntsche E. Development of alcohol expectancies and early alcohol use in children and adolescents: a systematic review. Clin Psychol Rev 2018;60:136-146. [doi: 10.1016/j.cpr.2018.02.002] [Medline: 29449029]
64. Georgie JM, Sean H, Deborah MC, Matthew H, Rona C. Peer-led interventions to prevent tobacco, alcohol and/or drug use among young people aged 11-21 years: a systematic review and meta-analysis. Addiction 2016;111(3):391-407 [FREE Full text] [doi: 10.1111/add.13224] [Medline: 26518976]
65. Gilmore AK, Leone RM, Oesterle DW, Davis KC, Orchowski LM, Ramakrishnan V, et al. Web-based alcohol and sexual assault prevention program with tailored content based on gender and sexual orientation: preliminary outcomes and usability study of positive change (+change). JMIR Form Res 2022;6(7):e23823 [FREE Full text] [doi: 10.2196/23823] [Medline: 35867393]

XSL•FO
RenderX

66. Heron J, Macleod J, Munafò MR, Melotti R, Lewis G, Tilling K, et al. Patterns of alcohol use in early adolescence predict problem use at age 16. Alcohol Alcohol 2012;47(2):169-177 [FREE Full text] [doi: 10.1093/alcalc/agr156] [Medline: 22215001]
67. Guttmannova K, Bailey JA, Hill KG, Lee JO, Hawkins JD, Woods ML, et al. Sensitive periods for adolescent alcohol use initiation: predicting the lifetime occurrence and chronicity of alcohol problems in adulthood. J Stud Alcohol Drugs 2011;72(2):221-231 [FREE Full text] [doi: 10.15288/jsad.2011.72.221] [Medline: 21388595]
68. Stockings E, Hall WD, Lynskey M, Morley KI, Reavley N, Strang J, et al. Prevention, early intervention, harm reduction, and treatment of substance use in young people. Lancet Psychiatry 2016;3(3):280-296. [doi: 10.1016/S2215-0366(16)00002-X] [Medline: 26905481]
69. Hughes K, Bellis MA, Hardcastle KA, Sethi D, Butchart A, Mikton C, et al. The effect of multiple adverse childhood experiences on health: a systematic review and meta-analysis. Lancet Public Health 2017;2(8):e356-e366 [FREE Full text] [doi: 10.1016/S2468-2667(17)30118-4] [Medline: 29253477]
70. König L, Kuhlmey A, Suhr R. Digital health literacy of the population in Germany and its association with physical health, mental health, life satisfaction, and health behaviors: nationally representative survey study. JMIR Public Health Surveill 2024;10:e48685 [FREE Full text] [doi: 10.2196/48685] [Medline: 38381497]
71. Kilian C, Liebig M, Manthey J. Education-based differences in alcohol health literacy in Germany. Drug Alcohol Rev 2025;44(2):555-562. [doi: 10.1111/dar.13985] [Medline: 39626769]
72. Hanewinkel R, Tomczyk S, Goecke M, Isensee B. Preventing binge drinking in adolescents. Dtsch Arztebl Int 2017 May 21;114(16):280-287 [FREE Full text] [doi: 10.3238/arztebl.2017.0280] [Medline: 28502313]
73. Wang X, Cheng Z. Cross-sectional studies: strengths, weaknesses, and recommendations. Chest 2020;158(1S):S65-S71. [doi: 10.1016/j.chest.2020.03.012] [Medline: 32658654]
74. Knight JR, Sherritt L, Shrier LA, Harris SK, Chang G. Validity of the CRAFFT substance abuse screening test among adolescent clinic patients. Arch Pediatr Adolesc Med 2002;156(6):607-614. [doi: 10.1001/archpedi.156.6.607] [Medline: 12038895]
75. Davis CG, Thake J, Vilhena N. Social desirability biases in self-reported alcohol consumption and harms. Addict Behav 2010;35(4):302-311. [doi: 10.1016/j.addbeh.2009.11.001] [Medline: 19932936]

Abbreviations

CRAFFT-d: German version of Car-, Relax-, Alone-, Forget-, Friends-, Trouble- questionnaire to assess problematic alcohol use
HLS-EU-Q16: European Health Literacy Survey instrument

Edited by A Mavragani; submitted 14.Nov.2025; peer-reviewed by L Hoffmann, A Poulton; comments to author 02.Dec.2025; revised version received 14.Jan.2026; accepted 19.Jan.2026; published 06.Feb.2026.

Please cite as:
Schröder R, Hamer T, Suhr R, König L
Problematic Alcohol Use Among Adolescents in Germany: Representative Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e87800
URL: https://publichealth.jmir.org/2026/1/e87800
doi:10.2196/87800
PMID:41650444

©Rebekka Schröder, Tim Hamer, Ralf Suhr, Lars König. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 06.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Association of Vitamin D and Weight Status With Neurodevelopmental Outcomes in a Large Pediatric Population: Cross-Sectional Study

Ying Shen[1*], BSc, MSc, MD; Meiying Gao[1*], BSc, MSc; Lejing Guan[1*], BSc, MSc; Lidan Sun[1], BA, MA; Yanfen Xu[1], BSc; Bethel Daniel[2], HS; Kyla Zhang[3], HS; Guannan Bai[4], BSc, MSc, PhD

[1]Children's Health Management Center, Children's Hospital, Zhejiang University School of Medicine, National Clinical Research Center for Children and Adolescents' Health and Diseases, Hangzhou, China

[2]School of Public Health, Brown University, Providence, RI, United States

[3]Warren Alpert Medical School, Brown University, Providence, RI, United States

[4]Department of Child Health Care, Children's Hospital, Zhejiang University School of Medicine, National Clinical Research Center for Children and Adolescents' Health and Diseases, Hangzhou, Zhejiang, China

[*]these authors contributed equally

Corresponding Author:
Guannan Bai, BSc, MSc, PhD
Department of Child Health Care
Children's Hospital
Zhejiang University School of Medicine, National Clinical Research Center for Children and Adolescents' Health and Diseases
3333 Binsheng Rd
Hangzhou, Zhejiang, 310052
China
Phone: 86 0571 8173 2404
Email: guannanbai@zju.edu.cn

Abstract

Background: Vitamin D and weight status are key determinants for neurodevelopmental outcomes in children. However, their independent and interactive associations with neurodevelopmental outcomes represent a critical yet insufficiently explored issue in pediatric health.

Objective: This study aims to evaluate the independent associations of vitamin D status and weight status with neurodevelopmental outcomes in a large pediatric cohort and to explore how these relationships are stratified by age and sex.

Methods: This retrospective study analyzed the electronic health records of 19,283 healthy children who visited a health management center in a tertiary children's hospital. Neurodevelopmental outcomes were assessed using the Ages and Stages Questionnaire, Third Edition (ASQ-3) (<6 years) and Conners Parent Rating Scale (CPRS) (≥6 years). Logistic regression analyses were applied to assess the independent and interactive impacts of vitamin D and weight status on neurodevelopmental outcomes.

Results: The study included 19,283 children, 10,065 (52.2%) of whom were under 6 years old and 9218 (47.8%) who were between 6 and 18 years old. Among them, 1275 (12.67%) children under 6 were at risk for neurodevelopmental delay, and 1483 (16.09%) children aged between 6 and 18 years had behavioral challenges. Multivariable logistic regression revealed that vitamin D insufficiency/deficiency was significantly associated with suboptimal neurodevelopment in both the younger (odds ratio [OR] 1.48, 95% CI 1.23-1.79; $P<.001$) and older cohorts (OR 1.56, 95% CI 1.39-1.74; $P<.001$. A subgroup analysis revealed that vitamin D insufficiency/deficiency significantly increased the risk of neurodevelopmental delay among normal-weight girls under 6 years (OR 2.24, 95% CI 1.60-3.15; $P<.001$; $q<.006$) after the Benjamini-Hochberg false discovery rate correction. In the 6-to-18-year-old cohort, similar robust associations with behavioral challenges were observed in normal-weight girls (OR 1.61, 95% CI 1.30-2.00; $P<.001$; $q<.006$) and boys (OR 1.64, 95% CI 1.36-1.96; $P<.001$; $q<.006$).

Conclusions: Vitamin D insufficiency/deficiency is associated with suboptimal neurodevelopmental outcomes in a sex-dependent manner modulated by weight status, challenging "one-size-fits-all" clinical models. We advocate for stratified, sex-specific screening and personalized pediatric care strategies.

(JMIR Public Health Surveill 2026;12:e89756) doi:10.2196/89756

XSL•FO
RenderX

KEYWORDS

vitamin D; weight status; developmental outcomes; children; pediatrics

Introduction

Optimal neurodevelopment during childhood and adolescence is fundamental for lifelong cognitive function, emotional regulation, and social adaptation. However, this critical developmental window is increasingly threatened by two concurrent global public health challenges: the escalating prevalence of pediatric obesity and widespread vitamin D insufficiency [1,2]. The rising rates of childhood overweight and obesity, observed in both limited-income and high-income countries, including China, are no longer viewed solely as a risk for metabolic disease but also as a potential detriment to neurological health [3]. Concurrently, vitamin D, once primarily associated with skeletal homeostasis, is now recognized for its vital extraskeletal role, with a high prevalence of insufficiency and deficiency reported across diverse pediatric populations, independent of their nutritional status [4]. The potential intersection of these two prevalent conditions presents a significant and underinvestigated area of concern for pediatric health.

A substantial body of evidence has independently linked both adiposity and vitamin D status to neurodevelopmental outcomes [5-7]. Overweight and obesity in children are associated with chronic low-grade systemic inflammation and metabolic dysregulation, which can create a neuroinflammatory microenvironment detrimental to neural circuit maturation and synaptic plasticity [8]. On a parallel track, the biological plausibility for vitamin D's role in brain health is compelling. The vitamin D receptor and its activating enzymes are widely expressed throughout the developing brain, including in neurocritical regions like the hippocampus and prefrontal cortex [9]. As a potent neurosteroid, vitamin D modulates the expression of neurotrophic factors, regulates neurotransmitter synthesis, and exerts neuroprotective anti-inflammatory effects, making it indispensable for processes such as neuronal proliferation and synaptogenesis [9,10]. Clinical observations corroborate these mechanisms, with studies reporting lower serum vitamin D levels in children with neurodevelopmental disorders such as autism spectrum disorder (ASD) and attention-deficit/hyperactivity disorder (ADHD) compared to their neurotypical peers [11].

Despite the evidence establishing these separate associations, a critical knowledge gap exists in understanding their combined and interactive effects. Few large-scale epidemiological studies have simultaneously investigated vitamin D status, weight status, and neurodevelopment within a single pediatric cohort. This is a significant oversight, particularly given the intricate physiological relationship between these two factors; adiposity is a well-established risk factor for vitamin D deficiency, primarily through sequestration of this fat-soluble vitamin in adipose tissue, thereby reducing its systemic bioavailability [12]. Furthermore, much of the existing research is based on specific clinical populations (eg, children already diagnosed with neurodevelopmental disorders) or smaller samples, and there is a particular paucity of data from large, community-based cohorts of children spanning the full developmental spectrum from infancy to adolescence.

Given these research gaps, a comprehensive investigation in a large, diverse pediatric sample is warranted to investigate the complex interplay between vitamin D, weight status, and neurodevelopment. Therefore, this study aimed to investigate the associations of vitamin D status and weight status with neurodevelopmental outcomes in a large pediatric population aged between 0 and 18 years, using data from a major children's health management center in East China.

Methods

Study Design and Participants

Our study retrospectively reviewed the electronic medical records database of participants who had health checkups at the Children's Hospital, Zhejiang University School of Medicine, from January 1, 2019, to June 30, 2025. This study included all children who met the following inclusion criteria: (1) aged ≤ 18 years old; (2) had data regarding serum levels of 25-hydroxyvitamin D, or 25(OH)D, and neurodevelopmental outcomes measured by the standard instruments; and (3) had no neurological disorders or physical disabilities. The exclusion criteria were as follows: (1) multiple records for one child, only retaining the earliest record; and (2) missing information on demographic characteristics such as age, sex, height, and weight.

Ethical Considerations

This study was conducted in accordance with the Declaration of Helsinki [13] and was approved by the Medical Ethics Committee of the Children's Hospital, Zhejiang University School of Medicine (2021-IRB-185). Informed consent was waived by the ethics committee because the research involved a retrospective review of deidentified electronic health records and posed minimal risk to the participants. To ensure privacy and confidentiality, all data were fully deidentified and anonymized prior to analysis. No financial compensation or incentives were provided to the participants, as the data were collected during routine health checkups.

Data Collection and Serum Levels 25(OH)D

Venous blood samples were collected from all participants in the morning after an overnight fast. Serum 25(OH)D levels were measured using a chemiluminescent immunoassay (25-Hydroxy Vitamin D Assay Kit, Orienter) on an automated analyzer. The laboratory strictly followed the manufacturer's protocols, with regular internal quality control to ensure that the intraassay and interassay coefficients of variation were within clinical standards. According to the Global Consensus Recommendations on Prevention and Management of Nutritional Rickets [14], the sufficiency levels are defined as follows: sufficiency (>50 nmol/l), insufficiency (30-50 nmol/l), and deficiency (<30 nmol/l).

XSL•FO
RenderX

Neurodevelopmental Outcomes

In alignment with standardized clinical screening protocols, neurodevelopmental outcomes were assessed using age-appropriate instruments. The Ages and Stages Questionnaire, Third Edition (ASQ-3) was employed for children under 6 years to capture early developmental milestones, while the Conners Parent Rating Scale (CPRS) was used for children aged between 6 and 18 years to evaluate behavioral and psychological symptoms. This transition between tools reflects the shifting clinical focus from motor and communication milestones in early childhood to behavioral and emotional regulation in school-aged children and adolescents. Age 6 serves as a critical clinical threshold in our health management center to ensure a streamlined workflow and avoid tool overlapping. The ASQ-3 was administered for neurodevelopmental screening in children under 6 years. This parent-completed tool, which takes approximately 10 to 15 minutes, consists of 30 items across 5 domains: communication, gross motor, fine motor, problem-solving, and personal-social. Each item is scored as 10 (yes), 5 (sometimes), or 0 (not yet). Domain-specific and total scores were calculated and compared to established age-specific cut-offs derived from the Chinese standardization of the ASQ-3 [15]. These Chinese norms were validated in a large-scale national sample to ensure cultural and linguistic appropriateness [15]. Children were categorized into 3 groups based on their scores relative to the mean (x) and standard deviation (SD): (1) "on schedule," where scores were above x−1SD, indicating typical development; (2) "monitored," where scores fell between x−2SD and x−1SD, indicating the monitoring zone; and (3) "possibly delayed," where scores were ≤x−2SD, indicating a high risk of developmental delay. For this study, children in both the "monitored" and "possibly delayed" categories (ie, any score ≤x−1SD in one or more domains) were classified as being at risk for neurodevelopmental delay to ensure a comprehensive screening of children requiring clinical attention. The CPRS (Chinese version) was used to evaluate the behavioral and psychological status of children aged 3 to 17 years. Our study population's actual age distribution fell entirely within its validated range; all participants were under the age of 18 (maximum age 17.9 years), and those aged ≥17.0 years represented a negligible fraction (54/19,283, 0.28%) of the total sample. Consequently, the use of the CPRS was both clinically appropriate and strictly aligned with its validated age parameters. The CPRS consists of 48 items rated on a 4-point Likert scale (0=not at all present to 3=very much present) and encompasses 5 subscales: oppositional (12 items), learning problems (4 items), psychosomatic problems (5 items), hyperactivity-impulsivity (4 items), and anxiety (4 items) [16]. The score for each subscale was determined by the mean of its component items. A participant was classified as having behavior problems if one or more subscale scores exceeded the Chinese standardized norms (ie, the mean score by more than two SDs) [17]. The questionnaires were completed by parents or primary guardians under the supervision of trained health professionals.

Covariate Evaluation

The selection of covariates for the multivariable models was primarily determined by the availability of variables in our health management database, further guided by clinical expertise and established literature on pediatric neurodevelopment. Consequently, we included all available clinical and demographic parameters from the electronic health records that were identified as potential confounders, including age, sex, and weight status. Age was calculated as the time interval from birth to the examination date. Height and weight were measured using the age-appropriate equipment and protocols. For children under 3 years of age, recumbent length was measured in the supine position using a baby height-measuring board, while for children aged 3 years and older, standing height was measured using a calibrated stadiometer. Weight was measured to the nearest 50 g using an infant electronic scale for children under 2 years of age and a lever scale for those older than 2 years. All measurements were taken with the children wearing minimal clothing and no shoes. Weight status across all age groups was classified according to the World Health Organization (WHO) growth standards [18]. Weight status for children under 2 years was defined according to weight-for-length *z* score criteria. Participants were classified as underweight (*z* score < –2), overweight (*z* score > 2), and obesity (*z* score > 3). For children and adolescents aged 2 to 18 years, BMI was calculated as weight (kg) divided by height squared (m^2). In accordance with WHO standards, participants were classified as underweight (BMI for age ≤5th percentile), overweight (≥85th percentile), or obese (≥95th percentile) [18].

Statistical Analysis

A complete case analysis was performed for this study. Participants with missing values for serum 25(OH)D, neurodevelopmental screening scores, or covariates (age, sex, and weight status) were excluded during the initial data cleaning phase, as detailed in our inclusion and exclusion criteria. This resulted in a final analytic sample of 19,283 children with complete information for all variables of interest. First, the normality of continuous variables was assessed using the Kolmogorov-Smirnov test. In this study, all continuous variables followed a normal distribution; therefore, the results are presented as mean and SD. Categorical variables are expressed as numbers and percentages. Differences in the distribution of the above dependent variables were compared between children with and without typical neurodevelopmental outcomes using *t* tests for continuous variables and chi-square tests for categorical variables. In addition, we assessed sex differences in weight status and vitamin D nutritional status in the 2 age groups (between 0 and 6 years and above 6 years). Because some vitamin D deficiency subgroups had small sample sizes, we combined the insufficiency and deficiency categories for the following analysis. Second, we evaluated the associated factors for suboptimal neurodevelopmental outcomes in the study population using binary logistic regression analysis. Further, a subgroup analysis was conducted using the binary logistic regression models to assess the association between vitamin D nutritional status and neurodevelopmental outcomes across different subgroups of weight status in both boys and girls. To account for the increased risk of type I errors due to multiple comparisons, we applied the Benjamini-Hochberg false discovery rate (FDR) correction. The findings were considered statistically robust if the FDR-adjusted *q* value was <.05. For

XSL•FO
RenderX

results where the original *P* value was <.05 but the *q* value fell between .05 and .10, the associations were interpreted as "marginally significant" or "suggestive trends" rather than definitive findings, ensuring a more conservative and prudent interpretation of our results.

All statistical analyses were conducted using SPSS Statistics (version 27.0; IBM Corp) and R software (version 4.4.1; R Foundation for Statistical Computing). Statistical significance was indicated when *P*<.05 (2 tails).

Results

General and Clinical Characteristics of the Study Population

Table 1 outlines the general and clinical characteristics of the study population. The study analyzed a total of 10,065 (52.2%) children aged below 6 years and 9218 (47.8%) children aged 6 years or above. Among children under the age of 6 years, 1275 (12.67%) were at risk for neurodevelopmental delay. Among children older than 6 years, 1483 (16.09%) were identified to have behavioral challenges. In both age groups, children with suboptimal neurodevelopmental outcomes (ie, being at risk for neurodevelopmental delay under age 6 and having behavioral challenges older than 6) were significantly younger, more often boys, and had higher BMI and lower 25(OH)D level (all *P* values <.05).

Table 1. General and clinical characteristics of the study population (N=19,283).

Characteristics	<6 years				≥6 years			
	Overall (n=10,065)	Typicality (n=8790)	At risk for delay (n=1275)	*P* value	Overall (n=9218)	Typicality (n=7,735)	Having behavioral challenges (n=1,483)	*P* value
Age (years), mean (SD)	2.76 (1.57)	2.81 (1.56)	2.48 (1.59)	<.001	9.76 (2.61)	9.79 (2.63)	9.58 (2.52)	.003
Sex, n (%)								<.001
Male	5794 (57.57)	4909 (55.85)	885 (69.41)	<.001	5771 (62.61)	4903 (63.39)	868 (58.53)	
Female	4271 (42.43)	3881 (44.15)	390 (30.59)		3447 (37.39)	2832 (36.61)	615 (41.47)	
Height (cm), mean (SD)	92.32 (14.73)	92.78 (14.60)	89.15 (15.24)	<.001	140.30 (16.37)	140.66 (16.49)	138.42 (15.60)	<.001
Weight (kg), mean (SD)	13.58 (3.98)	13.69 (3.98)	12.86 (3.90)	<.001	35.24 (14.61)	35.52 (14.72)	33.80 (13.92)	<.001
BMI (kg/m^2), mean (SD)	15.77 (1.71)	15.74 (1.68)	16.03 (1.86)	<.001	17.18 (3.66)	17.22 (3.66)	16.96 (3.61)	.01
Weight status, n (%)								.05
Underweight	834 (8.29)	723 (8.23)	111 (8.71)	.46	781 (8.47)	633 (8.18)	148 (9.98)	
Normal weight	8521 (84.66)	7456 (84.82)	1065 (83.53)		6299 (68.33)	5290 (68.39)	1009 (68.04)	
Overweight and obesity	710 (7.05)	611 (6.95)	99 (7.76)		2138 (23.19)	1812 (23.43)	326 (21.98)	
25(OH)D[a] (nmol/L), mean (SD)	98.37 (42.76)	98.82 (42.58)	95.28 (43.89)	.006	56.69 (23.47)	57.47 (23.78)	52.63 (21.33)	<.001
Vitamin D nutritional status, n (%)								<.001
Sufficiency	8954 (88.96)	7841 (89.20)	1113 (87.29)	.08	5020 (54.46)	4333 (56.02)	687 (46.33)	
Insufficiency	1021 (10.14)	875 (9.95)	146 (11.45)		3422 (37.12)	2766 (35.76)	656 (44.23)	
Deficiency	90 (0.89)	74 (0.84)	16 (1.25)		776 (8.42)	636 (8.22)	140 (9.44)	

[a]25(OH)D: 25-hydroxyvitamin D.

In addition, Multimedia Appendices 1 and 2 highlight the sex differences in the proportion of weight status and vitamin D nutritional status in the two age groups. Girls at risk for neurodevelopmental delay had significantly higher rates of vitamin D insufficiency/deficiency than their peers with typical neurodevelopment (15.64% vs 11.00%; *P*=.006). In children above 6 years old, we observe significantly higher rate of vitamin D insufficiency/deficiency in both boys and girls with behavioral challenges compared to their typical peers (boys: 50.35% vs 42.12%; *P*<.001 and girls: 58.37% vs 47.21%; *P*<.001). Multimedia Appendices 3 and 4 present the rates of being at risk for neurodevelopmental delay (<6 years) and behavioral challenges (>6 years) across the subgroups of weight status and vitamin D nutritional status, respectively. The rate of being at risk for neurodevelopmental delay was highest in the subgroup of underweight + vitamin D insufficiency/deficiency (18.75%), followed by that in the subgroup of overweight/obesity + vitamin D

XSL•FO
RenderX

insufficiency/deficiency (17.95%). The rate of behavioral challenges was highest in the subgroup of underweight + vitamin D insufficiency/deficiency (23.7%), followed by that in the subgroup of normal weight + vitamin D insufficiency/deficiency (19.38%), while the rate in the subgroup of overweight/obesity + vitamin D insufficiency/deficiency was relatively low (16.61%).

Associated Factors of Suboptimal Neurodevelopmental Outcomes by Multivariable Logistic Regression Analysis

Table 2 presents the results of the logistic regression analysis regarding the associated factors of being at risk of neurodevelopmental delay in children under 6 years old. Older age, being male, and having vitamin D insufficiency/deficiency were statistically significantly associated with being at risk for neurodevelopmental delay (all *P* values <.05). More specifically, when a child had vitamin D insufficiency or deficiency, the odds of being at risk of neurodevelopmental delay were 1.48 times higher than those who had sufficient vitamin D (95% CI 1.23-1.79; *P*<.001).

Table 2. Logistic regression analysis of factors associated with risk of neurodevelopmental delay in children under 6 years of age (n=10,065).

Characteristics	OR (95% CI)	*P* value
Age	0.84 (0.81-0.88)	<.001
Sex		
Female	reference	N/A[a]
Male	1.79 (1.57-2.03)	<.001
Weight status		
Normal weight	reference	N/A
Underweight	1.23 (0.99-1.53)	.06
Overweight and obesity	1.20 (0.96-1.50)	.12
Vitamin D nutritional status		
Sufficiency	reference	N/A
Insufficiency/Deficiency	1.48 (1.23-1.79)	<.001

[a]N/A: not applicable.

The associated factors for each specific domain of ASQ had different patterns (Multimedia Appendix 5). To be noted, underweight and overweight/obesity were significantly associated with delays in gross motor (underweight: OR 1.58, 95% CI 1.17-2.12; *P*=.003; overweight/obesity: OR 1.64, 95% CI 1.21-2.23; *P*=.002) and fine motor (underweight: OR 1.77, 95% CI 1.32-2.38; *P*<.001; overweight/obesity: OR 1.55, 95% CI 1.13-2.14; *P*=.007). Vitamin D insufficiency/deficiency was significantly associated with elevated odds of being at risk of delay at the overall level and each domain level (all *P*<.05). There was also a sex difference regarding the pattern of associated factors, as underscored in Multimedia Appendices 6 and 7. Boys with overweight/obesity had a higher risk of overall delay (odds ratio [OR] 1.30, 95% CI 1.01-1.68; *P*=.042), while vitamin D insufficiency or deficiency was not statistically significantly associated with overall delay or delay in a specific domain except for communication (OR 1.57, 95% CI 1.02-2.42; *P*=.042). However, in girls, the significant association between overweight/obesity and delay disappeared, and vitamin D insufficiency or deficiency had a statistically significant association with overall delay or delay in a specific domain (*P*<.05).

Table 3 presents the results of the logistic regression analysis in terms of the associated factors of having behavioral challenges in children above 6 years old. Older age, being female, being underweight, and having vitamin D insufficiency/deficiency were statistically significantly associated with having behavioral challenges (all *P*<.05). More specifically, vitamin D insufficiency or deficiency was significantly associated with an increased odds of having behavioral challenges (OR 1.56, 95% CI 1.39-1.74; *P*<.001).

XSL•FO
RenderX

Table 3. Logistic regression analysis of factors associated with behavioral challenges in children over 6 years (n=9218).

Characteristics	OR (95% CI)	*P* value
Age	0.95 (0.93-0.98)	<.001
Sex		
Female	reference	N/A[a]
Male	1.19 (1.06-1.33)	.003
Weight status		
Normal weight	reference	N/A
Underweight	1.25 (1.04-1.52)	.02
Overweight and obesity	0.95 (0.83-1.10)	.50
Vitamin D nutritional status		
Sufficiency	reference	N/A
Insufficiency/deficiency	1.56 (1.39-1.74)	<.001

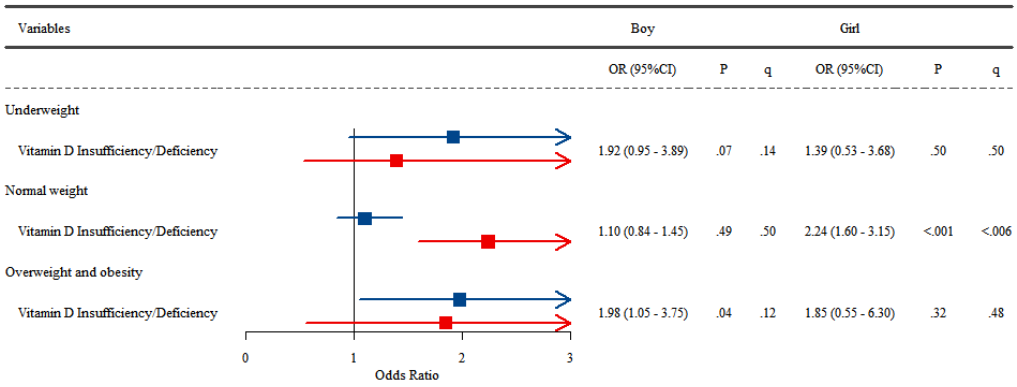
[a]N/A: not applicable.

The pattern of associated factors for each domain of CPRS differed (Multimedia Appendix 8). In particular, underweight was significantly associated with elevated risk of having psychosomatic problems (OR 1.31, 95% CI 1.01-1.71; *P*=.043), while overweight/obesity was significantly associated with lower odds of having psychosomatic problems (OR 0.68, 95% CI 0.55-0.85; *P*<.001). Vitamin D insufficiency/deficiency was significantly associated with elevated odds of behavioral challenges and odds of oppositional disorder, learning challenges, psychosomatic challenges, and hyperactivity-impulsivity at the domain level (all *P*<.05). Multimedia Appendices 9 and 10 display the profiles of associated factors in boys and girls, respectively.

Subgroup Analysis of the Association Between Vitamin D and Neurodevelopmental Outcomes Across Different Weight Subgroups

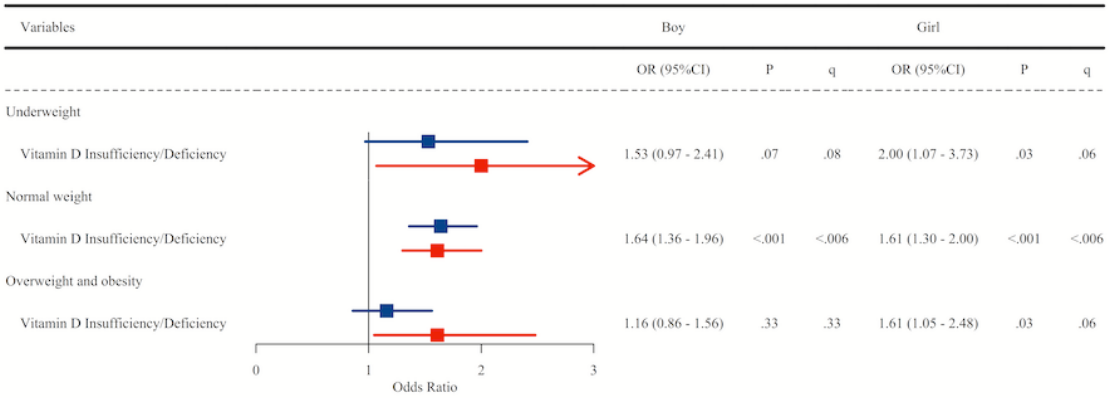
Figure 1 contains a forest plot showing the associations between vitamin D insufficiency/deficiency and neurodevelopmental delay in boys and girls under 6 years old in the underweight, normal weight, and overweight/obesity subgroups. Notably, after FDR correction, the most robust association was observed in the normal-weight subgroup, where girls with vitamin D insufficiency/deficiency had a significantly higher risk of neurodevelopmental delay (OR 2.24, 95% CI 1.60-3.15; *P*<.001, q<.006). Figure 2 demonstrates the associations between vitamin D insufficiency/deficiency and behavioral challenges in boys and girls above 6 years old in the underweight, normal weight, and overweight/obesity subgroups. Among girls, while the original *P* values indicated significant associations before FDR corrections across all weight categories (all *P*<.05), these associations remained statistically robust primarily in the normal-weight group (OR 1.61, 95% CI 1.30-2.00; *P*<.001; *q*<.006) and showed a marginal significance in the underweight and overweight/obesity groups (all *q*=.06). Among boys, the association remained significant only in the normal-weight subgroup, where vitamin D insufficiency/deficiency was linked to a higher chance of behavioral challenges (OR 1.64, 95% CI 1.36-1.96; *P*<.001; *q*<.006).

Figure 1. Associations between Vitamin D status and being at risk for neurodevelopmental delay in children aged under 6 years, stratified by sex and weight status.



XSL•FO
RenderX

Figure 2. Associations between Vitamin D status and behavior problems in children aged above 6 years, stratified by sex and weight status.



Discussion

Principal Findings

This large-scale study revealed the complex relationship of vitamin D and weight status with neurodevelopmental outcomes across a wide pediatric age range. Our principal and most significant finding is a profound sex-specific divergence in risk profiles: in children under 6 years, neurodevelopmental risk in boys was predominantly linked to overweight/obesity, whereas in girls, vitamin D insufficiency/deficiency emerged as the critical independent determinant. Furthermore, we uncovered an unexpected pattern in that the harmful effects of low vitamin D on neurodevelopment were surprisingly strongest in normal-weight children. Collectively, these findings challenge a simplistic risk model, suggesting that the pathways linking vitamin D and weight status to childhood neurodevelopment are fundamentally moderated by both sex and developmental context. This highlights the necessity for more tailored, sex- and context-specific strategies in both future research and clinical practice.

Our study first highlights the significant public health challenge of vitamin D insufficiency in Chinese children, revealing a dramatic age-related disparity. While the combined prevalence of vitamin D insufficiency and deficiency was 11.03% in children under 6 years, this figure notably increased to 45.54% in the 6-to-18 years age group. This significant increase is consistent with other reports from China [19] and likely reflects a critical gap in public health practices. The current supplementation guidelines and parental attention are predominantly focused on infancy and early childhood [20,21], after which routine supplementation often stops [22]. The prevalence in our cohort, particularly among the older group, appears considerably higher than that reported in children in the United States by the National Health and Nutrition Examination Survey (2011-2014) [23], underscoring the urgent need to revise and extend public health strategies in China.

A main finding of our study is the confirmation that vitamin D insufficiency/deficiency is a strong independent risk factor for suboptimal neurodevelopmental outcomes, after adjusting for a comprehensive set of critical covariates, including age, sex, and weight status. What makes this association particularly compelling is the distinct and pervasive nature of the risk profile associated with poor vitamin D status when compared to that of weight status.

Our analysis revealed that the impact of abnormal weight status was primarily domain-specific and manifested differently across developmental stages. While it was associated with gross and fine motor skill delays in the younger cohort (under 6 years), it was linked to psychosomatic disorders in the older cohort (between 6 and 18 years). These divergent patterns likely reflect both age-dependent biological vulnerabilities and the distinct clinical focus of the assessment tools optimized for each age group.

This association likely reflects different physiological mechanisms depending on weight status: underweight children may lack essential nutrients needed for brain growth, while those with overweight or obesity may experience chronic inflammation and metabolic disruption. In contrast, vitamin D insufficiency/deficiency demonstrated a remarkably global effect that transcended developmental stages. In our younger cohort, this manifested as pan-domain developmental delays, while in the older group, it presented as broad-spectrum behavioral and learning challenges. This broad-spectrum risk profile not only aligns with but also substantially extends the scope of previous research, which often linked low vitamin D to more isolated outcomes such as communication delays or an increased risk for ADHD [24,25]. The consistency of this widespread risk reinforces the hypothesis that vitamin D functions as a fundamental neurosteroid, a role supported by strong biological plausibility. Vitamin D acts via receptors that are ubiquitously expressed throughout the brain, where it modulates genes controlling fundamental neurodevelopmental processes, including neuronal differentiation, synaptic plasticity, and neuroprotection [9,11,26-30]. A deficit in this essential nutrient can thus lead to widespread disruption at any stage from early brain formation to later neural maturation. This highlights vitamin D deficiency as a fundamental threat to brain health throughout childhood.

Beyond the aforementioned main effects, our study reveals a complex, sex-specific interplay between vitamin D status and weight status that challenges simplistic, additive risk models. Following rigorous FDR correction, the most robust associations between vitamin D insufficiency/deficiency and adverse neurodevelopmental outcomes were consistently observed within the normal-weight subgroup across all age groups in girls and

XSL•FO
RenderX

among boys aged 6 years and older. In normal-weight children, the absence of major metabolic disturbances—often associated with extreme weight statuses—may provide a clearer "biological window" to observe the direct neurotropic impact of vitamin D. This suggests that when weight status is within the normal range, the brain's foundational development and behavioral regulation may be more directly vulnerable to the lack of this critical neurosteroid.

In children with overweight or obesity, neurodevelopment is influenced by a multifactorial array of risks, including chronic inflammation and metabolic dysregulation, which may independently impair cognitive and behavioral functions. These competing risks may "mask" or dilute the specific contribution of vitamin D in our statistical models, leading to the attenuated associations observed ($q>.05$). Interestingly, in girls aged 6 years and older, a consistent but marginally significant trend was observed in the underweight and overweight/obesity subgroups (both $q=.06$). This suggests that the neuroprotective role of vitamin D in girls above 6 years may be more pervasive across different body compositions than in other cohorts. The attenuation from significance ($P<.05$) to a marginal trend ($q=.06$) after FDR correction likely reflects the reduced statistical power within these smaller samples of weight-extreme strata, rather than a true absence of biological effects. Nevertheless, these findings should be considered exploratory, warranting further validation in even larger specialized cohorts of pediatric populations with underweight or overweight/obesity.

Our findings urge a shift toward a more precise, stratified approach in neurodevelopmental surveillance. From a public health perspective, the sharp rise in vitamin D inadequacy after infancy highlights a critical gap in health education, necessitating revised strategies that advocate for continued supplementation throughout childhood. In the clinical setting, our results call for more proactive and risk-based screening. Our data suggest that clinicians should view the combination of weight status and vitamin D insufficiency/deficiency as an identifiable risk signal for future neurodevelopmental and behavioral adversity. Our results suggest that girls may exhibit greater vulnerability to the neurodevelopmental impacts of vitamin D insufficiency/deficiency. For girls older than 6 years, the risks associated with vitamin D insufficiency/deficiency should be taken seriously, regardless of their weight status.

Strengths and Limitations

Our study has several key strengths. The main strength is its large sample size, which ensures sufficient statistical power and allows us to examine complex sex-specific relationships that differ between boys and girls. This kind of detailed analysis is often not possible in smaller studies. Second, the inclusion of a wide age range, combined with validated, age-appropriate neurodevelopmental assessment tools (ASQ-3 and CPRS), was essential for revealing how these risk factors manifest differently across distinct developmental stages. Furthermore, our rigorous adjustment for a comprehensive set of covariates strengthens the conclusion that vitamin D status and weight status are independently associated with neurodevelopmental outcomes.

Nevertheless, we must acknowledge several limitations. First, the cross-sectional design of this study establishes associations but precludes definitive causal inference. We cannot entirely rule out reverse causality, where children with subtle neurodevelopmental or behavioral challenges may have reduced outdoor activities or restrictive dietary habits, leading to secondary vitamin D deficiency. However, since our participants were primarily from the general pediatric population undergoing routine health screenings, the impact of severe clinical behaviors is likely limited. Future longitudinal studies are necessary to elucidate this potentially bidirectional relationship. Second, our outcomes were based on parent-reported screening questionnaires. While these tools are highly feasible for large-scale epidemiological research, they are not a substitute for gold-standard clinical diagnoses and may be subject to reporting bias. Third, despite the adjustment for essential clinical covariates, the potential for residual confounding remains. As our data were derived from routine electronic health records rather than structured research surveys, certain socioenvironmental and lifestyle factors, such as parental education level, family income, and outdoor physical activity, were not available in our database. These factors are known to influence both vitamin D levels and neurodevelopmental outcomes. Although we included all relevant clinical variables available in our records to mitigate this, our results should be interpreted with caution regarding these unmeasured confounders. Fourth, our study utilized different assessment tools for the two age cohorts (ASQ-3 and CPRS). While both are validated for their respective populations, their domains do not completely overlap (eg, hyperactivity is not captured by ASQ-3). Consequently, our findings should be interpreted as stage-specific associations rather than a continuous developmental trajectory across the pediatric lifespan.

Conclusions

Our study provides compelling evidence that vitamin D deficiency is an independent determinant of neurodevelopmental and behavioral adversities throughout childhood. The interaction with weight status is not uniform but sex dependent. These findings highlight the inadequacy of a one-size-fits-all approach. We advocate for a shift toward proactive, stratified screening, where neurodevelopmental risk is assessed through a dynamic, sex- and context-specific lens to enable truly personalized pediatric care.

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or nonprofit sectors. The authors declare the use of generative AI in the research and writing process. According to the 2025 Generative Artificial Intelligence Delegation taxonomy (GAIDeT), proofreading and editing tasks were delegated to GAI tools under full human supervision. The GAI tool used was Gemini 3.0. The responsibility for the final manuscript lies entirely with the authors. The GAI tools are not listed as authors and do not bear responsibility for the final outcomes.

XSL•FO
RenderX

Data Availability

The data that support the findings of this study are available from the corresponding author (guannanbai@zju.edu.cn) upon reasonable request.

Authors' Contributions

YS contributed to conceptualization, methodology, investigation, data curation, project administration, resources, writing the original draft, and writing, review, and editing. MG contributed to the methodology, formal analysis, data curation, writing the original draft, and writing, review, and editing. LG contributed to the methodology, formal analysis, writing the original draft, and writing, review, and editing. GB contributed to the study conceptualization, methodology, supervision, writing the original draft, and writing, reviewing, and editing. LS contributed to writing the original draft and writing, reviewing, and editing. YX contributed to data curation and writing, reviewing, and editing. BD contributed to writing, reviewing, and editing. KZ contributed to writing, reviewing, and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sex differences in the proportion of weight status and vitamin D nutritional status in children between 0 and 6 years old with typical neurodevelopment and those at risk for delay (n=10,065).
[DOCX File , 14 KB - publichealth_v12i1e89756_app1.docx]

Multimedia Appendix 2
Sex differences in the proportion of weight status and vitamin D nutritional status in children above 6 years old with and without behavioral challenges (n=9218).
[DOCX File , 14 KB - publichealth_v12i1e89756_app2.docx]

Multimedia Appendix 3
Rates of being at risk for neurodevelopmental delay in subgroups of weight status and vitamin D nutritional status in children under 6 years old (n=10,065).
[DOCX File , 14 KB - publichealth_v12i1e89756_app3.docx]

Multimedia Appendix 4
Rates of behavior problems in subgroups of weight status and vitamin D nutritional status in children above 6 years old (n=9218).
[DOCX File , 14 KB - publichealth_v12i1e89756_app4.docx]

Multimedia Appendix 5
Associated factors of different ASQ-3 (Ages and Stages Questionnaire, Third Edition) domains in children under 6 years old using logistic regression analysis (n=10,065).
[DOCX File , 14 KB - publichealth_v12i1e89756_app5.docx]

Multimedia Appendix 6
Associated factors of being at risk of neurodevelopmental delay in boys under 6 years old using logistic regression analysis (n=5794).
[DOCX File , 15 KB - publichealth_v12i1e89756_app6.docx]

Multimedia Appendix 7
Associated factors of being at risk of neurodevelopmental delay in girls under 6 years old according the logistic regression analysis (n=4271).
[DOCX File , 15 KB - publichealth_v12i1e89756_app7.docx]

Multimedia Appendix 8
Associated factors of different Conners Parent Rating Scale (CPRS) subscales in children above 6 years via logistic regression analysis (n=9218).
[DOCX File , 15 KB - publichealth_v12i1e89756_app8.docx]

XSL•FO
RenderX

Multimedia Appendix 9
Associated factors of behavioral challenges in boys above 6 years using logistic regression analysis (n=5771).
[DOCX File , 15 KB - publichealth_v12i1e89756_app9.docx]

Multimedia Appendix 10
Associated factors of behavioral challenges in girls above 6 years using logistic regression analysis (n=3447).
[DOCX File , 15 KB - publichealth_v12i1e89756_app10.docx]

Multimedia Appendix 11
STROBE (Strengthening the Reporting of Observational Studies in Epidemiology): checklist of items that should be included in reports of observational studies.
[DOC File , 74 KB - publichealth_v12i1e89756_app11.doc]

References

1. Obesity and overweight. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight [accessed 2025-12-04]
2. Roth DE, Abrams SA, Aloia J, Bergeron G, Bourassa MW, Brown KH, et al. Global prevalence and disease burden of vitamin D deficiency: a roadmap for action in low- and middle-income countries. Ann N Y Acad Sci 2018 Oct 08;1430(1):44-79 [FREE Full text] [doi: 10.1111/nyas.13968] [Medline: 30225965]
3. Zhao S, Semeia L, Veit R, Moser J, Preissl H, Kullmann S. Structural and functional brain changes in children and adolescents with obesity. Obes Rev 2025 Dec;26(12):e70001. [doi: 10.1111/obr.70001] [Medline: 40759532]
4. Bouillon R, Marcocci C, Carmeliet G, Bikle D, White JH, Dawson-Hughes B, et al. Skeletal and extraskeletal actions of vitamin d: current evidence and outstanding questions. Endocr Rev 2019 Aug 01;40(4):1109-1151 [FREE Full text] [doi: 10.1210/er.2018-00126] [Medline: 30321335]
5. Yau PL, Kang EH, Javier DC, Convit A. Preliminary evidence of cognitive and brain abnormalities in uncomplicated adolescent obesity. Obesity (Silver Spring) 2014 Aug;22(8):1865-1871 [FREE Full text] [doi: 10.1002/oby.20801] [Medline: 24891029]
6. Tian P, Zhu X, Liu Z, Bian B, Jia F, Dou L, et al. Effects of vitamin D on brain function in preschool children with autism spectrum disorder: a resting-state functional MRI study. BMC Psychiatry 2025 Mar 03;25(1):198 [FREE Full text] [doi: 10.1186/s12888-025-06534-8] [Medline: 40033268]
7. Khoshbakht Y, Bidaki R, Salehi-Abargouei A. Vitamin D status and attention deficit hyperactivity disorder: a systematic review and meta-analysis of observational studies. Adv Nutr 2018 Jan 01;9(1):9-20 [FREE Full text] [doi: 10.1093/advances/nmx002] [Medline: 29438455]
8. Filgueiras MS, Rocha NP, Novaes JF, Bressan J. Vitamin D status, oxidative stress, and inflammation in children and adolescents: A systematic review. Crit Rev Food Sci Nutr 2020;60(4):660-669. [doi: 10.1080/10408398.2018.1546671] [Medline: 30596263]
9. Eyles DW, Liu PY, Josh P, Cui X. Intracellular distribution of the vitamin D receptor in the brain: comparison with classic target tissues and redistribution with development. Neuroscience 2014 May 30;268:1-9. [doi: 10.1016/j.neuroscience.2014.02.042] [Medline: 24607320]
10. Mayne PE, Burne THJ. Vitamin D in synaptic plasticity, cognitive function, and neuropsychiatric illness. Trends Neurosci 2019 Apr;42(4):293-306. [doi: 10.1016/j.tins.2019.01.003] [Medline: 30795846]
11. Ye X, Zhou Q, Ren P, Xiang W, Xiao L. The synaptic and circuit functions of vitamin d in neurodevelopment disorders. Neuropsychiatr Dis Treat 2023;19:1515-1530 [FREE Full text] [doi: 10.2147/NDT.S407731] [Medline: 37424961]
12. Wortsman J, Matsuoka LY, Chen TC, Lu Z, Holick MF. Decreased bioavailability of vitamin D in obesity. Am J Clin Nutr 2000 Sep;72(3):690-693. [doi: 10.1093/ajcn/72.3.690] [Medline: 10966885]
13. World Medical Association. Declaration of Helsinki: Ethical principles for medical research involving human subjects. Bull World Health Organ 2001;79(4):373-374 [FREE Full text] [Medline: 11357217]
14. Munns CF, Shaw N, Kiely M, Specker BL, Thacher TD, Ozono K, et al. Global consensus recommendations on prevention and management of nutritional rickets. Horm Res Paediatr 2016;85(2):83-106 [FREE Full text] [doi: 10.1159/000443136] [Medline: 26741135]
15. Wei M, Bian X, Squires J, Yao G, Wang X, Xie H, et al. [Studies of the norm and psychometrical properties of the ages and stages questionnaires, third edition, with a Chinese national sample]. Zhonghua Er Ke Za Zhi 2015 Dec;53(12):913-918. [Medline: 26887546]
16. Conners CK, Sitarenios G, Parker JDA, Epstein JN. The revised Conners' Parent Rating Scale (CPRS-R): factor structure, reliability, and criterion validity. J Abnorm Child Psychol 1998 Aug;26(4):257-268. [doi: 10.1023/a:1022602400621] [Medline: 9700518]
17. Fan J, Du Y, Wang L. The norm and reliability of the Conners Parent Symptom Questionnaire in Chinese urban children. Shanghai Arch Psychiatry 2005;17(6):321-323. [doi: 10.3969/j.issn.1002-0829.2005.06.001]

XSL•FO
RenderX

18. WHO Multicentre Growth Reference Study Group. WHO Child Growth Standards based on length/height, weight and age. Acta Paediatr Suppl 2006 Apr 02;450(S450):76-85. [doi: 10.1111/j.1651-2227.2006.tb02378.x] [Medline: 16817681]
19. Zhu Z, Zhan J, Shao J, Chen W, Chen L, Li W, et al. High prevalence of vitamin D deficiency among children aged 1 month to 16 years in Hangzhou, China. BMC Public Health 2012 Mar 14;12:126 [FREE Full text] [doi: 10.1186/1471-2458-12-126] [Medline: 22330045]
20. Editorial Board of Chinese Journal of Pediatrics, Subspecialty Group of Child Health Care, Chinese Pediatric Society, Chinese Medical Association, Cooperation Group of PreventionTreatment of Rickets of Vitamin D Deficiency in Childhood. Recommendation for prevention and treatment of rickets of vitamin D deficiency in childhood. Zhonghua Er Ke Za Zhi 2008 Mar;46(3):190-191. [Medline: 19099706]
21. Subspecialty Group of Children Health, the Society of Pediatrics, Chinese Medical Association, Editorial Board, Chinese Journal of Pediatrics. Practical guidelines for clinical issues related to vitamin D nutrition in Chinese children. Zhonghua Er Ke Za Zhi 2022 May 02;60(5):387-394. [doi: 10.3760/cma.j.cn112140-20211230-01092] [Medline: 35488630]
22. Yu C, Cai J, Wang C, Luo Y, Fang J. Knowledge, attitude, and practice toward pediatric vitamin D deficiency among parents. Front Pediatr 2024;12:1393488 [FREE Full text] [doi: 10.3389/fped.2024.1393488] [Medline: 39005508]
23. Herrick KA, Storandt RJ, Afful J, Pfeiffer CM, Schleicher RL, Gahche JJ, et al. Vitamin D status in the United States, 2011-2014. Am J Clin Nutr 2019 Jul 01;110(1):150-157 [FREE Full text] [doi: 10.1093/ajcn/nqz037] [Medline: 31076739]
24. Li B, Xu Y, Zhang X, Zhang L, Wu Y, Wang X, et al. The effect of vitamin D supplementation in treatment of children with autism spectrum disorder: a systematic review and meta-analysis of randomized controlled trials. Nutr Neurosci 2022 Apr;25(4):835-845 [FREE Full text] [doi: 10.1080/1028415X.2020.1815332] [Medline: 32893747]
25. Bener A, Kamal M. Predict attention deficit hyperactivity disorder? Evidence -based medicine. Glob J Health Sci 2013 Nov 27;6(2):47-57 [FREE Full text] [doi: 10.5539/gjhs.v6n2p47] [Medline: 24576365]
26. De Marzio M, Lasky-Su J, Chu SH, Prince N, Litonjua AA, Weiss ST, et al. The metabolic role of vitamin D in children's neurodevelopment: a network study. Sci Rep 2024 Jul 23;14(1):16929 [FREE Full text] [doi: 10.1038/s41598-024-67835-8] [Medline: 39043876]
27. Sailike B, Onzhanova Z, Akbay B, Tokay T, Molnár F. Vitamin D in central nervous system: implications for neurological disorders. Int J Mol Sci 2024 Jul 17;25(14) [FREE Full text] [doi: 10.3390/ijms25147809] [Medline: 39063051]
28. Eyles DW. Vitamin D: Brain and Behavior. JBMR Plus 2021 Jan;5(1):e10419 [FREE Full text] [doi: 10.1002/jbm4.10419] [Medline: 33553986]
29. Kiraly SJ, Kiraly MA, Hawe RD, Makhani N. Vitamin D as a neuroactive substance: review. Sci World J 2006 Jan 26;6:125-139 [FREE Full text] [doi: 10.1100/tsw.2006.25] [Medline: 16493517]
30. Pérez-López FR. Vitamin D: the secosteroid hormone and human reproduction. Gynecol Endocrinol 2007 Jan;23(1):13-24 [FREE Full text] [doi: 10.1080/09513590601045629] [Medline: 17484507]

Abbreviations

25(OH)D: 25-hydroxyvitamin D
ADHD: attention-deficit/hyperactivity disorder
ASD: autism spectrum disorder
ASQ-3: Ages and Stages Questionnaire, Third Edition
CPRS: Conners Parent Rating Scale
OR: odds ratio
WHO: World Health Organization

Edited by F Pabon-Rodriguez; submitted 20.Dec.2025; peer-reviewed by L Xiao, K Ye, S Shao; comments to author 21.Jan.2026; revised version received 01.Feb.2026; accepted 03.Feb.2026; published 27.Feb.2026.

Please cite as:
Shen Y, Gao M, Guan L, Sun L, Xu Y, Daniel B, Zhang K, Bai G
Association of Vitamin D and Weight Status With Neurodevelopmental Outcomes in a Large Pediatric Population: Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e89756
URL: https://publichealth.jmir.org/2026/1/e89756
doi:10.2196/89756
PMID:

©Ying Shen, Meiying Gao, Lejing Guan, Lidan Sun, Yanfen Xu, Bethel Daniel, Kyla Zhang, Guannan Bai. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 27.Feb.2026. This is an open-access article distributed

XSL•FO
RenderX

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

XSL•FO
RenderX

Original Paper

Social Networks and Their Influence on the Choice of Unassisted Smoking Cessation: Cross-Sectional Study in Six Cities in China

Sihui Peng[1], PhD; Tingzhong Yang[2,3,4], MD; Lijing Li[5], MD; Weifang Zhang[6], PhD; Randall R Cottrell[7], DEd

[1]School of Medicine, Jinan University, Guangzhou, China
[2]Center for Tobacco Control Research, Zhejiang University, Hangzhou, Zhejiang, China
[3]Research Center for Digital Health Behavior Theory and Management, National Health Big Data Institute, Zhejiang University, Hangzhou, China
[4]Injury Control Research Center, West Virginia University, Morgantown, WV, United States
[5]Department of Obstetrics and Gynecology, Yongkang Women and Children's Health Hospital, Yongkang, Zhejiang, China
[6]Children's Hospital of Zhejiang University, Hangzhou, Zhejiang, China
[7]Public Health Studies Program, School of Health and Applied Human Sciences, University of North Carolina, Wilmington, NC, United States

Corresponding Author:
Tingzhong Yang, MD
Center for Tobacco Control Research
Zhejiang University
No.866 Yuhangtang Road
Hangzhou, Zhejiang, 310058
China
Phone: 86 571 88208219
Email: tingzhongyang@zju.edu.cn

Abstract

Background: Many countries, including China, have implemented nationwide smoking cessation programs in accordance with Article 14 of the World Health Organization (WHO) Framework Convention on Tobacco Control (FCTC). However, the use of assisted smoking cessation services remains low, while unassisted smoking cessation (USC) methods are widely preferred. Studies have shown that many individual and environmental factors are associated with USC adoption. However, no studies have examined the association between social networks and USC adoption.

Objective: This study examined the effects of social networks on USC choice among male smokers in China.

Methods: A cross-sectional multivariable sampling design was used to interview subjects from six selected cities in China. The study sample included only male participants. A standardized questionnaire was used to obtain information about sociodemographic characteristics, social networks, and USC choice. Multivariable logistic regression models were used to examine the association between social networks and USC choice. Furthermore, quantitative analysis was conducted to demonstrate the dose-response relationship between the city-level social networks and the prevalence of USC. Structural equation modeling (SEM) was used to establish the mechanisms by which reference group norms affect USC through social networks.

Results: We identified 2852 smokers, of whom 1647 (57.7%) had attempted to or had quit smoking. Among them, 91.6% (n=1509; 95% CI 90.9%-97.5%) reported quitting without assistance, and 42.1% (n=58; 95% CI 32.4%-61.3%) of the remaining 138 (8.4%) participants who used USC methods achieved abstinence. Multiple logistic regression analysis found that although the higher social network (HSN) is not significantly associated with USC adoption, both the ordinary social network (OSN) and the lower social network (LSN) are significantly negatively associated with USC adoption. Compared to the low-size group of the OSN, the odds ratios (ORs) for the medium-, high-, and very high-size groups were 0.23 (95% CI 0.10-0.80), 0.22 (95% CI 0.07-0.60), and 0.25 (95% CI 0.08-0.72), respectively. Similarly, compared to the low- and medium-size groups of the LSN, the OR for the high-size group was 0.32 (95% CI 0.23-0.45). The analysis revealed a significant positive dose-response relationship of city-level OSN and LSN sizes with the probability of USC adoption, as indicated by regression coefficients β=0.2784 (P<.01) and β=2.2269 (P<.01), respectively. SEM analysis indicated that the size of OSNs and LSNs exerts an indirect effect on the relationship between reference group norms and USC (β=0.2315, P<.01; β=0.4613, P<.01).

Conclusions: Both OSNs and LSNs are significantly negatively associated with USC adoption. The findings underscore the significance of USC and its potential role in reducing smoking prevalence within the Chinese population.

XSL•FO
RenderX

(JMIR Public Health Surveill 2026;12:e74147) doi:10.2196/74147

KEYWORDS

unassisted smoking cessation; smoking cessation; social network; tobacco control; smoking; smoke; ordinary social network; lower social network; higher social network

Introduction

Numerous studies have shown that the success rate of smoking cessation assisted by health professionals, which includes both pharmacological and nonpharmacological methods, is higher than that of unassisted smoking cessation (USC) [1-3]. Article 14 of the World Health Organization (WHO) Framework Convention on Tobacco Control (FCTC) requires parties to promote smoking cessation and enforce effective measures to assist tobacco users in quitting [4]. For decades, various countries worldwide have made efforts to assist people with smoking cessation. In China, nationwide smoking cessation clinics and hotlines have been established by the central and numerous local governments, but the number of smokers who use these smoking cessation services is small [5]. In contrast, USC methods are widely used by individuals seeking to quit smoking, as they believe they can achieve abstinence through their own efforts without the need for external assistance [6-8].

To understand why this situation occurs, a thorough understanding of the nature of USC is necessary. Studies have shown that some individual and environmental factors are associated with USC adoption [6,9-11]. When considering environmental influences on USC adoption, the environment in which people live is primarily shaped by the interactions between individuals, their social relationships, and social networks. Interpersonal relationships and social networks, as key environmental variables, are likely to have a significant impact on people's adoption of USC.

The theory of reasoned action (TRA) offers a valuable framework for understanding individual decision-making in health-related behaviors, including smoking cessation. According to TRA, behavioral intention is shaped by two primary factors, attitudes toward the behavior and subjective norms [12]. In the context of USC, these referent norms may encompass beliefs and social pressures from close contacts, such as family members, peers, or colleagues. TRA further suggests that normative beliefs, behavioral beliefs, and attitudes, especially those shaped by individuals with whom the person shares meaningful relationships, can significantly impact behavioral decisions [12]. Accordingly, the structure and composition of one's social network may play a pivotal role in shaping the intention to quit smoking without external aids. For example, individuals embedded in networks where smoking is discouraged or where others have successfully quit may be more inclined to attempt smoking cessation independently. In contrast, networks that normalize smoking or lack smoking cessation role models may diminish the likelihood of unassisted quitting. Building upon this theoretical foundation, our study examined how social network size and characteristics relate to the choice of USC. We hypothesized that broader or more supportive networks may facilitate autonomous quitting decisions, in line with the mechanisms proposed by TRA.

The concept of social networks emphasizes that each individual is connected to others, who, in turn, are connected to many more. Individuals linked through social ties influence one another's norms, attitudes, and behaviors. When people are embedded in networks with smokers, repeated exposure to smoking behavior can normalize it to the extent that smoking becomes a perceived social norm. This norm is further reinforced by others who smoke within the network. Social networks serve as contexts for communication, reinforcement of beliefs, and daily interactions. Scholars have argued that strong ties within a network possess a unique capacity to shape beliefs and behaviors, while weak ties facilitate the diffusion of diverse information and behavioral alternatives. Strong ties tend to promote conformity to existing norms [13,14]. Empirical studies have shown that adolescents who occupy central positions within social cliques tend to exhibit lower-than-average smoking rates, whereas those who are socially isolated often display higher-than-average smoking rates [15-18]. People's daily life involves dealing with various relationships, such as those with themselves, with nature, with the supernatural, and with others. Chinese culture places great emphasis on interpersonal relationships, considering them perhaps the most important aspect of their life [12]. With the rapid economic development in China over the past 40 years, managing social interactions and personal relationships has become increasingly important. It serves as a foundation for the commercial enterprise and business acumen that are deeply ingrained in society. Chinese culture regards smoking as a tool for social interaction [19]. Studies have reported that in China, the most important reasons for smoking are embedded in social interactions and interpersonal relationships [20,21]. In Chinese culture, offering a cigarette has been a custom to strengthen personal friendships and relationships. Chinese people share cigarettes with one another and offer them to guests as a sign of hospitality [19].USC is a self-determined, self-help smoking cessation behavior that is inherently closely linked to social networks. The influence of social networks on USC may be particularly prominent in Chinese culture.

Social network theory emphasizes the fact that the structural and compositional features of a social network are its essential functions. Social networks matter because different types of connections have different effects on behaviors. The impact on smoking by networks composed of different types of ties—family versus friends, distant versus proximate—is complicated and can be different for different individuals. Some studies have found smoking cessation among social networks members, including spouses, siblings, friends, and neighbors, are particularly relevant for smoking cessation [22-25]. Therefore, based on these findings, we hypothesized that the type of social network is positively associated with the adoption of USC.

XSL•FO
RenderX

No studies have examined the association between social networks and USC adoption. This study aimed to explore the impact of social networks on USC adoption. The findings should help in understanding the nature of USC and its role in reducing smoking prevalence among the Chinese population. As most people quit smoking through the use of USC, we believe that USC can serve as a proxy for the common characteristics of smoking cessation.

Evidence from several studies suggests that USC may function as a scalable public health strategy, capable of lowering smoking rates across the broader population [6]. A study [26] found that unassisted quitting contributed to successful quitting by approximately 70 million smokers in China. This effectively decreased the smoking prevalence in the population and yielded substantial health benefits. These findings contribute to a deeper understanding of a key factor affecting USC trends at the population level, particularly in settings where assisted methods are less accessible or culturally less adopted.

According to TRA, reference group norms and social expectations significantly affect behavioral intentions. In this study, we used structural equation modeling (SEM) to establish the mechanisms by which reference group norms affect USC through social networks. Reference group norms not only exert a direct effect on USC but also indirectly shape this behavior through the size of general and lower-level social networks. The more individuals one interacts with in these networks, the stronger the normative the pressures become, thereby amplifying the impact of subjective norms on smoking cessation outcomes. SEM is particularly appropriate in this context, as it allows for the simultaneous estimation of direct and indirect pathways, providing a more nuanced understanding of the interplay between individual perceptions and social structures. This approach enriches the explanatory power of TRA by integrating social network dynamics, and it offers practical insights for designing interventions: by identifying key network structures and normative influences, policymakers and health practitioners can more effectively promote smoking cessation behaviors.

This study sought to address the following research questions:

- How do variations in the social network size influence USC?
- What is the quantitative relationship between regional-level differences in the social network size and USC?
- Do different social network sizes serve as mediators between subjective norms and USC?

Methods

Study Design

This observational, cross-sectional study had a multistage cluster sampling design. First, we selected six sample cities from various regions in China and then two residential districts from the primary urban areas within each city. Next, we selected four communities from each district and then five building blocks within each community. Finally, within each building block, we selected 1 out of every 20 households on the family household registration list. All selection was random.

Data Collection

Participants were identified based on specific criteria set by the research team. Inclusion criteria were as follows: (1) the male resident in the household (if there were two or more male residents, the person whose birth date was closest to the contact date was selected); (2) age≥15 years; (3) living in the study city for at least 1 year; and (4) willingness to participate in the study and provide survey-based information. Individuals were excluded if they refused to provide survey-based information or had a medical condition that could limit or preclude their participation.

Potential participants were approached by trained interviewers to gauge their willingness to participate. Each interviewer received 1-day training on the survey protocol and interviewing procedures.

A self-administered questionnaire was scheduled once an individual agreed to participate in the survey. The questionnaire was administered by fourth-year medical students or graduates from a local medical college. The survey protocol was identical across all six cities to maintain consistency in interviews and data collection. Participants were privately administered the survey in their homes or at a quiet designated location, such as a backyard or a community park. The survey primarily took place on weekends, in the evening, or at other times convenient for the participants.

Ethical Considerations

This study adhered to the guidelines stated in the World Medical Association's Declaration of Helsinki. Ethical approval was obtained from the Ethics Committee at the Medical Center of Zhejiang University (approval number 2014:1-017). Verbal informed consent was obtained from all participants before data collection. Participants were provided with clear and detailed information about the study and informed that their participation was voluntary and that they could withdraw at any time without consequence. All responses were kept anonymous, and each participant got an opportunity to ask for clarification about the survey questions. The survey data were securely stored in accordance with institutional guidelines, with access restricted to authorized personnel only. Participants typically spent around 10 minutes completing the survey. After a qualification check, they received a gift of renminbi (RMB) 10 (~US $1.46) after questionnaire completion.

Measurements

Dependent Variables

Participants were asked whether they currently smoked. Response options included the following: "Yes, smoke every day; "Yes, smoke on one or more days but not every day"; and "No." Those who answered yes were categorized as current smokers. Those who quit smoking were defined as individuals with a continuous or cumulative smoking history of at least 6 months or more who had abstained from smoking at the time of the study. Smoking intensity consisted of smoking amount and smoking duration. Quitting attempts were defined as having made at least three attempts to quit smoking, each lasting a

XSL•FO
RenderX

minimum of 3 days, without achieving sustained abstinence at the time of the survey [10].

USC is defined as smokers quitting on their own without professionally or pharmacologically mediated assistance; however, USC includes the use of over-the-counter agents, such as nicotine gum and patches. USC attempts refer to smoking cessation efforts made without professional support. USC success denotes cases in which individuals achieve sustained abstinence without using interventions such as guidance from health professionals, quitting clinics, hotlines, or other structured support resources [11,27].

Independent Variables

In this study, we measured the social network size using the network scale-up method. Participants were asked via surveys how many people they know from different subpopulations [28,29]. We created our questionnaire by appropriately considering Chinese culture and societal norms and making modifications accordingly. The questionnaire included 16 items under "total social network" (TSN) categorized into three subquestionnaires: (1) ordinary social network (OSN), which consisted of 6 items and asked "How many people do you know with the surname Ma, Zhu, or another surname?"; (2) lower social network (LSN), which consisted of 3 items and asked "How many people do you know who are seriously ill, victims of domestic violence, or drug users?"; and (3) higher social network (HSN), which consisted of 5 items and asked "How many people do you know who are company managers, professors, engineers, party and government officials, or physicians?" The network sizes were estimated by directly adding the number individuals known to participants [12].

We conducted comprehensive reliability and validity analyses for the TSN and its three subcomponents: OSN, HSN, and LSN. The TSN demonstrated excellent internal consistency, with Cronbach α=.91. Exploratory factor analysis (EFA) was performed on 14 observed variables, resulting in the extraction of three distinct factors. The item groupings within each factor corresponded closely with the theoretical expectations. Specifically, the OSN factor explained 33.5% of the total variance; the HSN factor, 12.2%; and the LSN factor, 9.8%, yielding a cumulative explained variance of 55.5%. This level of explained variance is considered acceptable, particularly in exploratory studies, where a threshold of 50% or higher is commonly recommended [30,31]. All items exhibited loadings above 0.65 within the OSN factor, above 0.59 within the HSN factor, and above 0.42 within the LSN factor. These values met or surpassed the commonly accepted threshold of 0.40 for factor loadings, which is widely used to assess construct validity in EFA [31,32].

In addition, EFA was conducted separately for each subscale. Each subscale yielded a single underlying factor, which explained 59% of the variance for the OSN, 50% for the HSN, and 56% for the LSN. According to established methodological guidelines, a single factor that accounts for at least 40% of the total variance is generally considered acceptable in exploratory research within the social sciences [31]. All items exhibited factor loadings above 0.48 in the OSN measure, above 0.47 in the HSN measure, and above 0.44 in the LSN measure. Factor loadings above 0.40 are generally considered acceptable indicators of construct validity [31,32]. In this study, all retained items met or surpassed this threshold, indicating satisfactory construct representation across all three subscales. Cronbach α for the OSN, HSN, and LSN subscales was .82, .68, and .79, respectively. Although Cronbach α for the LSN fell slightly below the conventional threshold of .70, it was sufficiently close to be deemed acceptable, particularly in our exploratory research settings, where slightly lower reliability is often tolerated [31-33]. All social network measures demonstrated acceptable reliability and construct validity, supporting the robustness of the measurement framework.

Within the TRA framework, several variables were included: social norms for smoking cessation, OSNs, LSNs, and USC. The latter three have already been described in this paper; social norms for smoking cessation were measured with the question, "Have any of your five closest friends quit smoking?" with response options coded as "Yes" and "No."

Covariates

Sociodemographic characteristics were age, ethnicity, educational level, occupation, family location, and household income. Family location was where the participants' families were located and was categorized into three types: rural area or township, county town or county-level city, and medium or large city. Household income data were collected by requesting participants to report the average income per person in their household for the previous year.

Statistical Analysis

All data were entered into a database using Microsoft Excel. The dataset was then imported into SAS (9.4 version) for statistical analyses. Descriptive statistics were calculated for USC adoption and success prevalence. Unadjusted logistic models were built for each primary predictor. SAS survey logistic procedures were applied in the analysis. The district was used as the clustering unit in order to account for a within-clustering correlation, attributable to the complex sample for analysis. Multiple logistic regression analysis was conducted to examine the association between social networks and USC adoption.

We constructed several models for the multivariable logistic regression analyses. For USC adoption, the models were (1) the base model, which included individual demographic variables to form (2) the demographic model, in which we entered the number of cigarettes smoked and different types of social networks to form (3) the number-of-cigarettes-smoked model, (4) the OSN model, (5) the LSN model, and (6) a comprehensive model encompassing all these.

To evaluate multicollinearity among the independent variables, variance inflation factors (VIFs) were calculated using standard linear regression procedures. Each predictor was included in the comprehensive model, and VIF values were computed to quantify the extent to which the variance of each estimated coefficient was inflated due to correlations with other predictors. A VIF greater than 10 is generally considered indicative of severe multicollinearity, whereas values below 5 are typically

XSL•FO
RenderX

regarded as acceptable, suggesting no serious multicollinearity concerns [34].

Model fit was evaluated using multiple information criteria, including −2 log likelihood, the Akaike Information Criterion (AIC), and the Schwarz Criterion (SC). Lower values of these indices indicate better model fit and were used to compare both nested and nonnested models. To assess potential overdispersion in the multivariable logistic regression models, the Pearson chi-square statistic divided by its degrees of freedom (χ^2/df) was examined. A χ^2/df ratio substantially greater than 1 is commonly interpreted as evidence of overdispersion, suggesting that the observed variability exceeds expectations under the assumed binomial distribution [35]. Conversely, a χ^2/df ratio below 3 is generally considered indicative of acceptable model fit and reasonable residual variance [36].

Subsequently, we applied SEM using the CALIS procedure in SAS to investigate the effects of reference group norms and social networks on USC. The analysis specifically examined whether reference group norms have a direct impact on USC or whether their influence is mediated through OSNs and LSNs. SEM, as a second-generation multivariate technique, allows for the simultaneous estimation of causal relationships among multiple variables. Parameter estimates were obtained using the maximum likelihood method, and model fit was assessed with standard indices, including the goodness-of-fit index (GFI>0.90), the normed fit index (NFI>0.90), and the root mean square error of approximation (RMSEA<0.10) [37].

All analyses were weighted. The weights included (1) sampling weights, calculated as the inverse of the probability of selection at the city and district levels and then multiplied together; (2) nonresponse weights, which accounted for household and individual factors; and (3) poststratification weights, derived using age groups (<25, 25-34, 35-44, 45-54, and ≥55 years) based on the estimated distributions of these characteristics from a national survey [32]. The final overall weights were computed as the product of these three weight sets.

Results

Participant Details

A total of 6500 individuals were identified as potential participants for this study, of whom 6010 (93.9%) agreed to participate, when contacted. Of the 6010 questionnaires, 5782 (96.2%) valid records were obtained. Participants' demographic characteristics are presented in Table 1. Of the 5782 participants, 2852 (49.3%) were smokers (95% CI 41.1%-48.5%). Among current smokers, 1647 (57.7%) had attempted to or had quit smoking, of whom 91.6% (n=1509; 95% CI 90.9%-97.5%) reported quitting without assistance, and 42.1% (n=58; 95% CI 32.4%-61.3%) of the remaining 138 (8.4%) participants who used USC methods achieved abstinence. Unadjusted logistic regression analysis showed that age, education level, marital status, occupation, number of cigarettes smoked, smoking duration, OSNs, and LSNs are significantly associated with USC adoption (see Table 1).

XSL•FO
RenderX

Table 1. Characteristics of participants who attempted to or had quit smoking (N=1647) and USC[a] prevalence.

Characteristics	Participants, n (%)	USC prevalence (%)	Unadjusted OR[b] (95% CI)
Age (years)			
<25	155 (9.8)	83.6	1.00
25-34	315 (16.8)	88.1	1.45 (0.98-2.99)
35-44	427 (19.4)	90.0	1.78 (0.88-3.61)
45-54	406 (23.3)	92.8	2.51 (1.94-3.25)[c]
≥55	344 (30.7)	96.1	4.78 (2.16-10.15)[c]
Ethnicity			
Han	1566 (95.6)	92.0	1.00
Minority	81 (4.4)	81.5	0.38 (0.16-0.92)[d]
Education			
Elementary school or less	158 (17.0)	98.8	1.00
Junior high school	434 (29.9)	90.3	0.12 (0.03-0.50)[c]
High school	481 (17.4)	91.2	0.09 (0.02-0.42)[c]
Junior college	344 (17.4)	91.4	0.13 (0.05-0.37)[c]
College or more	230 (12.2)	91.2	0.13 (0.03-0.68)[d]
Marital status			
Unmarried	298 (18.1)	86.3	1.00
Married	1248 (75.6)	92.4	1.94 (1.31-2.88)[c]
Divorced or widowed	101 (6.4)	96.1	3.86 (0.99-14.89)
Occupation			
Manager and service	208 (8.7)	95.0	1.00
Professional	140 (8.3)	87.6	0.37 (0.19-0.78)[c]
Commercial and social service	321 (18.4)	91.3	0.55 (0.21-1.47)
Technical worker	492 (29.7)	91.5	0.57 (0.25-1.29)
Operations	188 (15.7)	93.5	1.47 (0.46-4.71)
Retired	59 (3.9)	92.5	0.54 (0.27-1.07)
Student	74 (4.9)	90.7	0.49(0.15-1.61)
Other	165 (10.2)	94.7	0.32(0.10-1.01)
Income/person/year (RMB[e]; US $[f])			
<20,000; 2923.38	489 (30.1)	92.3	1.00
20,000-39,999; 2923.38-5846.61	504 (31.6)	91.3	0.88 (0.39-2.02)
40,000-59,999; 5846.76-8769.99	288 (16.2)	91.0	0.84 (0.41-1.72)
≥60,000; 8770.13	366 (22.0)	91.4	0.89 (0.57-1.41)
Number of cigarettes smoked			
<10	875 (56.5)	95.0	1.00
10-19	372 (18.3)	92.7	0.67 (0.31-1.42)
≥20	400 (25.2)	83.1	0.26 (0.14-0.49)[c]
Smoking time(year)			
<10	764 (51.7	93.6	1.00
10-19	194 (14.7	86.4	0.43 (0.18-1.10)

XSL•FO
RenderX

Characteristics	Participants, n (%)	USC prevalence (%)	Unadjusted OR[b] (95% CI)
20-29	300 (14.8	90.3	0.64 (0.28-1.47)
≥30	289 (18.7	91.2	0.72 (0.41-1.25)
TSN[g] size			
<10	204 (14.8)	94.2	1.00
10-19	466 (26.6)	93.3	0.86 (0.36-2.05)
20-29	428 (25.3)	90.0	0.62 (0.34-1.11)
≥30	549 (33.3)	89.4	0.53 (0.27-1.04)
OSN[h] size			
<10	218 (15.1)	97.3	1.00
10-14	448 (25.5)	91.2	0.29 (0.08-0.96)[d]
15-19	483 (30.0)	89.5	0.23 (0.08-0.72)[c]
≥20	498 (29.4)	91.0	0.28 (0.09-0.84)[d]
LSN[i] size			
0	773 (48.6)	94.5	1.00
1-2	512 (29.0)	93.0	0.76 (0.43-1.40)
≥3	362 (22.4)	83.4	0.29 (0.18-0.46)[c]
HSN[j] size			
<5	339 (29.0)	92.4	1.00
5-9	309 (26.0)	92.9	1.07 (0.57-2.01)
10-14	224 (26.0)	89.8	0.71 (0.59-1.33)
≥15	240 (24.0)	90.2	0.87 (0.31-1.47)

[a]USC: unassisted smoking cessation.
[b]OR: odds ratio.
[c]$P<.01$.
[d]$P<.05$.
[e]RMB: renminbi.
[f]The currency exchange rate applied was RMB 1=US $0.15.
[g]TSN: total social network.
[h]OSN: ordinary social network.
[i]LSN: lower social network.
[j]HSN: higher social network.

VIF analysis revealed that all variables had values below 5, ranging from a minimum of 1.3 to a maximum of 4.7, which are generally considered acceptable. These results support the reliability of the regression estimates and indicate that multicollinearity is not a major concern in the current model specification.

Model fit statistics indicated progressive improvement from the demographic model to the OSN model and the LSN model. The −2 log likelihood decreased from 3534.6 in the demographic model to 3456.7 in the OSN model and 3461.5 in the LSN model, accompanied by corresponding reductions in the AIC and SC values, with the former outperforming the latter. These results suggest that the inclusion of social network covariates substantially enhanced model fit. Notably, the χ^2/df ratio decreased across models: 1.32 for the base model, 0.96 for the demographic model, and 0.93 for the comprehensive model. These values indicated satisfactory model fit and suggested that the overall residual variation was within an acceptable range.

Multiple logistic regression analysis revealed that older age was linked to increased adoption of USC in the 45-54–year age group (odds ratio [OR] 2.16, 95% CI 1.05-4.45) compared to the <25-year age group (see Table 2). Those in junior high school exhibited lower USC adoption (OR 0.57, 95% CI 0.39-0.85) compared to the reference group. The number of cigarettes smoked was significantly correlated with lower adoption of USC in the 10-19-cigarette (OR 0.66, 95% CI 0.44-0.97) and 20-cigarette (OR 0.51, 95% CI 0.41-0.63) groups. Higher OSN sizes were significantly associated with lower adoption of USC (OR 0.23, 95% CI 0.10-0.80), (OR 0.22, 95% CI 0.07-0.60), and (OR 0.25, 95% CI 0.08-0.72). Similarly, a

XSL•FO
RenderX

high LSN size was significantly associated with lower adoption of USC 0.32 (95% CI 0.23-0.45).

Table 2. Multivariable analysis of social net position and USC[a] adoption.

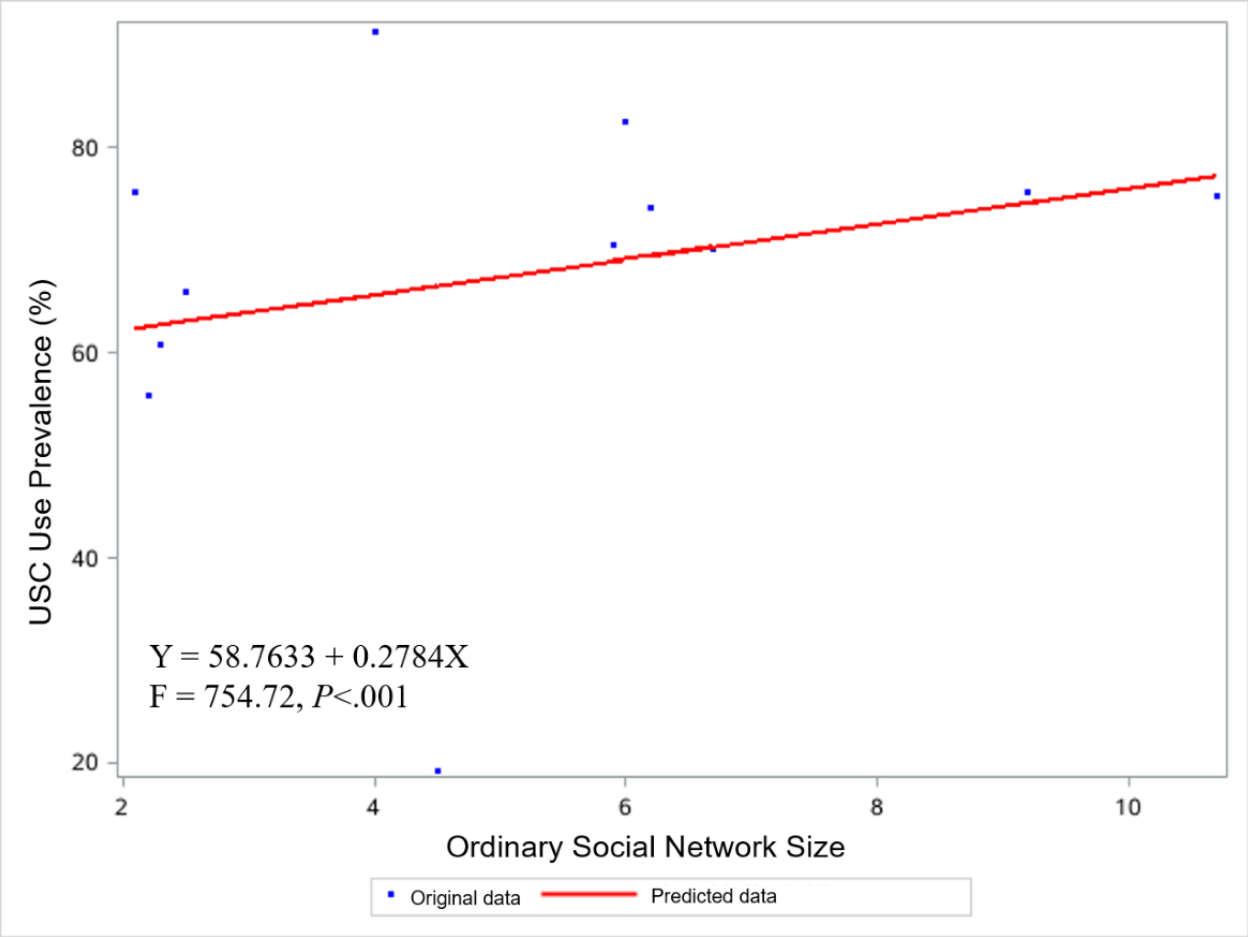
Characteristics	Demographic model	Number-of-cigarettes-smoked model, OR[b] (95% CI)	OSN[c] model, OR (95% CI)	LSN[d] model, OR (95% CI)	Comprehensive model, OR (95% CI)
Age (years)					
<25	1.00	1.00	1.00	1.00	1.00
25-34	1.43 (0.54-3.77)	1.91 (1.14-3.20)[e]	2.08 (1.28-3.39)[e]	1.88 (1.08-3.29)[f]	1.64 (1.03-2.52)[f]
35-44	0.89 (0.36-2.21)	1.23 (0.86-1.74)	2.43 (1.04-5.56)[f]	2.25 (0.92-5.20)	2.06 (1.20-3.53)[f]
45-54	2.16 (1.05-4.45)[f]	2.99 (1.68-5.33)[e]	3.78 (2.45-5.83)[e]	3.00 (1.84-4.74)[e]	3.15 (2.03-4.91)[e]
≥55	1.66 (0.81-3.43)	2.06 (1.51-2.83)[e]	4.52 (1.77-11.52)[e]	4.01 (1.47-10.89)[e]	4.43 (1.57-11.31)[e]
Education					
Elementary school or less	1.00	1.00	1.00	1.00	1.00
Junior high school	0.57 (0.39-0.85)[e]	0.60 (0.38-0.95)[f]	0.14 (0.03-0.64)[e]	0.14 (0.03-0.68)[f]	0.13 (0.02-0.76)[f]
High school	1.37 (0.60-3.15)	1.50 (0.55-4.09)	0.15 (0.03-0.75)[f]	0.15 (0.03-0.74)[f]	0.12 (0.02-0.78)[f]
Junior college	1.04 (0.30-3.65)	1.08 (0.27-4.30)	0.19 (0.05-0.69)[f]	0.20 (0.06-0.69)[f]	0.18 (0.04-0.71)[f]
College and more	0.96 (0.77-1.18)	0.90 (0.69-1.17)	0.21 (0.03-1.57)	0.22 (0.03-1.54)	0.19 (0.02-1.75)
Number of cigarettes smoked					
<10	__[g]	1.00	1.00	1.00	1.00
10-19	—	0.66 (0.44-0.97)[f]	0.81 (0.45-1.47)	0.74 (0.41-1.35)	0.67 (0.37-1.22)
≥20	—	0.51 (0.41-0.63)[e]	0.25 (0.12-0.52)[e]	0.26 (0.12-0.57)[e]	0.24 (0.11-0.59)[e]
OSN size					
<10 (low)	—	—	1.00	—	1.00
10-14 (medium)	—	—	0.23 (0.10-0.80)[f]	—	0.37 (0.13-1.06)
15-19 (high)	—	—	0.22 (0.07-0.60)[e]	—	0.25 (0.08-0.81)[f]
≥20 (very high)	—	—	0.25 (0.08-0.72)[e]	—	0.26 (0.08-0.82)[f]
LSN size					
0	—	—	—	1.00	1.00
1-2 (medium)	—	—	—	0.77 (0.39-1.37)	0.85 (0.47-1.54)
≥3 (high)	—	—	—	0.32 (0.23-0.45)[e]	0.36 (0.27-0.50)[e]

[a]USC: unassisted smoking cessation.
[b]OR: odds ratio.
[c]OSN: ordinary social network.
[d]LSN: lower social network.
[e]$P<.01$.
[f]$P<.05$.
[g]Not applicable.

The analysis revealed a significant positive dose-response relationship of city-level OSN and LSN sizes with the probability of USC adoption, as indicated by regression coefficients β=0.2784 ($P<.01$) and β=2.2269 ($P<.01$), respectively (see Figures 1 and 2).

XSL•FO
RenderX

Figure 1. Scatter plot showing the relationship between city-level OSN size and the prevalence of USC adoption. OSN: ordinary social network; USC: unassisted smoking cessation.



XSL•FO
RenderX

Figure 2. Scatter plot showing the relationship between city-level LSN size and the prevalence of USC adoption. OSN: lower social network; USC: unassisted smoking cessation.

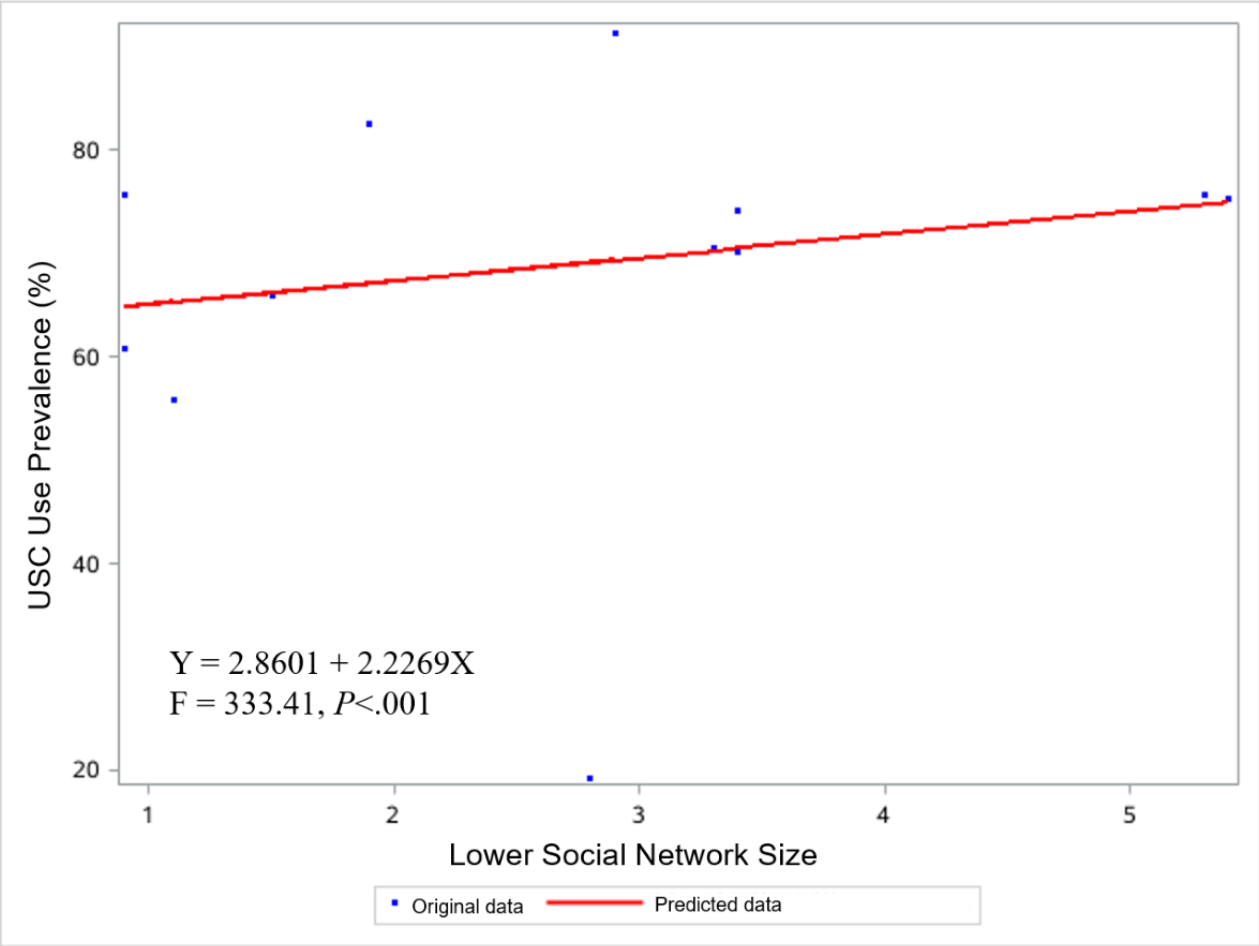


Table 3 presents the means (SDs) of the variables used in model construction. SEM analysis demonstrated good model fit, with GFI=0.976, NFI=0.954, and RMSEA=0.0063. Results also indicated that reference group norms had both direct effects on USC adoption (β=0.0669, *P*<.01) and indirect effects through OSNs and LSNs, with the indirect effect being β=0.2315×0.4613 (*P*<.01). In contrast, bigger OSN and LSN sizes also had a direct association with USC adoption (β=0.4613, *P*<.01), as illustrated in Figure 3.

Table 3. Means (SDs) of variables included in constructing models.

Variable	Mean (SD)
LSN[a] size	4.05 (1.36)
OSN[b] size	7.64 (2.41)
USC[c]	1.75 (1.69)
Norm[d]	1.55 (0.44)

[a]LSN: lower social network.

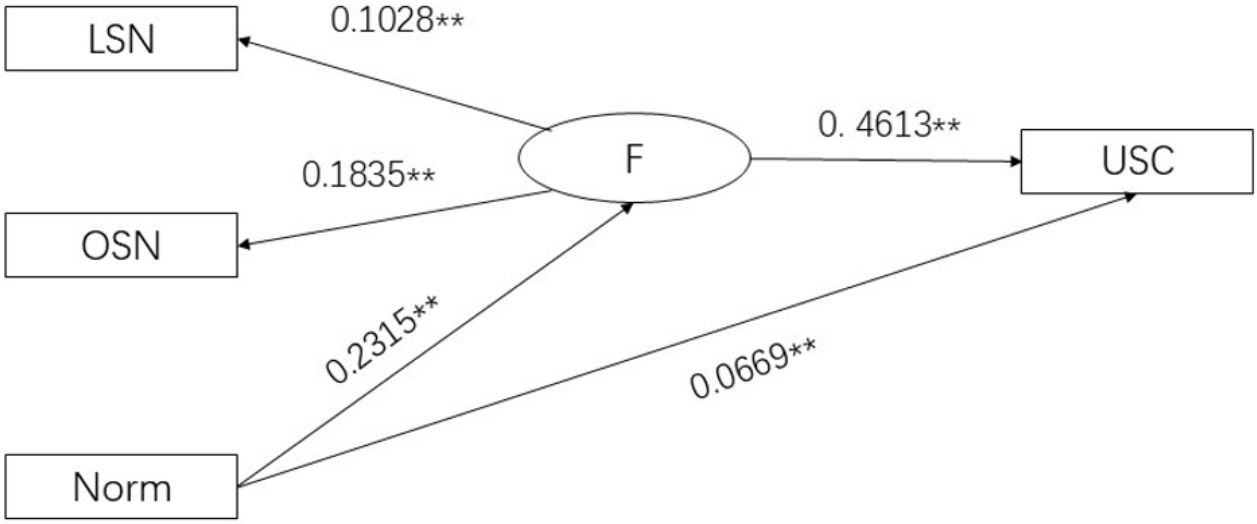
[b]OSN: ordinary social network.

[c]USC: unassisted smoking cessation.

[d]Norm: subjective norm, perceived social pressure to engage in or refrain from a behavior.

XSL•FO
RenderX

Figure 3. Results of SEM analysis. ***P*<.05. F: social network size; LSN: lower social network; Norm: subjective norm, perceived social pressure to engage in or refrain from a behavior; OSN: ordinary social network; SEM: structural equation modeling; USC: unassisted smoking cessation.



Discussion

Principal Findings

This study found that social networks are significantly related to USC adoption. China produces and consumes more tobacco than any other country in the world and as such experiences high levels of tobacco-related disease. Many studies have recently directly or indirectly referenced Chinese culture as a major contributor to China's high tobacco usage [5]. In contrast to Western culture, Chinese culture has historically seen smoking as a form of social communication. The practices of gifting and sharing cigarettes are well accepted and pervasive across China [5]. The Chinese place a high value on interpersonal relationships and connections. These connections can be mutually beneficial. Central to the idea of "relationship" is the concept of reciprocity. If someone receives a gift but fails to reciprocate appropriately, that will not only lead to a loss of face but also threaten that person's "relationship" and social capital. Accepting a cigarette can symbolize a person's willingness to engage in a future business partnership, illicit or otherwise [38]. Given the importance of social networks, they need to be associated with smoking or quitting.

The findings of this study indicate that although HSNs have no significant association with USC adoption, both OSNs and LSNs are negatively related to the likelihood of adopting USC. Furthermore, this study found a significant positive dose-response relationship of city-level OSN and LSN sizes with the probability of adopting USC. These results suggest that the structure and composition of a smoker's social network may play a crucial role in determining their smoking cessation strategies. According to TRA, reference group norms and social expectations significantly affect behavioral intentions. Social networks exert a strong effect on smoking behaviors and cessation outcomes, as individuals embedded in smoking-prevalent circles are more likely to maintain smoking habits or face challenges in quitting because of constant exposure to smoking cues, such as friends who smoke or social gatherings where tobacco use is normative. Those within OSN and LSN environments may experience stronger social reinforcement of smoking and fewer supportive cues for quitting, which, in turn, reduces their likelihood of successful USC. Peer influence in such networks often normalizes smoking and undermines self-initiated quitting efforts, especially when external interventions or alternative social incentives are lacking. Several mechanisms may further explain this pattern. First, the smoking cessation behaviors of socioeconomic groups are often more significantly affected by collective norms and the sharing of peer experiences. When an individual observes others around them successfully quitting smoking with the aid of cessation tools (eg, nicotine replacement products, quit lines, or the doctor's advice), they are more likely to imitate this behavior. Thus, within OSNs and LSNs, the diffusion of smoking cessation assistance behaviors tends to be more powerful than quitting through willpower alone [22,39]. Moreover, individuals from socioeconomically disadvantaged groups tend to report lower self-efficacy in managing health behaviors, making them more inclined to seek professional help instead of relying solely on self-control [40,41]. Finally, these individuals also tend to experience a higher burden of health problems and therefore have more frequent interactions with medical and public health services [42]. Such increased contact may enhance their exposure to assisted smoking cessation programs, which helps explain why smokers embedded in OSNs and LSNs are more likely to seek professional support for smoking cessation and less likely to engage in USC. Taken together, these findings suggest that low-status social environments may influence smoking cessation behaviors, not simply through socioeconomic disadvantage, but also through network-based pathways that shape access to institutional and social resources.

This study found that reference group norms exert both a direct effect on USC and indirect effects through OSNs and LSNs. By integrating social network dynamics into TRA, this approach enriches its explanatory power and highlights the importance of interpersonal influences in shaping behavioral intentions [12,39]. Academically, the findings contribute to the literature on social influence and health behavior by demonstrating that normative pressures are not only transmitted directly but also

XSL•FO
RenderX

mediated through network structures. Practically, the results provide valuable insights for intervention design: by identifying key network configurations and normative pathways, policymakers and health practitioners can more effectively promote smoking cessation behaviors. This underscores the role of social networks in embedding reference group norms within behavioral decision-making processes, suggesting that interventions targeting peer groups and social ties may yield greater success in reducing smoking prevalence.

In this study, 91.6% of those who quit smoking reported quitting without assistance, a percentage similar to that found in another study (93.1%) by the Chinese Center for Disease Control and Prevention (CDC) [26]. USC adoption is much higher in China than in the Western world. In Australia, Smith et al [43] found that between 54% and 69% of former smokers had quit without professional help, while 41%-58% of current smokers had tried to quit without assistance. In the United States, earlier research on successful attempts to quit smoking found that unassisted quit rates vary between 64% and 78% [44,45]. In Canada, a study found that Chinese immigrant smokers infrequently use cessation aids or services [46]. Furthermore, the prevalence of USC success was high (42.1%) in this study, which is also much higher than that in Western society [11,45]. This may reflect differences in cultural norms. Chinese culture, to a large extent, still adheres to agrarian social mores, which emphasize individual spirit and perseverance in coping with behavioral problems [12]. Thus, quitting smoking on their own rather than relying on professional assistance reflects an individual's strength and willpower. Based on this cultural model for smoking cessation, decision makers need to adapt to the situation and actively promote USC among the Chinse population.

The findings reveal that socioeconomic and demographic factors play a complex role in the adoption of USC. Although initial analyses identified ethnicity, education level, and marital status as significant correlates of USC adoption, only education remained an independent predictor when these factors were examined concurrently in a multivariate model. Specifically, compared to individuals with an elementary school education or below, those with junior high school education showed a significantly lower adoption rate of USC. This suggests that the apparent associations of ethnicity and marital status with USC adoption may be largely confounded by or mediated through educational attainment. Higher education likely confers advantages in health awareness and behavioral capacity that enable individuals to attempt and succeed in quitting smoking without formal assistance, regardless of their ethnic background or marital status. This study found an increase in USC adoption with age, which aligns with findings from other studies as well [11,38]. This may reflect the fact that as individuals age, the prevalence of health problems may increase, leading to a greater need and motivation to quit smoking [5]. Older smokers grew up in an era of relative scarcity of material and medical resources, making them more inclined to rely on personal willpower rather than external aids to solve problems. Nicotine replacement therapy (NRT), quit lines, and digital health tools have only become widespread in recent decades. Older adults may be less familiar with these newer methods or skeptical about their safety and effectiveness, leading them to place greater trust in traditional, self-driven approaches to quitting smoking. It is interesting to note that in this study, the smoking amount was linked to the choice of unassisted quitting, and heavy smokers were less likely to quit smoking independently. This aligns with findings from other studies [11,45]. Perhaps this is because heavy smokers typically have higher nicotine dependence, which can hinder their ability to quit smoking independently, leading to a lower prevalence of USC adoption.

This study revealed that the smoking prevalence was 44.8%. Among smokers, 57.8% quit smoking, with 91.6% of them not using assistance. Among the USC users, 42.1% successfully achieved abstinence. This indicates that the USC success prevalence among smokers was 22.3%. According to the Chinese CDC survey of 2018, more than 308 million adults in China were current smokers [47]. Using the USC success prevalence from this study, approximately 68.7 million Chinese smokers could be projected to achieve abstinence if they try USC. A study [26] using direct USC numbers found that USC contributed to successful smoking cessation by approximately 70 million smokers in China. It is worth noting that the numbers from both sources are similar, indicating their reliability. It is indeed a massive number of those who quit smoking, and if this were to happen, it would produce positive changes in their health, lifestyle, and overall well-being. It is widely recognized that clinical and other individual-level smoking cessation methods are insufficient to fundamentally reduce smoking prevalence at the population level. Although such interventions can achieve high success rates among participants, their overall impact on national smoking rates remains limited, as they fail to reach the majority of smokers. More importantly, individual-oriented strategies often do not adequately address broader socioeconomic and cultural factors that contribute to sustained smoking behavior in China. These include aggressive marketing by the tobacco industry; social norms that tolerate, and even encourage, smoking by males; the use of cigarettes in social rituals; and the affordability of tobacco products. Without complementary population-level measures, such as those recommended by the WHO FCTC, including substantial tobacco tax increases, comprehensive bans on tobacco advertising, and the implementation of smoke-free laws, the effectiveness of stand-alone clinical interventions remains significantly constrained. Such policies are essential to reshaping social, economic, and environmental conditions to make smoking less attractive and accessible. This study provides evidence that USC adoption can effectively reduce the smoking prevalence in the population. However, this approach seems to have been overlooked by many Chinese health professionals and government officials. The WHO FCTC requires parties to promote tobacco cessation and implement effective measures to help tobacco users quit [4]. In China, smoking cessation assistance by health care professionals is common and receives considerable attention from the government. However, USC is often overlooked by tobacco control agencies, as they tend to consider it nonprofessional. USC can, in fact, be strengthened through professional strategies, such as implementing smoke-free environments, launching large-scale health education campaigns, and using media-based cessation interventions. Notably, engagement in USC may also heighten individuals' awareness of, and interest in, professional smoking cessation

XSL•FO
RenderX

support services. What we urgently need to do is improve the understanding of the effectiveness of USC and enhance professional services for USC to increase the number of people trying to quit smoking and to enhance USC success rates.

Study Limitations

This study has a few limitations. First, as this study used a cross-sectional design, causal relationships cannot be established. Future longitudinal or experimental studies are needed to further explore causal mechanisms underlying these relationships. Although USC may promote attempts to quit smoking or short-term smoking cessation success, it does not necessarily lead to sustained long-term abstinence. The issue of long-term smoking cessation outcomes, along with comparisons between USC and other established smoking cessation support measures, particularly those aligned with Article 14 of the WHO FCTC, remains a critical area for future research.

Second, we assessed the social network size using the network scale-up method, which estimates the number of individuals a respondent knows. This approach has been widely applied in population studies and social network research to approximate the personal network size [28,29]. We acknowledge that this approach primarily captures the breadth of social connections rather than their depth, intensity, or emotional closeness. Although this method may not fully reflect the strength or quality of relationships, it provides a standardized and efficient means of estimating the network size across large populations. Given the scope and objectives of our study, this method was appropriate for our analytical framework. Nonetheless, we recognize the limitations of this approach and suggest that future research incorporate complementary measures, such as tie strength, frequency of interaction, or perceived support, to provide a more nuanced understanding of social networks.

Third, the occupational categories used in the LSN and HSN measures are relatively limited, as they were specifically designed to capture culturally recognized social groups within the Chinese context. Although this targeted approach facilitates the identification of low- and high-status networks, it may not fully encompass the broader diversity of participants' social connections. Although reliability and validity analyses indicate strong psychometric performance, it remains possible that certain subpopulations were insufficiently represented. To enhance representativeness, future research should consider broadening the occupational scope of these measures.

Fourth, the OSN measure is based on the network scale-up method, which estimates the social network size by asking respondents how many people they know with certain common surnames. This approach has been widely used in social network research and has demonstrated validity in estimating general network exposure [48]. In our adaptation, we selected the most prevalent Chinese surnames to ensure cultural relevance. This method provides a practical proxy for the network size, but we recognize that relying solely on surname recognition may underestimate the complexity and richness of actual social networks, especially in urban or highly mobile populations.

Conclusion

This study demonstrates that USC is the predominant method used by Chinese smokers. It further reveals that individuals embedded in OSNs or LSNs are less likely to adopt USC and that these networks exert indirect effects within the structure linking reference group norms and USC adoption, underscoring the influence of social environments on smoking cessation behavior. Mass media campaigns, public health initiatives, and the promotion of smoke-free environments are all effective ways to encourage the use of smoking cessation strategies, which lead to changes within social networks, ultimately resulting in shifts in network norms and USC adoption. For further tobacco control, policies should move beyond treatment provision toward reshaping social environments and network norms, as sustained reductions in smoking prevalence depend on cultural denormalization rather than clinical service expansion alone.

Acknowledgments

We would like to thank all the participants of this study and our partners in data collection.

Funding

The analysis resource of this research was partly funded by the Young Talent Training Program of the Guangdong Association for Science and Technology (grant number SKXRC2025158). The study design was partly funded by the National Natural Science Foundation of China (grant number 71473221). Data management also was partly funded by the Bloomberg Initiative to Reduce Tobacco Use Project (grant number China-37-06). The funders had no involvement in the interpretation and writing of the manuscript.

Data Availability

The datasets used and analyzed during the study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization, data curation, and project administration were handled by SP; conceptualization, methodology, resources, formal analysis, and writing—original draft by TY; data curation and project administration by LL; investigation and writing—review and editing by WZ; and validation and writing—review and editing by RRC. All authors have read and approved

XSL•FO
RenderX

the final version of the manuscript. The authors declare the use of generative artificial intelligence (AI) in the research and writing process. According to GAIDeT (Generative AI Delegation Taxonomy, 2025), the following tasks were delegated to generative AI (GAI) tools under full human supervision: proofreading and editing. The GAI tool used was ChatGpt-5.2. Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes. We used ChatGPT-5.2 solely for English language grammar check, under full human supervision. No scientific content was generated by AI.

Conflicts of Interest

None declared.

References

1. Schoberberger R, Böhm G, Schroeder Y. Heavy dependent nicotine smokers--newfound lifestyle appreciation after quitting successfully. Experiences from inpatient smoking cessation therapy. Public Health 2015 May;129(5):539-544 [FREE Full text] [doi: 10.1016/j.puhe.2015.02.011] [Medline: 25773313]
2. Tonstad S, Tønnesen P, Hajek P, Williams KE, Billing CB, Reeves KR, Varenicline Phase 3 Study Group. Effect of maintenance therapy with varenicline on smoking cessation: a randomized controlled trial. JAMA 2006 Jul 05;296(1):64-71 [FREE Full text] [doi: 10.1001/jama.296.1.64] [Medline: 16820548]
3. Hand S, Edwards S, Campbell IA, Cannings R. Controlled trial of three weeks nicotine replacement treatment in hospital patients also given advice and support. Thorax 2002 Aug;57(8):715-718 [FREE Full text] [doi: 10.1136/thorax.57.8.715] [Medline: 12149533]
4. World Health Organization. WHO Framework Convention on Tobacco Control. Geneva: WHO Document Production Services; 2005.
5. Yang T, Yang G. Smoking cessation strategies in tobacco control in China. J Tuberc Lung Health 2012;1(1):45-47.
6. Jiang S, Yang T, Bullen C, Chen J, Yu L, Peng S, et al. Real-world unassisted quit success and related contextual factors: a population-based study of Chinese male smokers. Tob Control 2021 Sep;30(5):498-504 [FREE Full text] [doi: 10.1136/tobaccocontrol-2019-055594] [Medline: 32675250]
7. Yan Y, Lin B, Xu Q, Xie H, Zeng X, Di X, et al. Utilization of smoking cessation support among adults - 18 PLADs, China, 2020. China CDC Wkly 2023 May 26;5(21):459-463 [FREE Full text] [doi: 10.46234/ccdcw2023.087] [Medline: 37405263]
8. Smith AL, Chapman S, Dunlop SM. What do we know about unassisted smoking cessation in Australia? A systematic review, 2005-2012. Tob Control 2015 Jan;24(1):18-27 [FREE Full text] [doi: 10.1136/tobaccocontrol-2013-051019] [Medline: 24026163]
9. Soulakova JN, Crockett LJ. Unassisted quitting and smoking cessation methods used in the United States: analyses of 2010-2011 tobacco use supplement to the current population survey data. Nicotine Tob Res 2017 Dec 13;20(1):30-39 [FREE Full text] [doi: 10.1093/ntr/ntw273] [Medline: 27798084]
10. Yang T, Zhu Z, Barnett R, Zhang W, Jiang S. Tobacco advertising, anti-tobacco information exposure, environmental smoking restrictions, and unassisted smoking cessation among Chinese male smokers: a population-based study. Am J Mens Health 2019 Jun 11;13(3):1557988319856152 [FREE Full text] [doi: 10.1177/1557988319856152] [Medline: 31185783]
11. Zhu S, Melcer T, Sun J, Rosbrook B, Pierce JP. Smoking cessation with and without assistance: a population-based analysis. Am J Prev Med 2000 May;18(4):305-311 [FREE Full text] [doi: 10.1016/s0749-3797(00)00124-0] [Medline: 10788733]
12. Yang T. Health Research: Social and Behavioral Theories and Methods, 3rd ed. Beijing: People's Medical Publishing House; 2025.
13. Baer M. The strength-of-weak-ties perspective on creativity: a comprehensive examination and extension. J Appl Psychol 2010 May;95(3):592-601 [FREE Full text] [doi: 10.1037/a0018761] [Medline: 20476837]
14. Yang XY, Kelly BC, Yang T. The influence of self-exempting beliefs and social networks on daily smoking: a mediation relationship explored. Psychol Addict Behav 2014 Sep;28(3):921-927 [FREE Full text] [doi: 10.1037/a0037176] [Medline: 25134037]
15. Abel G, Plumridge L, Graham P. Peers, networks or relationships: strategies for understanding social dynamics as determinants of smoking behaviour. Drugs: Educ Prev Policy 2009 Jul 10;9(4):325-338 [FREE Full text] [doi: 10.1080/09687630210157636]
16. Fang X, Li X, Stanton B, Dong Q. Social network positions and smoking experimentation among Chinese adolescents. Am J Health Behav 2003;27(3):257-267 [FREE Full text] [doi: 10.5993/ajhb.27.3.7] [Medline: 12751622]
17. Pearson M, Michell L. Smoke rings: social network analysis of friendship groups, smoking and drug-taking. Drugs: Educ Prev Policy 2000;7(1):21-37 [FREE Full text] [doi: 10.1080/dep.7.1.21.37]
18. Seo D, Huang Y. Systematic review of social network analysis in adolescent cigarette smoking behavior. J Sch Health 2012 Jan 06;82(1):21-27 [FREE Full text] [doi: 10.1111/j.1746-1561.2011.00663.x] [Medline: 22142171]
19. Ross B, Tingzhong Y, Xiaozhao Y. Smoking Environments in China: Challenges for Tobacco Control. Switzerland: Springer Nature; 2021.

XSL•FO
RenderX

20. Pan Z. Socioeconomic predictors of smoking and smoking frequency in urban China: evidence of smoking as a social function. Health Promot Int 2004 Sep;19(3):309-315 [FREE Full text] [doi: 10.1093/heapro/dah304] [Medline: 15306615]
21. Yang T, Fisher KJ, Li F, Danaher BG. Attitudes to smoking cessation and triggers to relapse among Chinese male smokers. BMC Public Health 2006 Mar 14;6(1):65 [FREE Full text] [doi: 10.1186/1471-2458-6-65] [Medline: 16533411]
22. Christakis NA, Fowler JH. The collective dynamics of smoking in a large social network. N Engl J Med 2008 May 22;358(21):2249-2258 [FREE Full text] [doi: 10.1056/NEJMsa0706154] [Medline: 18499567]
23. Hoffman BR, Sussman S, Unger JB, Valente TW. Peer influences on adolescent cigarette smoking: a theoretical review of the literature. Subst Use Misuse 2006;41(1):103-155 [FREE Full text] [doi: 10.1080/10826080500368892] [Medline: 16393739]
24. Montes F, Blanco M, Useche AF, Sanchez-Franco S, Caro C, Tong L, et al. Exploring the mechanistic pathways of how social network influences social norms in adolescent smoking prevention interventions. Sci Rep 2023 Feb 21;13(1):3017 [FREE Full text] [doi: 10.1038/s41598-023-28161-7] [Medline: 36810585]
25. Sohn M, Moon D, Kim J. A longitudinal cohort study of adolescent social network positions and lifetime daily smoking and nicotine dependence. Youth Soc 2023 Apr 05;56(1):217-238 [FREE Full text] [doi: 10.1177/0044118x231164742]
26. World Health Organization. Global adult tobacco survey (GATS) China 2010 country report. 2011. URL: https://cdn.who.int/media/docs/default-source/ncds/ncd-surveillance/data-reporting/china/gats/en-tfi-china-gats-factsheet-2010.pdf [accessed 2026-02-27]
27. Williams MB, Beebe LA, Neas BR. State-level correlates of unassisted quit attempts and success. J Okla State Med Assoc 2015 Nov;108(11):455-462 [FREE Full text] [Medline: 26817062]
28. Maltiel R, Raftery AE, McCormick TH, Baraff AJ. Estimating population size using the network scale up method. Ann Appl Stat 2015 Sep;9(3):1247-1277 [FREE Full text] [doi: 10.1214/15-AOAS827] [Medline: 26949438]
29. McCormick TH, Salganik MJ, Zheng T. How many people do you know?: efficiently estimating personal network size. J Am Stat Assoc 2010 Mar 01;105(489):59-70 [FREE Full text] [doi: 10.1198/jasa.2009.ap08518] [Medline: 23729943]
30. Fabrigar LR, Wegener DT, MacCallum RC, Strahan EJ. Evaluating the use of exploratory factor analysis in psychological research. Psychol Methods 1999 Sep;4(3):272-299 [FREE Full text] [doi: 10.1037/1082-989X.4.3.272]
31. Hair J. Multivariate data analysis: an overview. In: Lovric M, editor. International Encyclopedia of Statistical Science. Berlin, Heidelberg: Springer; 2011:904-907.
32. Stevens J. Applied Multivariate Statistics for the Social Sciences, 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1992.
33. Taber KS. The use of Cronbach's alpha when developing and reporting research instruments in science education. Res Sci Educ 2017 Jun 7;48(6):1273-1296 [FREE Full text] [doi: 10.1007/s11165-016-9602-2]
34. Gareth J, Daniela W, Trevor H, Tibshirani R. An Introduction to Statistical Learning. New York: Springer; 2021.
35. Allison P. Logistic Regression Using SAS: Theory and Application, 2nd ed. Cary, NC: SAS Institute; 2012.
36. Sommet N, Morselli D. Keep calm and learn multilevel logistic modeling: a simplified three-step procedure using Stata, R, Mplus, and SPSS. Int Rev Soc Psychol 2017;30(1):203-218 [FREE Full text] [doi: 10.5334/irsp.90]
37. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Model 1999 Jan;6(1):1-55 [FREE Full text] [doi: 10.1080/10705519909540118]
38. Wank D. Cigarettes and domination in Chinese business networks: institutional change during the market transition. In: Davis D, editor. The Consumer Revolution in Urban China. Berkeley, CA: University of California Press; 2000.
39. Valente TW. Social Networks and Health: Models, Methods, and Applications. New York: Oxford University Press; 2010.
40. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977;84(2):191-215 [FREE Full text] [doi: 10.1037/0033-295X.84.2.191]
41. Hiscock R, Bauld L, Amos A, Fidler JA, Munafò M. Socioeconomic status and smoking: a review. Ann N Y Acad Sci 2012 Feb;1248:107-123 [FREE Full text] [doi: 10.1111/j.1749-6632.2011.06202.x] [Medline: 22092035]
42. Brown T, Platt S, Amos A. Equity impact of European individual-level smoking cessation interventions to reduce smoking in adults: a systematic review. Eur J Public Health 2014 Aug;24(4):551-556 [FREE Full text] [doi: 10.1093/eurpub/cku065] [Medline: 24891458]
43. Smith AL, Carter SM, Chapman S, Dunlop SM, Freeman B. Why do smokers try to quit without medication or counselling? A qualitative study with ex-smokers. BMJ Open 2015 Apr 30;5(4):e007301 [FREE Full text] [doi: 10.1136/bmjopen-2014-007301] [Medline: 25933811]
44. Cokkinides VE, Ward E, Jemal A, Thun MJ. Under-use of smoking-cessation treatments: results from the National Health Interview Survey, 2000. Am J Prev Med 2005 Jan;28(1):119-122 [FREE Full text] [doi: 10.1016/j.amepre.2004.09.007] [Medline: 15626567]
45. Shiffman S, Brockwell SE, Pillitteri JL, Gitchell JG. Use of smoking-cessation treatments in the United States. Am J Prev Med 2008 Feb;34(2):102-111 [FREE Full text] [doi: 10.1016/j.amepre.2007.09.033] [Medline: 18201639]
46. Mao A, Bottorff JL. A qualitative study on unassisted smoking cessation among Chinese Canadian immigrants. Am J Mens Health 2017 Nov;11(6):1703-1712 [FREE Full text] [doi: 10.1177/1557988315627140] [Medline: 26819181]

XSL•FO
RenderX

47. Xiao L, Nan Y, Di XB, Meng ZD. Study on smoking behavior and its changes among Chinese people aged 15 years and above in 2018. Zhonghua Liu Xing Bing Xue Za Zhi 2022 Jun 10;43(6):811-817 [FREE Full text] [doi: 10.3760/cma.j.cn112338-20211130-00934] [Medline: 35725334]
48. Clay-Warner J, Kawashima T, Edgemon TG. Measure of personal network size using the known population method: a methodological guide. Am J Public Health 2022 May 03;112(5):747-753 [FREE Full text] [doi: 10.2105/AJPH.2022.306731] [Medline: 35298281]

Abbreviations

AIC: Akaike Information Criterion
CDC: Center for Disease Control and Prevention
EFA: exploratory factor analysis
FCTC: Framework Convention on Tobacco Control
GFI: goodness-of-fit index
HSN: higher social network
LSN: lower social network
NFI: normed fit index
OR: odds ratio
OSN: ordinary social network
RMSEA: root mean square error of approximation
SC: Schwarz Criterion
SEM: structural equation modeling
TRA: theory of reasoned action
TSN: total social network
USC: unassisted smoking cessation
VIF: variance inflation factor
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 18.Mar.2025; peer-reviewed by YTD Cheung, H Tateno; comments to author 01.Sep.2025; revised version received 12.Feb.2026; accepted 24.Feb.2026; published 25.Mar.2026.

Please cite as:
Peng S, Yang T, Li L, Zhang W, Cottrell RR
Social Networks and Their Influence on the Choice of Unassisted Smoking Cessation: Cross-Sectional Study in Six Cities in China
JMIR Public Health Surveill 2026;12:e74147
URL: https://publichealth.jmir.org/2026/1/e74147
doi:10.2196/74147
PMID:

©Sihui Peng, Tingzhong Yang, Lijing Li, Weifang Zhang, Randall R Cottrell. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 25.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Urinary Sodium Excretion and the Risk of Prevalent Anemia: Nationwide Population-Based Cross-Sectional Study

Sang Heon Suh[1,2*], MD, PhD; Dohyeon Lee[3*], MS; Seong Kwon Ma[1,2], MD, PhD; Sunyong Yoo[3,4*], PhD; Soo Wan Kim[1,2*], MD, PhD

[1]Department of Internal Medicine, Chonnam National University Medical School, Gwangju, Republic of Korea
[2]Department of Internal Medicine, Chonnam National University Hospital, Gwangju, Republic of Korea
[3]Department of Intelligent Electronics and Computer Engineering, Chonnam National University, Gwangju, Republic of Korea
[4]Research and Development Center, MATILO AI Incorporated, Gwangju, Republic of Korea
[*]these authors contributed equally

Corresponding Author:
Soo Wan Kim, MD, PhD
Department of Internal Medicine
Chonnam National University Medical School
42 Jebongro
Gwangju, 61469
Republic of Korea
Phone: 82 622206271
Email: skimw@chonnam.ac.kr

Abstract

Background: While excessive dietary sodium intake is an established risk factor for cardiovascular and renal complications, its potential association with anemia remains largely unexplored.

Objective: We hypothesized that, based on the observed benefits of sodium-glucose cotransporter 2 inhibitors in correcting anemia through alterations in renal tubular metabolism and oxygen homeostasis, elevated urinary sodium excretion, as a surrogate of dietary sodium intake, may increase the risk of prevalent anemia.

Methods: This nationwide cross-sectional study analyzed 54,802 adults from the Korea National Health and Nutrition Examination Survey (2014-2023). Participants were stratified by spot urine sodium-to-creatinine ratio (Na^+/Cr) quartiles (first quartile [Q1], second quartile [Q2], third quartile [Q3], and fourth quartile [Q4]). Anemia was defined as hemoglobin <13 g/dL for men and <12 g/dL for women.

Results: Anemia prevalence increased progressively across spot urine Na^+/Cr quartiles (Q1: 925/13,700, 6.8%, Q2: 1126/13,701, 8.2%, Q3: 1393/13,701, 10.2%, and Q4: 1893/13,700, 13.8%). Multivariable logistic regression demonstrated that participants in the highest quartile had 43% higher odds of anemia compared with the lowest quartile (adjusted odds ratio 1.429, 95% CI 1.269-1.610; P<.001). Each 1-unit log increase in spot urine Na^+/Cr conferred a 67% increase in odds of anemia (adjusted odds ratio 1.674, 95% CI 1.452-1.930; P<.001). Sensitivity analyses using tertiles, quintiles, estimated 24-hour sodium excretion, and restriction to preserved kidney function consistently confirmed these associations.

Conclusions: Higher urinary sodium excretion exhibits a robust, graded association with increased anemia prevalence in the general population. These findings suggest that dietary sodium restriction may provide additional benefits beyond cardiovascular protection.

(JMIR Public Health Surveill 2026;12:e88408) doi:10.2196/88408

KEYWORDS

anemia; sodium; natriuresis; kidney; urine

XSL•FO
RenderX

Introduction

Anemia affects more than 1.6 billion people globally and significantly impacts health care use and patient outcomes [1,2]. Beyond well-established causes, including nutritional deficiencies, chronic diseases, and genetic disorders, emerging evidence suggests that dietary factors may play broader roles than previously recognized [3].

Dietary sodium intake has received considerable attention primarily for its cardiovascular and renal implications. Current evidence firmly establishes excessive sodium consumption as a risk factor for hypertension, stroke, and chronic kidney disease [4,5]. However, potential relationships between sodium homeostasis and hematologic parameters have received limited investigation.

Recent observations from sodium-glucose cotransporter-2 (SGLT2) inhibitor trials provide intriguing insights. These medications, which increase urinary sodium excretion by inhibiting proximal tubular sodium reabsorption, consistently demonstrate increases in hemoglobin levels [6,7]. While traditionally attributed to hemoconcentration from volume contraction, this association raises questions about more direct relationships between sodium handling and erythropoiesis. Animal studies further suggest that sodium balance may influence iron metabolism and erythropoietin responsiveness [8,9].

Despite these observations, population-based evidence examining associations between sodium excretion and anemia remains scarce. Understanding such relationships could have important public health implications, particularly given ongoing efforts to reduce population sodium intake for cardiovascular protection. Therefore, we investigated the association between urinary sodium excretion and anemia prevalence using data from the Korea National Health and Nutrition Examination Survey (KNHANES), a nationally representative cross-sectional survey. We hypothesized that higher urinary sodium excretion would be associated with increased anemia prevalence, independent of traditional risk factors.

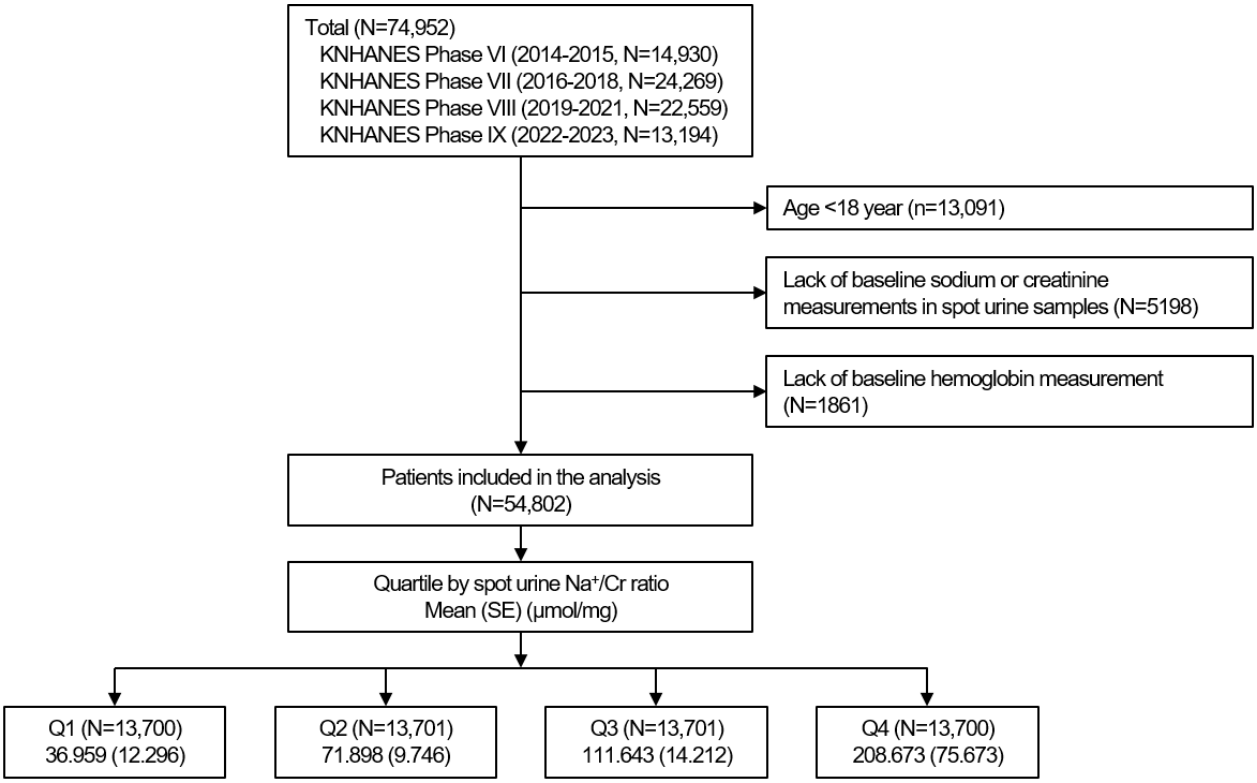
Methods

Study Design

We analyzed data from the KNHANES, an ongoing nationwide cross-sectional survey conducted by the Korea Centers for Disease Control and Prevention to monitor the health and nutritional status of the Korean population [10]. The KNHANES uses a stratified, multistage, probability sampling design based on household registries to select a representative sample of the noninstitutionalized Korean civilian population [11]. All analyses incorporated the survey sampling weights, strata, and primary sampling units. Sampling weights reflected the inverse probability of selection with adjustments for nonresponse and poststratification to produce nationally representative estimates of the noninstitutionalized Korean population. We applied the integrated sampling weights to account for participants who completed all 3 survey components, including the nutrition survey. This study analyzed data obtained from the KNHANES 2014 to 2023 (Figure 1), as these were the most recent 10-year data available and included the measurement of spot urine sodium and creatinine. After excluding those aged less than 18 years (13,091/74,952, 17.5%) and those lacking the baseline measurement of spot urine sodium or creatinine (5198/74,952, 7%) or hemoglobin (1861/74,952, 2.5%), a total of 54,802 (73.1%) were finally included for further analyses.

XSL•FO
RenderX

Figure 1. Study design. KNHANES: Korea National Health and Nutrition Examination Survey; Na+/Cr: sodium-to-creatinine ratio; Q1: first quartile; Q2: second quartile; Q3: third quartile; Q4: fourth quartile.



Ethical Considerations

This study protocol was approved by the Institutional Review Board of Chonnam National University Hospital (CNUH-EXP-2025-369) and was conducted in accordance with the Declaration of Helsinki and its later amendments or comparable ethical standards. All participants provided written informed consent. Identifying details of the participants were omitted. No additional compensation was provided for participation in this study.

Exposure Assessment: Urinary Sodium Excretion

Spot urine specimens were collected during health examination visits and immediately refrigerated until measurement of urinary sodium and creatinine concentrations. We calculated the urinary sodium-to-creatinine ratio (Na^+/Cr) as the primary exposure variable. Study participants were divided into quartiles by spot urine Na^+/Cr (first quartile [Q1], second quartile [Q2], third quartile [Q3], and fourth quartile [Q4]; Figure 1).

Outcome Definition: Anemia

Anemia was defined according to World Health Organization criteria: hemoglobin concentration <13.0 g/dL in men and <12.0 g/dL in nonpregnant women [12].

Covariates

We selected potential confounders based on established associations with both sodium excretion and anemia. Variables were categorized as follows.

Demographic factors included age, sex, educational status, and household income levels.

Lifestyle factors included smoking status, drinking history, daily iron and protein intake by 24-hour recall survey, and BMI.

Clinical factors included diabetes mellitus (DM), hypertension, dyslipidemia, fasting glucose, systolic blood pressure, estimated glomerular filtration rate (eGFR), and dipstick urine protein positivity.

Comorbid conditions were operationally defined, as previously described [13,14]. DM was defined as serum fasting glucose level ≥126 mg/dL, use of antidiabetic medication, or a physician diagnosis of DM; hypertension was defined as systolic blood pressure ≥140 mm Hg, diastolic blood pressure ≥90 mm Hg, or use of antihypertensive medication; and dyslipidemia was defined either by self-report or physician diagnosis. Laboratory tests were conducted in the central laboratory (Neodin Medical Institute, Seoul, Republic of Korea) using blood samples after at least an 8-hour overnight fasting. eGFR was calculated from serum creatinine level using the Chronic Kidney Disease Epidemiology Collaboration equation [15]. Proteinuria was defined as albuminuria (≥1+) determined by dipstick urine test [13].

Statistical Analysis

Continuous and categorical variables were expressed as the mean (SE) and n (%), respectively. Baseline characteristics across quartiles were compared using survey-weighted ANOVA for continuous variables and chi-square test for categorical variables. To assess the shape of the association between spot urine Na^+/Cr and hemoglobin, restricted cubic splines were fitted using multivariable linear regression [16]. Multivariable associations with anemia were estimated using survey-weighted logistic regression, accounting for strata, clusters, and sampling

XSL•FO
RenderX

weights, and are reported as adjusted odds ratios (ORs) with 95% CIs [17]. ORs were used as the primary measure of association, as they allow for robust adjustment of multiple confounders. Given that the prevalence of anemia in most of our study population was <10%, the OR served as a reliable approximation of the prevalence ratio according to the rare disease assumption. Participants with any missing values were excluded from the primary analysis, while missing values were handled using multiple imputation according to variable types in a sensitivity analysis. For each variable with missing data, 5 imputed datasets were generated. Categorical variables were imputed using multinomial logistic regression, and continuous variables were imputed using the predictive mean matching method. Statistical analyses were conducted separately on each imputed dataset, and the results were pooled using the Rubin rule [18]. This approach appropriately accounts for the uncertainty introduced by missing values while maintaining the validity of statistical inference [19]. All statistical analyses were performed in SAS (version 9.4; SAS Institute; survey procedures), and figures were generated in R (version 4.4.1; R Foundation for Statistical Computing), with statistical significance defined as $P<.05$.

Results

Baseline Characteristics

The baseline characteristics of the participants across the quartiles by spot urine Na^+/Cr demonstrated significant graded differences in demographic, anthropometric, and laboratory parameters (Table 1). Individuals in the highest quartile (Q4) were notably older (mean 58.3, SE 0.2 vs mean 39.1, SE 0.2 years; $P<.001$) and more likely to be female participants (4347/13,700, 31.7% vs 6315/13,700, 46.1%; $P<.001$) compared with the lowest quartile (Q1). Higher spot urine Na^+/Cr quartiles exhibited a stepwise increase in cardiovascular risk factors, including DM (2518/13,700, 18.4% in Q4 vs 953/13,700, 7% in Q1), hypertension (6363/13,700, 46.4% vs 2937/13,700, 21.4%), and dyslipidemia (4118/13,700, 30.1% vs 1604/13,700, 11.7%), with statistical significance in all the comparisons ($P<.001$). A consistent inverse relationship was observed between spot urine Na^+/Cr and hemoglobin concentration, declining from a mean of 14.5 (SE 0.02) g/dL in Q1 to a mean of 13.6 (SE 0.02) g/dL in Q4 ($P<.001$). Concurrently, eGFR decreased progressively across quartiles (mean 104.8, SE 0.2 in Q1 vs mean 96.6, SE 0.2 mL/min/1.73 m^2 in Q4; $P<.001$), while proteinuria prevalence exhibited a U-shaped distribution, being highest in Q1 (605/13,700, 4.4%) and lowest in Q3 (253/13,701, 1.8%).

XSL•FO
RenderX

Table 1. Baseline characteristics of participants by quartiles of spot urine sodium-to-creatinine ratio (Na+/Cr).

	Q1 (n=13,700)	Q2 (n=13,701)	Q3 (n=13,701)	Q4 (n=13,700)	*P* value
Age (y), mean (SE)	39.124 (0.156)	45.230 (0.177)	50.598 (0.185)	58.285 (0.198)	<.001
Male, n (%)	7385 (53.9)	6804 (49.7)	5994 (43.7)	4347 (31.7)	<.001
Education (years), n (%)					<.001
≤6	1180 (8.6)	1739 (12.7)	2583 (18.9)	4451 (32.5)	
7-9	868 (6.3)	1236 (9)	1482 (10.8)	1870 (13.6)	
10-12	4734 (34.6)	4440 (32.4)	4436 (32.4)	3702 (27)	
>12	6364 (46.4)	5681 (41.5)	4468 (32.6)	2780 (20.3)	
Household income, n (%)					<.001
<20%	1544 (11.3)	1730 (12.6)	2275 (16.6)	3476 (25.4)	
20%-39%	2174 (15.9)	2408 (17.6)	2545 (18.6)	3138 (22.9)	
40%-59%	2895 (21.1)	2874 (21)	2794 (20.4)	2450 (17.9)	
60%-79%	3369 (24.6)	3149 (23)	3017 (22)	2287 (16.7)	
≥80%	3673 (26.8)	3485 (25.4)	3022 (22.1)	2273 (16.6)	
Smoking history, n (%)					<.001
Current smoker	3241 (23.7)	2543 (18.6)	2049 (15)	1451 (10.6)	
Former smoker	2879 (21)	3255 (23.8)	3251 (v)	2697 (19.7)	
Nonsmoker	7580 (55.3)	7903 (57.7)	8401 (61.3)	9552 (69.7)	
Iron intake (mg/d), mean (SE)	11,995 (0.104)	12.314 (0.204)	12.093 (0.101)	11.107 (0.096)	<.001
Protein intake (g/kg/d), mean (SE)	1.157 (0.007)	1.139 (0.007)	1.109 (0.007)	1.036 (0.007)	<.001
Drinking history, n (%)					<.001
Nondrinking	1790 (13.1)	2102 (15.3)	2523 (18.4)	2885 (21.1)	
Occasional moderate drinking	3168 (23.1)	3443 (25.1)	3485 (25.4)	3534 (25.8)	
Regular moderate drinking	276 (2)	346 (2.5)	354 (2.6)	428 (3.1)	
Occasional binge drinking	2552 (18.6)	2175 (15.9)	1785 (13)	1266 (9.2)	
Regular binge drinking	4839 (35.3)	4319 (31.5)	3765 (27.5)	2776 (20.3)	
Medical history, n (%)					
DM[a]	953 (7)	1266 (9.2)	1648 (12)	2518 (18.4)	<.001
Hypertension	2937 (21.4)	3693 (27)	4546 (33.2)	6363 (46.4)	<.001
Dyslipidemia	1604 (11.7)	2318 (16.9)	2993 (21.8)	4118 (30.1)	<.001
BMI (kg/m^2), mean (SE)	23.852 (0.039)	24.047 (0.038)	24.156 (0.040)	24.110 (0.036)	<.001
SBP[b] (mm Hg), mean (SE)	114.128 (0.140)	116.357 (0.155)	118.856 (0.170)	123.884 (0.201)	<.001
Laboratory findings					
Hemoglobin (g/dL), mean (SE)	14.535 (0.016)	14.283 (0.015)	14.014 (0.016)	13.595 (0.016)	<.001
Fasting glucose (mg/dL), mean (SE)	97.495 (0.220)	99.207 (0.219)	101.437 (0.241)	103.852 (0.259)	<.001
Total cholesterol (mg/dL), mean (SE)	191.102 (0.384)	191.251 (0.392)	191.036 (0.395)	189.003 (0.412)	<.001
HDL-C[c] (mg/dL), mean (SE)	52.445 (0.138)	52.441 (0.142)	52.599 (0.137)	53.230 (0.145)	<.001
eGFR[d] (mL/min/1.73 m^2), mean (SE)	104.775 (0.170)	102.736 (0.181)	100.217 (0.186)	96.556 (0.202)	<.001
Dipstick urine protein ≥1+, n (%)	605 (4)	318 (2.2)	253 (1.6)	282 (1.9)	<.001

[a]DM: diabetes mellitus.

XSL•FO
RenderX

[b]SBP: systolic blood pressure.

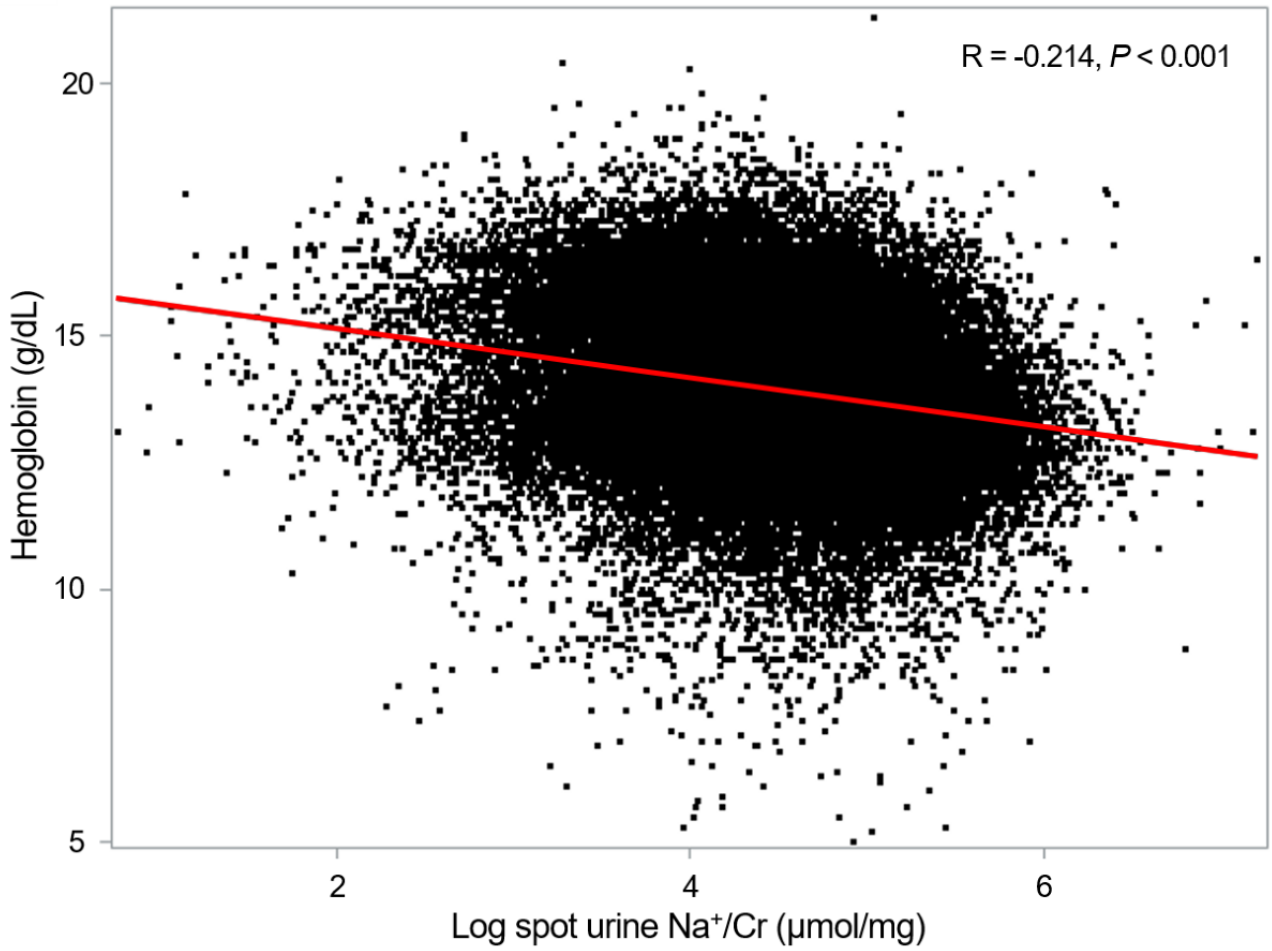
[c]eGFR: estimated glomerular filtration rate.

[d]HDL-C: high-density lipoprotein cholesterol.

Association of Urinary Sodium Excretion and the Risk of Prevalent Anemia

A clear dose-response relationship emerged between spot urine Na^+/Cr and anemia prevalence. The scatter plot analysis revealed a significant negative correlation between spot urine Na^+/Cr and hemoglobin concentration (r=−0.214; P<.001; Figure 2). The prevalence of anemia increased progressively across quartiles from 6.8% (925/13,700) in Q1 to 13.8% (1893/13,700) in Q4. Multivariable logistic regression analysis demonstrated graded associations that remained significant after adjustment for potential confounders (Table 2). In the fully adjusted model (model 4), which accounted for demographic variables, comorbidities, metabolic parameters, kidney function, and proteinuria, participants in Q4 exhibited 43% higher odds of prevalent anemia compared with Q1 (adjusted OR 1.429, 95% CI 1.269-1.610; P<.001). Intermediate quartiles showed incrementally increasing risk, with an adjusted OR of 1.207 (95% CI 1.073-1.357; P=.002) in Q2 and an adjusted OR of 1.306 (95% CI 1.167-1.462; P<.001) in Q3. When modeled continuously, each 1-unit log increase in spot urine Na^+/Cr was associated with 67% higher odds of anemia (adjusted OR 1.674, 95% CI 1.452-1.930; P<.001). The relationship between spot urine Na^+/Cr and hemoglobin concentration was further characterized using penalized spline curve analysis (Figure S1 in Multimedia Appendix 1), demonstrating a linear inverse association. Complementary spline analysis modeling the risk of anemia as the outcome confirmed these findings, revealing a monotonic increase in ORs across the spectrum of spot urine Na^+/Cr values after full covariate adjustment (Figure 3). The curves demonstrated no apparent inflection points, suggesting that the relationship between urinary sodium excretion and anemia risk operates continuously throughout the observed range.

Figure 2. Scatter plot of spot urine sodium-to-creatinine ratio (Na+/Cr) with hemoglobin levels. The correlation of spot urine Na+/Cr with hemoglobin levels was assessed by the Pearson correlation coefficient (r). A linear inverse correlation was observed between spot urine Na+/Cr and hemoglobin levels, though the correlation intensity is low.



XSL•FO
RenderX

Table 2. Association between spot urine sodium-to-creatinine ratio (Na+/Cr) and anemia prevalence: multivariable logistic regression analysis.

	Spot urine Na^+/Cr (μmol/mg)	Prevalence, n (%)	Model 1[a]		Model 2[b]		Model 3[c]		Model 4[d]	
			OR[e] (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Q1[f] (n=13,700)	2.1-54.1	925 (6.8)	Reference	—[g]	Reference	—	Reference	—	Reference	—
Q2[h] (n=13,701)	54.1-87.2	1126 (8.2)	1.311 (1.178-1.459)	<.001	1.092 (0.978-1.218)	.12	1.147 (1.022-1.287)	0.019	1.207 (1.073-1.357)	.002
Q3[i] (n=13,701)	87.2-136.7	1393 (10.2)	1.631 (1.479-1.798)	<.001	1.121 (1.012-1.242)	.03	1.210 (1.085-1.350)	<.001	1.306 (1.167-1.462)	<.001
Q4[j] (n=13,700)	136.7-1344.1	1893 (13.8)	2.415 (2.192-2.660)	<.001	1.225 (1.100-1.365)	.002	1.300 (1.159-1.457)	<.001	1.429 (1.269-1.610)	<.001
Per 1 log increase	—	—	3.237 (2.865-3.657)	<.001	1.346 (1.184-1.530)	<.001	1.466 (1.277-1.683)	<.001	1.674 (1.452-1.930)	<.001

[a]Model 1: unadjusted.

[b]Model 2: model 1 adjusted for age and sex.

[c]Model 3: model 2 adjusted for household income, education years, smoking history, medical history (diabetes mellitus, hypertension, and dyslipidemia), BMI, and systolic blood pressure.

[d]Model 4: model 3 adjusted for fasting glucose, estimated glomerular filtration rate, and dipstick urine protein positivity.

[e]OR: odds ratio.

[f]Q1: first quartile.

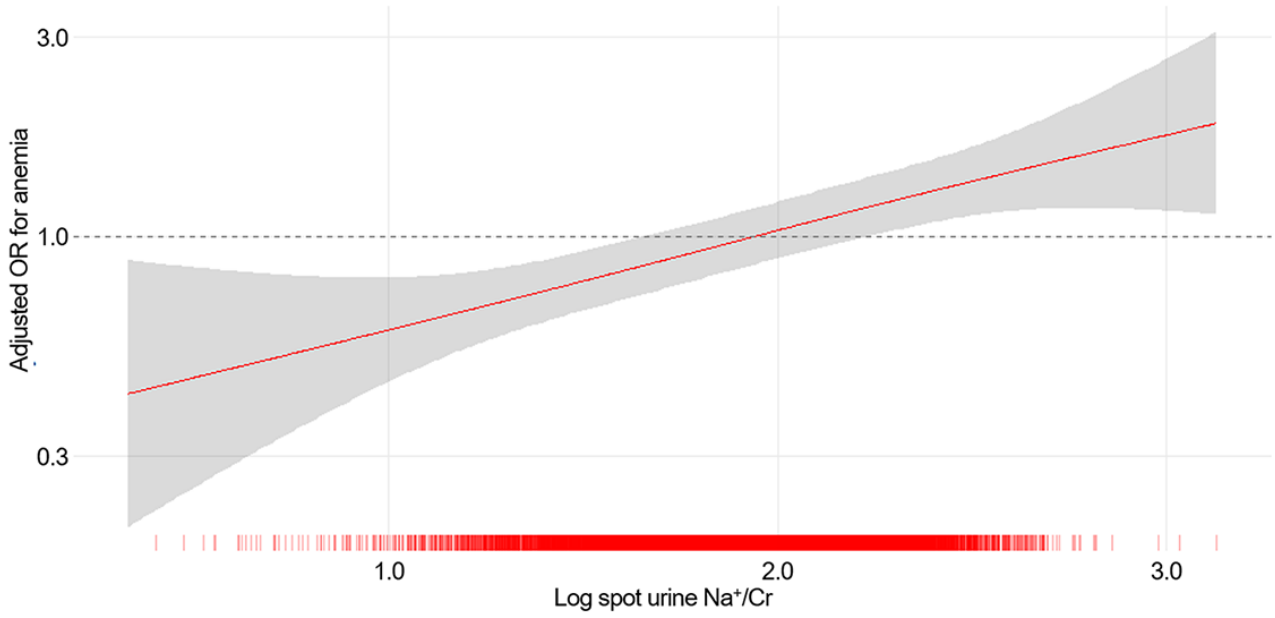
[g]Not applicable.

[h]Q2: second quartile.

[i]Q3: third quartile.

[j]Q4: fourth quartile.

Figure 3. Penalized spline curve of spot urine sodium-to-creatinine ratio (Na+/Cr) and the risk of prevalent anemia. Adjusted for age, sex, household income levels, education years, smoking history, medical history (diabetes mellitus, hypertension, and dyslipidemia), BMI, systolic blood pressure, fasting glucose, estimated glomerular filtration rate, and dipstick urine protein positivity. A linear association was observed between spot urine Na+/Cr and the risk of prevalent anemia. OR: odds ratio.



XSL•FO
RenderX

Sensitivity Analysis

Sensitivity analyses confirmed the robustness of the primary findings. Alternative categorization schemes using tertiles and quintiles of spot urine Na^+/Cr yielded consistent results (Table S1 in Multimedia Appendix 1). In tertile analysis, the highest tertile demonstrated an adjusted OR of 1.415 (95% CI 1.278-1.567; *P*<.001) compared with the lowest tertile in the fully adjusted model. Quintile stratification revealed a more granular dose-response gradient, with ORs progressively increasing from 1.218 (second quintile) to 1.547 (fifth quintile) relative to the lowest quintile (all *P*<.01). When using estimated 24-hour urinary sodium excretion by the Tanaka formula [20] as an alternative exposure measure, the association remained significant, with the highest quartile exhibiting an adjusted OR of 1.376 (95% CI 1.237-1.531; *P*<.001) in model 4 (Table S2 in Multimedia Appendix 1), though the effect size was modestly attenuated compared with spot urine Na^+/Cr. To address potential confounding by fluid retention related to high sodium intake in participants with impaired kidney function, we performed a restricted analysis excluding participants with eGFR <60 mL/min/1.73 m² (1513/54,802, 2.8%), where the association between spot urine Na^+/Cr and anemia risk persisted with minimal attenuation (Table S3 in Multimedia Appendix 1). The highest quartile retained 27% higher odds of anemia (adjusted OR 1.266, 95% CI 1.118-1.434; *P*<.001) compared with the reference quartile after full adjustment. The association remained unchanged even after additional adjustment for daily iron and protein intake (24-hour dietary recall survey) and alcohol consumption (Table S4 in Multimedia Appendix 1). Finally, multiple imputation for missing covariate data yielded nearly identical point estimates and CIs to the complete-case analysis, with Q4 demonstrating an adjusted OR of 1.429 (95% CI 1.269-1.610; *P*<.001) in the fully adjusted model (Table 3), confirming that missing data did not materially influence the observed associations.

Table 3. Binary logistic regression of spot urine sodium-to-creatinine ratio (Na+/Cr) levels for the risk of prevalent anemia after multiple imputation.

Spot urine Na^+/Cr level	Model 1[a]		Model 2[b]		Model 3[c]		Model 4[d]	
	OR[e] (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Q1[f]	Reference	__[g]	Reference	—	Reference	—	Reference	—
Q2[h]	1.321 (1.182-1.475)	<.001	1.107 (0.988-1.240)	.08	1.148 (1.023-1.283)	.02	1.207 (1.073-1.357)	.002
Q3[i]	1.617 (1.460-1.791)	<.001	1.137 (1.021-1.266)	.02	1.211 (1.085-1.351)	<.001	1.306 (1.167-1.462)	<.001
Q4[j]	2.313 (2.091-2.558)	<.001	1.215 (1.085-1.360)	<.001	1.301 (1.160-1.458)	<.001	1.429 (1.269-1.610)	<.001

[a]Model 1: unadjusted.

[b]Model 2: model 1 adjusted for age and sex.

[c]Model 3: model 2 adjusted for household income, education years, smoking history, medical history (diabetes mellitus, hypertension, and dyslipidemia), BMI, and systolic blood pressure.

[d]Model 4: model 3 adjusted for fasting glucose, total cholesterol, estimated glomerular filtration rate, and dipstick urine protein positivity.

[e]OR, odds ratio.

[f]Q1: first quartile.

[g]Not applicable.

[h]Q2: second quartile.

[i]Q3: third quartile.

[j]Q4: fourth quartile.

Discussion

Main Findings

In this nationwide cross-sectional study involving 54,802 Korean adults, we identified a robust and graded association between higher urinary sodium excretion and increased prevalence of anemia. This relationship persisted after comprehensive adjustment for demographic variables, comorbid conditions, eGFR, and proteinuria, with individuals in the highest quartile of spot urine Na^+/Cr demonstrating 43% higher odds of anemia compared with the lowest quartile. To our knowledge, this represents the first population-based investigation examining the link between dietary sodium intake and anemia risk in the general population.

Clinical Implications

Our findings extend beyond the established cardiovascular and renal consequences of excess sodium intake. The observed association between sodium excretion and anemia prevalence—with rates increasing from 6.8% (925/13,700) to 13.8% (1892/13,700) across quartiles—represents a clinically meaningful effect size comparable to traditional anemia risk factors. These data suggest that dietary sodium reduction may be an additional option for the prevention or correction of anemia.

XSL•FO
RenderX

Biological Plausibility

High sodium intake may increase anemia prevalence by altering cellular metabolism in renal proximal tubular cells. Sodium reabsorption constitutes the predominant energy-consuming process in the kidney, accounting for more than 90% of total renal oxygen consumption [21,22]. The proximal tubule, responsible for reabsorbing approximately 60% to 70% of filtered sodium, operates under relatively high metabolic demand despite its favorable oxygen supply from the renal cortex [23]. Studies examining high salt intake in experimental models have demonstrated progressive increases in whole-kidney oxygen consumption accompanied by upregulation of glycolytic pathway genes in proximal tubular segments, suggesting heightened metabolic stress [24]. Higher metabolic demand triggers tubulointerstitial hypoxia, compromising the function of erythropoietin-producing cells.

Our findings align conceptually with findings from previous clinical trials of SGLT2 inhibitors, in which these agents consistently increased hemoglobin levels by 0.6 to 0.7 g/dL with enhanced erythropoietin production and red blood cell mass expansion [6,7]. The proposed mechanism involved reduced proximal tubular workload through blockade of sodium-glucose reabsorption, thereby decreasing cellular ATP consumption and oxygen demand required for sodium handling [25]. This metabolic reprieve alleviates tissue hypoxia and preserves erythropoietin-producing interstitial fibroblasts by suppressing the transition to myofibroblasts under sustained tissue hypoxia [9]. Importantly, experimental studies have shown that dietary sodium restriction redistributes the tubular metabolic workload, shifting sodium reabsorption from less energy-efficient distal segments toward the more efficient proximal tubule [26]. Conversely, high dietary sodium intake shifts reabsorption distally, increasing overall oxygen consumption relative to sodium transport efficiency [26]. This metabolic framework suggests that lower dietary sodium may confer hematologic benefits through mechanisms similar to SGLT2 inhibition.

Alternative Explanation

Hemodilution related to volume expansion from high sodium intake should also be considered as an alternative explanation. Plasma volume expansion can reduce hemoglobin concentration without decreasing red blood cell mass, thereby causing apparent rather than true anemia. However, several lines of evidence argue against hemodilution as the primary mechanism for the principal findings of this study. First, the association between urinary sodium excretion and anemia persisted after adjustment for blood pressure and BMI, both of which reflect sodium-dependent volume expansion. Second, studies examining blood volume components in various patient populations demonstrate that, while hemodilution contributes to anemia in acute fluid overload states such as heart failure, the relationship between chronic dietary sodium intake and sustained plasma volume expansion in otherwise healthy individuals is more complex [27,28]. Significant acute volume changes are unlikely, given the community-dwelling nature of the KNHANES cohort. Third, the findings from the trials of SGLT2 inhibitors suggest that, although these agents result in transient natriuresis and plasma volume contraction, the increase in hemoglobin is sustained much longer, with enhanced erythropoiesis reflected by changes in iron metabolism markers [29]. Fourth, as most participants in the KNHANES dataset have normal kidney function (ie, eGFR ≥60 mL/min/1.73 m^2), dietary salt intake should be balanced with urinary sodium excretion rather than sodium retention in that population.

Limitations

Several limitations should be acknowledged. First, the cross-sectional design precludes causal inference, and residual confounding from unmeasured variables remains possible despite extensive adjustment. Second, we lacked data on dietary patterns, micronutrient intake, and other nutritional factors that influence anemia risk. Third, as hemoglobin measurements occurred at a single time point, we were unable to distinguish acute from chronic anemia or assess temporal relationships between sodium intake changes and hemoglobin dynamics. Fourth, spot urine specimens, while practical for population surveys, provide less accurate sodium excretion estimates than timed collections. As 24-hour urinary sodium excretion remains the gold standard of the surrogate for dietary sodium intake, spot urine sodium excretion is less accurate. Fifth, due to the intrinsic limitation of the dataset, only cross-sectional analysis was available, which was not sufficient to suggest a causal relation between excess sodium intake and the risk of prevalent anemia. Finally, our findings were derived from a Korean population, and their applicability to other ethnic groups requires validation, considering that population-specific dietary patterns significantly differ from those of a typical Western diet.

Conclusions

In conclusion, we demonstrate a novel association between urinary sodium excretion and the risk of prevalent anemia in a nationwide Korean population. These data suggest that, in the context of the findings from trials of SGLT2 inhibitors, excessive dietary sodium intake may impair the production of erythropoietin from a subset of kidney interstitial fibroblasts. Our findings indicate that dietary sodium restriction may reduce the risk of anemia, in addition to the established cardiovascular and renal benefits. While confirmatory studies are required, our findings support expanding the rationale for sodium restriction beyond traditional cardiovascular and renal indications to include potential hematologic benefits.

Acknowledgments

The authors declare that generative artificial intelligence (Gemini, Google) was used solely for refining the language and enhancing the clarity of the manuscript. The study's conceptualization and core analysis were conducted by the authors, with technical

XSL•FO
RenderX

methodology supported by researchers specialized in artificial intelligence. The final manuscript was meticulously reviewed by all authors to ensure accuracy and to prevent hallucination or misinformation.

Data Availability

The datasets generated or analyzed during this study are available in the Korea National Health and Nutrition Examination Survey (KNHANES) repository [30].

Funding

This work was supported by the Research Program funded by the National Research Foundation of Korea grant from the Korea government (Ministry of Science and ICT [MSIT]; RS-2023-00217317 and RS-2023-00278258), by a grant (HCRI26006) from Chonnam National University Hwasun Hospital Institute for Biomedical Science, and by a grant (RS-2024-00332003) from the Ministry of Food and Drug Safety in 2024.

Authors' Contributions

SHS was responsible for study conceptualization, statistical analysis, results interpretation, and manuscript writing. DL was responsible for data acquisition, data processing, statistical modeling, and figure preparation. SKM contributed to data interpretation, methodology validation, and critical revision of the manuscript. SY (co-corresponding author) supervised the computational analysis, provided expertise in data science methodology, and critically revised the manuscript. SWK (co-corresponding author) conceived the original research idea, supervised the overall study, secured research funding, and provided final approval. All authors reviewed and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional data on the different binary logistic regressions of spot urine Na+/Cr levels.
[DOCX File , 207 KB - publichealth_v12i1e88408_app1.docx]

References

1. Weiss G, Goodnough LT. Anemia of chronic disease. N Engl J Med 2005 Mar 10;352(10):1011-1023. [doi: 10.1056/NEJMra041809] [Medline: 15758012]
2. Ganz T. Anemia of inflammation. N Engl J Med 2019 Sep 19;381(12):1148-1157. [doi: 10.1056/NEJMra1804281] [Medline: 31532961]
3. Anand IS, Gupta P. Anemia and iron deficiency in heart failure: current concepts and emerging therapies. Circulation 2018 Jul 03;138(1):80-98. [doi: 10.1161/CIRCULATIONAHA.118.030099] [Medline: 29967232]
4. He FJ, Tan M, Ma Y, MacGregor GA. Salt reduction to prevent hypertension and cardiovascular disease: JACC state-of-the-art review. J Am Coll Cardiol 2020 Feb 18;75(6):632-647 [FREE Full text] [doi: 10.1016/j.jacc.2019.11.055] [Medline: 32057379]
5. McMahon EJ, Campbell KL, Bauer JD, Mudge DW, Kelly JT. Altered dietary salt intake for people with chronic kidney disease. Cochrane Database Syst Rev 2021 Jun 24;6(6):CD010070 [FREE Full text] [doi: 10.1002/14651858.CD010070.pub3] [Medline: 34164803]
6. Perkovic V, Jardine MJ, Neal B, Bompoint S, Heerspink HJ, Charytan DM, et al. Canagliflozin and renal outcomes in type 2 diabetes and nephropathy. N Engl J Med 2019 Jun 13;380(24):2295-2306 [FREE Full text] [doi: 10.1056/NEJMoa1811744] [Medline: 30990260]
7. The EMPA-KIDNEY Collaborative Group, Herrington WG, Staplin N, Wanner C, Green JB, Hauske SJ, et al. Empagliflozin in patients with chronic kidney disease. N Engl J Med 2023 Jan 12;388(2):117-127 [FREE Full text] [doi: 10.1056/NEJMoa2204233] [Medline: 36331190]
8. Sano M, Goto S. Possible mechanism of hematocrit elevation by sodium glucose cotransporter 2 inhibitors and associated beneficial renal and cardiovascular effects. Circulation 2019 Apr 23;139(17):1985-1987. [doi: 10.1161/CIRCULATIONAHA.118.038881] [Medline: 31009585]
9. Nangaku M. Chronic hypoxia and tubulointerstitial injury: a final common pathway to end-stage renal failure. J Am Soc Nephrol 2006 Jan;17(1):17-25. [doi: 10.1681/ASN.2005070757] [Medline: 16291837]
10. Oh K, Kim Y, Kweon S, Kim S, Yun S, Park S, et al. Korea National Health and Nutrition Examination Survey, 20th anniversary: accomplishments and future directions. Epidemiol Health 2021;43:e2021025 [FREE Full text] [doi: 10.4178/epih.e2021025] [Medline: 33872484]

XSL•FO
RenderX

11. Kweon S, Kim Y, Jang MJ, Kim Y, Kim K, Choi S, et al. Data resource profile: the Korea National Health and Nutrition Examination Survey (KNHANES). Int J Epidemiol 2014 Feb;43(1):69-77 [FREE Full text] [doi: 10.1093/ije/dyt228] [Medline: 24585853]
12. Cappellini MD, Motta I. Anemia in clinical practice-definition and classification: does hemoglobin change with aging? Semin Hematol 2015 Oct;52(4):261-269. [doi: 10.1053/j.seminhematol.2015.07.006] [Medline: 26404438]
13. Suh SH, Song SH, Choi HS, Kim CS, Bae EH, Ma SK, et al. Parental educational status independently predicts the risk of prevalent hypertension in young adults. Sci Rep 2021 Feb 12;11(1):3698 [FREE Full text] [doi: 10.1038/s41598-021-83205-0] [Medline: 33580117]
14. Chin SO, Rhee SY, Chon S, Hwang YC, Jeong IK, Oh S, et al. Sarcopenia is independently associated with cardiovascular disease in older Korean adults: the Korea National Health and Nutrition Examination Survey (KNHANES) from 2009. PLoS One 2013;8(3):e60119 [FREE Full text] [doi: 10.1371/journal.pone.0060119] [Medline: 23533671]
15. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF3, Feldman HI, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med 2009 May 05;150(9):604-612 [FREE Full text] [doi: 10.7326/0003-4819-150-9-200905050-00006] [Medline: 19414839]
16. Desquilbet L, Mariotti F. Dose-response analyses using restricted cubic spline functions in public health research. Stat Med 2010 Apr 30;29(9):1037-1057. [doi: 10.1002/sim.3841] [Medline: 20087875]
17. Lumley T. Analysis of complex survey samples. J Stat Softw 2004;9(8):1-19. [doi: 10.18637/jss.v009.i08]
18. Little RJ, Rubin DB. Statistical Analysis with Missing Data. Hoboken, NJ: John Wiley & Sons; 1987.
19. van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. J Stat Softw 2011;45(3):1-67. [doi: 10.18637/jss.v045.i03]
20. Tanaka T, Okamura T, Miura K, Kadowaki T, Ueshima H, Nakagawa H, et al. A simple method to estimate populational 24-h urinary sodium and potassium excretion using a casual urine specimen. J Hum Hypertens 2002 Feb;16(2):97-103. [doi: 10.1038/sj.jhh.1001307] [Medline: 11850766]
21. Palmer LG, Schnermann J. Integrated control of Na transport along the nephron. Clin J Am Soc Nephrol 2015 Apr 07;10(4):676-687 [FREE Full text] [doi: 10.2215/CJN.12391213] [Medline: 25098598]
22. Bhargava P, Schnellmann RG. Mitochondrial energetics in the kidney. Nat Rev Nephrol 2017 Oct;13(10):629-646 [FREE Full text] [doi: 10.1038/nrneph.2017.107] [Medline: 28804120]
23. Evans RG, Ince C, Joles JA, Smith DW, May CN, O'Connor PM, et al. Haemodynamic influences on kidney oxygenation: clinical implications of integrative physiology. Clin Exp Pharmacol Physiol 2013 Feb;40(2):106-122. [doi: 10.1111/1440-1681.12031] [Medline: 23167537]
24. Shimada S, Hoffmann BR, Yang C, Kurth T, Greene AS, Liang M, et al. Metabolic responses of normal rat kidneys to a high salt intake. Function (Oxf) 2023 Jun 22;4(5):zqad031 [FREE Full text] [doi: 10.1093/function/zqad031] [Medline: 37575482]
25. Vallon V, Thomson SC. The tubular hypothesis of nephron filtration and diabetic kidney disease. Nat Rev Nephrol 2020 Jun;16(6):317-336 [FREE Full text] [doi: 10.1038/s41581-020-0256-y] [Medline: 32152499]
26. Udwan K, Abed A, Roth I, Dizin E, Maillard M, Bettoni C, et al. Dietary sodium induces a redistribution of the tubular metabolic workload. J Physiol 2017 Nov 15;595(22):6905-6922 [FREE Full text] [doi: 10.1113/JP274927] [Medline: 28940314]
27. Androne AS, Katz SD, Lund L, LaManca J, Hudaihed A, Hryniewicz K, et al. Hemodilution is common in patients with advanced heart failure. Circulation 2003 Jan 21;107(2):226-229. [doi: 10.1161/01.cir.0000052623.16194.80] [Medline: 12538419]
28. Androne AS, Hryniewicz K, Hudaihed A, Mancini D, Lamanca J, Katz SD. Relation of unrecognized hypervolemia in chronic heart failure to clinical status, hemodynamics, and patient outcomes. Am J Cardiol 2004 May 15;93(10):1254-1259. [doi: 10.1016/j.amjcard.2004.01.070] [Medline: 15135699]
29. Packer M. Mechanisms of enhanced renal and hepatic erythropoietin synthesis by sodium-glucose cotransporter 2 inhibitors. Eur Heart J 2023 Dec 21;44(48):5027-5035 [FREE Full text] [doi: 10.1093/eurheartj/ehad235] [Medline: 37086098]
30. Are you subject to investigation? KNHANES. URL: https://knhanes.kdca.go.kr/knhanes/main.do [accessed 2025-05-25]

Abbreviations

DM: diabetes mellitus
eGFR: estimated glomerular filtration rate
KNHANES: Korea National Health and Nutrition Examination Survey
Na+/Cr: sodium-to-creatinine ratio
OR: odds ratio
Q1: first quartile
Q2: second quartile
Q3: third quartile
Q4: fourth quartile

XSL•FO
RenderX

SGLT2: sodium-glucose cotransporter 2

Edited by A Mavragani, T Sanchez; submitted 24.Nov.2025; peer-reviewed by M Tarabeih, M Kang; comments to author 22.Jan.2026; revised version received 26.Feb.2026; accepted 31.Mar.2026; published 21.Apr.2026.

Please cite as:
Suh SH, Lee D, Ma SK, Yoo S, Kim SW
Urinary Sodium Excretion and the Risk of Prevalent Anemia: Nationwide Population-Based Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e88408
URL: https://publichealth.jmir.org/2026/1/e88408
doi:10.2196/88408
PMID:

©Sang Heon Suh, Dohyeon Lee, Seong Kwon Ma, Sunyong Yoo, Soo Wan Kim. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 21.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Long-Term Exposure to Particulate Matter 2.5 and Ozone and the Risk of Acute Respiratory Infections: Community-Based Prospective Cohort Study

Xiaolei Duan[1*], MSc; Xiao Yu[2*], MSc; Hongyu Liang[1*], MSc; Zexuan Wen[1,3], BSc; Wenyong Zhou[1,3], BSc; Ying Ma[3], BSc; Weibing Wang[1,3,4,5,6], PhD; Yaxu Zheng[2], PhD; Huanyu Wu[7], MSc

[1]Department of Epidemiology, School of Public Health, Fudan University, Shanghai, China
[2]Institute for Surveillance and Early Warning, Shanghai Municipal Center For Disease Control Prevention, Shanghai, China
[3]Shanghai Institute of Infectious Disease and Biosecurity, Fudan University, Shanghai, China
[4]Key Laboratory of Public Health Safety of Ministry of Education, Fudan University, Shanghai, China
[5]Key Laboratory of Health Technology Assessment, Fudan University, Shanghai, China
[6]IRDR-ICoE on Risk Interconnectivity and Governance on Weather/Climate Extremes Impact and Public Health, Fudan University, Shanghai, China
[7]Shanghai Municipal Center for Disease Control and Prevention, No. 1399 Shenhong Road, Xinhong Subdistrict, Shanghai, China
[*]these authors contributed equally

Corresponding Author:
Huanyu Wu, MSc
Shanghai Municipal Center for Disease Control and Prevention, No. 1399 Shenhong Road, Xinhong Subdistrict, Shanghai, China

Abstract

Background: Acute respiratory infections (ARIs) remain a major global health concern. Although long-term air pollution exposure has been linked to ARIs, prospective evidence from community-based populations remains limited.

Objective: This study aimed to quantify the burden of ARIs in the community and examine the associations between long-term exposure to particulate matter 2.5 ($PM_{2.5}$) and ozone (O_3) and the risks of ARIs, with additional analyses using influenza-like illness (ILI) as a more specific outcome.

Methods: We conducted a prospective cohort study including 3617 residents in Shanghai, China, who were followed weekly for 1 year. Individual-level exposure to $PM_{2.5}$ and O_3 concentrations was estimated using high-resolution datasets. Cox proportional hazards models with shared frailty were applied to assess associations with ARIs. Exposure windows of 3, 6, 9, and 12 months were evaluated, and the optimal window was selected based on the Akaike information criterion. Effect estimates were reported per IQR increase. Dose-response relationships, subgroup analyses, and multiple sensitivity analyses were performed.

Results: During 3217 person-years of follow-up, 885 ARI events were documented (0.27 per person-year). In the fully adjusted model using the 12-month exposure window, each IQR increase in $PM_{2.5}$ was associated with higher risks of ARIs (hazard ratio [HR] 1.594, 95% CI 1.340 - 1.897), with stronger associations observed for ILI (HR 1.948, 95% CI 1.484 - 2.557). For O_3, the corresponding HRs were 1.510 (95% CI 1.135 - 2.007) for ARIs, with stronger associations for ILI (HR 2.229, 95% CI 1.385 - 3.588). $PM_{2.5}$ showed a nonlinear association with ARIs, whereas linear relationships were observed for $PM_{2.5}$ with ILI and O_3 with both outcomes. Evidence of effect modification was observed by age, residence, and season for $PM_{2.5}$ and by season for O_3. Results were robust across multiple sensitivity analyses.

Conclusions: Long-term exposure to $PM_{2.5}$ and O_3 is associated with increased risk of ARIs, with similar associations observed for ILI. These findings highlight the importance of long-term air pollution control and targeted interventions for susceptible populations, particularly during cold seasons.

International Registered Report Identifier (IRRID): RR2-10.1136/bmjopen-2025-103001

(JMIR Public Health Surveill 2026;12:e88045) doi:10.2196/88045

KEYWORDS

acute respiratory infections; particulate matter; ozone; influenza-like illness; cohort study

XSL•FO
RenderX

Introduction

Acute respiratory infections (ARIs) are among the most common infectious diseases worldwide and remain a serious public health issue [1]. In 2021, an estimated 12.8 billion upper respiratory infections and 344 million lower respiratory infections (LRIs), excluding COVID-19, occurred globally [2,3]. The etiology of ARIs is multifactorial, involving a broad spectrum of pathogens, including viruses, bacteria, and atypical organisms [4]. In addition, host factors such as age, immune status, socioeconomic conditions, and environmental determinants influence the susceptibility to and severity of ARIs [5-7]. Among environmental exposures, air pollution—particularly particulate matter (PM) and ozone (O_3)—has been widely recognized as an important external factor influencing the risk of ARIs [8,9]. According to the World Health Organization (WHO), nearly 6.67 million premature deaths worldwide each year are attributable to air pollution, including approximately 4.14 million and 0.37 million deaths attributed to PM and O_3, respectively [10].

Previous studies have shown that long-term exposure to air pollution is associated with an increased risk of ARIs [11-13]. In Spain, each IQR increase in $PM_{2.5}$ and O_3 was associated with 10% and 3% higher risks of LRIs, respectively [11]. However, most existing studies have relied on outpatient or inpatient data, which may underestimate the true burden of ARIs in the general population, as many mild cases do not seek medical care [14]. In addition, most studies have focused on first-onset events, while evidence on recurrent ARIs remains limited. Only a few studies have examined recurrent respiratory infections, often using clinical records, which may still miss a substantial proportion of community-level cases [15]. Furthermore, frailty models account for within-individual correlation and unobserved heterogeneity, providing more reliable risk estimates for recurrent events [16].

This study aimed to quantify the community burden of ARIs and to evaluate the associations between long-term exposure to $PM_{2.5}$ and O_3 and the risk of ARIs in a community-based cohort. By integrating high-frequency follow-up data with individual-level exposure assessment, this study provides insights into the community burden of ARIs and the potential health effects of long-term air pollution exposure. These findings may inform targeted interventions and public health strategies to reduce the burden of ARIs in community settings.

Methods

Study Population

A stratified random sampling design was used to ensure representative coverage of Shanghai's central urban and suburban regions. Two central urban districts (Jing'an and Xuhui) and 3 suburban districts (Jiading, Songjiang, and Fengxian) were selected based on geographic distribution and socioeconomic characteristics. Within each selected district, subdistricts (urban) or towns (suburban) were randomly sampled. Households were then recruited through local community health centers using a probability proportional to size method based on data from the 2020 Shanghai Census, and all household members were enrolled in the cohort. The detailed design has been described previously [17].

Participants were eligible if they had resided in Shanghai for ≥6 months and provided written informed consent. Individuals with ARIs within 2 weeks before enrollment or with a diagnosis of immunodeficiency or malignancy were excluded. The final sample was demographically representative of the general population, thereby minimizing potential selection bias (Table S1 in Multimedia Appendix 1).

Data Collection

In May 2024, a questionnaire survey was administered at community health service centers to collect individual-level information on sociodemographic characteristics and lifestyle factors, including age, sex, height, weight, smoking status, alcohol consumption, mask-wearing frequency, and vaccination history (Table S2 in Multimedia Appendix 1). Residential environmental data were also collected, including household address, fuel type, and exposure to cooking fumes.

In addition, city-level influenza-like illness (ILI) activity was obtained from sentinel hospital surveillance systems, defined as the proportion of ILI cases among all outpatient and emergency visits. For each participant, season was assigned based on the time of event occurrence (or the last follow-up date for noncases), with May to October defined as the warm season and November to April as the cold season.

Ethical Considerations

This study was approved by the Ethics Committee of the Shanghai Center for Disease Control and Prevention (KY-2024 - 19). Written informed consent was obtained from all participants before enrollment.

Environmental Exposure

Daily $PM_{2.5}$ concentrations and O_3 levels, defined as the maximum daily 8-hour average, were obtained from the Tracking Air Pollution in China dataset, with spatial resolutions of 1 km×1 km and 10 km×10 km, respectively [18]. Daily temperature and relative humidity data were obtained from the ERA5-Land reanalysis dataset provided by the European Centre for Medium-Range Weather Forecasts, with a spatial resolution of 0.1°×0.1° (approximately 10 km×10 km; [19]). Individual exposure levels were estimated using bilinear interpolation based on the 4 nearest grid points surrounding each participant's geocoded residential address [20].

Air pollution data were additionally obtained from the China National Environmental Monitoring Center (19 monitoring stations in Shanghai), and meteorological data were obtained from the Global Historical Climatology Network Daily dataset provided by the National Centers for Environmental Information (3 meteorological stations) for sensitivity analyses (Figure S1 in Multimedia Appendix 1).

Outcome Definition

ARIs were defined according to previous studies as the acute onset of at least 1 respiratory symptom (eg, cough, sputum, sore throat, nasal congestion, rhinorrhea, anosmia, ageusia, shortness

XSL•FO
RenderX

of breath, chest pain, tachypnea, or respiratory distress) together with at least 1 systemic symptom (fever, chills, hypothermia, or myalgia) [17,21]. Prior to data collection, general practitioners and community health workers received standardized training. Weekly telephone follow-ups were conducted every Thursday over a 52-week period (through May 2025). For each self-reported episode, trained staff recorded symptom type, onset date, recovery date, and health care use to ascertain ARIs.

For participants with recurrent episodes but missing recovery dates, recovery was assumed to occur 14 days after symptom onset, as ARIs are typically self-limited illnesses, with most symptoms resolving within 1 to 2 weeks [22]. Alternative recovery durations of 7 and 21 days were further evaluated in sensitivity analyses. Considering the pronounced seasonal pattern of ARIs, all participants were followed for 1 year. Participants who were unreachable for 2 consecutive follow-ups were classified as lost to follow-up. In addition, a stricter definition of ILI was applied, defined as ARIs with a body temperature ≥38.0 °C [23].

Statistical Analyses

Descriptive Analyses

Continuous variables with nonnormal distributions were summarized as medians with IQRs, and group differences were compared using the Wilcoxon rank-sum test. Categorical variables were presented as frequencies and percentages, and differences between groups were assessed using Pearson chi-square test. Missing covariate data were handled using multiple imputation by chained equations.

Exposure Modeling

Given the strong within-individual correlation of recurrent ARIs, Cox proportional hazards models with frailty terms were applied to account for unobserved individual heterogeneity, an approach considered appropriate for recurrent event data [24]. Participants who experienced an event remained at risk for subsequent events until censoring, allowing for recurrent event analysis. Schoenfeld residual tests were used to assess the proportional hazards assumption (Table S3 in Multimedia Appendix 1). Covariates were selected to minimize confounding while avoiding overadjustment, based on a priori knowledge and univariable analyses (Table S4 in Multimedia Appendix 1). The adjusted model included age, sex, residence, household size, BMI category, smoking status, alcohol consumption, mask-wearing frequency, underlying disease, vaccination history, household fuel type, cooking fume exposure, season, and city-level ILI activity.

As the optimal exposure window for long-term air pollution remains uncertain, we evaluated multiple exposure windows of 3, 6, 9, and 12 months based on previous studies [11,13]. Given that this study involved recurrent event data, C-index and receiver operating characteristic–based approaches were not applicable. Therefore, model comparison was primarily conducted using Akaike information criterion. The primary analysis was conducted using the optimal exposure window, and effect estimates were reported per IQR increase in exposure. Four sequential models were further constructed: a single-pollutant model, a copollutant model (mutually adjusting $PM_{2.5}$ and O_3), a model additionally adjusted for temperature, and a model further adjusted for relative humidity. Variance inflation factors were calculated to assess collinearity among pollutants and meteorological variables (Table S5 in Multimedia Appendix 1).

Dose-Response and Subgroup Analyses

Restricted cubic spline models were used to assess the dose-response relationships between long-term air pollution exposure and ARIs, as well as ILI, with reference values based on the WHO guidelines ($PM_{2.5}$: 15 μg/m^3; O_3: 100 μg/m^3). Subgroup analyses were conducted stratified by age (<18, 18 - 59, and ≥60 years), sex (male and female), residential area (central urban and suburban), smoking status (never and ever), and season (cold season and warm season) to explore potential effect modification.

Sensitivity Analyses

Several sensitivity analyses were conducted to assess the robustness of the results: (1) alternative assumptions for recovery time (7-day and 21-day definitions); (2) Andersen-Gill models for recurrent events, which provide robustness under different assumptions regarding within-individual correlation (Table S6 in Multimedia Appendix 1); (3) use of ground monitoring station data instead of satellite-based exposure estimates; and (4) exposure estimates weighted by individual time-activity patterns in a subcohort (n=1487), allowing for refined personal exposure assessment.

All analyses were performed using R software (version 4.5.0; R Foundation for Statistical Computing). A 2-sided *P* value <.05 was considered statistically significant.

Results

Descriptive Characteristics

A total of 3617 participants were included, with a median age of 44.1 (IQR 31.6 - 57.7) years and a balanced sex distribution. Most participants had normal BMI (n=2264, 62.6%), had never smoked (n=2881, 79.7%) or consumed alcohol (n=2790, 77.1%), and used clean household fuel (n=3295, 91.1%; Table 1). During the follow-up period spanning 3217 person-years, a total of 885 ARI events were documented, corresponding to an overall incidence rate of 0.27 per person-year. Compared with participants without ARIs, those who experienced ARIs tended to be younger, female, and reside in suburban areas. Participants with ILI showed similar demographic patterns.

XSL•FO
RenderX

Table . Baseline characteristics and acute respiratory infection (ARI) incidence among participants.

Characteristic	Overall (N=3617)	Non-ARI (n=3049)	ARI (n=568)	*P* value[a]	ILI[b] (n=189)
Age (years), median (IQR)	44.1 (31.6 - 57.7)	45.7 (33.0 - 59.3)	36.4 (25.7 - 48.2)	<.001	34.2 (16.9 - 48.1)
Age (years), n (%)				<.001	
<18	447 (12.4)	332 (10.9)	115 (20.2)		48 (25.4)
18 - 59	2377 (65.7)	1980 (64.9)	397 (69.9)		123 (65.1)
≥60	793 (21.9)	737 (24.2)	56 (9.9)		18 (9.5)
Sex, n (%)				<.001	
Male	1803 (49.8)	1573 (51.6)	230 (40.5)		74 (39.2)
Female	1814 (50.2)	1476 (48.4)	338 (59.5)		115 (60.8)
Residence, n (%)				.004	
Central urban	1075 (29.7)	935 (30.7)	140 (24.6)		42 (22.2)
Suburban	2542 (70.3)	2114 (69.3)	428 (75.4)		147 (77.8)
Household size, n (%)				.30	
1	330 (9.1)	275 (9.0)	55 (9.7)		12 (6.3)
2	954 (26.4)	818 (26.8)	136 (23.9)		44 (23.3)
≥3	2333 (64.5)	1956 (64.2)	377 (66.4)		133 (70.4)
BMI category (kg/m^2), n (%)				.06	
Underweight	293 (8.1)	232 (7.6)	61 (10.7)		26 (13.8)
Normal	2264 (62.6)	1923 (63.1)	341 (60.0)		112 (59.3)
Overweight	829 (22.9)	703 (23.1)	126 (22.2)		37 (19.6)
Obese	231 (6.4)	191 (6.3)	40 (7.0)		14 (7.4)
Smoking status, n (%)				<.001	
Never	2881 (79.7)	2386 (78.3)	495 (87.1)		173 (91.5)
Ever	736 (20.3)	663 (21.7)	73 (12.9)		16 (8.5)
Alcohol consumption, n (%)				.003	
Never	2790 (77.1)	2325 (76.3)	465 (81.9)		161 (85.2)
Ever	827 (22.9)	724 (23.7)	103 (18.1)		28 (14.8)
Mask-wearing frequency, n (%)				.006	
Never	508 (14.0)	429 (14.1)	79 (13.9)		24 (12.7)
Occasional	1587 (43.9)	1370 (44.9)	217 (38.2)		77 (40.7)
Frequent	1522 (42.1)	1250 (41.0)	272 (47.9)		88 (46.6)
Underlying disease, n (%)				<.001	
Present	701 (19.4)	637 (20.9)	64 (11.3)		20 (10.6)
Absent	2916 (80.6)	2412 (79.1)	504 (88.7)		169 (89.4)
Vaccination history, n (%)				<.001	
COVID-19 vaccine	2243 (62.0)	1967 (64.5)	276 (48.6)		93 (49.2)
Influenza vaccine	719 (19.9)	528 (17.3)	191 (33.6)		65 (34.4)
Other vaccine	295 (8.2)	226 (7.4)	69 (12.1)		24 (12.7)
Unvaccinated	360 (10.0)	328 (10.8)	32 (5.6)		7 (3.7)
Household fuel type, n (%)				.20	
Clean fuel	3295 (91.1)	2786 (91.4)	509 (89.6)		169 (89.4)
Traditional fuel	322 (8.9)	263 (8.6)	59 (10.4)		20 (10.6)

XSL•FO
RenderX

Characteristic	Overall (N=3617)	Non-ARI (n=3049)	ARI (n=568)	*P* value[a]	ILI[b] (n=189)
Cooking fume exposure, n (%)				.06	
Yes	3035 (83.9)	2573 (84.4)	462 (81.3)		152 (80.4)
No	582 (16.1)	476 (15.6)	106 (18.7)		37 (19.6)

[a]*P* values compare non–acute respiratory infection (ARI) and ARI groups.

[b]ILI: influenza-like illness.

Time-Varying Exposures

The overall median annual mean $PM_{2.5}$ concentration was 19.3 (IQR 14.3 - 31.4) μg/m^3, while the annual mean O_3 concentration was 99.5 (IQR 98.4 - 100.5) μg/m^3(Table 2). Compared with participants without ARIs, those who developed ARIs were exposed to higher levels of $PM_{2.5}$ but lower levels of O_3. This pattern is likely attributable to seasonal variation, as ARI events predominantly occurred during the cold season when O_3 concentrations are typically lower (Figure S2 in Multimedia Appendix 1). A similar pattern was observed for ILI, which also showed higher $PM_{2.5}$ exposure and lower O_3 levels.

Table . Distribution of time-varying environmental exposures and contextual factors by acute respiratory infections (ARIs) during follow-up.

Variables	Overall (N=3617)	Non-ARI (n=3049)	ARI (n=568)	*P* value[a]	ILI[b] (n=189)
$PM_{2.5}$[c] (μg/m^3), median (IQR)	19.3 (14.3 - 31.4)	18.7 (14.2 - 28.7)	24.3 (16.2 - 36.9)	<.001	27.4 (16.4 - 39.8)
O_3[d] (μg/m^3), median (IQR)	99.5 (98.4 - 100.5)	99.9 (98.8 - 100.6)	96.2 (94.8 - 97.3)	<.001	96.3 (94.9 - 97.3)
Temperature (°C), median (IQR)	17.6 (17.6 - 17.7)	17.6 (17.6 - 17.7)	17.5 (16.9 - 17.7)	<.001	17.5 (16.9 - 17.7)
Relative humidity (%), median (IQR)	73.3 (72.6 - 74.5)	73.1 (72.5 - 74.0)	76.8 (75.7 - 77.4)	<.001	76.9 (76.0 - 77.4)
City-level ILI activity (%), median (IQR)	4.2 (4.2 - 4.2)	4.2 (4.2 - 4.2)	4.3 (3.7 - 5.1)	<.001	4.7 (3.9 - 5.3)
Season, n (%)				<.001	
Cold season	1470 (40.6)	1178 (38.6)	292 (51.4)		93 (49.2)
Warm season	2147 (59.4)	1871 (61.4)	276 (48.6)		96 (50.8)

[a]*P* values compare non–acute respiratory infections and ARIs groups.

[b]ILI: influenza-like illness.

[c]$PM_{2.5}$: particulate matter 2.5.

[d]O_3: ozone.

Exposure Window Selection

The associations between air pollution exposure and ARIs or ILI across different exposure windows are presented in Table 3. For $PM_{2.5}$, hazard ratios (HRs) tended to increase with longer exposure windows. For ARIs, the HR increased from 1.018 (95% CI 1.011-1.033) for the 3-month window to 1.031 (95% CI 1.023-1.039) for the 12-month window. For ILI, the HR increased from 1.021 (95% CI 1.011-1.031) for the 3-month window to 1.042 (95% CI 1.030-1.054) for the 12-month window. A similar pattern was observed for O_3, with generally stronger associations observed under longer exposure windows and for ILI. On the basis of the lowest Akaike information criterion, the 12-month exposure window provided the best model fit for both pollutants and was selected for subsequent analyses.

XSL•FO
RenderX

Table . Associations of air pollutants with acute respiratory infections (ARIs) and influenza-like illness (ILI) across different exposure windows and comparison of model fit using the Akaike information criterion [AIC]).

Pollutant and exposure window (months)	ARI		ILI	
	HR[a] (95% CI)[b]	AIC	HR (95% CI)[b]	AIC
$PM_{2.5}$[c]				
3	1.018 (1.011 - 1.033)	14,547.9	1.021 (1.011 - 1.031)	6365.3
6	1.026 (1.018 - 1.033)	14,505.9	1.031 (1.019 - 1.042)	6312.0
9	1.029 (1.021 - 1.037)	14,483.3	1.038 (1.026 - 1.050)	6254.1
12	1.031 (1.023 - 1.039)	14,479.0[d]	1.042 (1.030 - 1.054)	6240.3[d]
O_3[e]				
3	1.044 (1.008 - 1.083)	14,666.3	1.087 (1.022 - 1.157)	6499.8
6	1.123 (1.065 - 1.183)	14,641.4	1.231 (1.127 - 1.343)	6430.2
9	1.191 (1.115 - 1.274)	14,617.3	1.372 (1.231 - 1.530)	6346.5
12	1.262 (1.168 - 1.363)	14,576.7[d]	1.446 (1.279 - 1.636)	6313.7[d]

[a]HR: hazard ratio.

[b]Models were adjusted for age, sex, residence, household size, BMI, smoking status, alcohol consumption, mask-wearing frequency, underlying disease, vaccination history, household fuel, cooking fume exposure, season, and city-level influenza-like illness activity; hazard ratios represent per 1 $\mu g/m^3$ increase.

[c]$PM_{2.5}$: particulate matter 2.5.

[d]Indicates the lowest Akaike information criterion (best model fit) within each pollutant.

[e]O_3: ozone.

Main Associations

On the basis of the 12-month exposure window, both $PM_{2.5}$ and O_3 were associated with increased risks of ARIs, and consistent associations were observed for ILI across all exposure models (Table 4). For $PM_{2.5}$, in the single-pollutant model, each IQR increase was associated with a significantly higher risk of ARIs (HR 1.691, 95% CI 1.428 - 2.002) and ILI (HR 2.188, 95% CI 1.678 - 2.852). After full adjustment for copollutants, temperature, and relative humidity, the associations remained statistically significant, with HRs of 1.594 (95% CI 1.340 - 1.897) for ARIs and 1.948 (95% CI 1.484 - 2.557) for ILI. For O_3, in the single-pollutant model, each IQR increase was associated with HRs of 1.826 (95% CI 1.387 - 2.405) for ARIs and 2.942 (95% CI 1.866 - 4.640) for ILI. After full adjustment, the associations were slightly attenuated but remained statistically significant, with HRs of 1.510 (95% CI 1.135 - 2.007) for ARIs and 2.229 (95% CI 1.385 - 3.588) for ILI. Results per 1 $\mu g/m^3$ increase were consistent in direction and statistical significance (Table S7 in Multimedia Appendix 1).

XSL•FO
RenderX

Table . Associations of air pollutants with acute respiratory infections (ARIs) and influenza-like illness (ILI) under different exposure adjustment models.

Pollutant and model	ARIs		ILI	
	HR[a,b] (95%CI)[c]	*P* value	HR (95%CI)[c]	*P* value
$PM_{2.5}$[d]				
Single pollutant	1.691 (1.428 - 2.002)	<.001	2.188 (1.678 - 2.852)	<.001
+O_3[e]	1.587 (1.334 - 1.887)	<.001	1.942 (1.481 - 2.549)	<.001
+Temperature	1.592 (1.339 - 1.894)	<.001	1.943 (1.481 - 2.550)	<.001
+Humidity	1.594 (1.340 - 1.897)	<.001	1.948 (1.484 - 2.557)	<.001
O_3				
Single pollutant	1.826 (1.387 - 2.405)	<.001	2.942 (1.866 - 4.640)	<.001
+$PM_{2.5}$	1.521 (1.144 - 2.021)	.004	2.219 (1.379 - 3.571)	.001
+Temperature	1.509 (1.135 - 2.006)	.005	2.226 (1.383 - 3.583)	.001
+Humidity	1.510 (1.135 - 2.007)	.005	2.229 (1.385 - 3.588)	.001

[a]HR: hazard ratio.

[b]HRs represent per IQR increase ($PM_{2.5}$=19.4 μg/m^3; O_3=2.5 μg/m^3).

[c]Models were adjusted for age, sex, residence, household size, BMI, smoking status, alcohol consumption, mask-wearing frequency, underlying disease, vaccination history, household fuel, cooking fume exposure, season, and city-level ILI activity, with sequential additional adjustment for copollutant, temperature, and relative humidity.

[d]$PM_{2.5}$: particulate matter 2.5.

[e]O_3: ozone.

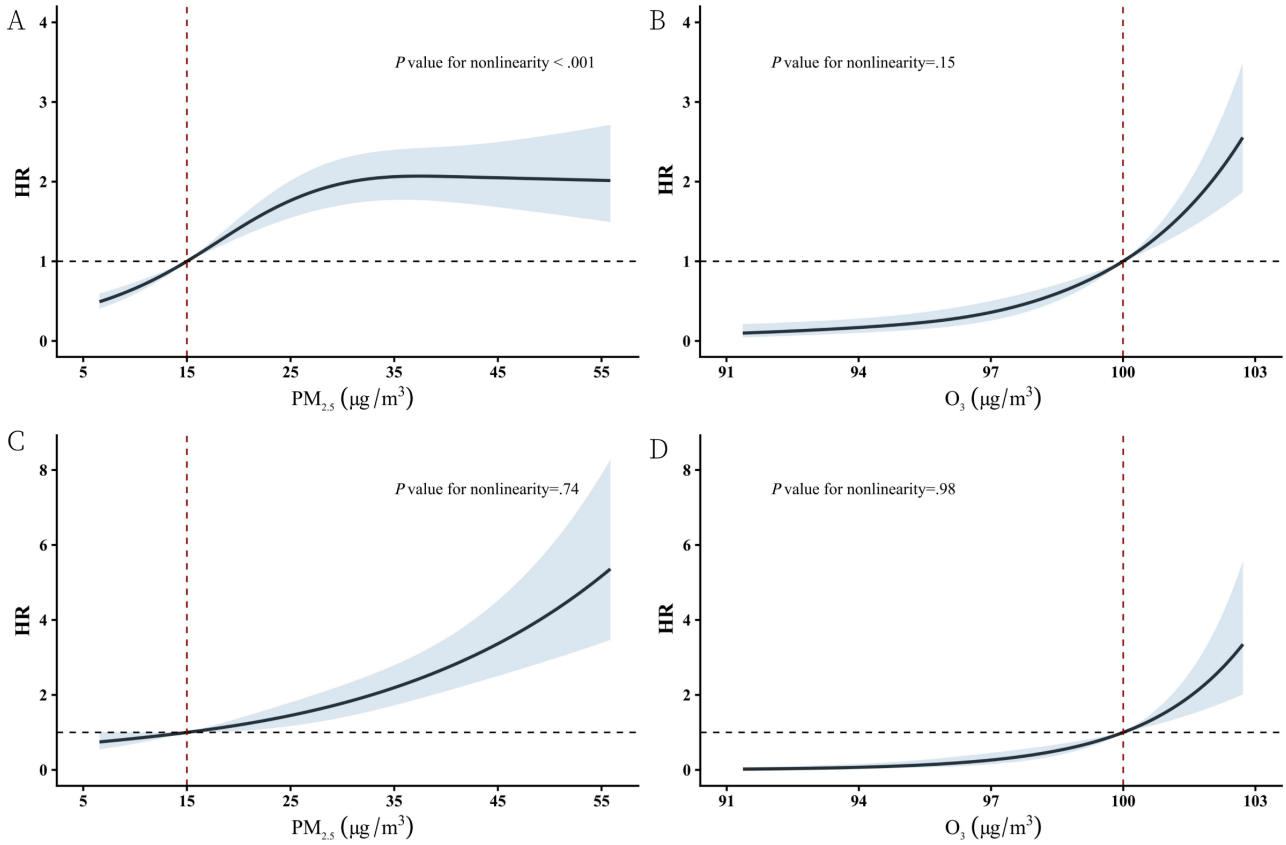
Dose-Response and Subgroup Analyses

On the basis of the WHO air quality guidelines, $PM_{2.5}$ exhibited a nonlinear association with ARIs, characterized by a steeper increase in risk at lower concentrations followed by a plateau at higher concentrations. In contrast, linear exposure-response relationships were observed for $PM_{2.5}$ with ILI and for O_3 with both ARIs and ILI (Figure 1).

Subgroup analyses showed significant effect modification by age, residence, and season in the associations between $PM_{2.5}$ and ARIs, with stronger associations observed among children, older adults, and suburban residents and during the cold season. No significant interactions were observed for sex or smoking status. For O_3, a significant interaction with season was identified for ILI, with stronger associations in the cold season (Figure 2).

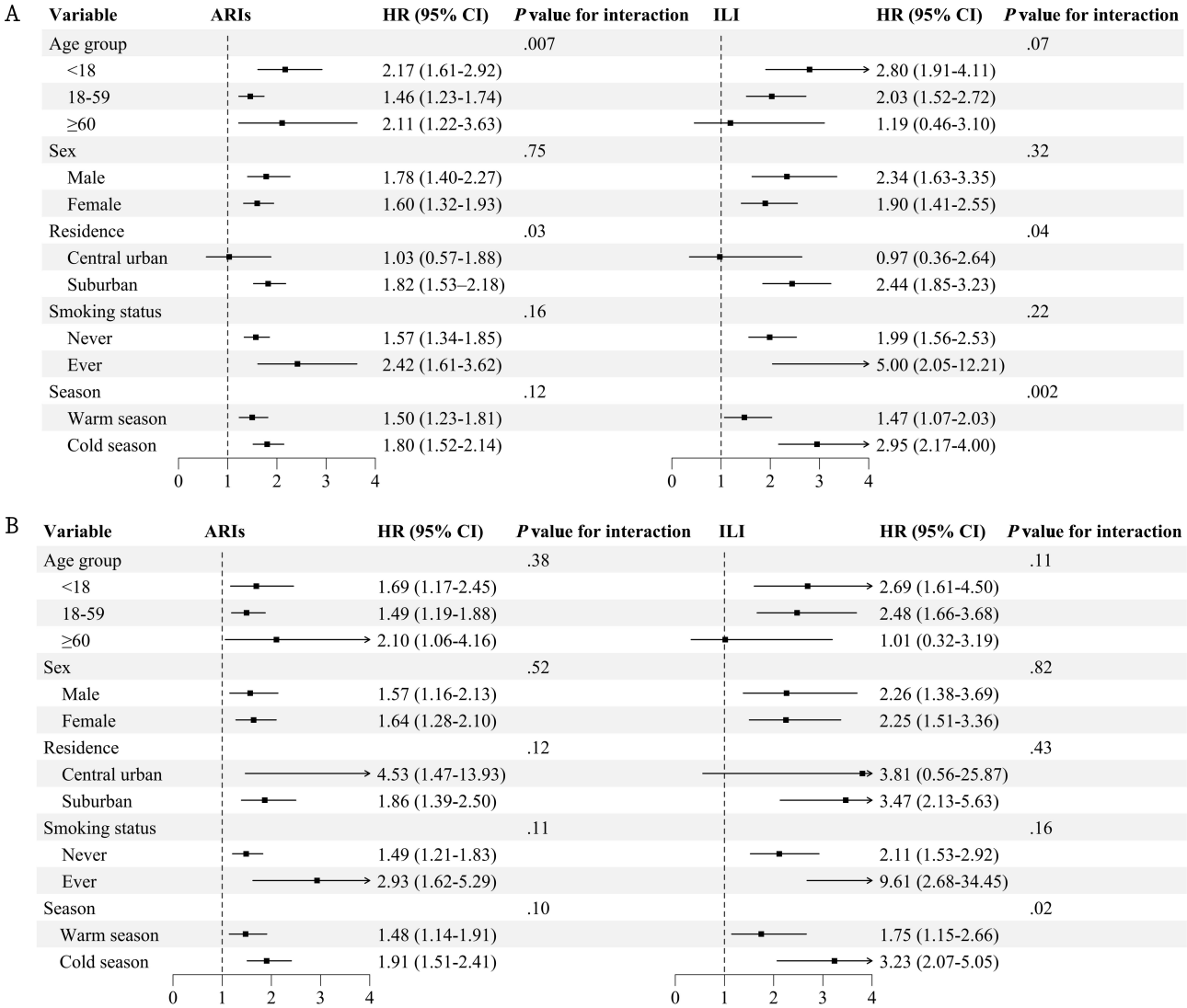
XSL•FO
RenderX

Figure 1. Dose-response relationships of long-term air pollutant exposure with acute respiratory infections (ARIs) and influenza-like illness (ILI): (A) and (B): ARIs and (C) and (D): ILI. Models were adjusted for age, sex, residence, household size, BMI, smoking status, alcohol consumption, mask-wearing frequency, underlying disease, vaccination history, household fuel, cooking fume exposure, season, and city-level ILI activity. Solid lines represent hazard ratios (HRs), and shaded areas indicate 95% CIs. The dashed horizontal line indicates HR=1. The red vertical line represents the World Health Organization (2021) guideline. O_3: ozone; $PM_{2.5}$: particulate matter 2.5.



XSL•FO
RenderX

Figure 2. Subgroup analyses of the associations between long-term air pollution exposure and acute respiratory infections (ARIs) and influenza-like illness (ILI): (A) particulate matter 2.5 ($PM_{2.5}$) and (B) ozone (O_3). Models were adjusted for age, sex, residence, household size, BMI category, smoking status, alcohol consumption, mask-wearing frequency, underlying disease, vaccination history, household fuel, cooking fume exposure, season, and city-level ILI activity. Hazard ratios (HRs) represent per IQR increase ($PM_{2.5}$=19.4 μg/m^3; O_3=2.5 μg/m^3).



Sensitivity Analyses

The associations of $PM_{2.5}$ and O_3 with ARIs and ILI were consistent across different modeling assumptions. Results remained robust when alternative recovery time definitions (7-day and 21-day definitions) and Andersen-Gill models were applied. Using ground-monitoring–based exposures yielded slightly stronger associations. In the subcohort, time-activity–weighted exposure strengthened $PM_{2.5}$ associations, while O_3 showed largely null effects (Tables S8-S12 in Multimedia Appendix 1).

Discussion

Principal Findings

On the basis of individual-level exposure assessment and high-frequency follow-up data, we found that long-term exposure to $PM_{2.5}$ and O_3 was associated with increased risks of ARIs, and similar associations were observed for ILI. These associations remained robust across different exposure windows, model adjustments, and multiple sensitivity analyses. $PM_{2.5}$ showed a nonlinear association with ARIs, while linear relationships were observed for $PM_{2.5}$ with ILI and O_3 with both outcomes. In addition, effect modification was observed by age, residence, and season for $PM_{2.5}$, while stronger effects of O_3 were observed in the cold season.

Comparison With Previous Studies

The incidence density of ARIs in this community-based population was approximately 0.27 per person-year, which was lower than that reported in similar community surveillance studies [25,26]. A 5-year community study in New York reported an incidence density of 0.62 per person-year, which may be partly explained by its broader case definition based on a single symptom [25]. In Hong Kong, the incidence density among older adults was approximately 0.9 per person-year, likely due to the older study population and a higher proportion of winter observations [26]. In some low-income countries, higher incidence rates have been reported, such as 2.66 per person-year in Uganda and up to 3.62 per person-year among

XSL•FO
RenderX

older adults in India [27,28]. Differences in incidence density across studies may reflect variations in case definitions, participant age, follow-up duration, and health care accessibility.

We found that long-term exposure to $PM_{2.5}$ and O_3 was associated with increased risks of ARIs and ILI, consistent with previous epidemiological evidence. A nationwide cohort study in Denmark showed that $PM_{2.5}$ exposure over different averaging periods (with incidence rate ratios ranging from 1.52 for a 1-month averaging period to 1.25 for a 12-month averaging period) was significantly associated with increased ARI risk [13]. However, O_3 was negatively associated with infection risk in that study, which may be explained by its inverse spatial correlation with combustion-related pollutants such as nitrogen oxides. Another Danish prospective cohort further confirmed that long-term $PM_{2.5}$ exposure over 3 years was associated with increased risks of both incident and recurrent LRIs, providing evidence across broader exposure windows [15]. A large Spanish cohort also demonstrated that long-term exposure to $PM_{2.5}$ and O_3 was associated with increased risks of LRIs, and these associations remained robust after accounting for residential mobility [11].

Biological Mechanisms

Existing evidence suggests that the biological mechanisms underlying the effects of air pollution on respiratory infections mainly involve oxidative stress, inflammatory responses, and immune dysregulation. As fine PM capable of reaching the terminal bronchioles and alveolar regions, $PM_{2.5}$ can deposit on the airway epithelium, induce reactive oxygen species production, and activate inflammatory signaling pathways such as nuclear factor kappa-light-chain-enhancer of activated B cells. This may promote the release of cytokines, including interleukin 6 and tumor necrosis factor [29]. Long-term exposure may also disrupt immune regulation, reduce the clearance of viruses and bacteria, and increase susceptibility to ARIs [30]. As a strong oxidant, O_3 can deplete antioxidant defenses in the respiratory tract, induce chronic oxidative stress, damage ciliary structure, and impair mucociliary clearance. It may also contribute to airway remodeling and reduced lung function [31].

Limitations

Several limitations should be acknowledged. First, exposure assessment was based on residential address–matched outdoor air pollution concentrations and did not account for indoor sources or individual time-activity patterns. Second, outcomes were based on symptom definitions without clinical confirmation, which may have introduced outcome misclassification and prevented differentiation between upper and lower respiratory tract infections. Finally, key confounders related to ARI risk, such as hand hygiene practices and exposure to health care settings, were not included in the covariate set, and residual confounding cannot be excluded.

Conclusions

In conclusion, this prospective community-based study provides evidence that long-term exposure to $PM_{2.5}$ and O_3 is independently associated with increased risk of ARIs, with consistently stronger associations observed for ILI. Children and older adults emerged as particularly vulnerable subgroups, and the effects are more pronounced during cold seasons, underscoring the importance of targeted interventions for high-risk populations. These findings highlight the importance of long-term air pollution control and support targeted public health interventions for high-risk populations to reduce the burden of respiratory infections in the community. Individual mobility and indoor exposure should be considered in future exposure assessments, and the biological mechanisms linking long-term air pollution to ARIs should be further explored.

Acknowledgments

The authors would like to thank all study participants, as well as general practitioners and community health service staff for their valuable support. WW, YZ, and HW contributed equally as co-corresponding authors.

Funding

This work was supported by the Shanghai Municipal Science and Technology Major Project (grant ZD2021CY001), the Shanghai New Three-year Action Plan for Public Health (grant GWVI-1, GWVI-11.1-03 and GWVI-11.1-01), the Major Program of National Social Science Foundation of China (grant 22&ZD142), the Eastern Talent Plan Youth Project (2023), and the National Natural Science Foundation of China (grant 82073612).

Data Availability

The data and R code supporting the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization, supervision, and project administration: WW, YZ, and HW.
Methodology and formal analysis: XD, XY, and HL.
Data curation and investigation: ZW, WZ, and YM.
Funding acquisition: HW.
Writing–original draft: XD, XY, and HL.
Writing–review & editing: XD, XY, HL, ZW, WZ, YM, WW, YZ, and HW.

XSL•FO
RenderX

All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables and figures for sampling, variables, model diagnostics, and sensitivity analyses.
[DOCX File, 1993 KB - publichealth_v12i1e88045_app1.docx]

References

1. Monoson A, Schott E, Ard K, et al. Air pollution and respiratory infections: the past, present, and future. Toxicol Sci 2023 Mar 20;192(1):3-14. [doi: 10.1093/toxsci/kfad003] [Medline: 36622042]
2. GBD 2021 Lower Respiratory Infections and Antimicrobial Resistance Collaborators. Global, regional, and national incidence and mortality burden of non-COVID-19 lower respiratory infections and aetiologies, 1990-2021: a systematic analysis from the Global Burden of Disease Study 2021. Lancet Infect Dis 2024 Sep;24(9):974-1002. [doi: 10.1016/S1473-3099(24)00176-2] [Medline: 38636536]
3. GBD 2021 Upper Respiratory Infections Otitis Media Collaborators. Global, regional, and national burden of upper respiratory infections and otitis media, 1990-2021: a systematic analysis from the Global Burden of Disease Study 2021. Lancet Infect Dis 2025 Jan;25(1):36-51. [doi: 10.1016/S1473-3099(24)00430-4] [Medline: 39265593]
4. Vieceli T, Tejada S, Martinez-Reviejo R, et al. Impact of air pollution on respiratory microbiome: a narrative review. Intensive Crit Care Nurs 2023 Feb;74:103336. [doi: 10.1016/j.iccn.2022.103336] [Medline: 37440188]
5. Petat H, Corbet S, Leterrier B, Vabret A, Ar Gouilh M. Unravelling the acute respiratory infection landscape: virus type, viral load, health status and coinfection do matter. Front Cell Infect Microbiol 2024;14:1380855. [doi: 10.3389/fcimb.2024.1380855] [Medline: 38803572]
6. Hassen S, Getachew M, Eneyew B, et al. Determinants of acute respiratory infection (ARI) among under-five children in rural areas of Legambo District, South Wollo Zone, Ethiopia: a matched case-control study. Int J Infect Dis 2020 Jul;96:688-695. [doi: 10.1016/j.ijid.2020.05.012] [Medline: 32413607]
7. Yang S, Tong T, Wang H, Li Z, Wang M, Ni K. Causal relationship between air pollution and infections: a two-sample Mendelian randomization study. Front Public Health 2024;12:1409640. [doi: 10.3389/fpubh.2024.1409640] [Medline: 39148655]
8. Collaborators GH. Global, regional, and national burden of household air pollution, 1990-2021: a systematic analysis for the Global Burden of Disease Study 2021. Lancet 2025 Apr 5;405(10485):1167-1181. [doi: 10.1016/S0140-6736(24)02840-X] [Medline: 40118081]
9. Danesh Yazdi M, Wang Y, Di Q, Zanobetti A, Schwartz J. Long-term exposure to PM2.5 and ozone and hospital admissions of Medicare participants in the Southeast USA. Environ Int 2019 Sep;130:104879. [doi: 10.1016/j.envint.2019.05.073] [Medline: 31238267]
10. Fuller R, Landrigan PJ, Balakrishnan K, et al. Pollution and health: a progress update. Lancet Planet Health 2022 Jun;6(6):e535-e547. [doi: 10.1016/S2542-5196(22)00090-0] [Medline: 35594895]
11. Alari A, Ranzani O, Milà C, et al. Long-term exposure to air pollution and lower respiratory infections in a large population-based adult cohort in Catalonia. Environ Int 2025 Jan;195:109230. [doi: 10.1016/j.envint.2024.109230] [Medline: 39732111]
12. Gandini M, Scarinzi C, Bande S, et al. Long term effect of air pollution on incident hospital admissions: results from the Italian Longitudinal Study within LIFE MED HISS project. Environ Int 2018 Dec;121(Pt 2):1087-1097. [doi: 10.1016/j.envint.2018.10.020] [Medline: 30366659]
13. Kaspersen KA, Antonsen S, Horsdal HT, et al. Exposure to air pollution and risk of respiratory tract infections in the adult Danish population-a nationwide study. Clin Microbiol Infect 2024 Jan;30(1):122-129. [doi: 10.1016/j.cmi.2023.10.013] [Medline: 37858866]
14. Hayward AC, Fragaszy EB, Bermingham A, et al. Comparative community burden and severity of seasonal and pandemic influenza: results of the Flu Watch cohort study. Lancet Respir Med 2014 Jun;2(6):445-454. [doi: 10.1016/S2213-2600(14)70034-7] [Medline: 24717637]
15. Zhang J, Lim YH, So R, et al. Long-term exposure to air pollution and risk of acute lower respiratory infections in the Danish Nurse Cohort. Ann Am Thorac Soc 2024 Aug;21(8):1129-1138. [doi: 10.1513/AnnalsATS.202401-074OC] [Medline: 38513223]
16. Ozga AK, Kieser M, Rauch G. A systematic comparison of recurrent event models for application to composite endpoints. BMC Med Res Methodol 2018 Jan 4;18(1):2. [doi: 10.1186/s12874-017-0462-x] [Medline: 29301487]
17. Yu X, Wen Z, Zhou W, et al. Respiratory syndromic disease study in Shanghai community population. BMJ Open 2026 Mar 18;16(3):e103001. [doi: 10.1136/bmjopen-2025-103001] [Medline: 41857851]

XSL•FO
RenderX

18. Geng G, Xiao Q, Liu S, et al. Tracking air pollution in China: near real-time PM2.5 retrievals from multisource data fusion. Environ Sci Technol 2021 Sep 7;55(17):12106-12115. [doi: 10.1021/acs.est.1c01863] [Medline: 34407614]
19. Muñoz-Sabater J, Dutra E, Agustí-Panareda A, et al. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. Earth Syst Sci Data 2021;13(9):4349-4383. [doi: 10.5194/essd-13-4349-2021]
20. Zhou W, Wen Z, Sun B, Chen R, Xiong H, Wang W. Global analysis links ambient fine particulate matter to bacterial antimicrobial resistance based on 110,000 + genomes. Environ Int 2025 Aug;202:109696. [doi: 10.1016/j.envint.2025.109696] [Medline: 40714380]
21. Cui AL, Xia BC, Zhu Z, et al. Epidemiological characteristics of human respiratory syncytial virus among acute respiratory infection cases in 16 provinces of China from 2009 to 2023 [Article in Chinese]. Zhonghua Yu Fang Yi Xue Za Zhi 2024 Jul 6;58(7):945-951. [doi: 10.3760/cma.j.cn112150-20231213-00440] [Medline: 38403282]
22. Howard LM, Liu Y, Zhu Y, et al. Assessing the impact of acute respiratory illnesses on the risk of subsequent respiratory illness. J Infect Dis 2022 Jan 5;225(1):42-49. [doi: 10.1093/infdis/jiab313] [Medline: 34120189]
23. WHO Regional Office for Europe guidance for sentinel influenza surveillance in humans. WHO Institutional Repository for Information Sharing (IRIS). 2011. URL: https://iris.who.int/handle/10665/349780 [accessed 2026-06-24]
24. Amorim LD, Cai J. Modelling recurrent events: a tutorial for analysis in epidemiology. Int J Epidemiol 2015 Feb;44(1):324-333. [doi: 10.1093/ije/dyu222] [Medline: 25501468]
25. Stockwell MS, Reed C, Vargas CY, et al. Five-year community surveillance study for acute respiratory infections using text messaging: findings from the MoSAIC study. Clin Infect Dis 2022 Sep 29;75(6):987-995. [doi: 10.1093/cid/ciac027] [Medline: 35037056]
26. Han L, Ran J, Chan KH, et al. Indoor environmental factors and acute respiratory illness in a prospective cohort of community-dwelling older adults. J Infect Dis 2020 Aug 17;222(6):967-978. [doi: 10.1093/infdis/jiaa188] [Medline: 32297941]
27. Musiime AK, Krezanoski PJ, Smith DL, et al. House design and risk of malaria, acute respiratory infection and gastrointestinal illness in Uganda: a cohort study. PLOS Glob Public Health 2022;2(3):e0000063. [doi: 10.1371/journal.pgph.0000063] [Medline: 36962263]
28. Kumar R, Dar L, Amarchand R, et al. Incidence, risk factors, and viral etiology of community-acquired acute lower respiratory tract infection among older adults in rural north India. J Glob Health 2021 Apr 3;11:04027. [doi: 10.7189/jogh.11.04027] [Medline: 33880179]
29. Priyadarshini NP, Gopamma D, Srinivas N, Malla RR, Kumar KS. Particulate matter and its impact on macrophages: unraveling the cellular response for environmental health. Crit Rev Oncog 2024;29(4):33-42. [doi: 10.1615/CritRevOncog.2024053305] [Medline: 38989736]
30. Qin T, Zhang F, Zhou H, et al. High-level PM2.5/PM10 exposure is associated with alterations in the human pharyngeal microbiota composition. Front Microbiol 2019;10:54. [doi: 10.3389/fmicb.2019.00054] [Medline: 30804895]
31. Celebi Sözener Z, Cevhertas L, Nadeau K, Akdis M, Akdis CA. Environmental factors in epithelial barrier dysfunction. J Allergy Clin Immunol 2020 Jun;145(6):1517-1528. [doi: 10.1016/j.jaci.2020.04.024] [Medline: 32507229]

Abbreviations

ARI: acute respiratory infection
HR: hazard ratio
ILI: influenza-like illness
LRI: lower respiratory infection
O_3: ozone
PM: particulate matter
WHO: World Health Organization

Edited by C Argyropoulos; submitted 19.Nov.2025; peer-reviewed by E Etafo, F Mallapiang, H Qian, W Zhong, Y Hu; revised version received 06.Jun.2026; accepted 09.Jun.2026; published 13.Jul.2026.

Please cite as:
Duan X, Yu X, Liang H, Wen Z, Zhou W, Ma Y, Wang W, Zheng Y, Wu H
Long-Term Exposure to Particulate Matter 2.5 and Ozone and the Risk of Acute Respiratory Infections: Community-Based Prospective Cohort Study
JMIR Public Health Surveill 2026;12:e88045
URL: https://publichealth.jmir.org/2026/1/e88045
doi:10.2196/88045

XSL•FO
RenderX

©Xiaolei Duan, Xiao Yu, Hongyu Liang, Zexuan Wen, Wenyong Zhou, Ying Ma, Weibing Wang, Yaxu Zheng, Huanyu Wu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 13.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Longitudinal Population-Level Shifts in Infectious Disease–Specific Health Literacy During and After the COVID-19 Pandemic in Zhejiang, China: Six-Round Cross-Sectional Study

Zhao Yusui; Dingming Yao*; Yue Xu; Xiujing Hu; Qingqing Wu; Lei Wang; Xuehai Zhang*; Zhen Jiang*

Zhejiang Provincial Center for Disease Control and Prevention, 3399 Binsheng Road, Hangzhou, Zhejiang, China

*these authors contributed equally

Corresponding Author:
Zhen Jiang
Zhejiang Provincial Center for Disease Control and Prevention, 3399 Binsheng Road, Hangzhou, Zhejiang, China

Abstract

Background: Infectious disease–specific health literacy (IDSHL)—the capacity to obtain, process, and understand basic infectious disease information needed to make appropriate health decisions—is a critical determinant of public health outcomes. The COVID-19 pandemic created an unprecedented global public health education context; yet, its long-term association with population-level IDSHL beyond the pandemic remains unclear. In China, stringent public health measures and digital information campaigns were implemented, yet their lasting impact on IDSHL across different disease domains has not been systematically evaluated over a full pandemic cycle.

Objective: This study aimed to evaluate the longitudinal trends in IDSHL among residents of Zhejiang Province, China, from 2019 (prepandemic) to 2024 (postpandemic), and to assess whether the pandemic period was temporally associated with sustained improvements in IDSHL.

Methods: Six annual cross-sectional surveys were conducted from 2019 to 2024 across 30 counties in Zhejiang Province using consistent multistage stratified random sampling. A validated 12-item IDSHL questionnaire assessed knowledge, behavior, and skills. Annual sample sizes ranged from 17,131 to 19,257 (N=112,917). Data were weighted to the 2020 census population by age and urban or rural residence. Joinpoint regression identified inflection points in the annual proportion of residents with adequate IDSHL (score≥12). Multivariate logistic regression estimated adjusted odds ratios (ORs) for adequate IDSHL by year, with 2021 as the reference (the peak acute pandemic phase), controlling for sociodemographics. All analyses incorporated sampling weights.

Results: Population-weighted overall IDSHL scores increased from 2019 (mean 10.46 SD 3.09) to 2024 (mean 11.40 SD 2.88; $P<.001$). All subscale scores (knowledge, behavior, and skills) showed upward trends (all $P<.001$). Joinpoint regression revealed a rapid annual increase in adequate IDSHL during 2019-2021 (annual percent change 9.42%, 95% CI 6.10 - 13.45; $P<.001$), which slowed post-2021 (annual percent change 2.28%, 95% CI –0.38 to 3.73; $P=.07$). Correct response rates surged for pandemic-salient items (eg, respiratory etiquette from 28.44% to 39.67%), while hepatitis B transmission knowledge fluctuated (63.28% in 2019, 60.71% in 2022, 63.84% in 2024), indicating that not all non–COVID-19 content improved monotonically. Adjusted logistic regression showed the lowest odds of adequate IDSHL in 2019 (OR 0.74, 95% CI 0.71 - 0.78) and highest in 2024 (OR 1.10, 95% CI 1.05 - 1.15) compared to 2021.

Conclusions: The COVID-19 pandemic was temporally associated with significant and sustained population-level improvements in IDSHL, particularly for pandemic-relevant knowledge and behaviors. However, gains were time-sensitive and not uniform across all disease domains or demographic groups. These findings underscore the need for postpandemic health strategies that reinforce comprehensive IDSHL through sustained education and address digital divides to bridge persistent equity gaps. Policymakers should prioritize low digital-literacy populations through hybrid (offline+online) educational interventions, and sustain routine surveillance of IDSHL to detect declines in nonpandemic disease knowledge. This study is the first to document six-year population-level IDSHL trends across the full pandemic cycle, providing evidence for future crisis communication and routine health promotion.

(JMIR Public Health Surveill 2026;12:e89958) doi:10.2196/89958

KEYWORDS

IDSHL; COVID-19; longitudinal study; public health; China; infectious disease-specific health literacy

XSL•FO
RenderX

Introduction

Infectious disease–specific health literacy (IDSHL)—defined as the cognitive, behavioral, and skills-based capacity to obtain, process, and understand basic infectious disease information needed to make appropriate health decisions—is a critical social determinant of health outcomes [1,2]. It empowers individuals to engage in preventive behaviors, adhere to treatment regimens, and navigate health systems effectively [3,4]. The COVID-19 pandemic, caused by SARS-CoV-2, placed unprecedented emphasis on public understanding of transmission dynamics, nonpharmaceutical interventions (NPIs), and vaccination [5-7]. For nearly three years, governments worldwide, including China, implemented rigorous containment strategies and launched intensive public health communication campaigns [8,9]. The pandemic evolution in China was broadly characterized by experts into several distinct phases: the prepandemic era (2019), the acute pandemic phase (2020‐2022, characterized by stringent NPIs and intensive public health messaging), and the initial postpandemic period (2023‐2024, following the transition to endemic management in January 2023) [10-13]. This phased approach provides a structured framework to examine the differential impacts of varying intensities of public health interventions on IDSHL over time.

This period offered a unique opportunity to examine population-level changes in IDSHL in response to a major, sustained public health crisis. By capitalizing on this unique context, we aimed to assess whether a shift in IDSHL was observed during and after the pandemic. It is particularly noteworthy that during the COVID-19 pandemic, digital platforms, and social media (eg, WeChat, Weibo, and news applications) became the central channels through which the public accessed authoritative pandemic updates, health guidance, and subsequently formed health-related behaviors [14,15]. This shift created a highly digitalized health information environment. Consequently, the observed changes in public IDSHL captured by our consecutive surveys coincide with the period of intense digital information exposure, during which public cognition and behavior were shaped by large-scale digital health messaging.

It is crucial to emphasize that this study uses an observational, repeated cross-sectional design without a control group. Therefore, while we document strong temporal associations between the pandemic phases and changes in IDSHL, we cannot infer causality. All interpretations of “association,” “temporal relationship,” or “observed change” refer to statistical associations at the population level.

Prior to the pandemic, studies indicated suboptimal levels of IDSHL in various populations, with significant gaps in knowledge about both common and emerging infections [16]. Research from our previous work using 2019 and 2022 data suggested potential improvements in IDSHL, especially for respiratory pathogens, but was limited to two time points [17]. Understanding whether the seismic event of COVID-19 was temporally associated with a transient or lasting shift in the public’s IDSHL is crucial [18]. A sustained elevation would suggest that major public health crises may coincide with durable positive changes in population health competence. Conversely, a reversion to baseline or uneven progress would highlight the need for more targeted, persistent educational efforts beyond crisis periods.

Based on a unique longitudinal dataset comprising six consecutive annual surveys (2019‐2024) from Zhejiang Province, China, we propose a temporal reconceptualization for analysis. The timeline is divided into three distinct phases aligned with fundamental shifts in the pandemic response focus: stage 1: the prepandemic era (2019), stage 2: the acute pandemic phase (2020‐2022), and stage 3: the initial postpandemic period (2023‐2024). This phased framework corresponds directly with the annual measurement points, thereby enabling a clearer examination of the longitudinal patterns associated with a major public health crisis.

Our primary objectives were to (1) describe longitudinal trends in overall and domain-specific IDSHL scores, (2) examine changes in correct response rates for specific IDSHL items, and (3) assess temporal changes in the proportion and adjusted odds of having adequate IDSHL across the prepandemic, acute pandemic, and postpandemic phases. We hypothesized that the pandemic would be temporally associated with a stepwise and sustained increase in IDSHL, with larger gains for pandemic-salient content (eg, respiratory etiquette and vaccination) and persistent disparities across sociodemographic groups.

Methods

Study Design

This study used a repeated cross-sectional design, conducting independent surveys annually from 2019 to 2024. Given the observational, repeated cross-sectional design without a control group, we interpret our findings as population-level associations over time, not causal effects.

Setting

The surveys were conducted in Zhejiang Province, a major coastal province in eastern China with a diverse population of over 64 million. The formal survey was conducted annually between August 1 and October 31 of each year from 2019 to 2024. Recruitment and data collection occurred simultaneously during this period. The surveys were carried out within 30 representative counties or districts surveillance sites across the province [19]. These sites were selected using stratified cluster random sampling based on per capita gross domestic product (tertiles) to ensure representativeness. The methods for sampling, sample size determination, and data collection were consistent with those described in our previously published two-round cross-sectional study using the 2019 and 2022 data from the same surveillance system [17]. This study extends this work by including four additional survey rounds (2020, 2021, 2023, and 2024) to cover the full pandemic cycle from prepandemic to postpandemic periods.

XSL•FO
RenderX

Participants

Inclusion and Exclusion Criteria

Eligibility criteria were (1) age 15‑69 years, (2) ability to communicate in Mandarin or local dialect, and (3) permanent resident status (accessibility). No additional exclusion criteria were applied. There were no restrictions based on sex, ethnicity, education, or other demographic characteristics.

Sampling Procedures

A multistage stratified random sampling method was consistently applied. The sampling procedure followed five stages, as detailed in our previous publication [17]: (1) county selection: 30 counties and districts were selected from the 90 counties in Zhejiang Province using stratified cluster random sampling based on per capita gross domestic product (tertiles); (2) township selection: within each selected county, 4 townships were randomly selected; (3) segment (residential block) selection: within each township, 2 residential segments were randomly selected; (4) household selection: approximately 100 households were randomly selected from a complete address list within each segment; and (5) individual selection: one permanent resident aged 15‑69 years per household was selected using a Kish grid.

Participant Characteristics

Participants were permanent residents of Zhejiang Province aged 15‑69 years. Across the six annual surveys (N=112,917), approximately 52% (n=58,717) were female and 48% (n=54,200) were male each year. The sample included participants with education levels ranging from primary school or lower (about 27%, 30,487/112,917 to 36%, 40,650/112,917 across years) to undergraduate or higher (less than 1% in each year). Farmers were the largest occupational group (35,004/112,917, 31%-51,941/112,917, 46% across years), with a decreasing trend over time.

Study Size, Power, and Precision

The target sample size was determined using the same method as our previous study [17]. Briefly, the sample size per county was calculated using the formula:

$$N=\mu\alpha 2\times p(1-p)\delta 2\times deff$$

We set α=.05 (hence μ_α=1.96), used the 2018 health literacy rate for Zhejiang (26.24%) as the expected prevalence p, allowed a maximum error δ=0.03936, and took a design effect deff=1 for simple random sampling. After accounting for an anticipated 25% rate of invalid questionnaires or refusals, the final target sample size per county was 640. With 30 counties, the total annual target sample size was approximately 19,200. No formal a priori power analysis for detecting specific effect sizes was conducted, as the primary goal was precise annual surveillance rather than hypothesis testing of a single a priori effect. The achieved annual sample sizes ranged from 17,131 to 19,257 across years. Precision was quantified using 95% CI for all point estimates (eg, proportions, odds ratios [ORs], annual percent changes [APC]).

Measures and Covariates

Outcome, Exposures, and Covariates

The primary outcome was adequate IDSHL, defined as a score of ≥12 on the IDSHL scale (see Instrumentation; Multimedia Appendix 1). The main exposure of interest was survey year (2019‑2024), which served as a proxy for the pandemic phase (prepandemic, acute pandemic, and postpandemic). Covariates included sex, age group (15-29, 30-39, 40-49, 50-59, and 60-69), ethnicity (Han vs minority), education level (primary school or lower, middle school, high school, technical school or college, undergraduate or higher), marital status (married vs unmarried, divorced, or widowed), occupation (farmer, worker, agency or institutional personnel, student, and other), and urban or rural residence. All covariates were self-reported and treated as potential confounders in multivariable analyses.

Data Collection

Data were collected via household visits using tablet computers. Prior to household visits, community workers or community health physicians (investigators) contacted the sampled household head by telephone, informed them of the selected family member, and scheduled a face-to-face visit. If the appointment failed, another was scheduled. If three consecutive attempts to schedule a successful visit failed, the household was replaced with the next household on the sampling list, and the recruitment process was restarted. Prior to the formal survey, both the surveyor and the quality control supervisor verified whether the participating family member matched the randomly selected subject. If not, the correct member was requested. If the selected member was unavailable, a new appointment was scheduled via telephone.

Questionnaire administration followed 2 modes: if the respondent was capable of self-completion, the surveyor remained present to provide assistance and correct obvious errors in basic information. If the respondent could not complete the questionnaire independently, the surveyor administered it through face-to-face interviewing.

Quality of Measurements

To ensure data authenticity and quality, the following measures were implemented.

Training

All surveyors, supervisors, and quality control personnel underwent a standardized training program consisting of collective lectures, on-site demonstrations, role-playing or simulation exercises, and assessments based on presurvey pilot tests. Unified technical manuals and standardized teaching materials were used. Trainers were staff members from the Zhejiang Provincial Center for Disease Control and Prevention, each with over 5 years of experience in IDSHL monitoring.

Audio Recording

Audio recording was activated throughout each survey session.

Photography

After the survey, the surveyor took photographs of the setting for subsequent verification.

XSL•FO
RenderX

Multidimensional Quality Checks

Supervisors reviewed audio recordings, survey duration, and on-site photos. A questionnaire was deemed invalid and replaced with a new household if (1) leading questions were present in the audio recording, (2) the survey duration was less than 10 minutes (a threshold established during pilot testing as the minimum reasonable time), and (3) there was an inconsistency between on-site photos and actual survey conditions.

Electronic Information System

The entire process—questionnaire administration, data collection, storage, and quality control—was managed through an integrated electronic information system, ensuring data consistency and integrity.

Instrumentation

IDSHL was assessed using a validated 12-item subscale from the National Health Literacy Surveillance Survey in China [17]. The scale covers 3 dimensions: knowledge (6 items), behavior (4 items), and skills (2 items). It includes one true or false item, 8 single-choice items, and 3 multiple-choice items. Correct answers to multiple-choice questions were awarded 2 points, while single-choice and true or false items received 1 point each, for a total score of 15. Participants scoring ≥12 points (≥80%) were classified as having "adequate IDSHL" [19].

This cutoff has been widely used in China's National Health Literacy Surveillance System and is recommended in the official surveillance guidelines [17].

Masking

Given the observational, cross-sectional survey design, no formal blinding or masking procedures were implemented. Surveyors were trained to administer the questionnaire in a standardized manner without revealing the study hypotheses or influencing responses. Participants were not informed of specific research hypotheses beyond the general purpose of the health literacy survey. No experimental manipulation or group assignment occurred, so masking of condition assignments was not applicable.

Psychometrics

All 12 items demonstrated satisfactory content validity, with indices greater than 0.8. The overall internal consistency of the scale was acceptable (Cronbach α=0.67). For the 3 subscales, Cronbach α values were 0.71 for knowledge, 0.65 for behavior, and 0.63 for skills. Given the scale's multidimensionality and diverse item content (covering multiple diseases and response formats), these coefficients are considered adequate for population-level surveillance [17]. Reliability estimates are based on the current sample (N=112,917). No additional validity coefficients from external samples are reported, as the scale is a national surveillance instrument with established content validity.

Conditions and Design

This was a nonexperimental (observational) design with no manipulated conditions. The study used a repeated cross-sectional design (six annual waves, 2019 - 2024). No randomization or intervention assignment occurred. The primary comparison was across time (survey year) as a proxy for pandemic phases.

Data Diagnostics

Planned data diagnostics included the following prespecified criteria: (1) post data collection exclusion of participants was based solely on the 3 quality control criteria (leading questions, duration <10 min, and photo inconsistency); no additional exclusion criteria were applied. (2) Missing data: less than 2% of sampled households were replaced due to quality issues; no imputation was performed for missing item responses as the questionnaire was reviewed for completeness on-site. We conducted Little Missing Completely At Random (MCAR) test on the full dataset (N=112,917) using full information maximum likelihood. The result was nonsignificant (χ^2_{145}=156.73; P=.24), indicating that missing data (at the household replacement level) were missing completely at random (degrees of freedom derived from the number of variables and observed missing patterns). (3) Outliers: no statistical outlier removal was performed; all valid responses were retained. (4) Data distributions: normality of IDSHL scores was assessed visually using histograms and Q-Q plots; for nonnormal distributions, nonparametric tests (Kruskal-Wallis) were used as specified. (5) Data transformations: no data transformations were applied.

Analytic Strategy

Data were analyzed using SPSS (version 27.0; IBM Corp) and Joinpoint Regression Program (version 5.4.0; National Cancer Institute). Descriptive statistics were presented as frequencies (percentages) for categorical variables and means (SDs) for continuous variables. No qualitative analysis was performed.

Weighting

To ensure the representativeness of the provincial estimates, all analyses (descriptive, ANOVA, Kruskal-Wallis, Joinpoint, and logistic regression) incorporated sampling weights. Weights were calculated as the inverse probability of selection, poststratified to the 2020 Zhejiang census population distribution by age group (15-29, 30-39, 40-49, 50-59, and 60-69 years) and urban or rural residence [20].

Group Comparisons

The Kruskal-Wallis H test was used to compare distributions of sociodemographic characteristics and correct response rates for individual items across the 6 years due to nonnormality. One-way ANOVA was used to compare weighted mean IDSHL scores (overall and subscales) across years, with post hoc least significant difference (LSD) tests for pairwise comparisons. No adjustment for multiple comparisons was applied to least significant difference tests, as these were considered exploratory; however, the overall ANOVA F test protected the type I error rate at the .05 level for the primary hypothesis of any yearly difference.

Joinpoint Regression

Joinpoint regression was used to model the temporal trend in the weighted annual proportion of residents with adequate IDSHL from 2019 to 2024 and to identify significant inflection points. The maximum number of joinpoints was set to 1 given

XSL•FO
RenderX

the limited number of time points. The APC and average annual percent change (AAPC) with 95% CIs were calculated.

Logistic Regression

Multivariate logistic regression was performed with "adequate IDSHL" (score≥12) as the dependent variable. The year 2021 was chosen as the reference category because it represents the peak of the acute pandemic phase in Zhejiang Province (characterized by the "Dynamic COVID-zero" strategy and intensive public health messaging). Joinpoint regression identified 2021 as a significant inflection point. All models adjusted for sex, age group, ethnicity, education, marital status, and occupation. Sampling weights were applied using SPSS Complex Samples module. A two-tailed *P* value of <.05 was considered statistically significant.

Relationship Between Joinpoint and Logistic Regression Results

The Joinpoint results describe the trajectory of the population prevalence of adequate IDSHL, while the logistic regression quantifies the individual-level adjusted odds for each year relative to 2021. Together, they show a consistent pattern: a sharp initial increase from 2019 to 2021 (lowest OR in 2019, steep positive APC), followed by a plateau from 2021 onwards (nonsignificant ORs for 2022 - 2024 compared to 2021, and a nonsignificant APC after 2021). Missing data were minimal (<2% of sampled households replaced due to quality issues). No imputation was performed.

Ethical Considerations

Ethics approval

The study protocol was approved by the Ethics Committee of the Zhejiang Provincial Center for Disease Control and Prevention (approval No. 2022-027-01 for the 2019 and 2022 surveys; subsequent annual surveys were conducted under the same ethical approval with protocol amendments approved by the committee). The study was conducted in accordance with the principles of the Declaration of Helsinki.

Informed consent

Written informed consent was obtained from all participants prior to the survey. Participants were informed about the purpose of the study, the procedures (including audio recording and photography for quality control), the voluntary nature of participation, and their right to withdraw at any time. The consent form explicitly covered the use of anonymized data for research and publication.

Privacy and confidentiality

All data were anonymized prior to analysis. No personal identifiers (eg, names, ID numbers, and exact addresses) were retained in the analytic dataset. Audio recordings and photographs were used solely for real-time quality control and were deleted upon completion of data validation. Access to the deidentified data was restricted to the research team.

Compensation

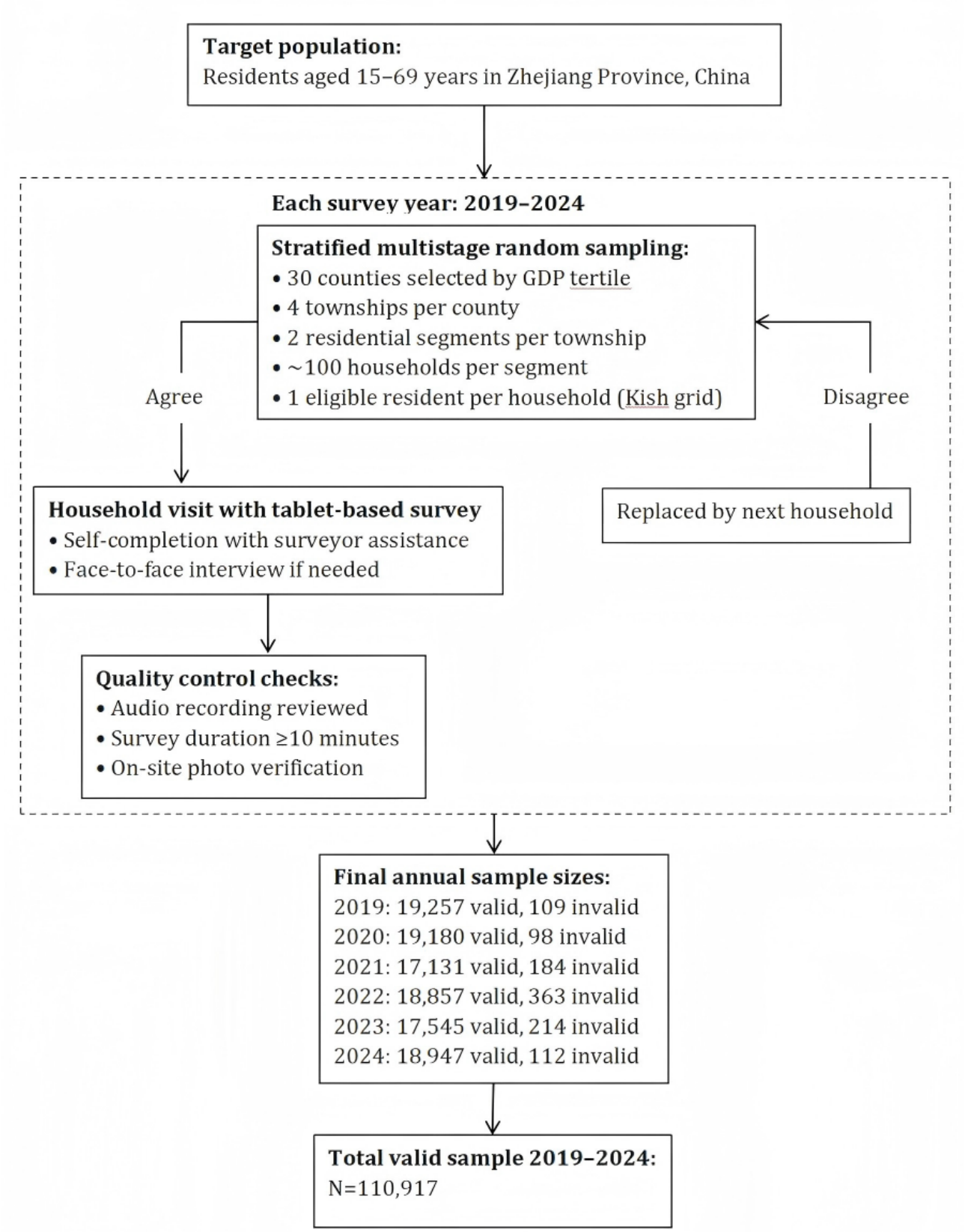
Upon completion of the survey, each participant received a small gift valued at approximately 50 RMB (about US $7) as compensation for their time.

Identifiable images

Figure 1 (participant flow diagram) contains no identifiable individual images. No photographs or images of individual participants are included in the paper or supplementary materials.

XSL•FO
RenderX

Figure 1. Participant flow diagram, showing the number of households contacted, eligible, and finally included for each survey year. GDP: gross domestic product.



XSL•FO
RenderX

Results

Participant Flow

A total of 112,917 valid questionnaires were analyzed across 6 years. Figure 1 presents the participant flow diagram, showing the number of households contacted, eligible, and finally included for each survey year. Across the 6 annual surveys (2019 - 2024), the number of valid participants ranged from 17,131 (2021) to 19,257 (2019). Replacement of invalid questionnaires due to quality issues (leading questions, duration<10 min, and photo inconsistency) occurred in less than 2% of sampled households per year; no further exclusions were applied.

Recruitment

Recruitment and data collection occurred annually between August 1 and October 31 of each year from 2019 to 2024, as described in the Setting subsection.

Participant Characteristics

Table 1 presents the detailed sociodemographic characteristics of participants across the six annual surveys (2019 - 2024). Significant differences were observed across years in sex, age, education, and occupation distributions (all $P<.001$), highlighting the need for weighting. Weighting was applied to all subsequent analyses.

XSL•FO
RenderX

Table . Sociodemographic characteristics of participants in six annual cross-sectional surveys (2019 - 2024), Zhejiang Province, China (N=110,917).

Content and group	2019 Group (n=19,257), n (%)	2020 Group (n=19,180), n (%)	2021 Group (n=17,131), n (%)	2022 Group (n=18,857), n (%)	2023 Group (n=17,545), n (%)	2024 Group (n=18,947), n (%)	H[a] (*df*)	*P* value
Sex							23.45 (6)	<.001
Male	9175 (47.65)	9298 (48.48)	8020 (46.82)	8762 (46.47)	8237 (46.95)	8807 (46.48)		
Female	10,082 (52.35)	9882 (51.52)	9111 (53.18)	10,095 (53.53)	9308 (53.05)	10,140 (53.52)		
Ethnicity							9.32 (6)	.10
Han	19,050 (98.93)	18,917 (98.63)	16,909 (98.70)	18,634 (98.82)	17,342 (98.84)	18,702 (98.71)		
Minority	207 (1.07)	263 (1.37)	222 (1.30)	223 (1.18)	203 (1.16)	245 (1.29)		
Marital status							256.14 (6)	<.001
Unmarried, divorced, or widowed	3214 (16.69)	3062 (15.96)	2652 (15.48)	3162 (16.77)	3374 (19.23)	3898 (20.57)		
Married	16,043 (83.31)	16,118 (84.04)	14,479 (84.52)	15,695 (83.23)	14,171 (80.77)	15,049 (79.43)		
Age group (years)							78.05 (20)	<.001
15 - 29	1370 (7.11)	1534 (8.00)	1278 (7.46)	1529 (8.11)	1568 (8.94)	1624 (8.57)		
30 - 39	2313 (12.01)	2486 (12.96)	2470 (14.42)	2626 (13.93)	2354 (13.42)	2540 (13.41)		
40 - 49	2749 (14.28)	3890 (20.28)	3502 (20.44)	3218 (17.07)	3361 (19.16)	3301 (17.42)		
50 - 59	6517 (33.84)	6001 (31.29)	5508 (32.15)	6268(33.24)	5245 (29.89)	5546 (29.27)		
60 - 69	6308(32.76)	5269 (27.47)	4373 (25.53)	5216(27.66)	5017 (28.60)	5936 (31.33)		
Education							751.61 (20)	<.001
Primary school or lower	7020 (36.45)	6643 (34.64)	5394 (31.49)	5887 (31.22)	4863 (27.72)	5197 (27.43)		
Middle school	6196 (32.18)	6580 (34.31)	5735 (33.48)	6254 (33.17)	5974 (34.05)	6313 (33.32)		
High school	3130 (16.25)	3000 (15.64)	2906 (16.96)	3134 (16.62)	3057 (17.42)	3399 (17.94)		
Technical school or college	2848 (14.79)	2888 (15.06)	3015 (17.60)	3480 (18.45)	3539 (20.17)	3937 (20.78)		
Undergraduate or higher	63 (0.33)	69 (0.36)	81 (0.47)	102 (0.54)	112 (0.64)	101 (0.53)		
Occupation							931.19 (20)	<.001
Farmers	8738 (45.38)	8979 (46.81)	6667 (38.92)	7600 (40.30)	6466 (36.85)	5945 (31.38)		
Workers	2581 (13.40)	2451 (12.78)	2409 (14.06)	2681 (14.22)	2840 (16.19)	3326 (17.55)		
Agency or Institutional personnel[b]	4565 (23.71)	4786 (24.95)	4855 (28.34)	5175 (27.44)	5179 (29.52)	5873 (31.00)		
Students	671 (3.48)	530 (2.76)	251 (1.47)	442 (2.34)	586 (3.34)	514 (2.71)		
Other[c]	2702 (14.03)	2434 (12.69)	2949 (17.21)	2959 (15.69)	2474 (14.10)	3289 (17.36)		
Total	19,257(100)	19,180(100)	17,131(100)	18,857(100)	17,545(100)	18,947(100)	—[d]	—

[a]The test statistic H refers to the Kruskal-Wallis H statistic.

[b]"Agency or Institutional personnel" refers to people working in state organs, state-owned enterprises, institutions, and other public roles.

[c]The "Other" category includes unemployed people and those with occupations other than those already listed.

[d]Not applicable.

Statistics and Data Analysis

Item-Specific Correct Response Rates

Table 2 presents correct response rates for each IDSHL item. Most items showed significant increases over time (*P*<.001). The most dramatic rise was for respiratory etiquette (cough or sneeze), from 28.44% (n=5476) in 2019 to 39.67% (n=7516) in 2024, with a sharp jump in 2020 (n=7438, 38.78%). Hepatitis B transmission knowledge followed a U-shaped pattern: 63.28% (n=12,186; 2019), 60.71% (n=11,448; 2022), and 63.84% (n=12,096; 2024). Other items, such as fever management and

XSL•FO
RenderX

vaccination contraindications, also improved steadily. For example, correct responses for vaccination contraindications increased from 80.69% (n=15,538) in 2019 to 85.35% (n=16,171) in 2024.

XSL•FO
RenderX

Table . Correct response rates for individual IDSHL[a] items across six survey years, 2019 - 2024.

Types and question	2019 Year N(%)	2020 Year N(%)	2021 Year N(%)	2022 Year N(%)	2023 Year N(%)	2024 Year N(%)	*t* test (*df*)	*P* value
True or false question								<.001
The best way to prevent flu is to take antibiotics (anti-inflammatories).	10,760 (55.88)	11,570 (60.32)	10,226 (59.69)	10,912 (57.87)	10,100 (57.57)	10,738 (56.67)	113.45 (5)	
Single-choice questions								
In which of the following ways can hepatitis B be transmitted to others?	12,186 (63.28)	11,923 (62.16)	10,684 (62.37)	11,448 (60.71)	11,128 (63.43)	12,096 (63.84)	52.11 (5)	
For the treatment of tuberculosis patients, which of the following statements is correct?	12,051 (62.58)	12,470 (65.02)	10,918 (63.73)	12,164 (64.51)	11,762 (67.04)	12,590 (66.45)	112.26 (5)	
In which of the following situations should vaccination of children be suspended?	15,538 (80.69)	15,803 (82.39)	14,399 (84.05)	15,960 (84.64)	14,946 (85.19)	16,171 (85.35)	231.72 (5)	
If you have a fever, which of the following is correct?	15,097 (78.40)	15,737 (82.05)	14,333 (83.67)	15,755 (83.55)	14,209 (80.99)	15,172 (80.08)	259.78 (5)	
If a virulent infectious disease occurs in a certain place, which of the following practices is correct?	15,812 (82.11)	16,510 (86.08)	14,936 (87.19)	16,464 (87.31)	15,286 (87.12)	16,458 (86.86)	320.03 (5)	
Open windows frequently for ventilation during flu season. Regarding window ventilation, which of the following statements is correct?	14,188 (73.68)	14,138 (73.71)	12,673 (73.98)	14,017 (74.33)	13,182 (75.13)	14,603 (77.07)	84.58 (5)	
What is the correct way to read body temperature with a glass thermometer?	10,060 (52.24)	10,564 (55.08)	9505 (55.48)	10,618 (56.31)	10,383 (59.18)	11,564 (61.03)	375.09 (5)	

XSL•FO
RenderX

Types and question	2019 Year N(%)	2020 Year N(%)	2021 Year N(%)	2022 Year N(%)	2023 Year N(%)	2024 Year N(%)	*t* test (*df*)	*P* value
If you are bitten by a dog but not seriously, what is the right thing to do?	18,669 (96.95)	18,555 (96.74)	16,646 (97.17)	18,315 (97.13)	17,108 (97.51)	18,506 (97.67)	41.31 (5)	
Multiple-choice questions								
What should parents do when their children have symptoms such as fever and a rash?	13,171 (68.40)	13,389 (69.81)	12,179 (71.09)	13,629 (72.28)	12,919 (73.63)	14,418 (76.10)	356.29 (5)	
When sick and dead livestock are found, which of the following practices is correct?	14,897 (77.36)	14,961 (78.00)	13,417 (78.32)	15,128 (80.22)	14,244 (81.19)	15,863 (83.72)	336.76 (5)	
When coughing or sneezing, which of the following is correct?	5476 (28.44)	7438 (38.78)	6830 (39.87)	7306 (38.74)	6820 (38.87)	7516 (39.67)	793.88 (5)	

[a]IDSHL: infectious disease-specific health literacy.

Overall and Domain-Specific IDSHL Scores

Weighted mean total IDSHL score increased from 10.46 (SD 3.09) in 2019 to 11.40 (SD 2.88) in 2024 (*P*<.001). All subscales (knowledge, behavior, and skills) showed significant upward trends (all *P*<.001), with behavior exhibiting the largest increase, particularly in 2020 (Table 3).

Table . Weighted mean (SD) scores for IDSHL[a] overall and subscales, 2019 - 2024.

Groups	Knowledge, mean (SD)	Behavior, mean (SD)	Skills, mean (SD)	Overall, mean (SD)
2019	4.41 (1.54)	4.62 (1.67)	1.44 (0.65)	10.46 (3.09)
2020	4.49 (1.52)	4.92 (1.69)	1.49 (0.62)	10.90 (3.09)
2021	4.53 (1.49)	5.02 (1.67)	1.51 (0.61)	11.07 (3.07)
2022	4.56 (1.47)	5.04 (1.63)	1.52 (0.61)	11.12 (2.99)
2023	4.64 (1.45)	5.05 (1.62)	1.55 (0.60)	11.24 (2.95)
2024	4.68 (1.40)	5.15 (1.60)	1.57 (0.59)	11.40 (2.88)
F test (*df*)	47.566 (5)	157.740 (5)	99.397 (5)	141.230 (5)
P value	<.001	<.001	<.001	<.001

[a]IDSHL: infectious disease-specific health literacy.

Joinpoint Regression of Adequate IDSHL Proportion

Joinpoint regression describes population-level annual prevalence trends, whereas logistic regression estimates individual-level adjusted odds for each year relative to 2021. Joinpoint regression identified a significant inflection point in 2021 (Table 4). From 2019 to 2021, the proportion of residents with adequate IDSHL increased sharply at an APC of 9.42% (95% CI 6.10 - 13.45; *P*<.001). From 2021 to 2024, the increase slowed to an APC of 2.28% (95% CI –0.38 to 3.73; *P*=.07). Over the entire period, AAPC was 5.08% (95% CI 3.98 - 6.44; *P*<.001).

XSL•FO
RenderX

Table . Joinpoint regression analysis of trends in weighted proportion of residents with adequate IDSHL[a], Zhejiang, 2019 - 2024.

Model and time period (year)	APC[b] or AAPC[c] (%)	95% CI	*P* value
Segment trends, APC			
2019 - 2021	9.42	6.10 to 13.45	<.001
2021 - 2024	2.28	-0.38 to 3.73	.07
Overall trend, AAPC			
2019 - 2024	5.08	3.98 to 6.44	<.001

[a]IDSHL: infectious disease-specific health literacy.

[b]APC: annual percent change.

[c]AAPC: average annual percent change.

Multivariate Logistic Regression

Table 5 shows adjusted ORs for adequate IDSHL. Higher education was the strongest predictor, while older age (50+ years) and minority ethnicity were associated with lower odds. Compared to 2021 (reference), the odds were significantly lower in 2019 (OR 0.74, 95% CI 0.71 - 0.78) and significantly higher in 2024 (OR 1.10, 95% CI 1.05 - 1.15). No significant differences were found for 2020, 2022, or 2023.

XSL•FO
RenderX

Table . Multivariate logistic regression of factors associated with adequate IDSHL[a].

Variables and categories	B	SE	Wald	*P* value	OR[b]	95% CI
Sex						
Male (reference)	0.05	0.01	15.00	<.001	1.06	1.03 to 1.08
Age group (years)[c]						
30 - 39	0.02	0.04	0.20	.66	1.02	0.95 to 1.09
40 - 49	0.04	0.04	1.00	.32	1.04	0.97 to 1.11
50 - 59	−0.34	0.04	92.70	<.001	0.72	0.67 to 0.77
60 - 69	−0.68	0.04	350.98	<.001	0.51	0.47 to 0.55
Ethnicity						
Minority (reference)	−0.20	0.06	10.41	.001	0.82	0.73 to 0.93
Education[d]						
Middle school	0.68	0.02	1371.63	<.001	1.98	1.91 to 2.05
High school	1.26	0.02	2964.03	<.001	3.52	3.36 to 3.68
Technical school or college	1.98	0.03	4797.55	<.001	7.23	6.84 to 7.65
Undergraduate or higher	2.49	0.13	394.10	<.001	12.10	9.46 to 15.47
Marital status						
Unmarried, divorced, or widowed (reference)	0.22	0.02	111.50	<.001	1.25	1.20 to 1.30
Occupation[e]						
Workers	0.05	0.02	5.40	.020	1.05	1.01 to 1.09
Agency or Institutional personnel	0.22	0.02	111.95	<.001	1.25	1.20 to 1.30
Students	0.36	0.05	48.17	<.001	1.43	1.29 to 1.58
Other	0.16	0.02	58.08	<.001	1.17	1.12 to 1.22
Year[f]						
2019	−0.30	0.02	158.01	<.001	0.74	0.71 to 0.78
2020	0.002	0.02	0.007	.94	1.00	0.96 to 1.05
2022	0.02	0.02	0.66	.42	1.02	0.97 to 1.07
2023	0.04	0.02	2.13	.14	1.04	0.99 to 1.09
2024	0.09	0.02	15.58	<.001	1.10	1.05 to 1.15
Constant	−0.98	0.08	164.74	<.001	0.38	—[g]

[a]IDSHL: infectious disease–specific health literacy.
[b]OR: odds ratio.
[c]Reference range for age: 15-29 years.
[d]Primary school or lower (reference).
[e]Farmers (reference).
[f]2021 (reference).
[g]Not applicable.

Discussion

Support of Original Hypotheses

This 6-year study, leveraging repeated cross-sectional data from before, during, and after the COVID-19 pandemic, demonstrates that the pandemic period was temporally associated with a significant and sustained population-level elevation in IDSHL among residents of Zhejiang Province, China. We observed a sustained population-level elevation in IDSHL, with the lowest likelihood of adequate literacy occurring in the prepandemic year (2019) and the highest in the postpandemic year (2024). The most rapid gains were temporally associated with the acute pandemic phase (2019 - 2021), after which the increase

XSL•FO
RenderX

decelerated. The temporal pattern showed a sharper increase during the acute pandemic phase (2019 - 2021) followed by a slower increase thereafter. Improvements were domain- and disease-specific: the largest gains occurred for pandemic-salient behaviors (eg, respiratory etiquette), while nonrespiratory diseases such as hepatitis B exhibited more complex, fluctuating patterns. Persistent disparities by age, education, and ethnicity indicated that the benefits associated with the pandemic were not uniformly distributed. Crucially, our observational design cannot establish causality; all interpretations refer to population-level temporal associations.

Similarity of Results

Our findings align with previous research showing reduced incidence of other infectious diseases during the pandemic [16] and improvements in respiratory-related health literacy [14,21]. The sharp increase in adequate IDSHL during 2019 - 2021 coincides with the period of most intensive public health messaging, stringent nonpharmaceutical interventions, and peak societal attention to infectious disease prevention [12]. This pattern is consistent with evidence that large-scale health communication campaigns can be temporally associated with rapid improvements in targeted knowledge and behaviors [22]. The subsequent deceleration after 2021, as the pandemic transitioned to endemic management, suggests that crisis-driven momentum is not self-sustaining—a temporal pattern not previously documented in the health literacy literature. The U-shaped pattern for hepatitis B knowledge is suggestive of a potential "crowding-out" effect described in prior literature [17,23]. However, unlike earlier work, our six-wave design revealed that not all non–COVID-19 content stagnated; items with stable, long-standing messaging (eg, vaccination contraindications) improved steadily [24].

Interpretation

The dramatic increase in correct responses for respiratory etiquette, a direct target of COVID-19 messaging, illustrates the focused impact of crisis-driven education [25]. In contrast, the U-shaped pattern for hepatitis B knowledge implies that intense focus on a novel pathogen may temporarily divert public attention from endemic threats, though knowledge may recover as the crisis abates [26]. Notably, items unrelated to respiratory illnesses but with consistent messaging (eg, vaccination contraindications) improved steadily, indicating that generalizable knowledge can still accrue during a pandemic. These nuanced patterns highlight a key limitation of crisis-driven education: it powerfully enhances literacy in highlighted areas but does not automatically generalize across all infectious diseases.

Despite overall improvements, preexisting disparities by age, education, and ethnicity persisted. This finding echoes evidence that health literacy initiatives often benefit more advantaged groups to a greater extent [27,28]. The pandemic's heavy reliance on digital platforms (eg, WeChat, news apps) for information dissemination may have exacerbated these gaps, as older adults and less-educated groups face documented barriers to digital access and literacy [29]. Although this study did not directly measure digital health literacy, prior literature has established its mediating role [30]. Therefore, postpandemic strategies should consider addressing digital divides through age-friendly interfaces, community-based outreach, and tailored messaging. Our data do not directly examine causal mechanisms, but personal experience of COVID-19 infection, including long COVID, could have contributed to knowledge and behavior change [31], a hypothesis warranting future investigation

Generalizability (External Validity)

The study was conducted in a single developed Chinese province (Zhejiang) with a population of over 64 million, which limits direct generalizability to other regions or countries with different pandemic responses, health systems, or cultural contexts. However, the sampling strategy (multistage stratified random sampling, weighting to census) supports generalizability within Zhejiang Province. The findings may be most applicable to other economically developed regions in China or similar settings with intensive digital information environments. Ecological validity is supported by the real-world pandemic context and the use of routine surveillance methods.

Limitations

Several limitations should be considered. First, the repeated cross-sectional design captures population trends but cannot follow individuals over time or establish causality. Unmeasured confounding factors (eg, other policy changes and media coverage) may have influenced IDSHL. Second, the study was conducted in a single developed Chinese province, limiting generalizability to other regions or countries with different pandemic responses and health systems. Third, the IDSHL scale primarily measures functional literacy and does not assess critical or interactive dimensions, nor digital health literacy directly. Fourth, self-reported data may be subject to social desirability bias, particularly during the pandemic. Fifth, while the Cronbach α values for the behavior (0.65) and skills (0.63) subscales are within the acceptable range for population surveillance given the scale's multidimensionality [32], they are modest; observed changes in these domains should therefore be interpreted with some caution. Finally, the weighting procedure reduces but does not eliminate potential bias from sample composition changes.

Implications

The findings carry several practical implications. First, public health authorities should leverage the postpandemic elevated baseline to broaden comprehensive infectious disease education, ensuring that respiratory-focused gains do not come at the expense of other threats (eg, hepatitis B and rabies). Second, targeted interventions are needed to reach older adults, less-educated populations, and ethnic minorities using hybrid (digital + traditional) channels. For example, in Zhejiang, community health workers could deliver face-to-face sessions on nonrespiratory diseases in rural senior centers, complemented by WeChat mini-programs with large-font, audio-assisted content, and monthly live question-and-answer sessions via local health official accounts. Third, routine surveillance of IDSHL should be sustained to detect declines in nonpandemic disease knowledge. By learning from the patterns observed during this global health crisis, short-term crisis responses can

XSL•FO
RenderX

be translated into long-term, equitable public health preparedness [33].

Conclusions

This 6-year study offers the first population-level evidence spanning the full COVID-19 pandemic cycle that a major public health crisis can be temporally associated with substantial and sustained gains in IDSHL. However, the observed improvements were not self-sustaining: they accelerated during the acute crisis phase and decelerated thereafter, and they did not automatically generalize to nonrespiratory diseases. Moreover, preexisting disparities by age, education, and ethnicity persisted. These findings underscore a core tension: crises can rapidly elevate targeted literacy, but maintaining and broadening those gains requires deliberate, ongoing investment. Policymakers should therefore view crisis-driven literacy improvements as a foundation for, not a replacement of, routine health education and equity-focused communication strategies.

Acknowledgments

Generative AI tools (specifically DeepSeek 3.2) were used under full human supervision during the research and writing process. According to the GAIDeT taxonomy (2025), delegated tasks included proofreading and editing, formulation of conclusions, and translation. The authors take full responsibility for the content of this manuscript. AI tools are not listed as authors and bear no responsibility for the final outcomes. Submitted under collective responsibility.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by the Zhejiang Provincial Science and Technology Program for Disease Prevention and Control [grant number 2025JK162].

Data Availability

The data from this study have been anonymized and are available for research purposes upon request from the corresponding author. However, owing to the restrictions set by the ethics committee of the Zhejiang Provincial Center for Disease Control and Prevention, broader public data sharing has been limited. The consent obtained from the participants did not cover the unlimited public sharing of the data.

Authors' Contributions

ZJ and XZ conceived the study; XZ and YX participated in the design; LW, XH, ZY, and DY collected the data; ZY and DY analyzed and interpreted the data; and ZJ, XZ, and YX helped draft, read, and approve the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questionnaire on infectious disease–specific health literacy.
[DOC File, 63 KB - publichealth_v12i1e89958_app1.doc]

References

1. Sørensen K, Van den Broucke S, Fullam J, et al. Health literacy and public health: a systematic review and integration of definitions and models. BMC Public Health 2012 Jan 25;12:80. [doi: 10.1186/1471-2458-12-80] [Medline: 22276600]
2. Tian X, Di Z, Cheng Y, et al. Study on the development of an infectious disease-specific health literacy scale in the Chinese population. BMJ Open 2016 Aug 5;6(8):e012039. [doi: 10.1136/bmjopen-2016-012039] [Medline: 27496240]
3. Seng JJB, Yeam CT, Huang CW, Tan NC, Low LL. Pandemic-related health literacy: a systematic review of literature in COVID-19, SARS and MERS pandemics. Singapore Med J 2025 May 1;66(5):244-255. [doi: 10.4103/singaporemedj.SMJ-2021-026] [Medline: 37459004]
4. Cai YX, Shen Y, Yao H, et al. Overview of and insights from the COVID-19 surveillance systems in major countries and regions worldwide. Int J Virol 2024;31(1):79-83. [doi: 10.3760/cma.j.issn.1673-4092.2024.01.017]
5. Advice for the public: coronavirus disease (COVID-19). World Health Organization. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public [accessed 2023-03-18]
6. Talic S, Shah S, Wild H, et al. Effectiveness of public health measures in reducing the incidence of covid-19, SARS-CoV-2 transmission, and covid-19 mortality: systematic review and meta-analysis. BMJ 2021 Nov 17;375:e068302. [doi: 10.1136/bmj-2021-068302] [Medline: 34789505]
7. Vasileia E, Koulierakis G, Fouskas T, Liarigkovinou A. Health literacy and acceptance of COVID-19 preventive measures and vaccination in the European Union: a scoping review. Health Lit Res Pract 2025 Jan;9(1):e46-e55. [doi: 10.3928/24748307-20250219-01] [Medline: 40064011]

XSL•FO
RenderX

8. Liu J, Liu M, Liang W. The dynamic COVID-zero strategy in China. China CDC Wkly 2022 Jan 28;4(4):74-75. [doi: 10.46234/ccdcw2022.015] [Medline: 35186372]
9. Zhu N, Zhang D, Wang W, et al. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med 2020 Feb 20;382(8):727-733. [doi: 10.1056/NEJMoa2001017] [Medline: 31978945]
10. Pan A, Liu L, Wang C, et al. Association of public health interventions with the epidemiology of the COVID-19 outbreak in Wuhan, China. JAMA 2020 May 19;323(19):1915-1923. [doi: 10.1001/jama.2020.6130] [Medline: 32275295]
11. Liang WN, Liu M, Liu J, Wang YD, Wu J, Liu X. The dynamic COVID-zero strategy on prevention and control of COVID-19 in China. Zhonghua Yi Xue Za Zhi 2022 Jan 25;102(4):239-242. [doi: 10.3760/cma.j.cn112137-20211205-02710] [Medline: 35073672]
12. Shi T, Zhang X, Meng L, et al. Immediate and long-term changes in infectious diseases in China at the “First-level-response”, “Normalized-control” and “Dynamic-COVID-zero” stages from 2020 to 2022: a multistage interrupted-time-series-analysis. BMC Public Health 2023 Jul 18;23(1):1381. [doi: 10.1186/s12889-023-16318-y] [Medline: 37464368]
13. Erzhen C. Expert consensus on the treatment of severe/critical cases of adult coronavirus disease 2019 in Shanghai. Chin J Infect Dis 2023;41(12):754-767. [doi: 10.3760/cma.j.cn311365-20230605-00168]
14. Yuan T, Li XD, Zhang M, Tao XB, Xu SJ, Liu H. Impact of the eHealth literacy, knowledge and attitudes on COVID-19 prevention behavior among residents in the second year of the COVID-19 pandemic: a cross-sectional study in Anhui Province, China. Front Public Health 2022;10:1015803. [doi: 10.3389/fpubh.2022.1015803] [Medline: 36408007]
15. McCaffery KJ, Dodd RH, Cvejic E, et al. Health literacy and disparities in COVID-19-related knowledge, attitudes, beliefs and behaviours in Australia. Public Health Res Pract 2020 Dec 9;30(4):30342012. [doi: 10.17061/phrp30342012] [Medline: 33294907]
16. Zhao Y, Wu Q, Xu S, et al. Understanding the risk factors for infectious diseases, their prevention, and control, among residents of Zhejiang Province. Chin J Prev Med 2016;50(9):806-810 [FREE Full text] [doi: 10.3760/cma.j.issn.0253-9624.2016.09.011]
17. Zhao Y, Xu Y, Yao D, et al. Changes in infectious disease-specific health literacy in the post-COVID-19 pandemic period: two-round cross-sectional survey study. JMIR Public Health Surveill 2024 Aug 30;10:e52666. [doi: 10.2196/52666] [Medline: 39213137]
18. Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health concern. Lancet 2020 Feb 15;395(10223):470-473. [doi: 10.1016/S0140-6736(20)30185-9] [Medline: 31986257]
19. Hu X, Yao D, Xu Y, Chen H, Zhang X. Health literacy and its influencing factors among residents aged 15-69 years with different chronic disease in Zhejiang Province. Chin J Health Manage 2024;18(2):107-113. [doi: 10.3760/cma.j.cn115624-20231022-00211]
20. Communiqué of the Seventh National Population Census (no. 3). National Bureau of Statistics of China. 2021. URL: http://www.stats.gov.cn/english/PressRelease/202105/t20210510_1817188.html [accessed 2026-07-10]
21. Hu CY, Tang YW, Su QM, et al. Public health measures during the COVID-19 pandemic reduce the spread of other respiratory infectious diseases. Front Public Health 2021;9:771638. [doi: 10.3389/fpubh.2021.771638] [Medline: 34858936]
22. Wakefield MA, Loken B, Hornik RC. Use of mass media campaigns to change health behaviour. Lancet 2010 Oct 9;376(9748):1261-1271. [doi: 10.1016/S0140-6736(10)60809-4] [Medline: 20933263]
23. Chen B, Wang M, Huang X, et al. Changes in incidence of notifiable infectious diseases in China under the prevention and control measures of COVID-19. Front Public Health 2021;9:728768. [doi: 10.3389/fpubh.2021.728768] [Medline: 34722440]
24. Vojtek I, van Wouw M, Thomson A. Impact of COVID-19 on vaccine confidence and uptake: a systematic literature review. Hum Vaccin Immunother 2024 Dec 31;20(1):2384180. [doi: 10.1080/21645515.2024.2384180] [Medline: 39106971]
25. Kite J, Chan L, MacKay K, et al. A model of social media effects in public health communication campaigns: systematic review. J Med Internet Res 2023 Jul 14;25:e46345. [doi: 10.2196/46345] [Medline: 37450325]
26. Patel MP, Kute VB, Agarwal SK, COVID-19 Working Group of Indian Society of Nephrology. “Infodemic” COVID 19: more pandemic than the virus. Indian J Nephrol 2020;30(3):188-191. [doi: 10.4103/ijn.IJN_216_20] [Medline: 33013069]
27. Svendsen MT, Bak CK, Sørensen K, et al. Associations of health literacy with socioeconomic position, health risk behavior, and health status: a large national population-based survey among Danish adults. BMC Public Health 2020 Apr 28;20(1):565. [doi: 10.1186/s12889-020-08498-8] [Medline: 32345275]
28. de Araújo GR, de Castro P, Ávila IR, Bezerra JMT, Barbosa DS. Effects of public health emergencies of international concern on disease control: a systematic review. Rev Panam Salud Publica 2023;47:e74. [doi: 10.26633/RPSP.2023.74] [Medline: 37089787]
29. Nguyen MH, Hargittai E, Marler W. Digital inequality in communication during a time of physical distancing: the case of COVID-19. Comput Human Behav 2021 Jul;120:106717. [doi: 10.1016/j.chb.2021.106717] [Medline: 34751201]
30. Xie B, Charness N, Fingerman K, Kaye J, Kim MT, Khurshid A. When going digital becomes a necessity: ensuring older adults’ needs for information, services, and social inclusion during COVID-19. J Aging Soc Policy 2020;32(4-5):460-470. [doi: 10.1080/08959420.2020.1771237] [Medline: 32507061]
31. Carrouel F, Lvovschi VE, du Sartz de Vigneulles B, et al. Prevalence, risk factors, disease-related knowledge, and vaccination attitudes and behaviors for long COVID among French civil servants: cross-sectional survey. JMIR Public Health Surveill 2025 Dec 5;11:e83323. [doi: 10.2196/83323] [Medline: 41348962]

XSL•FO
RenderX

32. Taber KS. The use of Cronbach's alpha when developing and reporting research instruments in science education. Res Sci Educ 2018 Dec;48(6):1273-1296. [doi: 10.1007/s11165-016-9602-2]
33. Okan O, Messer M, Levin-Zamir D, Paakkari L, Sørensen K. Health literacy as a social vaccine in the COVID-19 pandemic. Health Promot Int 2023 Aug 1;38(4):daab197. [doi: 10.1093/heapro/daab197] [Medline: 35022721]

Abbreviations

AAPC: average annual percent change
APC: annual percent changes
IDSHL: Infectious disease-specific health literacy
LSD: least significant difference
MCAR: Missing Completely At Random
NPI: nonpharmaceutical intervention
OR: odds ratio

Edited by S Brini; submitted 20.Dec.2025; peer-reviewed by F Carrouel, Y Fukuda; revised version received 12.Jun.2026; accepted 26.Jun.2026; published 05.Aug.2026.

Please cite as:
Yusui Z, Yao D, Xu Y, Hu X, Wu Q, Wang L, Zhang X, Jiang Z
Longitudinal Population-Level Shifts in Infectious Disease–Specific Health Literacy During and After the COVID-19 Pandemic in Zhejiang, China: Six-Round Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e89958
URL: https://publichealth.jmir.org/2026/1/e89958
doi:10.2196/89958

© Zhao Yusui, Dingming Yao, Yue Xu, Xiujing Hu, Qingqing Wu, Lei Wang, Xuehai Zhang, Zhen Jiang. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 5.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Socioeconomic and Regional Disparities in Long-Term Mortality Among Intensive Care Unit Survivors in South Korea: Nationwide Cohort Study

Nayeon Choi[1,2], MS; Jaehoon Oh[1,2], MD, PhD; Kyung Hun Yoo[1,2], MD, PhD; Juncheol Lee[1,2], MD, PhD; Sang Hwan Lee[3], MD

[1]Department of Emergency Medicine, College of Medicine, Hanyang University, 222 Wangsimni-ro, Seongdong-Gu, Seoul, Republic of Korea

[2]Hanyang Emergency Medicine AI & Data Science Lab (HEADS), Department of Emergency Medicine, College of Medicine, Hanyang University, Seongdong-Gu, Seoul, Republic of Korea

[3]Department of Emergency Medicine, Hanyang University Seoul Hospital, Seongdong-Gu, Seoul, Republic of Korea

Corresponding Author:
Jaehoon Oh, MD, PhD
Department of Emergency Medicine, College of Medicine, Hanyang University, 222 Wangsimni-ro, Seongdong-Gu, Seoul, Republic of Korea

Abstract

Background: Long-term outcomes of intensive care unit (ICU) survivors remain a major public health concern. Despite the existence of a universal health care system in South Korea, socioeconomic and regional disparities may affect postcritical illness survival.

Objective: In this nationwide cohort study, we aimed to examine the associations of socioeconomic and regional factors with long-term mortality among ICU survivors.

Methods: We conducted a retrospective population-based cohort study using data from the Korean National Health Insurance Service (NHIS). We included adult patients (≥20 y) admitted to ICUs between 2015 and 2019 who survived to hospital discharge. Patients admitted for injury-related or external causes were excluded. Postdischarge socioeconomic status (SES) indicators included insurance type (National Health Insurance [NHI] vs Medical Aid [MA]), NHIS eligibility status, and income level quartiles. Residential area was analyzed separately as a regional contextual factor. The primary outcome was all-cause mortality for up to 8 years after discharge, which was estimated using the Kaplan-Meier method. Adjusted hazard ratios (aHRs) and 95% CIs were estimated using multivariable time-dependent Cox proportional hazards models stratified by age (20‑39 y, 40‑64 y, and ≥65 y)

Results: Among 871,366 adult ICU survivors, 9.6% (n=83,766) were MA beneficiaries. Lower SES—defined by MA coverage, nonemployee NHIS eligibility categories, and lower income level—was consistently associated with higher long-term mortality. Compared with NHI enrollees, MA recipients had an aHR of 1.63 (95% CI 1.60‑1.66) among middle-aged adults and 1.48 (95% CI 1.38‑1.60) among younger adults. Dependents and self-employed insured household members had 2-fold to 3-fold higher mortality risk than employee-insured beneficiaries. A stepwise income-mortality gradient was evident across all age groups. Residence in other regions was also associated with higher mortality (aHR 1.06‑1.11). Younger individuals were particularly vulnerable to SES-related disparities.

Conclusions: Despite South Korea's universal insurance system, socioeconomic and regional disparities in long-term mortality persist among ICU survivors. Ensuring equitable access to post-ICU care, income security, and targeted social support may help promote sustainable recovery.

(JMIR Public Health Surveill 2026;12:e89127) doi:10.2196/89127

KEYWORDS

socioeconomic status; geographical regional factors; intensive care unit; long-term mortality; health disparities

Introduction

With the growing global population, the demand for intensive care services continues to increase. While advancements in critical care have reduced short-term mortality rates among critically ill patients, long-term outcomes remain a major concern, especially for those discharged from intensive care units (ICUs) [1-3]. ICU survivors frequently experience physical, cognitive, and mental health deficits, resulting in diminished quality of life and increased long-term mortality [4]. To enhance care for ICU survivors and their families,

XSL•FO
RenderX

practitioners and researchers must collaborate in both inpatient and outpatient environments [5,6]. However, socioeconomic status (SES) and age may affect access to post-ICU care, continuity of follow-up care, and long-term recovery [7-9].

Recent evidence has revealed that both SES-related factors, such as insurance type and income level, and regional factors may substantially influence long-term outcomes [7,8,10,11]. In South Korea, individuals receiving Medical Aid (MA), which is indicative of low-income and socially disadvantaged populations, have higher postdischarge mortality rates than do National Health Insurance (NHI) beneficiaries [7,12]. Additionally, regional disparities, reflected in urban morphology, health care access, and postacute care resources, may influence recovery pathways and long-term outcomes [11]. Social determinants such as unemployment, declining marriage rates, and social isolation, which are more common among individuals with low SES, have been associated with negative health outcomes and may impede recovery following treatment in ICUs [9,13]. Although awareness of these associations is increasing, there remains insufficient evidence, particularly in high-income Asian countries with different health care and welfare systems.

In this study, we aimed to investigate the associations of SES indicators (health insurance type, National Health Insurance Service [NHIS] eligibility status, and income level) and residential area, as a regional contextual factor, with long-term mortality among adult ICU survivors in South Korea. Since older adults are particularly susceptible and age markedly influences post-ICU prognosis, we examined whether these associations differed across age groups to enhance our understanding of age-related disparities in postdischarge outcomes.

Methods

Data Sources and Setting

This population-based cohort study used data from the NHIS database of South Korea. The NHIS operates as a government-administered, single-payer system under the National Health Insurance Act, which mandates universal participation. Consequently, the system covers nearly the entire Korean population [14]. The NHIS database contains comprehensive information on both inpatient and outpatient medical services, including demographic characteristics (eg, age, sex, residence, and smoking history), hospital visit details (department visited, visit dates, and patient status), diagnoses coded according to the *International Classification of Diseases, 10th Revision* (*ICD-10*) [15], prescriptions, medical procedures, surgeries, medical costs, and mortality. Socioeconomic variables such as income level and insurance type (NHI or MA) were also included, whereas cosmetic and nonevidence-based treatments were excluded [14]. Given its nationwide scope and data integrity, the NHIS database is a reliable source of data for large-scale population-based research.

NHIS in South Korea

In South Korea, universal health coverage comprises 2 programs: the NHI and MA. The NHI, introduced in 1977 and expanded to full coverage by 1989, was unified into a single-payer model in 2000 and currently insures about 98% of the population. Similar to Medicaid in the United States, the MA program provides public assistance to low-income individuals and includes beneficiaries of the National Basic Livelihood Security System [16]. Under the NHI, insured individuals are classified as either employee-insured or self-employed. Employees, public officials, school staff, and their dependents are classified as employee-insured. In contrast, dependents are not separately recognized among the self-employed; the household head is responsible for paying the entire premium for all family members, including spouses, children, and parents. Premiums for employee-insured individuals are based on income and shared between employees and employers, whereas those for the self-employed are determined by household income, assets, vehicle ownership, age, and sex [17].

Study Design and Population

This was a retrospective longitudinal cohort study designed to evaluate long-term mortality. We extracted data from patients admitted to ICUs between January 2015 and December 2019. ICU admission was determined by the attending physicians, and patients were identified using health insurance claim codes AJ100 and AJ390. Only adults aged 20 years or older were included in the study. ICU survivors were defined as patients discharged from the hospital. Patients admitted for injury-related or external causes (eg, burns, trauma, poisoning, asphyxia, and anaphylaxis) were excluded using *ICD-10* codes beginning with “S” or “T.” For patients with multiple ICU admissions during the study period, only the first admission was analyzed. The ICU admission date was defined as the cohort entry date, and the hospital discharge date of the first admission was defined as the index date. Demographic variables (sex, age, smoking status, comorbidities, interventions, hospitalization type, and residential area) and socioeconomic variables (health insurance type, NHIS eligibility status, and income level) were obtained from the NHIS database. We analyzed the associations of postdischarge socioeconomic indicators and residential area with long-term mortality, stratified by age group (20 - 39 y, 40 - 64 y, and ≥65 y). Multimedia Appendix 1 depicts the data selection process, exclusion criteria, enrollment period, and follow-up duration.

Socioeconomic Indicators and Regional Factors

Individual-level SES is commonly assessed using indicators such as income, educational attainment, occupation, and insurance-related measures [18]. SES was assessed using the following three individual-level indicators: (1) insurance type, (2) NHIS eligibility status, and (3) income level. Residential area was analyzed as an independent regional factor.

Substantial disparities have been reported in these SES indicators between beneficiaries of the MA and NHI programs. For example, the proportions of individuals with low income (95% vs 34%), those with education levels at or below the elementary school level (61% vs 40%), and those not participating in economic activity (82% vs 57%) differed markedly between MA and NHI recipients, respectively [19]. Medicaid eligibility in the United States has been validated as

XSL•FO
RenderX

a reliable proxy for individual-level SES, and insurance type has therefore been widely used as a surrogate measure of SES [12,18-20]. Given the substantial socioeconomic differences between MA and NHI beneficiaries in South Korea, we used MA status as an operational indicator of lower SES.

NHIS eligibility status was used as an SES-related administrative indicator in this study. Because the NHIS database does not directly capture employment status, unemployment, occupation, or labor force participation, this variable was interpreted as an administrative health insurance eligibility category rather than as a direct measure of employment or unemployment. Based on administrative eligibility data, NHIS eligibility status was categorized as employee-insured beneficiaries, dependents of employee-insured beneficiaries, self-employed insured householders, self-employed insured household members, and MA beneficiaries. Although employment and labor market participation are closely related to SES in the prior literature [21], this "NHIS eligibility status" variable should not be interpreted as a direct measure of employment. Employed-insured beneficiaries are generally workplace enrollees in the Korean health insurance system, whereas self-employed insured householders are responsible for paying insurance premiums based on income and assets. Dependents of employee-insured beneficiaries and self-employed insured household members may include individuals with limited income or without independent premium-paying status. MA beneficiaries represent a socioeconomically vulnerable group, and previous data have shown that 82% do not participate in economic activity [19]. However, individuals outside the employee-insured category may still derive income from self-employment, pensions, property, and other sources. Therefore, the NHIS eligibility status variable was used as an SES-related administrative indicator rather than as a direct classification of individuals as employed or unemployed [16,17].

The income level was inferred from health insurance premiums. MA beneficiaries are exempt from paying premiums, whereas NHI subscribers are categorized into quartiles according to their premium levels. MA eligibility requires a household income within 40% of the national median household income [17]. Accordingly, the lowest-income category in this study comprised MA beneficiaries and NHI enrollees within the lowest premium quartile, representing the lowest socioeconomic stratum.

The residential area was analyzed as a regional contextual factor rather than as a proxy for individual-level SES. Regional disparities may affect health outcomes through mechanisms distinct from those associated with socioeconomic disadvantage. Residential regions were classified into the Seoul metropolitan area (Seoul, Incheon, and Gyeonggi Province), noncapital metropolitan cities (excluding the Seoul metropolitan area), and other regions.

Outcomes

The primary outcome was long-term mortality over an 8-year follow-up period (until 2022). Mortality rates were stratified by SES indicators, including insurance type, NHIS eligibility status, income level, and residential area. Long-term mortality according to SES was evaluated by classifying ICU survivors into 3 age groups: 20 to 39 years, 40 to 64 years, and 65 years and older. Each age group was separately evaluated. The follow-up period began on the index date and ended on December 31, 2022, or the date of death, whichever occurred first. Mortality rates were assessed at 1, 2, 4, 6, and 8 years, as well as overall.

Furthermore, we evaluated long-term mortality by examining residential area in conjunction with other socioeconomic factors (insurance, NHIS eligibility status, and income level). For the additional descriptive analysis, the low-income group was defined as comprising MA beneficiaries and NHI enrollees in the lowest premium quartile. The characteristics of the deceased and survivors in the low-income group were compared.

Statistical Analysis

Baseline characteristics of the study population are presented as mean (SD) for continuous variables and as frequencies (%) for categorical variables. Differences among age groups were assessed using ANOVA for continuous variables and the chi-square test for categorical variables. The 8-year cumulative mortality was estimated using the Kaplan-Meier method. Analyses were stratified by age group, individual SES indicators, and residential area. Group differences were assessed using log-rank tests. The absolute mortality risk (AMR) difference across socioeconomic indicators and residential area categories was calculated as the difference in the 8-year cumulative mortality rates (eg, MA vs NHI, quartile 1 vs quartile 4, or dependents vs employee-insured beneficiaries). AMR is expressed in percentage points (%), providing an intuitive measure of absolute inequality in long-term mortality. Because socioeconomic factors and residential area vary over time, time-dependent Cox proportional hazards regression models were constructed to evaluate their independent associations with long-term mortality. Residential area was also included as an independent regional factor.

To avoid multicollinearity, in the time-dependent Cox models that included covariates, each SES indicator (insurance type, NHIS eligibility status, and income level) was included in a separate model along with residential area. Adjusted hazard ratios (aHRs) with 95% CIs were reported after controlling for age, sex, smoking status, the Charlson Comorbidity Index (CCI) score, and intervention status (mechanical ventilation, renal replacement, extracorporeal membrane oxygenation, and vasopressor or inotropic drugs), indicating disease severity.

Because residential areas may reflect contextual disparities that interact with, but are not fully captured by, individual-level SES indicators, we conducted an additional analysis to examine whether the association between residential area and long-term mortality persisted after adjusting for insurance type, NHIS eligibility status, or income level. In addition, to better understand mortality patterns among the most socioeconomically vulnerable groups, we compared the clinical and sociodemographic characteristics of survivors and decedents among MA recipients and individuals in the lowest-income quartile using independent 2-tailed *t* tests or chi-square tests. Because the Korean NHIS database is a nationwide administrative claims database covering virtually the entire Korean population, the present analysis was not based on a

XSL•FO
RenderX

complex survey sample. Therefore, survey-sampling weights were not applied. The exclusions shown in the flow diagram were applied to define the prespecified target population of adult ICU survivors with noninjury or nonexternal-cause admissions and complete essential covariate information, rather than to construct a sampled subset of the source population. Data analyses were performed using SAS Enterprise Guide (version 7.1; SAS Institute Inc) and R (version 4.0.3; R Foundation for Statistical Computing). All statistical tests were 2-sided, and significance was set at $P<.05$.

Ethical Considerations

The study design was approved by the Institutional Review Board of Hanyang University Hospital (HYUH 2023-04-049) and the Health Insurance Review and Assessment Service (NHIS-2024-1-315). Informed consent was waived as data analyses were performed retrospectively using deidentified data generated by an independent technician at the NHIS facility who had no association with this project.

Results

Study Population

Using the NHIS database, we identified 1,566,331 adult patients admitted to ICUs between January 1, 2015, and December 31, 2019. After excluding patients aged younger than 20 years of age; those who died during hospitalization; individuals with injury-related or external causes (*ICD-10* codes beginning with "S" or "T"); and 111,963 patients lacking complete data on age, sex, insurance, or premium percentile, a total of 871,366 patients were included in the final analysis (Figure 1).

Figure 1. Flow diagram of participant selection. ICU: intensive care unit; *ICD: International Classification of Diseases*.

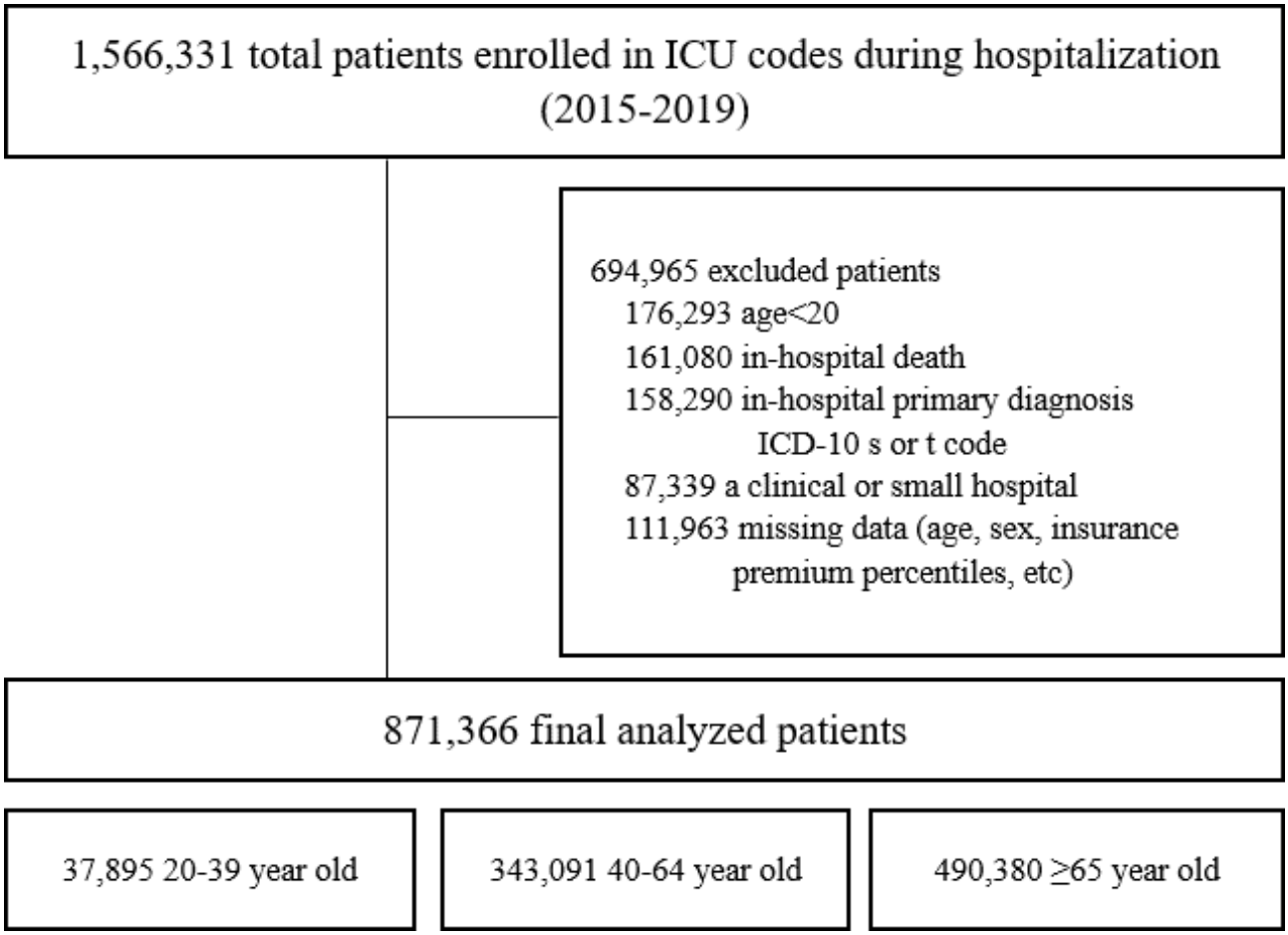


Table 1 summarizes the baseline characteristics of patients. The mean age was 66.0 (SD 14.4) years, and among 871,366 patients, 56.3% (n=490,380) were aged 65 years or older. Male patients accounted for 58.1% (n=506,307) of the cohort, with a more pronounced male sex predominance observed among younger adults. Most patients were covered by the NHI (n=787,600, 90.4%), while 9.6% (n=83,766) received MA. The mean CCI was 3.1 (SD 2.5), increasing significantly with age ($P<.001$). The proportion of MA beneficiaries increased with age (n=2460, 6.5% in the 20-year to 39-year group vs n=49,531, 10.1% in those ≥65 y). Regarding residence, 45.9% (n=399,804) and 36.8% (n=320,536) lived in the Seoul metropolitan area and other regions, respectively. The most frequent comorbidities were hypertension (n=539,389, 61.9%), dyslipidemia (n=537,561, 61.7%), and diabetes (n=375,003, 43.0%). Major ICU interventions included mechanical ventilation (n=167,287, 19.2%), use of vasopressors or inotropes (n=383,293, 43.9%), and renal replacement therapy (n=24,132, 2.8%). Notably, the overall mortality increased markedly with age ($P<.001$). The 1-year, 4-year, and 8-year mortality rates were 15.7% (n=136,379), 30.9% (n=269,382), and 39.6% (n=344,722), respectively. Long-term mortality was significantly higher among older adults ($P<.001$) than among younger adults. The main diagnoses for ICU patients were classified by organ

XSL•FO
RenderX

system. Most diagnoses were circulatory system disorders (I00-I99; 418,886 cases), neoplasms (C00-D48; 144,719 cases), and digestive system diseases (K00-K93; 79,171 cases; Multimedia Appendix 2).

XSL•FO
RenderX

Table . Baseline characteristics of patients.

Characteristics	Total	20 - 39 y	40 - 64 y	≥65 y	*P* value
Number of patients, n (%)	871,366 (100)	37,895 (4.3)	343,091 (39.4)	490,380 (56.3)	__[a]
Age (y), mean (SD)	66.0 (14.4)	32.2 (5.6)	54.9 (6.5)	76.4 (7.3)	<.001
Sex, n (%)					<.001
Male	506,307 (58.1)	22,605 (59.7)	228,938 (66.7)	254,764 (51.9)	
Female	365,059 (41.9)	15,290 (40.3)	114,153 (33.3)	235,616 (48.0)	
Insurance status, n (%)					<.001
NHI[b]	787,600 (90.4)	35,435 (93.5)	311,316 (90.7)	440,849 (89.9)	
MA[c]	83,766 (9.6)	2460 (6.5)	31,775 (9.3)	49,531 (10.1)	
NHIS[d] eligibility status, n (%)					<.001
Employee-insured	158,133 (18.1)	15,858 (41.8)	109,621 (31.9)	32,654 (6.7)	
Dependents	360,738 (41.4)	8710 (22.9)	81,238 (23.7)	270,790 (55.2)	
Self-employed-insured	174,541 (20.0)	4707 (12.4)	83,409 (24.3)	86,425 (17.6)	
Household members	94,188 (10.8)	6160 (16.3)	37,048 (10.8)	50,980 (10.4)	
MA	83,766 (9.6)	2460 (6.5)	31,775 (9.3)	49,531 (10.1)	
Income level, n (%)					<.001
MA	83,766 (9.6)	2460 (6.5)	31,775 (9.3)	49,531 (10.1)	
Quartile 1 (lowest)	134,793 (15.5)	6287 (16.6)	61,059 (17.8)	67,447 (13.8)	
Quartile 2	217,131 (24.9)	13,918 (36.7)	101,252 (29.5)	101,961 (20.8)	
Quartile 3	217,937 (25.0)	10,269 (27.1)	83,674 (24.4)	123,994 (25.3)	
Quartile 4 (highest)	217,739 (24.9)	4961 (13.1)	65,331 (19.0)	147,447 (30.1)	
Residential area, n (%)					<.001
Seoul metropolitan area	399,804 (45.9)	19,840 (52.4)	165,356 (48.2)	214,608 (43.8)	
Noncapital metropolitan cities	151,026 (17.3)	6777 (17.9)	63,637 (18.5)	80,612 (16.4)	
Other regions	320,536 (36.8)	11,278 (29.8)	114,098 (33.3)	195,160 (39.8)	
Smoking status, n (%)					<.001
Nonsmoker	334,314 (38.4)	11,095 (29.3)	119,875 (34.9)	203,344 (41.5)	
Ex-smoker	157,722 (18.1)	4974 (13.1)	75,276 (21.9)	77,472 (15.8)	
Current smoker	101,868 (11.7)	5269 (13.9)	64,673 (18.9)	31,926 (6.5)	
Unknown	277,462 (31.8)	16,557 (43.7)	83,267 (24.3)	177,638 (36.2)	
Comorbidities, n (%)[e]					
CCI[f] score, mean (SD)	3.1 (2.5)	1.4 (1.7)	2.5 (2.3)	3.6 (2.5)	<.001
Hypertension	539,389 (61.9)	7037 (18.6)	162,449 (47.3)	369,903 (75.4)	<.001
Dyslipidemia	537,561 (61.7)	11,877 (31.3)	189,637 (55.3)	336,047 (68.5)	<.001
CAD[g]	42,013 (4.8)	367 (1.0)	12,008 (3.5)	29,638 (6.0)	<.001
Diabetes	375,003 (43.0)	5972 (15.8)	119,906 (34.9)	249,125 (50.8)	<.001
Stroke	117,265 (13.5)	1019 (2.7)	24,652 (7.2)	91,594 (18.7)	<.001

XSL•FO
RenderX

Characteristics	Total	20 - 39 y	40 - 64 y	≥65 y	*P* value
Acute myocardial infarction	36,369 (4.2)	296 (0.8)	10,470 (3.1)	25,603 (5.2)	<.001
Congestive heart failure	118,867 (13.6)	1717 (4.5)	28,357 (8.3)	88,793 (18.1)	<.001
Peripheral vascular disease	169,758 (19.5)	1715 (4.5)	43,781 (12.8)	124,262 (25.3)	<.001
Cerebral vascular accident	225,806 (25.9)	2734 (7.2)	58,707 (17.1)	164,365 (33.5)	<.001
Dementia	92,191 (10.6)	120 (0.3)	5202 (1.5)	86,869 (17.7)	<.001
Pulmonary disease	412,710 (47.4)	12,264 (32.4)	127,131 (37.1)	273,315 (55.7)	<.001
Connective tissue disorder	47,606 (5.5)	1062 (2.8)	14,924 (4.3)	31,620 (6.4)	<.001
Peptic ulcer	337,030 (38.7)	8922 (23.5)	117,574 (34.3)	210,534 (42.9)	<.001
Liver disease	74,420 (8.5)	2117 (5.6)	35,470 (10.3)	36,833 (7.5)	<.001
Diabetes complications	111,407 (12.8)	1465 (3.9)	34,279 (9.9)	75,663 (15.4)	<.001
Paraplegia	18,731 (2.1)	379 (1.0)	5099 (1.5)	13,253 (2.7)	<.001
Renal disease	76,937 (8.8)	1863 (4.9)	22,671 (6.6)	52,403 (10.7)	<.001
Cancer	156,399 (17.9)	3729 (9.8)	59,471 (17.3)	93,199 (19.0)	<.001
Metastatic cancer	24,803 (2.8)	800 (2.1)	11,312 (3.3)	12,691 (2.6)	<.001
Severe liver disease	10,418 (1.2)	331 (0.9)	5695 (1.7)	4392 (0.9)	<.001
HIV	428 (0.0)	52 (0.1)	249 (0.1)	127 (0.0)	<.001
Intervention, n (%)[e]					
Mechanical ventilation	167,287 (19.2)	9690 (25.6)	69,610 (20.2)	87,987 (17.9)	<.001
Renal replacement	24,132 (2.8)	1247 (3.3)	9240 (2.7)	13,645 (2.8)	<.001
Vasopressor or inotropic drugs	383,293 (43.9)	17,452 (46.1)	155,457 (45.3)	210,384 (42.9)	<.001
ECMO[h]	3861 (0.4)	448 (1.2)	2220 (0.6)	1193 (0.2)	<.001
Hospitalization type, n (%)					<.001
Tertiary hospital[i]	370,599 (42.5)	19,463 (51.4)	162,182 (47.3)	188,954 (38.5)	
General hospital	500,767 (57.5)	18,432 (48.6)	180,909 (52.7)	301,426 (61.5)	
Year of ICU[j] admission, n (%)					<.001
2015	188,583 (21.6)	7715 (20.4)	73,418 (21.4)	107,450 (21.9)	
2016	180,033 (20.7)	7651 (20.2)	71,249 (20.8)	101,133 (20.6)	
2017	173,897 (19.9)	7514 (19.8)	68,636 (20.0)	97,747 (19.9)	
2018	163,623 (18.8)	7320 (19.3)	64,601 (18.8)	91,702 (18.7)	
2019	165,230 (18.9)	7695 (20.3)	65,187 (18.9)	92,348 (18.8)	
Long-term death (y), n (%)[k]					
1	136,379 (15.7)	2238 (5.9)	28,351 (8.3)	105,790 (21.6)	<.001
2	190,367 (21.8)	3184 (8.4)	41,640 (12.1)	145,543 (29.7)	<.001
4	269,382 (30.9)	4419 (11.7)	59,839 (17.4)	205,124 (41.8)	<.001
6	321,146 (36.8)	5129 (13.5)	71,693 (20.9)	244,324 (49.8)	<.001

XSL•FO
RenderX

Characteristics	Total	20 - 39 y	40 - 64 y	≥65 y	*P* value
8	344,722 (39.6)	5434 (14.3)	76,963 (22.4)	262,325 (53.5)	<.001

[a]Not applicable.

[b]NHI: National Health Insurance.

[c]MA: Medical Aid.

[d]NHIS: National Health Insurance Service.

[e]Comorbidities and interventions are not mutually exclusive; n and % reflect patients with each condition or intervention and do not sum to N or 100%.

[f]CCI: Charlson Comorbidity Index.

[g]CAD: coronary artery disease.

[h]ECMO: extracorporeal membrane oxygenation.

[i]Ministry-designated tertiary hospitals in Korea.

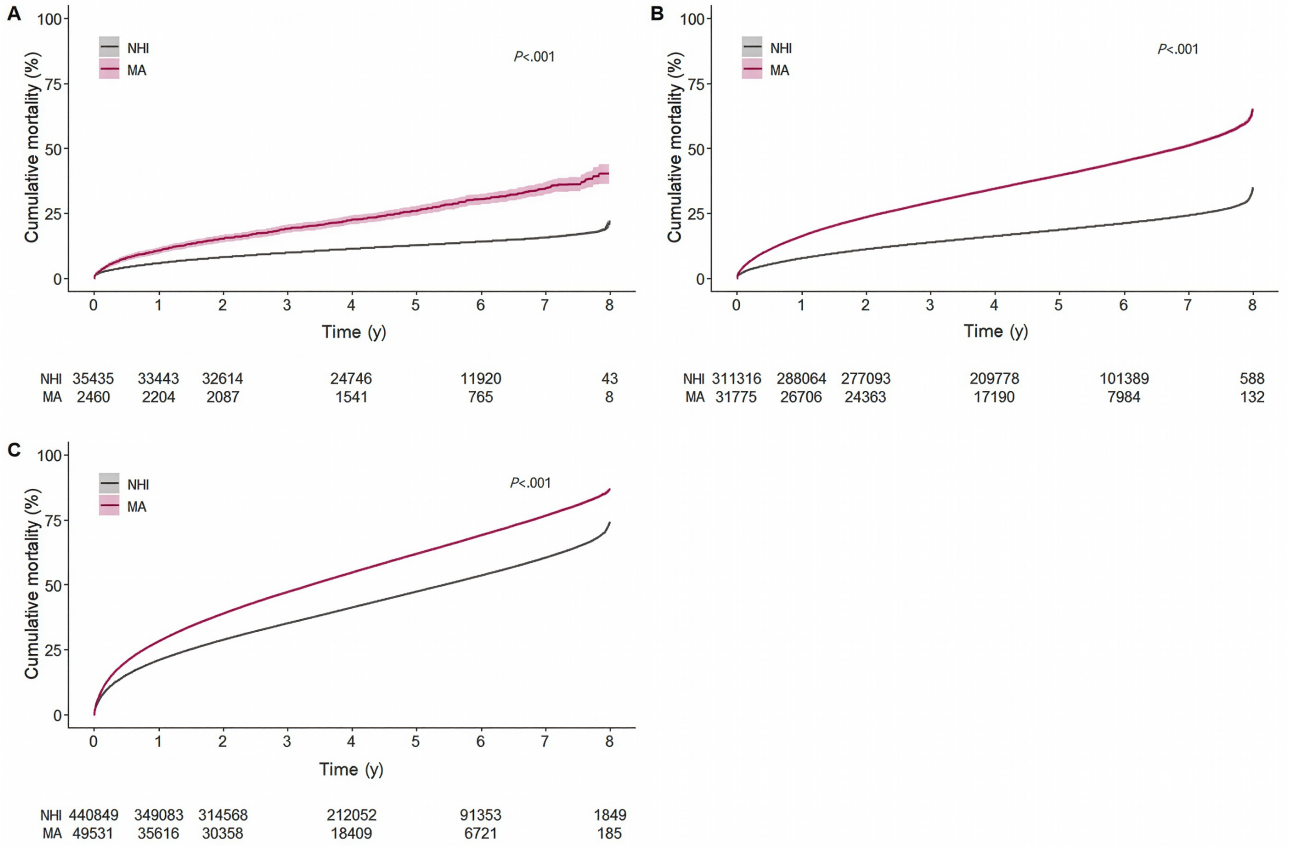
[j]ICU: intensive care unit.

[k]Long-term death values are cumulative across time points; n and % overlap across subcategories and do not sum to N or 100%.

SES and Regional Disparities in Long-Term Mortality

In all age groups, the cumulative mortality curves differed significantly according to SES indicators (Figures 2-5; log-rank test, *P*<.001). Specifically, MA beneficiaries had higher mortality than did NHI beneficiaries (Figure 2); dependents and self-employed household members had poorer survival than did employee-insured individuals (Figure 3); and mortality increased stepwise with decreasing income, with MA beneficiaries exhibiting the worst outcomes, even when compared with those in the lowest-income quartile (Figure 4). Patients residing in other regions also exhibited higher mortality rates than those residing in the Seoul metropolitan area or noncapital metropolitan cities (Figure 5). The 8-year cumulative mortality (%) and AMR differences by SES indicators and age groups are presented in Multimedia Appendix 3. The 8-year AMR differences between MA and NHI were 29.8% in middle-aged adults and 17.9% in young adults. Across all socioeconomic indicators, the absolute mortality gap increased as SES decreased, with the steepest gradients observed for insurance type, NHIS eligibility status, and income level, particularly among middle-aged adults (40 - 64 y old).

Figure 2. Cumulative mortality curves by insurance status, stratified by age group: (A) 20 to 39 years; (B) 40 to 64 years; and (C) 65 years and older. NHI: National Health Insurance; MA: Medical Aid.



XSL•FO
RenderX

Figure 3. Cumulative mortality curves by National Health Insurance Service (NHIS) eligibility status, stratified by age group: (A) 20 to 39 years; (B) 40 to 64 years; and (C) 65 years and older. MA: Medical Aid.

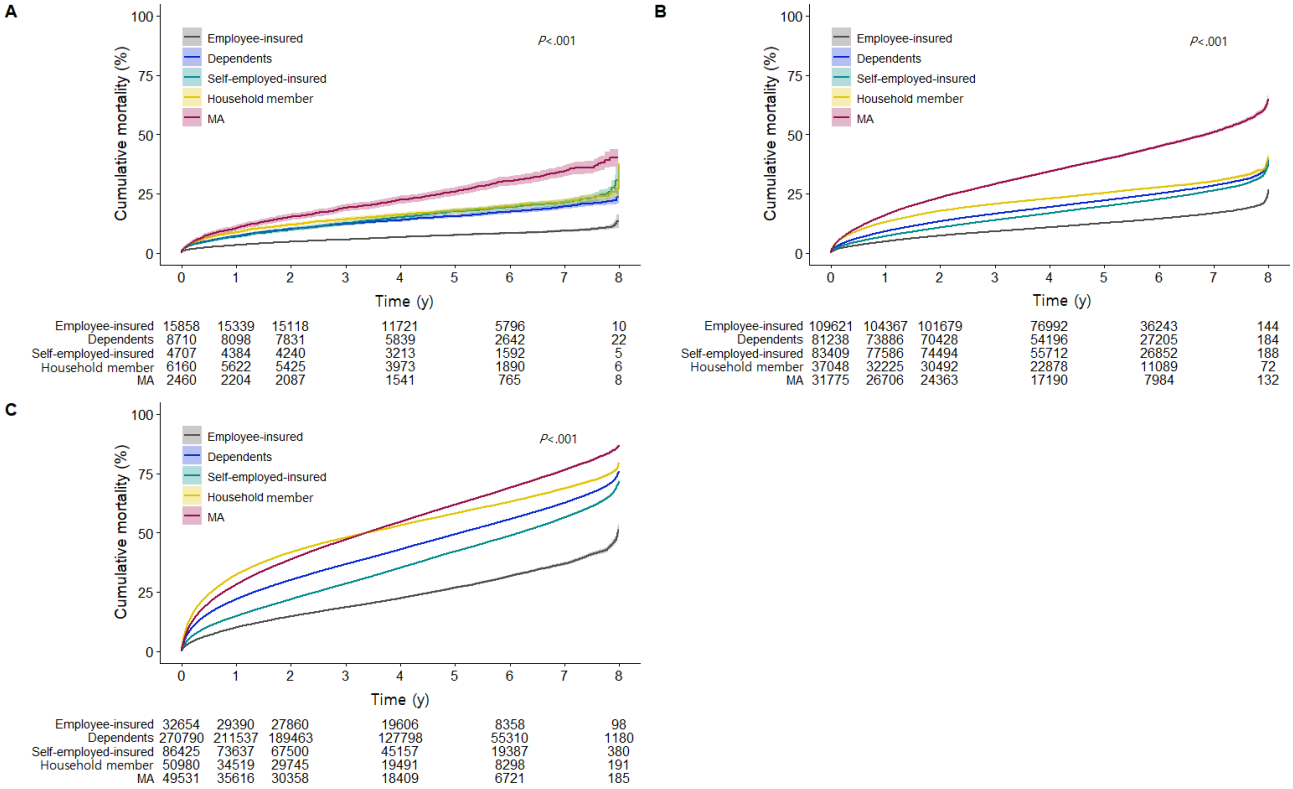
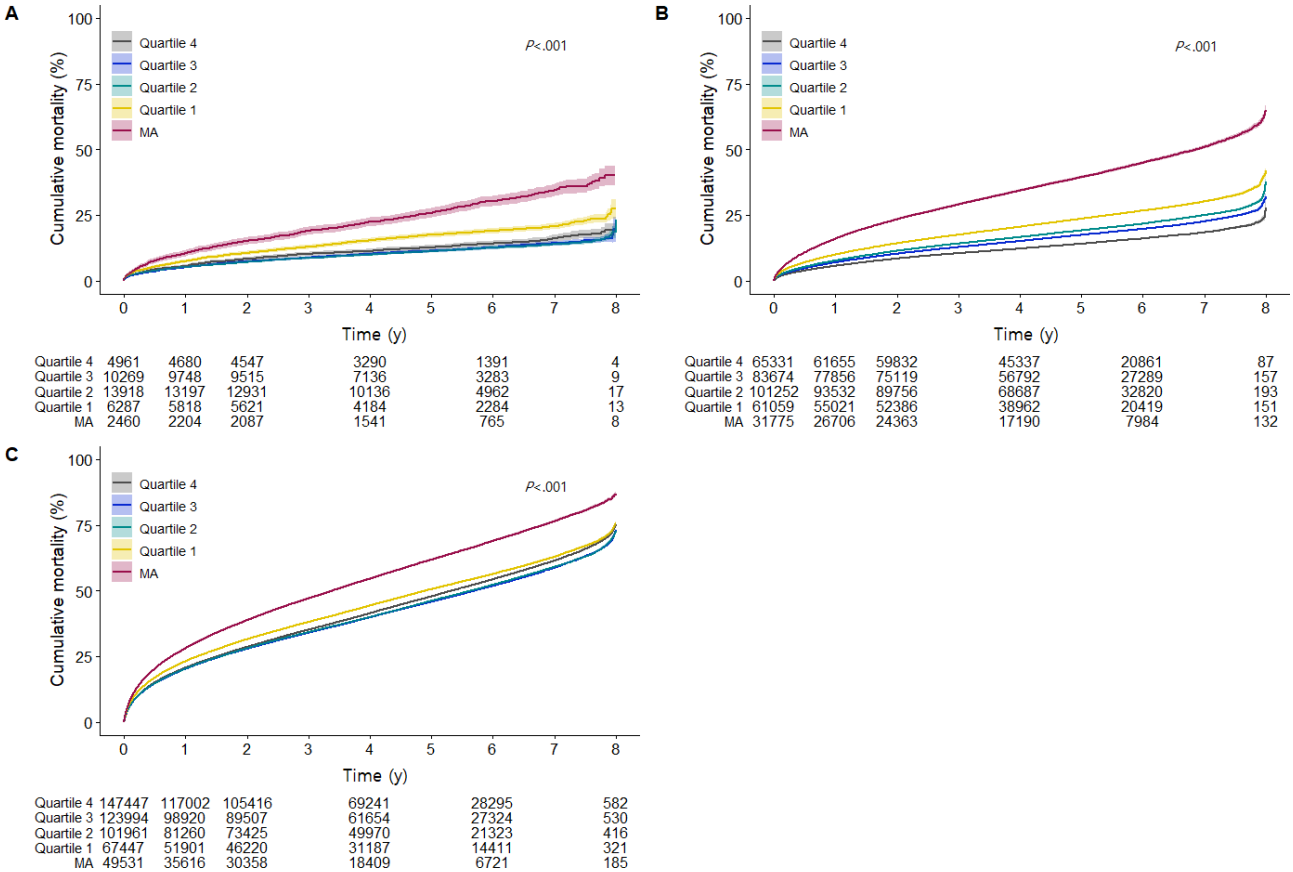
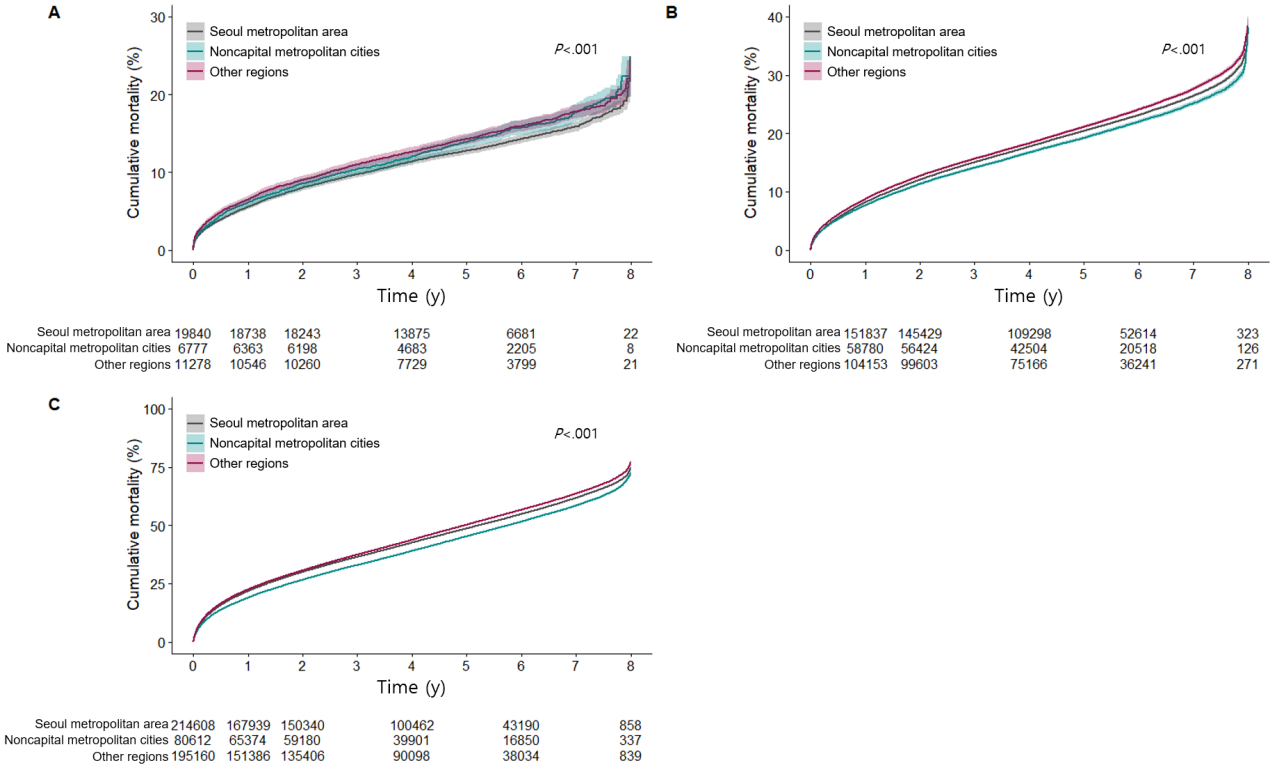


Figure 4. Cumulative mortality curves by income level, stratified by age group: (A) 20 to 39 years; (B) 40 to 64 years; and (C) 65 years and older. MA: Medical Aid.



XSL•FO
RenderX

Figure 5. Cumulative mortality curves by residential area, stratified by age group: (A) 20 to 39 years; (B) 40 to 64 years; and (C) 65 years and older.



Conversely, regional disparities were relatively small, with absolute differences of less than 3% across age categories, although differences by residential area remained statistically significant.

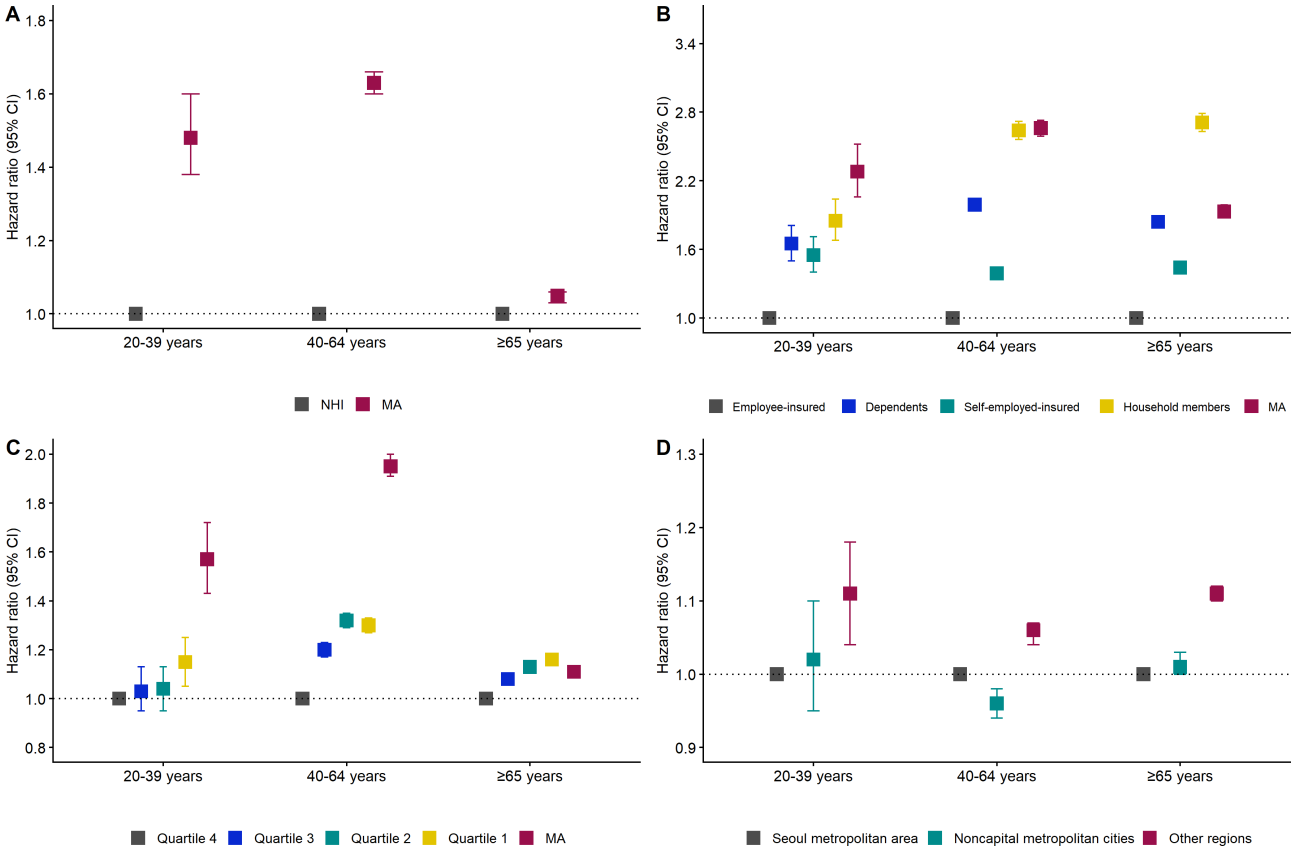
Multivariable Analysis of the Association Between SES, Residential Area, and Long-Term Mortality

Socioeconomic Indicators and Residential Area

Figure 6 presents forest plots of the aHRs for mortality according to socioeconomic indicators and residential area stratified by age group. Multimedia Appendix 4 summarizes the time-dependent Cox regression analyses evaluating the associations of socioeconomic indicators and residential area with long-term mortality among ICU survivors stratified by age. After adjusting for age, sex, smoking status, CCI score, and intervention status, both socioeconomic indicators and residential area were independently associated with long-term postdischarge mortality.

XSL•FO
RenderX

Figure 6. Forest plots of adjusted hazard ratios for mortality according to socioeconomic factors and residential area, stratified by age group: (A) insurance status; (B) National Health Insurance Service (NHIS) eligibility status; (C) income level; and (D) residential area. Adjusted for age, sex, smoking status, the Charlson Comorbidity Index, and interventions. NHI: National Health Insurance; MA: Medical Aid.



Insurance Status

In patients aged 65 years or older, the risk of mortality after full adjustment was marginally higher among MA beneficiaries than among NHI enrollees (8-y aHR 1.05, 95% CI 1.03 - 1.06). In contrast, younger age groups exhibited progressively higher mortality risk associated with MA coverage. Among individuals aged 40 to 64 years, the aHR increased from 1.28 (95% CI 1.24 - 1.32) at 1 year to 1.63 (95% CI 1.60 - 1.66) at 8 years. Among those aged 20 to 39 years, the aHR increased from 1.12 (95% CI 0.98 - 1.28) at 1 year to 1.48 (95% CI 1.38 - 1.60) at 8 years.

NHIS Eligibility Status

NHIS eligibility categories were associated with mortality. Compared with employee-insured beneficiaries, dependents of employee-insured beneficiaries, self-employed-insured householders, and self-employed-insured household members had significantly higher long-term mortality in all age groups. Among individuals aged 40 to 64 years, the 8-year aHRs were 1.99 (95% CI 1.94 - 2.04) for dependents of employee-insured beneficiaries and 2.64 (95% CI 2.56 - 2.72) for self-employed-insured household members. Among those aged 65 years or older, 8-year aHRs reached 1.84 (95% CI 1.80 - 1.89) for dependents of employee-insured beneficiaries and 2.71 (95% CI 2.63 - 2.79) for self-employed-insured household members.

Income Level

A distinct income-mortality gradient was observed across all age groups. When MA beneficiaries were analyzed along with income quartiles, their mortality risks were comparable to or exceeded those of the lowest-income quartile. Compared with the highest-income group, the 8-year mortality hazard increased progressively from the fourth quartile to the lowest-income group and to MA beneficiaries:

1. 20 to 39 years: 8-year aHR for the first quartile 1.15 (95% CI 1.05-1.25) and MA 1.57 (95% CI 1.43-1.72)
2. 40 to 64 years: 8-year aHR for the first quartile 1.30 (95% CI 1.27-1.33) and MA 1.95 (95% CI 1.91-2.00)
3. 65 years and older: 8-year aHR for the first quartile 1.16 (95% CI 1.14-1.17) and MA 1.11 (95% CI 1.10-1.13)

Residential Area

Compared with residents of the Seoul metropolitan area, those living in other regions consistently had higher mortality risks (20 - 39 y: 8-y aHR 1.11, 95% CI 1.04 - 1.18; 40 - 64 y: 8-y aHR 1.06, 95% CI 1.04 - 1.07; and ≥65 y: 8-y aHR 1.11, 95% CI 1.10 - 1.12). This difference remained significant even after adjusting for individual-level socioeconomic indicators (Multimedia Appendix 5).

Characteristics of the Survivors and Decedents Within the Low-Income Group (the MA and First Quartile Groups)

Among patients with low SES (n=218,559), the mean age was 65.9 (SD 14.8) years, and 55.3% (n=120,863) were male

XSL•FO
RenderX

patients. Compared with survivors (mean 60.9, SD 13.5 y), decedents were significantly older (mean 71.4, SD 14.2 y; *P*<.001) and had a higher comorbidity burden (mean CCI 3.9, SD 2.7 vs mean 2.7, SD 2.3; *P*<.001). The proportion of MA beneficiaries was greater among decedents (48,050/103,190, 46.6% vs 35,716/115,369, 30.9%; *P*<.001), and the proportion of residents in other regions was higher among decedents than among survivors (41,713/103,190, 40.4% vs 41,860/115,369, 36.3%; *P*<.001; Multimedia Appendix 6).

Discussion

Age-Related Differences in Socioeconomic Disparities

This nationwide cohort study demonstrated substantial socioeconomic and regional disparities in long-term mortality among ICU survivors in South Korea. Using time-dependent Cox models that accounted for temporal changes in SES and residential area, and that adjusted for clinical severity, we found that lower SES, defined by MA coverage, nonemployee NHIS eligibility categories, and low income, was independently associated with higher postdischarge mortality. In addition, residence in other regions was associated with an increased mortality risk.

The Kaplan-Meier survival curves clearly demonstrated these differences across all SES indicators and residential area categories, and the log-rank tests (all *P*'s<.001) confirmed that the curves differed significantly. This disparity was most evident among younger and middle-aged adults, with cumulative mortality rising more sharply in socioeconomically disadvantaged groups. One notable finding of this study was that socioeconomic gradients were most pronounced among younger adults, suggesting that SES-related factors may exert a greater proportional influence among populations with a lower baseline mortality risk. This pattern aligns with the "relative deprivation hypothesis," which posits that socioeconomic disadvantage exerts a more pronounced proportional impact on mortality among populations characterized by a low baseline risk [22]. In particular, younger MA beneficiaries or individuals with low incomes may encounter greater obstacles to returning to work, social engagement, or psychological recovery, which are crucial for long-term health following critical illness [23]. Building on this, the more pronounced SES-related disparities observed among younger age groups, especially among MA beneficiaries, may reflect the heterogeneous nature of this population. Young MA recipients may include individuals with severe chronic conditions, disabilities, mental health disorders, or marked social vulnerability, all of which may contribute to poor long-term outcomes after ICU discharge. Therefore, the observed disparities may not solely reflect economic disadvantages but rather the cumulative effect of multiple medical and social risk factors within these subgroups.

Socioeconomic Inequality and Long-Term Mortality After Critical Illness

In the United States, Medicaid users and uninsured patients are less likely to receive critical care services and are more likely to be withdrawn from life support, indicating that lack of insurance independently increases the risk of mortality [24]. ICU patients with public or no insurance had higher short-term and long-term mortality rates than did those with private insurance [25]. Similar evidence from Australia demonstrates that socioeconomic disadvantages negatively affect long-term survival after ICU admission [26]. Consistent with these findings, Cha et al [7] reported that the 1-year mortality was 1.31 times higher among MA recipients than among NHI enrollees in a propensity-matched cohort of 2495 ICU survivors in South Korea.

SES is well-established as a key determinant of cardiovascular health, particularly in acute coronary syndrome [27,28]. A 14-year Korean registry study involving 4873 cardiac arrest survivors revealed a statistically significant reduction in long-term survival rates among patients with MA (aHR 1.52) compared with NHI beneficiaries [12]. Nevertheless, a study conducted in Singapore revealed no significant correlation between SES and out-of-hospital cardiac arrest survival outcomes, likely attributable to the availability of affordable and equitable health care for all socioeconomic strata [29].

These studies corroborate our findings that socioeconomic factors influence survival, even within a universal health care framework. Specifically, compared with NHI enrollees, younger MA beneficiaries in our cohort exhibited up to a 1.5-fold increase in 8-year mortality risk, whereas the disparity was negligible among older adults (aHR 1.05, 95% CI 1.03 - 1.06). The persistence of SES-related disparities in a system characterized by nearly universal insurance coverage highlights that nonfinancial barriers, such as social isolation, restricted access to rehabilitation, limited continuity of care, and insufficient postacute care, remain crucial factors influencing post-ICU recovery [30-32].

NHIS Eligibility Status and Income as Markers of Post-ICU Vulnerability

In our study, differences across the NHIS eligibility categories were particularly evident. Compared with individuals with employee insurance, dependent and self-employed household members had a 2-fold to 3-fold increased mortality risk even after multivariable adjustment. A consistent income-mortality gradient was observed across all age groups, with mortality rates increasing from the highest to the lowest-income quartiles. MA beneficiaries, who were the most economically disadvantaged group, had an even higher risk of mortality than did those in the lowest-income NHI quartile, suggesting that MA status may capture broader socioeconomic and clinical vulnerabilities beyond income level alone. Our results indicate that clinical and sociodemographic characteristics vary considerably even within low-SES groups. This finding underscores the importance of developing more refined and tailored post-ICU discharge management strategies rather than applying a one-size-fits-all approach to socioeconomic risk.

Oh et al [33] analyzed a cohort of over 6000 adults from a tertiary hospital in South Korea and found that, compared with office workers, unemployed patients experienced markedly higher mortality at both 30 days and 1-year post-ICU admission (83%). In a substantial Taiwanese cohort representing the general population (not restricted to ICU patients), involuntary unemployment was found to be correlated with an increased

XSL•FO
RenderX

mortality risk (aHR 1.99) [34]. A meta-analysis of 52 studies revealed that approximately 60% of previously employed ICU survivors had resumed work 1 year postdischarge, and 68% after 5 years. The low rate of return to work was attributed to preexisting comorbidities and the ongoing physical, cognitive, and psychological effects of the critical illness. Even among individuals who return to work, many encounter job transitions, diminished working hours, or subsequent unemployment [35]. In Canada, a national cohort study of adults with severe trauma, including ICU patients, revealed a substantial return-to-work gap and a 19% reduction in average wages 3 years postinjury, with an employment rate of 79.3% compared with 91.7% in matched controls [36].

Collectively, these findings demonstrate that differences in workplace-based health surveillance, social capital, and continuity of care may affect long-term outcomes after serious illness. Job security and stable income seem to facilitate follow-up care, medication adherence, and the maintenance of healthy behaviors during long-term recovery.

Geographic Disparities in Access to Post-ICU Care

Patients living in other regions had modest but consistently higher mortality risks across all age groups (8-year aHR 1.06 - 1.11). In an additional analysis, residence in other regions was consistently associated with higher long-term mortality than with residence in the Seoul metropolitan area across all age groups, even after adjustment for individual socioeconomic indicators, such as insurance type, NHIS eligibility status, and income level. These findings suggest that residential areas may capture regional contextual disparities in post-ICU outcomes that are not fully explained by individual-level socioeconomic disadvantages alone.

A population-based study comparing rural and urban residents hospitalized for COVID-19 found that rural residents had a 22% higher risk of long-term mortality than did their urban counterparts [11]. Likewise, a retrospective analysis of Medicare beneficiaries in the United States, analogous to the NHI in Korea, revealed a persistent rural-urban disparity in mortality after discharge from acute care hospitals [37].

Communities in other regions may experience multiple structural disadvantages, such as increased distances to tertiary hospitals, constrained ICU capacity, shortages of specialists, and limited access to advanced medical equipment. Consequently, patients in these areas often have poorer postdischarge survival rates. Moreover, elevated long-term mortality may indicate both economic disparities and nonfinancial obstacles, including insufficient post-ICU rehabilitation resources, the absence of family or caregiver support, nursing workforce shortages, inadequate compliance with follow-up appointments, and difficulties in remote health monitoring.

Although the Korean health care system offers universal insurance coverage, health care resources and postacute care infrastructure are predominantly located in urban areas. These results highlight the necessity for region-specific post-ICU follow-up programs and telemedicine-based continuity of care to mitigate geographic disparities in long-term outcomes after critical illness.

Limitations

Some limitations of this study should be acknowledged. First, the NHIS database lacks detailed clinical information, such as physiological and laboratory variables (eg, Acute Physiology and Chronic Health Evaluation II, Sequential Organ Failure Assessment scores, or vital parameters). Although surrogate indicators (mechanical ventilation, renal replacement therapy, vasopressor use, and extracorporeal membrane oxygenation) were used to approximate illness severity, residual confounding owing to unmeasured clinical heterogeneity may persist. Second, SES indicators, including insurance type, NHIS eligibility status, and income level, were inferred from administrative and premium-based data rather than direct measures of education, occupation, household wealth, or actual labor force participation. While these proxies have been validated in prior population-based research, they may not fully capture multidimensional socioeconomic contexts, such as education or social capital. Third, behavioral and psychosocial variables (eg, marital status, mental health, family or social support, physical activity, lifestyle factors, and health literacy) were not included in the NHIS data. These unmeasured confounders may partly explain the observed SES-mortality association, particularly among younger and working-age adults. Additionally, variations in hospital resources and local post-ICU care infrastructure (eg, rehabilitation and nursing facilities) were not directly measured. Furthermore, although sampling weights were not applied because this study analyzed the complete eligible nationwide NHIS cohort rather than a sampled population, potential selection bias due to exclusions or missing data cannot be completely ruled out. Future studies could assess the impact of such bias using sensitivity analyses, including weighting methods or multiple imputation, where appropriate. Finally, owing to the retrospective observational design, causal interpretations should be made with caution, and unmeasured confounding may remain despite extensive covariate adjustment. Future studies integrating detailed clinical data, patient-reported outcomes, and regional health care resource mapping are required to elucidate the causal pathways linking socioeconomic and regional factors to long-term post-ICU mortality.

Conclusions

In this national cohort of ICU survivors, disparities in individual SES and residential region were independently associated with long-term mortality, even within a universal health care system. The impact of low SES was particularly strong among younger and middle-aged individuals. For these groups, low income and vulnerable NHIS eligibility categories were more strongly associated with long-term mortality, even after accounting for age and comorbidities. Residence in other regions was consistently associated with higher long-term mortality across all age groups. These findings suggest that socioeconomic disadvantage and regional disparities influence survival trajectories beyond acute illnesses, highlighting the persistent social gradient in health outcomes. Future policies should prioritize post-ICU transitional care networks, community-based rehabilitation, and risk screening during discharge planning to address socioeconomic and regional vulnerabilities.

XSL•FO
RenderX

Acknowledgments

We would like to thank Editage [38] for the English-language editing. During the preparation of this manuscript, the authors used ChatGPT (OpenAI) and QuillBot for language editing, grammar refinement, paraphrasing, and readability improvement. The authors did not use these tools for data analysis, statistical modeling, interpretation of findings, or the generation of scientific conclusions. After using these tools, the authors have reviewed and edited the content as needed and take full responsibility for the final content of the manuscript.

Funding

This work was partially supported by the National Research Foundation of Korea (NRF-2022R1A2C1012627) and by an Institute of Information & Communications Technology Planning & Evaluation (IITP) grant from the Korean government (Ministry of Science and Information and Communications Technology [MSIT]; number RS-2020-II201373, Artificial Intelligence Graduate School Program at Hanyang University). The funders had no role in the conceptualization, design, data collection, analysis, the decision to publish, or preparation of the manuscript.

Data Availability

The data that support the findings of this study are available from the National Health Insurance Service (NHIS) of South Korea; however, restrictions apply to the availability of these data, which were used under license for the current study and are therefore not publicly available. The data are available from the authors upon reasonable request and with permission from the NHIS.

Authors' Contributions

NC and JO were responsible for the conceptualization, methodology, analysis, writing the original draft, and review and editing. JO oversaw project administration and supervision and contributed to the review and editing. KHY and JL contributed to methodology, analysis, and data curation. SHL contributed to review and editing. All authors have read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study design and assessment windows.
[DOCX File, 73 KB - publichealth_v12i1e89127_app1.docx]

Multimedia Appendix 2
Primary diagnoses at intensive care unit admission classified by system.
[DOCX File, 24 KB - publichealth_v12i1e89127_app2.docx]

Multimedia Appendix 3
Eight-year cumulative mortality (%) and difference in absolute mortality risk by socioeconomic indicators, residential area, and age groups.
[DOCX File, 28 KB - publichealth_v12i1e89127_app3.docx]

Multimedia Appendix 4
Adjusted associations between age (y), socioeconomic factors, residential area, and long-term mortality among intensive care unit survivors.
[DOCX File, 36 KB - publichealth_v12i1e89127_app4.docx]

Multimedia Appendix 5
Age group-specific associations between residential area and long-term mortality among intensive care unit survivors after adjustment for individual socioeconomic indicators.
[DOCX File, 31 KB - publichealth_v12i1e89127_app5.docx]

Multimedia Appendix 6
Comparison of characteristics between survivors and decedents in the low-income group (Medical Aid and first quartile).
[DOCX File, 29 KB - publichealth_v12i1e89127_app6.docx]

XSL•FO
RenderX

References

1. Roller-Wirnsberger R, Thurner B, Pucher C, Lindner S, Wirnsberger GH. The clinical and therapeutic challenge of treating older patients in clinical practice. Br J Clin Pharmacol 2020 Oct;86(10):1904-1911. [doi: 10.1111/bcp.14074] [Medline: 31321798]
2. Herridge MS, Tansey CM, Matté A, et al. Functional disability 5 years after acute respiratory distress syndrome. N Engl J Med 2011 Apr 7;364(14):1293-1304. [doi: 10.1056/NEJMoa1011802] [Medline: 21470008]
3. Dowdy DW, Eid MP, Dennison CR, et al. Quality of life after acute respiratory distress syndrome: a meta-analysis. Intensive Care Med 2006 Aug;32(8):1115-1124. [doi: 10.1007/s00134-006-0217-3] [Medline: 16783553]
4. Rousseau AF, Prescott HC, Brett SJ, et al. Long-term outcomes after critical illness: recent insights. Crit Care 2021 Mar 17;25(1):108. [doi: 10.1186/s13054-021-03535-3] [Medline: 33731201]
5. Needham DM, Davidson J, Cohen H, et al. Improving long-term outcomes after discharge from intensive care unit: report from a stakeholders' conference. Crit Care Med 2012 Feb;40(2):502-509. [doi: 10.1097/CCM.0b013e318232da75] [Medline: 21946660]
6. Davidson JE, Jones C, Bienvenu OJ. Family response to critical illness: postintensive care syndrome-family. Crit Care Med 2012 Feb;40(2):618-624. [doi: 10.1097/CCM.0b013e318236ebf9] [Medline: 22080636]
7. Cha JK, Oh TK, Song IA. Impacts of financial coverage on long-term outcome of intensive care unit survivors in South Korea. Yonsei Med J 2019 Oct;60(10):976-983. [doi: 10.3349/ymj.2019.60.10.976] [Medline: 31538433]
8. Chetty R, Stepner M, Abraham S, et al. The association between income and life expectancy in the United States, 2001–2014. JAMA 2016 Apr 26;315(16):1750-1766. [doi: 10.1001/jama.2016.4226] [Medline: 27063997]
9. Ferrante LE, Pisani MA, Murphy TE, Gahbauer EA, Leo-Summers LS, Gill TM. Functional trajectories among older persons before and after critical illness. JAMA Intern Med 2015 Apr;175(4):523-529. [doi: 10.1001/jamainternmed.2014.7889] [Medline: 25665067]
10. Park HC, Kwon YE, Choi HY, et al. Health insurance status is related to risk of mortality and hospitalization in Korean maintenance hemodialysis patients: a longitudinal cohort study. Am J Nephrol 2020;51(12):975-981. [doi: 10.1159/000512855] [Medline: 33440390]
11. Yousufuddin M, Mahmood M, Barkoudah E, et al. Rural-urban differences in long-term mortality and readmission following COVID-19 hospitalization, 2020 to 2023. Open Forum Infect Dis 2024;11(5):ofae197. [doi: 10.1093/ofid/ofae197] [Medline: 38698896]
12. Yoo KH, Cho Y, Oh J, et al. Association of socioeconomic status with long-term outcome in survivors after out-of-hospital cardiac arrest: nationwide population-based longitudinal study. JMIR Public Health Surveill 2023 Jul 11;9:e47156. [doi: 10.2196/47156] [Medline: 37432716]
13. Marmot M, Allen JJ. Social determinants of health equity. Am J Public Health 2014 Sep;104(Suppl 4):S517-S519. [doi: 10.2105/AJPH.2014.302200] [Medline: 25100411]
14. Kyoung DS, Kim HS. Understanding and utilizing claim data from the Korean National Health Insurance Service (NHIS) and Health Insurance Review & Assessment (HIRA) database for research. J Lipid Atheroscler 2022 May;11(2):103-110. [doi: 10.12997/jla.2022.11.2.103] [Medline: 35656154]
15. Park J, Kwon S, Choi EK, et al. Validation of diagnostic codes of major clinical outcomes in a National Health Insurance database. Int J Arrhythm 2019;20:5. [doi: 10.1186/s42444-019-0005-0]
16. Lee DW, Jang J, Choi DW, Jang SI, Park EC. The effect of shifting medical coverage from National Health Insurance to Medical Aid type I and type II on health care utilization and out-of-pocket spending in South Korea. BMC Health Serv Res 2020 Oct 27;20(1):979. [doi: 10.1186/s12913-020-05778-2] [Medline: 33109176]
17. Cheol Seong S, Kim YY, Khang YH, et al. Data resource profile: the National Health Information Database of the National Health Insurance Service in South Korea. Int J Epidemiol 2017 Jun 1;46(3):799-800. [doi: 10.1093/ije/dyw253] [Medline: 27794523]
18. Chamberlain RC, Barnetson C, Clegg GR, Halbesma N. Association of measures of socioeconomic position with survival following out-of-hospital cardiac arrest: a systematic review. Resuscitation 2020 Dec;157:49-59. [doi: 10.1016/j.resuscitation.2020.09.025] [Medline: 33010372]
19. Kim JH, Lee KS, Yoo KB, Park EC. The differences in health care utilization between Medical Aid and Health Insurance: a longitudinal study using propensity score matching. PLoS One 2015;10(3):e0119939. [doi: 10.1371/journal.pone.0119939] [Medline: 25816234]
20. Foraker RE, Rose KM, Whitsel EA, Suchindran CM, Wood JL, Rosamond WD. Neighborhood socioeconomic status, Medicaid coverage and medical management of myocardial infarction: atherosclerosis risk in communities (ARIC) community surveillance. BMC Public Health 2010 Oct 21;10:632. [doi: 10.1186/1471-2458-10-632] [Medline: 20964853]
21. Yang Y, Niu L, Amin S, Yasin I. Unemployment and mental health: a global study of unemployment's influence on diverse mental disorders. Front Public Health 2024;12:1440403. [doi: 10.3389/fpubh.2024.1440403] [Medline: 39735766]
22. Mackenbach JP, Stirbu I, Roskam AJR, et al. Socioeconomic inequalities in health in 22 European countries. N Engl J Med 2008 Jun 5;358(23):2468-2481. [doi: 10.1056/NEJMsa0707519] [Medline: 18525043]

XSL•FO
RenderX

23. Lusk JB, Blass B, Mahoney H, et al. Neighborhood socioeconomic deprivation, healthcare access, and 30-day mortality and readmission after sepsis or critical illness: findings from a nationwide study. Crit Care 2023 Jul 15;27(1):287. [doi: 10.1186/s13054-023-04565-9] [Medline: 37454127]
24. Fowler RA, Noyahr LA, Thornton JD, et al. An official American Thoracic Society systematic review: the association between health insurance status and access, care delivery, and outcomes for patients who are critically ill. Am J Respir Crit Care Med 2010 May 1;181(9):1003-1011. [doi: 10.1164/rccm.200902-0281ST] [Medline: 20430926]
25. Barnato AE, Alexander SL, Linde-Zwirble WT, Angus DC. Racial variation in the incidence, care, and outcomes of severe sepsis: analysis of population, patient, and hospital characteristics. Am J Respir Crit Care Med 2008 Feb 1;177(3):279-284. [doi: 10.1164/rccm.200703-480OC] [Medline: 17975201]
26. Doherty Z, Kippen R, Bevan D, et al. Long-term outcomes of hospital survivors following an ICU stay: a multi-centre retrospective cohort study. PLoS One 2022;17(3):e0266038. [doi: 10.1371/journal.pone.0266038] [Medline: 35344543]
27. Schultz WM, Kelli HM, Lisko JC, et al. Socioeconomic status and cardiovascular outcomes: challenges and interventions. Circulation 2018 May 15;137(20):2166-2178. [doi: 10.1161/CIRCULATIONAHA.117.029652] [Medline: 29760227]
28. Beard JR, Earnest A, Morgan G, et al. Socioeconomic disadvantage and acute coronary events: a spatiotemporal analysis. Epidemiology 2008 May;19(3):485-492. [doi: 10.1097/EDE.0b013e3181656d7f] [Medline: 18414088]
29. Lee DYX, Yau CE, Pek MPP, et al. Socioeconomic disadvantage and long-term survival duration in out-of-hospital cardiac arrest patients: a population-based cohort study. Resusc Plus 2024;18:100610. [doi: 10.1016/j.resplu.2024.100610] [Medline: 38524148]
30. Chao PW, Shih CJ, Lee YJ, et al. Association of postdischarge rehabilitation with mortality in intensive care unit survivors of sepsis. Am J Respir Crit Care Med 2014 Nov 1;190(9):1003-1011. [doi: 10.1164/rccm.201406-1170OC] [Medline: 25210792]
31. Lai HH, Tseng PY, Wang CY, Wang JY. Long-term survival and medical costs of patients with prolonged mechanical ventilation and tracheostomy: a nationwide cohort study. Int J Environ Res Public Health 2021 Sep 29;18(19):10272. [doi: 10.3390/ijerph181910272] [Medline: 34639575]
32. Ikumi S, Shiga T, Ueda T, et al. Intensive care unit mortality and cost-effectiveness associated with intensivist staffing: a Japanese nationwide observational study. J Intensive Care 2023 Dec 4;11(1):60. [doi: 10.1186/s40560-023-00708-w] [Medline: 38049894]
33. Oh TK, Jo J, Jeon YT, Song IA. Impact of socioeconomic status on 30-day and 1-year mortalities after intensive care unit admission in South Korea: a retrospective cohort study. Acute Crit Care 2018 Nov;33(4):230-237. [doi: 10.4266/acc.2018.00514] [Medline: 31723890]
34. Tsai SL, Lan CF, Lee CH, Huang N, Chou YJ. Involuntary unemployment and mortality in Taiwan. J Formos Med Assoc 2004 Dec;103(12):900-907. [Medline: 15624038]
35. Kamdar BB, Suri R, Suchyta MR, et al. Return to work after critical illness: a systematic review and meta-analysis. Thorax 2020 Jan;75(1):17-27. [doi: 10.1136/thoraxjnl-2019-213803] [Medline: 31704795]
36. Haas B, Jeon SH, Rotermann M, et al. Association of severe trauma with work and earnings in a national cohort in Canada. JAMA Surg 2021 Jan 1;156(1):51-59. [doi: 10.1001/jamasurg.2020.4599] [Medline: 33112383]
37. Kosar CM, Loomer L, Ferdows NB, Trivedi AN, Panagiotou OA, Rahman M. Assessment of rural-urban differences in postacute care utilization and outcomes among older US adults. JAMA Netw Open 2020 Jan 3;3(1):e1918738. [doi: 10.1001/jamanetworkopen.2019.18738] [Medline: 31913495]
38. Editage. URL: https://www.editage.co.kr/ [accessed 2026-08-12]

Abbreviations

aHR: adjusted hazard ratio
AMR: absolute mortality risk
CCI: Charlson Comorbidity Index
HYUH: Hanyang University Hospital
***ICD-10*:** *International Classification of Diseases, 10th Revision*
ICU: intensive care unit
MA: Medical Aid
NHI: National Health Insurance
NHIS: National Health Insurance Service
SES : socioeconomic status

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 07.Dec.2025; peer-reviewed by J Choi, J Li; revised version received 02.Aug.2026; accepted 03.Aug.2026; published 31.Aug.2026.

Please cite as:
Choi N, Oh J, Yoo KH, Lee J, Lee SH
Socioeconomic and Regional Disparities in Long-Term Mortality Among Intensive Care Unit Survivors in South Korea: Nationwide Cohort Study
JMIR Public Health Surveill 2026;12:e89127
URL: https://publichealth.jmir.org/2026/1/e89127
doi:10.2196/89127

© Nayeon Choi, Jaehoon Oh, Kyung Hun Yoo, Juncheol Lee, Sang Hwan Lee. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 31.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Triglyceride-Glucose Index and Risk of End-Stage Kidney Disease in Young Adults: National Population-Based Cohort Study

Jin Sug Kim[1], MD, PhD; Yu Ho Lee[1], MD, PhD; Sungjin Chung[2], MD, PhD; Hyeon Seok Hwang[1], MD, PhD; Kyungdo Han[3], PhD; Hye Eun Yoon[4,5], MD, PhD

[1]Division of Nephrology, Department of Internal Medicine, Kyung Hee University College of Medicine, Kyung Hee University Medical Center, Seoul, Republic of Korea

[2]Division of Nephrology, Department of Internal Medicine, Yeouido St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea

[3]Department of Statistics and Actuarial Science, Soongsil University, Seoul, Republic of Korea

[4]Division of Nephrology, Department of Internal Medicine, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, 222 Banpo-daero, Seocho-gu, Seoul, Republic of Korea

[5]Transplantation Research Center, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea

Corresponding Author:
Hye Eun Yoon, MD, PhD
Division of Nephrology, Department of Internal Medicine, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, 222 Banpo-daero, Seocho-gu, Seoul, Republic of Korea

Abstract

Background: The triglyceride-glucose (TyG) index, a simple marker of insulin resistance, is associated with chronic kidney disease and end-stage kidney disease (ESKD). However, its association with ESKD in young adults remains unexplored.

Objective: This study aimed to investigate whether the TyG index increases the risk of ESKD in the young adult population.

Methods: This nationwide retrospective, population-based cohort study analyzed data from 6,555,863 individuals aged 20 - 39 years between January 2009 and December 2012 using the Korean National Health Insurance Database. Participants were categorized into quartiles and deciles based on the TyG index. The cumulative incidence and adjusted hazard ratios for ESKD were assessed.

Results: Over a median follow-up period of 10.61 years, 4910 (0.75%) participants developed ESKD. The incidence of ESKD increased with a higher TyG index in quartiles and deciles. The adjusted hazard ratio for ESKD in the highest quartile was 3.438 (95% CI 3.096 - 3.818) compared to the lowest quartile, and that in the highest decile was 5.944 (95% CI 5.022 - 7.035) compared to the lowest decile. These associations remained consistent across subgroup analyses by age, sex, obesity, dyslipidemia, and chronic kidney disease.

Conclusions: The TyG index is significantly associated with ESKD risk in young adults.

(JMIR Public Health Surveill 2026;12:e73750) doi:10.2196/73750

KEYWORDS

triglyceride-glucose index; chronic kidney disease; end-stage kidney disease; dialysis; young adults

Introduction

Chronic kidney disease (CKD) is a significant global health issue and a leading cause of mortality and morbidity. Its prevalence is rapidly increasing, with approximately 843.6 million patients affected globally in 2017 [1]. Additionally, CKD and its associated comorbidities impose a significant financial burden on the health care systems [2]. Although CKD is predominantly diagnosed in older adults, its global incidence among adolescents and young adults is increasing [3]. This trend is attributed to the increasing rates of obesity, lipid abnormalities, diabetes, and hypertension in childhood or adolescence [3-5]. Consequently, young adults are more susceptible to developing CKD than in the past. Moreover, early onset CKD is associated with greater disease severity and cardiovascular complications because of prolonged disease duration [6], as well as low quality of life and socioeconomic burden [7].

Insulin resistance is an early metabolic alteration in CKD [8], and its association with CKD is independent of diabetes [9,10]. It contributes to progressive kidney damage by causing glomerular hyperfiltration, sodium retention, tubular dysfunction, and kidney inflammation and fibrosis [11]. Although the hyperinsulinemic-euglycemic glucose clamp technique is considered the gold standard for measuring insulin resistance, its clinical utility is limited because of the time and

XSL•FO
RenderX

resources required. The homeostatic model assessment of insulin resistance is a more commonly used method [8] and has been reported to increase the risk of incident CKD development in individuals with normal kidney function [12]. The triglyceride-glucose (TyG) index, calculated as the logarithmic product of fasting triglyceride and glucose levels, correlates well with both the hyperinsulinemic-euglycemic clamp and homeostatic model assessment of insulin resistance indices [13]. Owing to its simplicity and reliability, the TyG index is widely used in large-scale epidemiological studies to assess insulin resistance.

Previous studies have established an association between the TyG index and incident CKD in the general population [14]. Recently, an association has also been reported between the TyG index and end-stage kidney disease (ESKD) in the general population [15] and in patients with diabetes [16]. However, no studies have examined the relationship between the TyG index and the development of ESKD, specifically in young adults. Therefore, this study aimed to investigate the association between the TyG index and the risk of incident ESKD in this population.

Methods

Database

This study retrospectively analyzed a nationwide cohort of patients who underwent health checkups between 2009 and 2012, as provided by the National Health Insurance Service (NHIS) of South Korea. The NHIS database includes individual health checkup results, along with claims data containing diagnostic, procedural, and prescription codes. As the NHIS serves as the sole public insurer in South Korea, its database is representative of the Korean population. The health checkup database comprises demographic data, including body measurements, blood and urine test results, and survey results on health-related behaviors.

Study Design and Population

This study identified 6,891,401 individuals aged 20 - 39 years who underwent a national health checkup between 2009 and 2012. We excluded 330,334 individuals with insufficient data to calculate the TyG index; 2412 individuals with ESKD identified using rare incurable disease registry codes for hemodialysis (V001), peritoneal dialysis (V003), or kidney transplantation (V005); and 2732 individuals with insurance claims related to ESKD or who died within 1 year after the index date. The remaining 6,555,863 eligible individuals were included and followed up until December 2021. Regarding the time frame, the start of follow-up (time zero) was defined as 1 year after the index date (lag period). Follow-up was stopped at death, censorship (eg, immigration), or ESKD development. The mean follow-up duration was 10.34 (SD 1.18) years. The participants were divided into 4 groups based on their TyG index, with quartiles 1 and 4 representing the lowest and the highest values, respectively. The mean (SD) values of the TyG index for each quartile were as follows: quartile 1: 7.68 (SD 0.29), quartile 2: 8.16 (SD 0.27), quartile 3: 8.52 (SD 0.32), and quartile 4: 9.15 (SD 0.51).

Clinical and Laboratory Measurements

All participants completed a questionnaire detailing their medical history and lifestyle habits, including smoking, alcohol consumption, and exercise. Smoking status was categorized as nonsmoker, former smoker, or current smoker. Alcohol consumption was classified as nondrinker, mild drinker (<30 g/day), or heavy drinker (≥30 g/day). Regular exercise was defined as vigorous-intensity exercise performed 3 or more times per week or moderate-intensity exercise performed 5 or more times per week. Routine physical examinations included body weight, height, waist circumference, and blood pressure. BMI was calculated as body weight (kg) divided by height squared (m^2). A low income was defined as being in the lowest quartile of the NHIS population or receiving government-funded medical aid. Blood samples were collected after overnight fasting to assess the serum levels of glucose, creatinine, total cholesterol, triglycerides, high-density lipoprotein-cholesterol (HDL-C), and low-density lipoprotein-cholesterol. The estimated glomerular filtration rate (eGFR) was calculated using the Modification of Diet in Renal Disease equation [17].

Definitions of Variables and Study Outcome

The TyG index was calculated using the formula log (fasting triglyceride [mg/dL] × fasting blood glucose [mg/dL]/2) [13,18]. Baseline comorbidities were defined as follows: hypertension (*ICD-10* [*International Statistical Classification of Diseases and Related Health Problems, 10th Revision*] codes I10-I13 or I15 and treatment with antihypertensive medications, systolic blood pressure [SBP] ≥ 140 mmHg, or diastolic blood pressure [DBP] ≥ 90 mmHg), diabetes (*ICD-10* codes E11-E14 and antidiabetic treatment or fasting glucose level ≥ 126 mg/dL), dyslipidemia (*ICD-10* code E78 with lipid-lowering agents or serum total cholesterol ≥240 mg/dL), and CKD (eGFR<60 mL/min/1.73 m^2). Antihypertensive medications included angiotensin type II receptor blockers, angiotensin-converting enzyme inhibitors, calcium channel blockers, beta-blockers, diuretics, or alpha-blockers. Antidiabetic medications included metformin, meglitinide, sulfonylurea, dipeptidyl peptidase IV inhibitor, alpha-glucosidase inhibitor, thiazolidines, or insulin. Lipid-lowering agents (dyslipidemia medication) included statins, ezetimibe, or fibrates.

In this study, obesity was defined as a BMI ≥25 kg/m^2, consistent with Asian-specific criteria [19]. Metabolic syndrome (MetS) was diagnosed based on the revised National Cholesterol Education Program Adult Treatment Panel III criteria [20,21]. MetS was diagnosed when three or more of the following five criteria were met: (1) elevated blood pressure (SBP ≥130 mmHg, DBP ≥85 mmHg, use of antihypertensive medications at baseline, and/or a history of hypertension), (2) elevated triglycerides (fasting triglycerides ≥ 150 mg/dL or use of dyslipidemia medications at baseline), (3) reduced HDL-C (<40 mg/dL for men and <50 mg/dL for women), (4) elevated fasting plasma glucose level (≥100 mg/dL or use of medications for elevated glucose levels), and (5) abdominal obesity (waist circumference ≥90 cm for men or ≥85 cm for women).

The primary outcome of this study was newly diagnosed ESKD, defined as the initiation of kidney replacement therapy and/or

XSL•FO
RenderX

kidney transplantation. Kidney replacement therapy or kidney transplantation was defined based on the rare incurable disease registry codes for hemodialysis (V001), peritoneal dialysis (V003), or kidney transplantation (V005).

Statistics Analysis

Categorical variables are presented as frequencies and percentages, whereas continuous variables are expressed as mean (SD). For continuous variables not following a normal distribution, log transformation was applied, and the results were presented as geometric means with 95% CIs. Baseline characteristics across the TyG index quartiles were compared using one-way ANOVA for continuous variables and the chi-square test for categorical variables. The follow-up duration was calculated for each TyG index quartile group. The incidence rates of ESKD were determined for the TyG index, fasting glucose, and triglyceride quartiles over the total follow-up period for the study outcome. Incidence rates were expressed as events per 1000 person-years. Kaplan-Meier curves were generated to estimate incidence trends, and the log-rank test was used to compare groups. All outcomes were analyzed using Cox proportional hazards regression analysis while controlling for baseline covariates. Sensitivity analyses were performed to reduce potential bias and demonstrate the consistency of our findings. Incidence rates of ESKD were calculated for deciles of the TyG index, fasting glucose, and triglycerides. Subgroup analyses were performed to explore whether the association between the TyG index and ESKD differed by age, sex, obesity, MetS, diabetes, diabetes medication, dyslipidemia, dyslipidemia medication, and CKD using Cox proportional hazard regression analysis adjusted for age, sex, BMI, smoking, drinking, regular exercise, low income, diabetes medication, dyslipidemia medication, hypertension, eGFR, and CKD. Additionally, 2-year follow-up values of the TyG index, fasting glucose, and triglyceride were used to analyze the incidence rates of ESKD and the association between the TyG index quartiles and ESKD. The incidence rates and associations were evaluated both in terms of the 2-year follow-up values and the time-averaged values of baseline and 2-year follow-up. The risks of ESKD associated with the TyG index groups were expressed as hazard ratios (HRs) with 95% CIs, using the first quartile or decile group as the reference. A two-tailed $P<.05$ was considered statistically significant. All analyses were performed using SAS (version 9.4; SAS Institute) and R (version 3.4.1; R Foundation for Statistical Computing).

Ethical Considerations

This study adhered to the Declaration of Helsinki (2013 revision) and was approved by the Institutional Review Board of CHA Bundang Medical Center (CHAMC 2024-04-034). The requirement for written informed consent was waived by the Institutional Review Board because of the retrospective nature of the study.

Results

Baseline Characteristics

Baseline demographic and clinical characteristics were compared across the TyG index quartiles (Table 1). An increase in the TyG index quartile was associated with significant increases in age, BMI, waist circumference, SBP, DBP, and levels of fasting plasma glucose, total cholesterol, and triglycerides, whereas HDL-C and eGFR levels decreased. Similarly, higher TyG index quartiles were associated with significantly greater proportions of individuals with high BMI, current smoking status, heavy alcohol consumption, and use of diabetes or dyslipidemia medication. The proportion of individuals with low income varied across the TyG index quartiles, with lower prevalence in the middle quartiles. Additionally, the prevalence of hypertension, diabetes, and dyslipidemia increased across quartiles, whereas the proportion of individuals engaging in regular exercise decreased. Elevated low-density lipoprotein-cholesterol levels and a higher prevalence of CKD were also observed in the higher TyG index quartiles.

XSL•FO
RenderX

Table . Baseline characteristics of quartiles of the triglyceride-glucose index.

Variable	Quartile 1 (n= 1,638,852)	Quartile 2 (n= 1,639,754)	Quartile 3 (n= 1,637,725)	Quartile 4 (n= 1,639,532)	*P* value
TyG[a] index, mean (SD)	7.68 (0.29)	8.16 (0.27)	8.52 (0.32)	9.15 (0.51)	—[b]
Age (years), mean (SD)	29.47 (4.92)	30.5 (4.94)	31.26 (4.91)	32.24 (4.75)	<.001
Sex, n (%)					.35
Male	975,747 (59.54)	975,000 (59.46)	975,117 (59.54)	975,291 (59.49)	
Female	663,105 (40.46)	664,754 (40.54)	662,608 (40.46)	664,241 (40.51)	
BMI (kg/m^2), mean (SD)	21.65 (2.86)	22.38 (3.22)	23.24 (3.56)	24.74 (3.94)	<.001
BMI category, n (%)					<.001
<18.5	182,525 (11.14)	144,074 (8.79)	109,906 (6.71)	59,424 (3.62)	
18.5≤ BMI<23	983,454 (60.01)	859,700 (52.43)	716,050 (43.72)	502,144 (30.63)	
23≤ BMI<25	273,285 (16.68)	315,712 (19.25)	338,387 (20.66)	333,491 (20.34)	
25≤ BMI<30	183,066 (11.17)	284,045 (17.32)	403,750 (24.65)	588,208 (35.88)	
≥30	16,522 (1.01)	36,223 (2.21)	69,632 (4.25)	156,265 (9.53)	
Smoking, n (%)					<.001
Non	981,155 (59.87)	925,478 (56.44)	877,695 (53.59)	806,929 (49.22)	
Former	158,638 (9.68)	169,509 (10.34)	175,628 (10.72)	175,936 (10.73)	
Current	499,059 (30.45)	544,767 (33.22)	584,402 (35.68)	656,667 (40.05)	
Drinking, n (%)					<.001
Non	650,144 (39.67)	627,375 (38.26)	608,019 (37.13)	586,979 (35.8)	
Mild	884,508 (53.97)	887,051 (54.1)	880,779 (53.78)	852,384 (51.99)	
Heavy	104,200 (6.36)	125,328 (7.64)	148,927 (9.09)	200,169 (12.21)	
Regular exercise, n (%)	244,252 (14.9)	212,375 (12.95)	198,244 (12.1)	187,284 (11.42)	<.001
Low income, n (%)	363,926 (22.21)	342,111 (20.86)	335,569 (20.49)	352,120 (21.48)	<.001
Hypertension, n (%)	57,896 (3.53)	86,336 (5.27)	124,874 (7.62)	212,860 (12.98)	<.001
Diabetes, n (%)	3085 (0.19)	7909 (0.48)	17,994 (1.1)	98,135 (5.99)	<.001
Dyslipidemia, n (%)	25,780 (1.57)	57,167 (3.49)	109,247 (6.67)	254,963 (15.55)	<.001
Diabetes medication, n (%)	1496 (0.09)	2292 (0.14)	4368 (0.27)	20,832 (1.27)	<.001
Dyslipidemia medication, n (%)	3686 (0.22)	7218 (0.44)	13,155 (0.8)	33,633 (2.05)	<.001
CKD[c], n (%)	43,725 (2.67)	41,914 (2.56)	42,859 (2.62)	47,590 (2.9)	<.001
Waist circumference (cm), mean (SD)	74.12 (8.39)	76.06 (9.29)	78.17 (10.03)	81.75 (10.55)	<.001
SBP[d], mmHg, mean (SD)	114.99 (12.13)	116.52 (12.58)	118.2 (13.1)	121.21 (14.03)	<.001
DBP[e], mmHg, mean (SD)	71.71 (8.68)	72.92 (9)	74.16 (9.38)	76.32 (10.07)	<.001
Fasting glucose (mg/dL), mean (SD)	85.07 (9.47)	88.62 (10.14)	91.16 (11.44)	98.71 (26.15)	<.001
Total cholesterol (mg/dL), mean (SD)	169.55 (27.9)	179.54 (29.72)	187.9 (31.95)	201.35 (36.74)	<.001

XSL•FO
RenderX

Variable	Quartile 1 (n= 1,638,852)	Quartile 2 (n= 1,639,754)	Quartile 3 (n= 1,637,725)	Quartile 4 (n= 1,639,532)	*P* value
HDL-C[f] (mg/dL), mean (SD)	61.3 (16.18)	58.68 (17.17)	56.06 (17.76)	53.14 (31.87)	<.001
LDL-C[g] (mg/dL), mean (SD)	97.55 (28.41)	104.67 (31.25)	108.84 (33.18)	107.78 (42.38)	<.001
eGFR[h] (mL/min/1.73 m^2), mean (SD)	98.03 (55.31)	96.4 (50.47)	95.26 (46.46)	94.67 (44.64)	<.001
Triglyceride (mg/dL), median (IQR)	51.45 (51.43-51.47)	79.15 (79.12-79.18)	111.27 (111.22-111.32)	194.1 (193.96-194.24)	<.001

[a]TyG: triglyceride-glucose.

[b]Not available.

[c]CKD: chronic kidney disease.

[d]SBP: systolic blood pressure.

[e]DBP: diastolic blood pressure.

[f]HDL-C: high-density lipoprotein-cholesterol.

[g]LDL-C: low-density lipoprotein-cholesterol.

[h]eGFR: estimated glomerular filtration rate.

Incidence Rates, Absolute Incidence Difference, Cumulative Incidence, and HRs of ESKD by TyG Index, Fasting Glucose, and Triglyceride Quartiles

The incidence rates, absolute incidence difference, 5- and 10-year cumulative incidence, and HRs of ESKD were analyzed across quartiles of the TyG index, fasting glucose, and triglycerides (Table 2). The incidence of ESKD increased progressively with higher TyG index quartiles (Figure 1). In both the unadjusted and adjusted models, the HRs of ESKD also increased significantly with increasing TyG index quartiles (all *P*<.001). In the fully adjusted Model 3, individuals in the highest TyG index quartile had an HR of 3.438 (95% CI 3.096 - 3.818) compared with those in the lowest quartile.

Regarding fasting glucose, ESKD incidence rates and HRs were significantly higher in the highest quartile than in the lower quartiles (*P*<.001). The HR for ESKD in the highest fasting glucose quartile was 1.393 (95% CI 1.284 - 1.511) compared with that in the lowest quartile in fully adjusted Model 3. However, the second (HR 0.836, 95% CI 0.761 - 0.918) and third (HR 0.891, 95% CI 0.810 - 0.979) quartiles showed significantly lower HRs compared to the lowest quartile group in adjusted Model 3.

For triglycerides, ESKD incidence rates and HRs showed patterns similar to those of the TyG index. The incidence rates increased progressively with the triglyceride quartiles. HRs for ESKD also increased significantly with increasing the triglyceride quartiles (*P*<.001). In the fully adjusted Model 3, the HR of the highest triglyceride quartile was 2.675 (95% CI 2.423 - 2.953) compared to that of the lowest quartile. The second (HR 1.463, 95% CI 1.314 - 1.628) and third (HR 1.693, 95% CI 1.527 - 1.876) quartiles also showed significantly higher HRs than the lowest quartile in adjusted Model 3.

XSL•FO
RenderX

Table . Incidence rates, absolute incidence difference, cumulative incidence, and adjusted hazard ratios of end-stage kidney disease according to quartiles of the triglyceride-glucose index, fasting glucose, and triglyceride levels.

TyG[a] index quartile	N	Event, n	Duration, PY[b]	IR[c] per 1000 PY	Absolute incidence difference (per 1000 PY)	Cumulative incidence (%)		HR[d] (95% CI)		
						5-year (95% CI)	10-year (95% CI)	Model 1[e]	Model 2[f]	Model 3[g]
Quartile 1	1,638,852	479	16,846,725.4	0.0284	0 (ref)	0.011 (0.010-0.013)	0.028 (0.025-0.030)	1 (ref)	1 (ref)	1 (ref)
Quartile 2	1,639,754	711	16,949,275.7	0.0420	0.013 (0.010 to 0.018)	0.017 (0.015-0.019)	0.042 (0.039-0.045)	1.472 (1.311-1.653)	1.425 (1.269-1.600)	1.400 (1.246-1.573)
Quartile 3	1,637,725	1014	16,976,870.8	0.0597	0.031 (0.027 to 0.036)	0.025 (0.023-0.028)	0.060 (0.056-0.064)	2.095 (1.879-2.335)	1.980 (1.775-2.209)	1.814 (1.624-2.026)
Quartile 4	1,639,532	2706	1,7002,806.5	0.1592	0.131 (0.124 to 0.137)	0.060 (0.057-0.064)	0.156 (0.150-0.162)	5.578 (5.062-6.148)	5.127 (4.646-5.658)	3.438 (3.096-3.818)
P value	—[h]	—	—	—	—	—	—	<.001	<.001	<.001
Fasting glucose										
Quartile 1	1,671,235	925	17,343,951.1	0.0533	0 (ref)	0.023 (0.021-0.026)	0.053 (0.049-0.056)	1 (ref)	1 (ref)	1 (ref)
Quartile 2	1,605,884	801	16,569,731.7	0.0483	−0.005 (−0.010 to 0.000)	0.021 (0.019-0.023)	0.049 (0.045-0.052)	0.908 (0.826-0.998)	0.883 (0.804-0.971)	0.891 (0.810-0.979)
Quartile 3	1,646,701	852	16,982,610.7	0.0502	−0.003 (−0.008 to 0.002)	0.022 (0.020-0.024)	0.050 (0.047-0.054)	0.943 (0.859-1.035)	0.886 (0.807-0.973)	0.836 (0.761-0.918)
Quartile 4	1,632,043	2332	16,879,384.8	0.1382	0.085 (0.078 to 0.091)	0.048 (0.045-0.052)	0.135 (0.129-0.141)	2.593 (2.403-2.799)	2.385 (2.209-2.576)	1.393 (1.284-1.511)
P value	—	—	—	—	—	—	—	<.001	<.001	<.001
Triglyceride										
Quartile 1	1,672,748	560	17,186,565.3	0.0326	0 (ref)	0.012 (0.011-0.014)	0.032 (0.029-0.035)	1 (ref)	1 (ref)	1 (ref)
Quartile 2	1,591,981	852	16,452,225.3	0.0518	0.019 (0.015 to 0.024)	0.019 (0.017-0.022)	0.051 (0.047-0.054)	1.586 (1.426-1.764)	1.521 (1.367-1.693)	1.463 (1.314-1.628)
Quartile 3	1,651,795	1154	17,122,569.9	0.0674	0.035 (0.030 to 0.040)	0.025 (0.023-0.028)	0.067 (0.063-0.072)	2.062 (1.864-2.281)	1.929 (1.743-2.135)	1.693 (1.527-1.876)
Quartile 4	1,639,339	2344	17,014,317.9	0.1378	0.105 (0.099 to 0.111)	0.057 (0.054-0.061)	0.136 (0.130-0.142)	4.212 (3.841-4.619)	3.830 (3.488-4.204)	2.675 (2.423-2.953)
P value	—	—	—	—	—	—	—	<.001	<.001	<.001

[a]TyG: triglyceride-glucose.

[b]PY: person-years.

[c]IR: incidence rate.

[d]HR: hazard ratio.

[e]Model 1: nonadjusted.

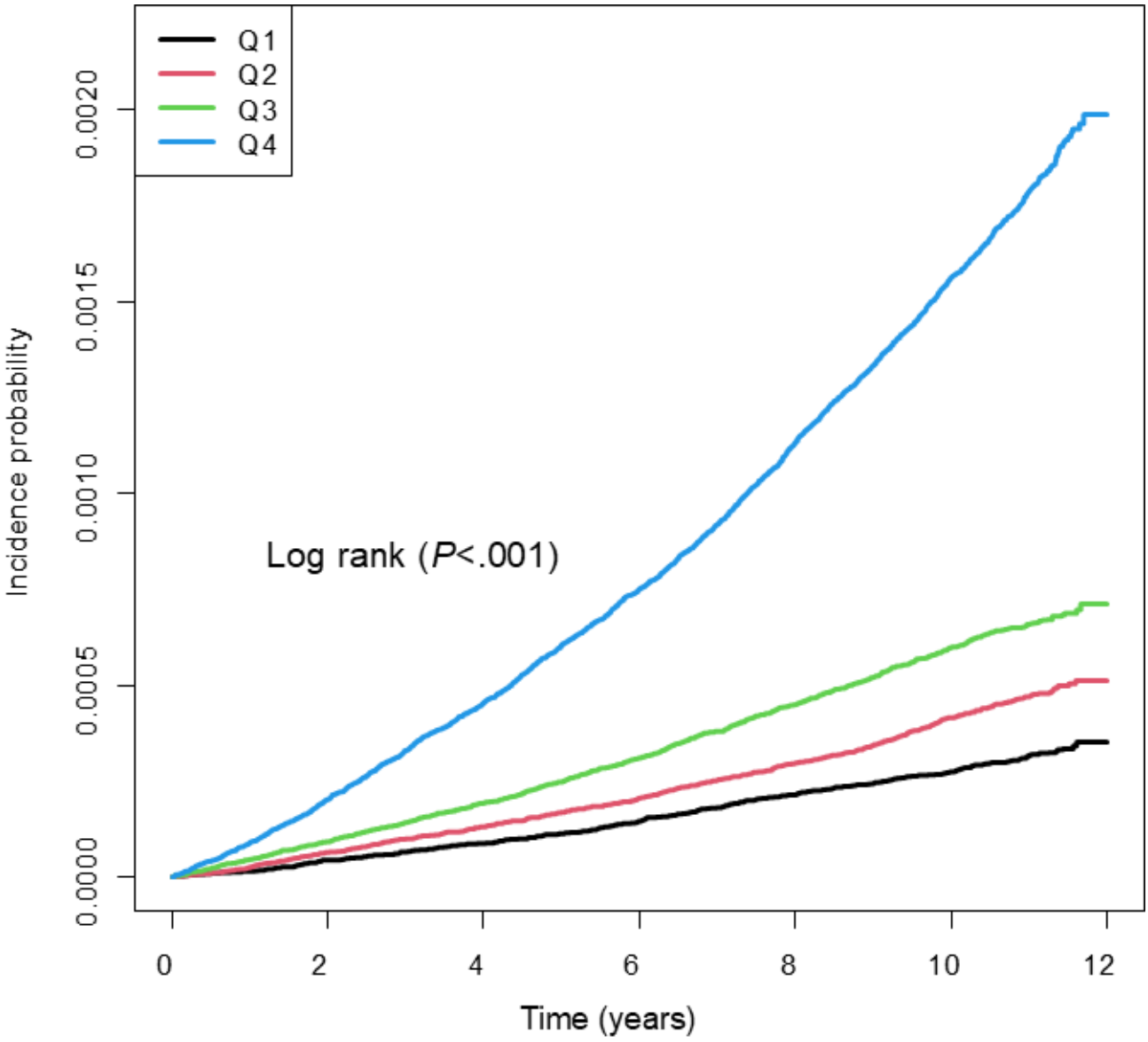
XSL•FO
RenderX

[f]Model 2: adjusted for age and sex.

[g]Model 3: adjusted for age, sex, BMI, smoking, drinking, regular exercise, diabetes medication, dyslipidemia medication, hypertension, low income, estimated glomerular filtration rate, and chronic kidney disease.

[h]Not available.

Figure 1. Cumulative incidence of end-stage kidney disease by the triglyceride-glucose (TyG) index quartiles: Kaplan-Meier analysis. Q: quartile.



Incidence Rates and HRs of ESKD by Deciles of the TyG Index, Fasting Glucose, and Triglycerides

ESKD incidence rates and HRs were assessed across deciles of the TyG index, fasting glucose, and triglyceride (Table 3). The incidence rates of ESKD increased progressively with increasing TyG index deciles. Similarly, HRs for ESKD showed a significant upward trend as TyG index deciles increased in both the unadjusted and adjusted models (all *P*<.001). In the fully adjusted Model 3, individuals in the highest TyG index decile had an HR of 5.944 (95% CI 5.022 - 7.035) compared to those in the lowest decile.

Regarding fasting glucose, the incidence rates and HRs of ESKD were significantly higher in the highest decile than in the lower deciles (*P*<.001). In the fully adjusted Model 3, the HR for ESKD in the highest fasting glucose decile was 2.029 (95% CI 1.805 - 2.282) compared to the lowest decile. However, the second through ninth deciles did not show statistically significant HRs compared to the lowest decile in the same model.

For triglycerides, ESKD incidence rates and HRs showed similar patterns to those of the TyG index. The incidence rates of ESKD increased progressively with increasing triglyceride deciles. Similarly, HRs for ESKD gradually increased with higher triglyceride deciles in both the unadjusted and adjusted models (all *P*<.001). In the fully adjusted Model 3, the HR for ESKD in the highest triglyceride decile was 4.042 (95% CI 3.433 - 4.759) compared to that in the lowest triglyceride decile.

XSL•FO
RenderX

Table . Incidence rates and adjusted hazard ratios of end-stage kidney disease according to deciles of the triglyceride-glucose index, fasting glucose, and triglyceride levels.

TyG[a] index decile	N	Event, n	Duration, PY[b]	IR[c] per 1000 PY	HR[d] (95% CI)		
					Model 1[e]	Model 2[f]	Model 3[g]
D1[h]	656,042	163	6,726,362.6	0.02	1 (ref)	1 (ref)	1 (ref)
D2	655,996	202	6,751,405.2	0.03	1.233 (1.003-1.515)	1.210 (0.984-1.487)	1.239 (1.008-1.524)
D3	655,258	235	6,757,963.2	0.03	1.432 (1.173-1.749)	1.390 (1.138-1.698)	1.425 (1.166-1.740)
D4	655,066	293	6,768,960.3	0.04	1.781 (1.471-2.157)	1.713 (1.414-2.075)	1.729 (1.427-2.095)
D5	656,244	297	6,791,309.8	0.04	1.799 (1.486-2.177)	1.716 (1.417-2.078)	1.715 (1.416-2.078)
D6	654,802	365	6,782,423.6	0.05	2.213 (1.839-2.661)	2.095 (1.741-2.521)	2.048 (1.700-2.466)
D7	655,955	395	6,801,963.6	0.06	2.387 (1.989-2.864)	2.242 (1.867-2.692)	2.135 (1.775-2.567)
D8	655,728	518	6,803,949.5	0.08	3.128 (2.623-3.730)	2.915 (2.443-3.479)	2.634 (2.203-3.149)
D9	655,160	652	6,797,631.0	0.10	3.941 (3.319-4.679)	3.639 (3.062-4.325)	3.038 (2.549-3.622)
D10	655,612	1790	6,793,709.5	0.26	10.827 (9.223-12.710)	9.850 (8.378-11.581)	5.944 (5.022-7.035)
P value	__[i]	—	—	—	<.001	<.001	<.001
Fasting glucose							
D1	676,873	395	7,052,603.2	0.06	1 (ref)	1 (ref)	1 (ref)
D2	604,458	337	6,263,814.7	0.05	0.963 (0.832-1.113)	0.971 (0.840-1.123)	1.014 (0.877-1.173)
D3	739,028	367	7,642,742.4	0.05	0.860 (0.746-0.991)	0.838 (0.727-0.966)	0.886 (0.769-1.022)
D4	621,214	331	6,413,596.6	0.05	0.925 (0.799-1.070)	0.883 (0.763-1.022)	0.923 (0.797-1.068)
D5	635,546	296	6,540,926.0	0.05	0.812 (0.698-0.944)	0.839 (0.722-0.976)	0.836 (0.718-0.972)
D6	646,017	320	6,659,217.4	0.05	0.861 (0.743-0.998)	0.817 (0.705-0.947)	0.816 (0.704-0.946)
D7	688,862	355	7,100,424.4	0.05	0.896 (0.777-1.034)	0.874 (0.757-1.009)	0.838 (0.725-0.967)
D8	658,165	382	6,796,143.4	0.06	1.007 (0.875-1.159)	0.943 (0.819-1.086)	0.866 (0.751-0.997)
D9	664,302	429	6,874,473.0	0.06	1.117 (0.974-1.281)	1.052 (0.918-1.207)	0.876 (0.763-1.006)
D10	621,398	1698	6,431,737.3	0.26	4.724 (4.234-5.270)	4.305 (3.855-4.808)	2.029 (1.805-2.282)
P value	—	—	—	—	<.001	<.001	<.001
Triglyceride							
D1	645,990	178	6,616,922.9	0.03	1 (ref)	1 (ref)	1 (ref)
D2	688,090	242	7,076,082.7	0.03	1.270 (1.046-1.541)	1.238 (1.020-1.502)	1.262 (1.040-1.531)

XSL•FO
RenderX

TyG[a] index decile	N	Event, n	Duration, PY[b]	IR[c] per 1000 PY	HR[d] (95% CI)		
					Model 1[e]	Model 2[f]	Model 3[g]
D3	638,511	266	6,586,656.6	0.04	1.497 (1.238-1.810)	1.440 (1.191-1.741)	1.440 (1.191-1.741)
D4	657,214	350	6,791,725.6	0.05	1.909 (1.593-2.286)	1.787 (1.492-2.141)	1.754 (1.464-2.102)
D5	634,924	376	6,567,402.8	0.06	2.120 (1.773-2.533)	1.996 (1.669-2.386)	1.918 (1.603-2.294)
D6	675,312	429	6,992,308.3	0.06	2.271 (1.907-2.704)	2.126 (1.784-2.532)	1.981 (1.661-2.362)
D7	647,027	457	6,711,946.2	0.07	2.518 (2.118-2.994)	2.294 (1.928-2.729)	2.034 (1.707-2.424)
D8	659,743	586	6,845,357.4	0.09	3.165 (2.677-3.743)	2.897 (2.448-3.428)	2.424 (2.044-2.874)
D9	653,633	713	6,784,455.6	0.11	3.886 (3.298-4.580)	3.490 (2.959-4.117)	2.732 (2.309-3.233)
D10	655,419	1313	6,802,820.2	0.19	7.138 (6.104-8.347)	6.331 (5.407-7.412)	4.042 (3.433-4.759)
P value	—	—	—	—	<.001	<.001	<.001

[a]TyG: triglyceride-glucose.

[b]PY: person-years.

[c]IR: incidence rate.

[d]HR: hazard ratio.

[e]Model 1: nonadjusted.

[f]Model 2: adjusted for age and sex.

[g]Model 3: adjusted for age, sex, BMI, smoking, drinking, regular exercise, diabetes medication, dyslipidemia medication, hypertension, low income, estimated glomerular filtration rate, and chronic kidney disease.

[h]D: decile.

[i]Not available.

Subgroup Analyses of Incidence Rates and HRs of ESKD by TyG Index Quartiles

The incidence rates and adjusted hazard ratios (aHRs) of ESKD were evaluated across TyG index quartiles within various subgroups (Table 4). Interaction tests revealed that age, sex, obesity, MetS, diabetes, use of medications for diabetes or dyslipidemia, dyslipidemia, low income, eGFR, and CKD significantly influenced the association between the TyG index and ESKD (all *P* for interaction<.0001).

In subgroups stratified by age (20 - 29 years and 30 - 39 years), sex (male and female), obesity status (with or without), and CKD status (with or without), both the incidence rates and aHRs for ESKD significantly increased with higher TyG index quartiles. For participants without MetS, diabetes, or dyslipidemia and those not receiving medications for diabetes or dyslipidemia, the incidence rates and aHRs of ESKD also rose consistently across increasing TyG index quartiles.

In participants with dyslipidemia, the aHR for ESKD was significantly higher in the highest TyG index quartile than in the lowest quartile; however, the aHRs for the second and third quartiles were not significant. Among participants with diabetes, the aHRs were lower in the second and third quartiles, and the highest quartile did not show a significant difference compared with the first quartile. For participants with MetS and those receiving medications for diabetes or dyslipidemia, the aHRs for ESKD were not significantly different between the TyG index quartiles.

XSL•FO
RenderX

Table . Adjusted hazard ratios of end-stage kidney disease according to various subgroups. Adjusted for age, sex, BMI, smoking, drinking, regular exercise, diabetes medication, dyslipidemia medication, hypertension, low income, estimated glomerular filtration rate (eGFR), and chronic kidney disease (CKD).

TyG[a] index	Number of patients	Event, n	Duration, PY[b]	IR[c] per 1000 PY	HR[d] (95% CI)	*P* for interaction
Age (years)						<.001
<30						
Quartile 1	894,184	214	9,102,734.5	0.0235	1 (ref)	
Quartile 2	743,906	238	7,604,879.7	0.0313	1.328 (1.104-1.597)	
Quartile 3	630,104	275	6,460,758.7	0.0426	1.763 (1.473-2.109)	
Quartile 4	487,580	701	5,011,557.2	0.1399	5.038 (4.311-5.888)	
≥30						
Quartile 1	744,668	265	7,743,990.9	0.0342	1 (ref)	
Quartile 2	895,848	473	9,344,396.1	0.0506	1.323 (1.157-1.513)	
Quartile 3	1,007,621	739	10,516,112.1	0.0703	1.562 (1.373-1.776)	
Quartile 4	1,151,952	2005	11,991,249.3	0.1672	2.703 (2.397-3.049)	
Sex						<.001
Male						
Quartile 1	975,747	366	10,059,535.3	0.0364	1 (ref)	
Quartile 2	975,000	529	10,135,210.0	0.0522	1.323 (1.157-1.513)	
Quartile 3	975,117	713	10,175,276.9	0.0701	1.562 (1.373-1.776)	
Quartile 4	975,291	1828	10,189,244.3	0.1794	2.703 (2.397-3.049)	
Female						
Quartile 1	663,105	113	6,787,190.0	0.0167	1 (ref)	
Quartile 2	664,754	182	6,814,065.7	0.0267	1.601 (1.266-2.025)	
Quartile 3	662,608	301	6,801,593.9	0.0443	2.582 (2.079-3.206)	
Quartile 4	664,241	878	6,813,562.2	0.1289	5.947 (4.876-7.253)	
Obesity						<.001
No						
Quartile 1	1,439,264	409	14,800,776.4	0.0276	1 (ref)	
Quartile 2	1,319,486	513	13,641,935.1	0.0376	1.292 (1.134-1.471)	
Quartile 3	1,164,343	662	12,070,099.3	0.0549	1.765 (1.559-1.998)	
Quartile 4	895,059	1326	9,288,242.4	0.1428	3.679 (3.284-4.123)	
Yes						
Quartile 1	199,588	70	2,045,949.0	0.0342	1 (ref)	
Quartile 2	320,268	198	3,307,340.7	0.0599	1.506 (1.147-1.978)	
Quartile 3	473,382	352	4,906,771.5	0.0717	1.508 (1.166-1.950)	
Quartile 4	744,473	1380	7,714,564.1	0.1789	2.425 (1.902-3.090)	
Metabolic syndrome						<.001
No						
Quartile 1	1,629,671	430	16,752,762.2	0.0257	1 (ref)	
Quartile 2	1,612,113	600	16,664,157.8	0.0360	1.411 (1.246-1.598)	
Quartile 3	1,513,148	707	15,683,481.0	0.0451	1.789 (1.585-2.019)	
Quartile 4	1,099,603	1008	11,397,312.7	0.0884	3.677 (3.274-4.129)	
Yes						

XSL•FO
RenderX

TyG[a] index	Number of patients	Event, n	Duration, PY[b]	IR[c] per 1000 PY	HR[d] (95% CI)	*P* for interaction
Quartile 1	9181	49	93,963.2	0.5215	1 (ref)	
Quartile 2	27,641	111	285,117.9	0.3893	0.920 (0.660-1.283)	
Quartile 3	124,577	307	1,293,389.8	0.2374	0.845 (0.627-1.139)	
Quartile 4	539,929	1698	5,605,493.8	0.3029	1.135 (0.856-1.505)	
Diabetes						<.001
No						
Quartile 1	1,635,767	435	16,814,934.1	0.0259	1 (ref)	
Quartile 2	1,631,845	649	16,867,366.5	0.0385	1.453 (1.287-1.640)	
Quartile 3	1,619,731	887	16,790,386.1	0.0528	1.914 (1.705-2.149)	
Quartile 4	1,541,397	1471	15,993,541.3	0.0920	3.020 (2.701-3.377)	
Yes						
Quartile 1	3085	44	31,791.2	1.3840	1 (ref)	
Quartile 2	7909	62	81,909.2	0.7569	0.650 (0.434-0.975)	
Quartile 3	17,994	127	186,484.7	0.6810	0.613 (0.427-0.880)	
Quartile 4	98,135	1235	1,009,265.2	1.2237	1.163 (0.843-1.605)	
Diabetes medication						<.001
No						
Quartile 1	1,637,356	438	16,831,389.4	0.0260	1 (ref)	
Quartile 2	1,637,462	657	16,925,781.9	0.0388	1.438 (1.274-1.622)	
Quartile 3	1,633,357	917	16,932,102.4	0.0542	1.892 (1.686-2.122)	
Quartile 4	1,618,700	2043	16,790,857.2	0.1217	3.759 (3.373-4.190)	
Yes						
Quartile 1	1496	41	15,336.0	2.6735	1 (ref)	
Quartile 2	2292	54	23,493.8	2.2985	0.692 (0.450-1.065)	
Quartile 3	4368	97	44,768.4	2.1667	0.617 (0.419-0.910)	
Quartile 4	20,832	663	211,949.3	3.1281	0.884 (0.631-1.240)	
Dyslipidemia						<.001
No						
Quartile 1	1,613,072	414	16,580,324.7	0.0250	1 (ref)	
Quartile 2	1,582,587	568	16,355,863.2	0.0347	1.384 (1.219-1.571)	
Quartile 3	1,528,478	728	15,840,855.1	0.0460	1.781 (1.577-2.012)	
Quartile 4	1,384,569	1476	14,358,412.0	0.1028	3.456 (3.084-3.874)	
Yes						
Quartile 1	25,780	65	266,400.6	0.2440	1 (ref)	
Quartile 2	57,167	143	593,412.5	0.2410	1.013 (0.753-1.363)	
Quartile 3	109,247	286	1,136,015.7	0.2518	1.021 (0.777-1.342)	
Quartile 4	254,963	1230	2,644,394.5	0.4651	1.486 (1.153-1.916)	
Dyslipidemia medication						<.001
No						
Quartile 1	1,613,072	414	16,580,324.7	0.0250	1 (ref)	
Quartile 2	1,582,587	568	16,355,863.2	0.0347	1.403 (1.241-1.587)	
Quartile 3	1,528,478	728	15,840,855.1	0.0460	1.798 (1.599-2.023)	

XSL•FO
RenderX

TyG[a] index	Number of patients	Event, n	Duration, PY[b]	IR[c] per 1000 PY	HR[d] (95% CI)	*P* for interaction
Quartile 4	1,384,569	1476	14,358,412.0	0.1028	3.992 (3.580-4.452)	
Yes						
Quartile 1	25,780	65	266,400.6	0.2440	1 (ref)	
Quartile 2	57,167	143	593,412.5	0.2410	0.926 (0.647-1.325)	
Quartile 3	109,247	286	1,136,015.7	0.2518	0.963 (0.695-1.335)	
Quartile 4	254,963	1230	2,644,394.5	0.4651	0.799 (0.589-1.084)	
CKD						<.001
No						
Quartile 1	1,595,127	333	16,368,449.1	0.0203	1 (ref)	
Quartile 2	1,597,840	453	16,493,096.1	0.0275	1.301 (1.129-1.500)	
Quartile 3	1,594,866	586	16,511,679.3	0.0355	1.564 (1.365-1.792)	
Quartile 4	1,591,942	1891	16,489,958.6	0.1147	3.703 (3.275-4.187)	
Yes						
Quartile 1	43,725	146	478,276.3	0.3053	1 (ref)	
Quartile 2	41,914	258	456,179.6	0.5656	1.721 (1.404-2.109)	
Quartile 3	42,859	428	465,191.5	0.9201	2.429 (2.010-2.935)	
Quartile 4	47,590	815	512,847.9	1.5892	2.909 (2.428-3.486)	

[a]TyG: triglyceride-glucose.

[b]PY: person-years.

[c]IR: incidence rate.

[d]HR: hazard ratio.

Sensitivity Analysis of Incidence Rates and HRs of ESKD by Follow-Up and Time-Averaged TyG Index Quartiles

Sensitivity analysis was performed using TyG index quartiles at 2-year follow-up and time-averaged values from baseline and 2-year follow-up (Table 5). The incidence rates and aHRs of ESKD across TyG index quartiles at 2-year follow-up and for time-averaged values showed findings consistent with those of the baseline TyG index quartiles.

The incidence of ESKD increased progressively with higher TyG index quartiles at 2-year follow-up. In both the unadjusted and adjusted models, the HRs of ESKD also increased significantly with increasing TyG index quartiles (all *P*<.001). In the fully adjusted Model 3, individuals in the highest TyG index quartile had an HR of 2.860 (95% CI 2.466 - 3.317) compared with those in the lowest quartile.

Regarding fasting glucose at 2-year follow-up, ESKD incidence rates and HRs were significantly higher in the highest quartile than in the lower quartiles (*P*<.001). The HR for ESKD in the highest fasting glucose quartile was 1.306 (95% CI 1.156 - 1.475) compared to that in the lowest quartile in the fully adjusted Model 3. However, the second (HR 0.845, 95% CI 0.737 - 0.973) and third (HR 0.838, 95% CI 0.729 - 0.963) quartiles showed significantly lower HRs compared with the lowest quartile group in adjusted Model 3.

For triglycerides at 2-year follow-up, ESKD incidence rates and HRs showed patterns similar to those of the TyG index. The incidence rates increased progressively with the triglyceride quartiles. HRs for ESKD also increased significantly with increasing the triglyceride quartiles (*P*<.001). In the fully adjusted Model 3, the HR of the highest triglyceride quartile was 2.174 (95% CI 1.884 - 2.508) compared to that of the lowest quartile. The second (HR 1.253, 95% CI 1.074 - 1.461) and third (HR 1.445, 95% CI 1.246 - 1.677) quartiles also showed significantly higher HRs than the lowest quartile in adjusted Model 3.

The incidence of ESKD increased progressively with higher time-averaged TyG index quartiles. In both the unadjusted and adjusted models, the HRs of ESKD also increased significantly with increasing TyG index quartiles (all *P*<.001). In the fully adjusted Model 3, individuals in the highest TyG index quartile had an HR of 3.449 (95% CI 2.937 - 4.050) compared with those in the lowest quartile.

Regarding time-averaged fasting glucose, ESKD incidence rates and HRs were significantly higher in the highest quartile than in the lower quartiles (*P*<.001). The HR for ESKD in the highest fasting glucose quartile was 1.565 (95% CI 1.389 - 1.763) compared to that in the lowest quartile in fully adjusted Model 3. However, the second (HR 0.906, 95% CI 0.788 - 1.042) and third (HR 0.770, 95% CI 0.669 - 0.886) quartiles showed significantly lower HRs compared with the lowest quartile group in adjusted Model 3.

For time-averaged triglycerides, ESKD incidence rates and HRs showed patterns similar to those of the TyG index. The incidence rates increased progressively with higher triglyceride quartiles.

XSL•FO
RenderX

HRs for ESKD also increased significantly with increasing the triglyceride quartiles (*P*<.001). In fully adjusted Model 3, the HR of the highest triglyceride quartile was 2.092 (95% CI 1.810 - 2.417) compared to that of the lowest quartile. The second (HR 1.320, 95% CI 1.131 - 1.542) and third (HR 1.475, 95% CI 1.268 - 1.715) quartiles also showed significantly higher HRs than the lowest quartile in adjusted Model 3.

XSL•FO
RenderX

Table . Incidence rates and adjusted hazard ratios of end-stage kidney disease according to quartiles of follow-up and time-averaged values of the triglyceride-glucose index, fasting glucose, and triglyceride levels.

2-year follow-up TyG[a] index quartile	Number of patients	Event	Duration, PY[b]	IR[c] per 1000 PY	HR[d] (95% CI)		
					Model 1[e]	Model 2[f]	Model 3[g]
Quartile 1	966,738	247	8,117,937.8	0.03043	1 (ref)	1 (ref)	1 (ref)
Quartile 2	965,852	331	8,155,391.1	0.04059	1.332 (1.130-1.571)	1.288 (1.092-1.519)	1.188 (1.007-1.402)
Quartile 3	966,636	460	8,190,218.9	0.05616	1.842 (1.578-2.150)	1.737 (1.487-2.028)	1.428 (1.219-1.672)
Quartile 4	966,389	1145	8,190,761.2	0.13979	4.584 (3.995-5.260)	4.205 (3.661-4.831)	2.860 (2.466-3.317)
P value	__[h]	—	—	—	<.001	<.001	<.001
Fasting glucose							
Quartile 1	918,400	388	7,806,410.5	0.0497	1 (ref)	1 (ref)	1 (ref)
Quartile 2	983,768	389	8,315,230	0.04678	0.943 (0.819-1.085)	0.917 (0.797-1.056)	0.845 (0.734-0.973)
Quartile 3	952,913	419	8,028,657.7	0.05219	1.053 (0.917-1.209)	1.003 (0.874-1.152)	0.838 (0.729-0.963)
Quartile 4	1,010,534	987	8,504,010.9	0.11606	2.343 (2.083-2.635)	2.140 (1.901-2.409)	1.306 (1.156-1.475)
P value	—	—	—	—	<.001	<.001	<.001
Triglyceride							
Quartile 1	966,138	278	8,103,153.6	0.03431	1 (ref)	1 (ref)	1 (ref)
Quartile 2	976,553	402	8,243,730.1	0.04876	1.419 (1.218-1.654)	1.367 (1.173-1.593)	1.253 (1.074-1.461)
Quartile 3	953,353	527	8,078,334.6	0.06524	1.897 (1.640-2.193)	1.779 (1.538-2.058)	1.445 (1.246-1.677)
Quartile 4	969,571	976	8,229,090.8	0.1186	3.447 (3.017-3.938)	3.163 (2.766-3.617)	2.174 (1.884-2.508)
P value	—	—	—	—	<.001	<.001	<.001
Time-average of baseline and 2-year follow-up							
TyG index							
Quartile 1	966,403	205	8,094,556.1	0.02533	1 (ref)	1 (ref)	1 (ref)
Quartile 2	966,403	329	8,157,710.3	0.04033	1.589 (1.335-1.892)	1.536 (1.290-1.829)	1.432 (1.201-1.706)
Quartile 3	966,405	444	8,196,304.3	0.05417	2.133 (1.808-2.517)	2.011 (1.703-2.375)	1.685 (1.423-1.996)
Quartile 4	966,404	1205	8,205,738.4	0.14685	5.779 (4.983-6.701)	5.287 (4.550-6.143)	3.449 (2.937-4.050)
P value	—	—	—	—	<.001	<.001	<.001
Fasting glucose							
Quartile 1	982,056	404	8,324,979.1	0.04853	1 (ref)	1 (ref)	1 (ref)
Quartile 2	956,138	383	8,049,552.7	0.04758	0.983 (0.854-1.130)	0.955 (0.831-1.099)	0.906 (0.788-1.042)
Quartile 3	982,130	383	8,271,683.5	0.0463	0.956 (0.832-1.100)	0.900 (0.782-1.035)	0.770 (0.669-0.886)

XSL•FO
RenderX

2-year follow-up TyG[a] index quartile	Number of patients	Event	Duration, PY[b]	IR[c] per 1000 PY	HR[d] (95% CI)		
					Model 1[e]	Model 2[f]	Model 3[g]
Quartile 4	945,291	1013	8,008,093.7	0.1265	2.608 (2.324-2.927)	2.399 (2.136-2.696)	1.565 (1.389-1.763)
P value	—	—	—	—	<.001	<.001	<.001
Triglyceride							
Quartile 1	968,392	268	8,108,399.1	0.03305	1 (ref)	1 (ref)	1 (ref)
Quartile 2	975,884	407	8,232,289.3	0.04944	1.493 (1.280-1.742)	1.442 (1.236-1.683)	1.320 (1.131-1.542)
Quartile 3	953,357	511	8,082,020.2	0.06323	1.907 (1.645-2.211)	1.783 (1.537-2.068)	1.475 (1.268-1.715)
Quartile 4	967,982	997	8,231,600.4	0.12112	3.650 (3.189-4.177)	3.307 (2.885-3.791)	2.092 (1.810-2.417)
P value	—	—	—	—	<.001	<.001	<.001

[a]TyG: triglyceride-glucose.

[b]PY: person-years.

[c]IR: incidence rate.

[d]HR: hazard ratio.

[e]Model 1: nonadjusted.

[f]Model 2: adjusted for age and sex.

[g]Model 3: adjusted for age, sex, BMI, smoking, drinking, regular exercise, diabetes medication, dyslipidemia medication, hypertension, low income, estimated glomerular filtration rate, and chronic kidney disease.

[h]Not available.

Discussion

Principal Findings

This observational, population-based cohort study found that a higher TyG index was significantly associated with an increased risk of incident ESKD in young adults. This association was independent of traditional CKD risk factors and remained robust across subgroups based on age, sex, obesity, dyslipidemia, and CKD. These findings suggest that the TyG index, a marker of insulin resistance, may be useful for identifying young adults at risk of developing ESKD, and early intervention targeting insulin resistance may mitigate the pathogenesis of ESKD in young adults.

CKD is a leading global cause of mortality, affecting over 10% of the general population owing to the rising risk factors such as hypertension, diabetes, and obesity [1]. CKD progression to ESKD, which requires kidney replacement therapy, represents a significant global health issue. According to the International Society of Nephrology's 2023 Global Kidney Health Atlas, the median incidence of treated ESKD was 146 individuals per million population in 84 countries, while the median prevalence of treated ESKD was 823 per million population in 95 countries [22]. Although CKD and ESKD are more prevalent in older adults, the global incidence of CKD is increasing in adolescents and young adults [3], who experience 30 times higher mortality rates than their healthy peers [23]. They face accumulating physiological complications and socioeconomic burdens from early life, predisposing them to poor long-term outcomes [7,24]. Therefore, identifying CKD risk factors early and implementing prompt interventions are critical for reducing the disease burden among young adults.

In this study, young adults with a higher TyG index were older, had a lower eGFR, and exhibited poorer metabolic health profiles than those with a lower TyG index. Moreover, they had higher BMI, waist circumference, blood pressure, and fasting plasma glucose, and poorer lipid profiles. Additionally, individuals with a higher TyG index showed higher rates of obesity, smoking, and alcohol consumption, as well as lower rates of regular exercise. They also had a greater prevalence of hypertension, diabetes, and dyslipidemia and were more likely to be on medications for diabetes or dyslipidemia than those with a lower TyG index. These findings align with previous studies [15,25,26]. During follow-up, ESKD incidence rates increased progressively with increasing TyG index, and the risk of ESKD significantly increased across both quartiles and deciles, independent of various ESKD risk factors. The individual contributions of the 2 TyG index components (ie, fasting plasma glucose and triglyceride levels) were separately investigated by analyzing the incidence rates and HRs across quartiles and deciles. Although fasting plasma glucose showed no significant effect, triglyceride levels demonstrated a weak effect in the decile-based analysis, which is consistent with a previous report [15]. Additionally, the impact of a high TyG index on ESKD pathogenesis was robust in subgroup analyses based on age, sex, obesity, dyslipidemia, and CKD. Sensitivity analyses using the 2-year follow-up TyG index and the time-averaged TyG index also showed significant associations

XSL•FO
RenderX

with ESKD development. These findings suggest that the TyG index serves as more than a simple sum of its components (fasting plasma glucose and triglyceride levels) and has the potential to be a marker for assessing ESKD risk in young adults.

Previous studies have shown that the TyG index mediates the relationship between BMI and ESKD risk in middle-aged adults [26] and in the general population [15]. Among adolescents and young adults, higher SBP, fasting plasma glucose levels, and BMI have been identified as important factors contributing to CKD risk [3]. In this study, an elevated TyG index was associated with incident ESKD in young adults with both normal and high BMI, suggesting that a high TyG index, reflective of insulin resistance, plays a role in ESKD development both in connection with BMI and independently. Insulin resistance is known to cause glomerular hyperfiltration, sodium retention, tubular dysfunction, kidney inflammation, and fibrosis [11], all of which may contribute to ESKD progression through pathways independent of BMI. Additionally, previous research has linked the TyG index to the progression of coronary artery calcification in patients with CKD [27], suggesting that insulin resistance may further exacerbate ESKD by promoting atherosclerosis and vascular calcification. Insulin resistance contributes to atherogenesis and atherosclerosis progression by inducing macrophage apoptosis and causing damage to endothelial and vascular smooth muscle cells [28,29]. Hyperinsulinemia also promotes oxidative stress, structural and functional vessel injury [30], and osteogenic differentiation resulting in vascular calcification [31]. These findings highlight the importance of managing metabolic factors, in addition to weight reduction, to mitigate the detrimental effects of insulin resistance on progressive kidney injury.

Unexpectedly, the subgroup analysis showed that the TyG index was not associated with the risk of ESKD in young adults with diabetes or MetS, or those on medications for diabetes or dyslipidemia. Recent research has linked the TyG index to an increased risk of ESKD in patients with diabetes [16]. The reason for the discrepancy in our study is unclear. We speculate that there are several possible explanations. First, the population in this study differed from that in previous research, as this analysis focused exclusively on young adults aged 20 - 39 years. Similarly, a recent study analyzing the National Health and Nutrition Examination Survey found no significant association between the TyG index and low eGFR (<60 mL/min/1.73 m^2) in individuals aged 20 - 40 years [25]. Second, among young adults with diabetes or those taking medications for diabetes, the association between the TyG index and ESKD may be weak. A previous study on CKD progression in type 2 diabetes suggested that the association between the TyG index and CKD was more pronounced in the early stages of the disease (eGFR >60 mL/min/1.73 m^2) and in individuals without a history of oral diabetes medications [32]. Other studies have reported that while the TyG index correlates significantly with microalbuminuria, its association with CKD [33] or eGFR [34] is less evident. These findings suggest that the TyG index has a stronger influence in the early stages of CKD but a diminished effect in later stages, such as ESKD. Insulin resistance, which is characterized by hyperglycemia and dysfunctional lipid oxidation and usage, contributes to early glomerular damage in patients with diabetes. However, as dysglycemia progresses, multifactorial processes, including inflammation, oxidative stress, and metabolic acidosis, exacerbate kidney damage [14,33]. Third, among young adults with MetS or those taking medications for dyslipidemia, MetS and dyslipidemia status can change over time due to lifestyle modifications, medications, or illness [35], potentially weakening the association between the TyG index and ESKD risk during follow-up. Additionally, as the 2 components of the TyG index (ie, fasting glucose and triglyceride levels) overlap with the 5 diagnostic criteria for MetS, this may diminish the effect of the TyG index in patients with MetS. Fourth, in high-risk populations with established metabolic abnormalities (eg, diabetes, MetS, or abnormal glucose metabolism), the relationship between the TyG index and risk of CKD may be nonlinear or modified by concomitant therapies. A restricted cubic spline analysis from a cohort of hypertensive patients with abnormal glucose metabolism demonstrated a U-shaped association between the TyG index and incident CKD (eGFR <60 mL/min/1.73 m^2 and/or proteinuria). The risk of CKD fell first and started to increase after the TyG index became higher than 8.94. Consistently, the risk of low eGFR also appeared to decrease first and then increase as the TyG quartile increased [36]. Taken together, these data partially support our finding that the TyG index predicted ESKD development mainly in metabolically healthier young adults, whereas its prognostic value was attenuated in those with preexisting diabetes or MetS, or those on medications for diabetes or dyslipidemia. Such patients have relatively higher baseline renal risk, and progression to ESKD is strongly influenced by disease duration, glycemic control, and renoprotective medications such as renin-angiotensin system inhibitors. Further studies are needed to elucidate the association between the TyG index and risk of ESKD in metabolically unhealthy young adults.

Limitations

This study had several limitations. First, its observational design and limited scope of the database prevented the inclusion of various confounding factors that contribute to the pathogenesis of ESKD, such as inflammation, nutrition, comorbidities, other kidney diseases, and additional metabolic risk factors. Particularly, the lack of consideration of genetic or congenital kidney diseases is a limitation in this cohort of young adults. Second, as the TyG index does not adequately reflect changes in response to drug treatment or lifestyle modification, the evidence supporting its role in the development of ESKD remains limited. Third, the main analysis did not adjust for other metabolic risk factors due to the lack of such information in the NHIS database. Fourth, there is no biological validation supporting the association between the TyG index and ESKD. Finally, the insignificant association between the TyG index and ESKD in subgroup analyses of young adults with diabetes, MetS, and those taking medications for diabetes or dyslipidemia is also a limitation that needs further research. Despite these limitations, the findings of our study remain significant, as the study was based on a large population-based cohort with a longitudinal database. The study population uniquely focused on young adults, a group that has been underrepresented in

XSL•FO
RenderX

ESKD research. Additionally, adjustments were made for established ESKD risk factors, eGFR, and income, and the TyG index was examined in quartiles and deciles, using follow-up and time-averaged values, enabling a comprehensive evaluation of its association with ESKD risk.

Conclusions

In conclusion, this study demonstrated that a higher TyG index is significantly associated with an increased risk of incident ESKD in young adults. These findings suggest that the TyG index may have potential utility in identifying young adults at risk of developing ESKD. Further comprehensive studies are warranted to confirm the relationship and clarify the role of the TyG index in the development of ESKD in young adult populations.

Acknowledgments

Generative artificial intelligence tools were not used.

Funding

This research was supported by a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health and Welfare, Republic of Korea (grant RS-2025-02215809). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflicts of Interest

None declared.

References

1. Kovesdy CP. Epidemiology of chronic kidney disease: an update 2022. Kidney Int Suppl (2011) 2022 Apr;12(1):7-11. [doi: 10.1016/j.kisu.2021.11.003]
2. Scurt FG, Ganz MJ, Herzog C, Bose K, Mertens PR, Chatzikyrkou C. Association of metabolic syndrome and chronic kidney disease. Obes Rev 2024 Jan;25(1):e13649. [doi: 10.1111/obr.13649] [Medline: 37783465]
3. Deng H, Zou Q, Chen Z, Hu B, Liao X. Global burden and risk factors of chronic kidney disease in adolescents and young adults: a study from 1990 to 2019. Kidney Res Clin Pract 2025 Jul;44(4):588-601. [doi: 10.23876/j.krcp.23.331] [Medline: 39034860]
4. Steinberger J, Daniels SR, Eckel RH, et al. Progress and challenges in metabolic syndrome in children and adolescents: a scientific statement from the American Heart Association Atherosclerosis, Hypertension, and Obesity in the Young Committee of the Council on Cardiovascular Disease in the Young; Council on Cardiovascular Nursing; and Council on Nutrition, Physical Activity, and Metabolism. Circulation 2009 Feb 3;119(4):628-647. [doi: 10.1161/CIRCULATIONAHA.108.191394] [Medline: 19139390]
5. Tsur AM, Akavian I, Landau R, et al. Adolescent body mass index and early chronic kidney disease in young adulthood. JAMA Pediatr 2024 Feb 1;178(2):142-150. [doi: 10.1001/jamapediatrics.2023.5420] [Medline: 38079159]
6. Kula AJ, Prince DK, Flynn JT, Bansal N. BP in young adults with CKD and associations with cardiovascular events and decline in kidney function. J Am Soc Nephrol 2021 May 3;32(5):1200-1209. [doi: 10.1681/ASN.2020081156] [Medline: 33692088]
7. Hamilton AJ, Clissold RL, Inward CD, Caskey FJ, Ben-Shlomo Y. Sociodemographic, psychologic health, and lifestyle outcomes in young adults on renal replacement therapy. Clin J Am Soc Nephrol 2017 Dec 7;12(12):1951-1961. [doi: 10.2215/CJN.04760517] [Medline: 29051144]
8. Spoto B, Pisano A, Zoccali C. Insulin resistance in chronic kidney disease: a systematic review. Am J Physiol Renal Physiol 2016 Dec 1;311(6):F1087-F1108. [doi: 10.1152/ajprenal.00340.2016] [Medline: 27707707]
9. Mykkänen L, Zaccaro DJ, Wagenknecht LE, Robbins DC, Gabriel M, Haffner SM. Microalbuminuria is associated with insulin resistance in nondiabetic subjects: the insulin resistance atherosclerosis study. Diabetes 1998 May;47(5):793-800. [doi: 10.2337/diabetes.47.5.793] [Medline: 9588452]
10. Chen J, Muntner P, Hamm LL, et al. The metabolic syndrome and chronic kidney disease in U.S. adults. Ann Intern Med 2004 Feb 3;140(3):167-174. [doi: 10.7326/0003-4819-140-3-200402030-00007] [Medline: 14757614]
11. Whaley-Connell A, Sowers JR. Insulin resistance in kidney disease: is there a distinct role separate from that of diabetes or obesity? Cardiorenal Med 2017 Dec;8(1):41-49. [doi: 10.1159/000479801] [Medline: 29344025]
12. Song SH, Goo YJ, Oh TR, et al. Insulin resistance is associated with incident chronic kidney disease in population with normal renal function. Kidney Res Clin Pract 2025 May;44(3):491-499. [doi: 10.23876/j.krcp.22.298] [Medline: 38148129]
13. Guerrero-Romero F, Simental-Mendía LE, González-Ortiz M, et al. The product of triglycerides and glucose, a simple measure of insulin sensitivity. Comparison with the euglycemic-hyperinsulinemic clamp. J Clin Endocrinol Metab 2010 Jul;95(7):3347-3351. [doi: 10.1210/jc.2010-0288] [Medline: 20484475]

XSL•FO
RenderX

14. Okamura T, Hashimoto Y, Hamaguchi M, Obora A, Kojima T, Fukui M. Triglyceride-glucose index is a predictor of incident chronic kidney disease: a population-based longitudinal study. Clin Exp Nephrol 2019 Jul;23(7):948-955. [doi: 10.1007/s10157-019-01729-2] [Medline: 30955189]
15. Fritz J, Brozek W, Concin H, et al. The triglyceride-glucose index and obesity-related risk of end-stage kidney disease in Austrian adults. JAMA Netw Open 2021 Mar 1;4(3):e212612. [doi: 10.1001/jamanetworkopen.2021.2612] [Medline: 33787913]
16. Gao YM, Chen WJ, Deng ZL, Shang Z, Wang Y. Association between triglyceride-glucose index and risk of end-stage renal disease in patients with type 2 diabetes mellitus and chronic kidney disease. Front Endocrinol 2023;14:1150980. [doi: 10.3389/fendo.2023.1150980]
17. Levey AS, Bosch JP, Lewis JB, Greene T, Rogers N, Roth D. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Modification of Diet in Renal Disease Study Group. Ann Intern Med 1999 Mar 16;130(6):461-470. [doi: 10.7326/0003-4819-130-6-199903160-00002] [Medline: 10075613]
18. Simental-Mendía LE, Rodríguez-Morán M, Guerrero-Romero F. The product of fasting glucose and triglycerides as surrogate for identifying insulin resistance in apparently healthy subjects. Metab Syndr Relat Disord 2008 Dec;6(4):299-304. [doi: 10.1089/met.2008.0034] [Medline: 19067533]
19. Nam GE, Park HS. Perspective on diagnostic criteria for obesity and abdominal obesity in Korean adults. J Obes Metab Syndr 2018 Sep 30;27(3):134-142. [doi: 10.7570/jomes.2018.27.3.134] [Medline: 31089555]
20. Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults. Executive summary of the third report of the National Cholesterol Education Program (NCEP) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (Adult Treatment Panel III). JAMA 2001 May 16;285(19):2486-2497. [doi: 10.1001/jama.285.19.2486]
21. Grundy SM, Cleeman JI, Daniels SR, et al. Diagnosis and management of the metabolic syndrome: an American Heart Association/National Heart, Lung, and Blood Institute scientific statement. Circulation 2005 Oct 25;112(17):2735-2752. [doi: 10.1161/CIRCULATIONAHA.105.169404] [Medline: 16157765]
22. Global kidney health atlas. International Society of Nephrology. 2023. URL: https://www.theisn.org/initiatives/global-kidney-health-atlas [accessed 2026-05-21]
23. Ferris ME, Gipson DS, Kimmel PL, Eggers PW. Trends in treatment and outcomes of survival of adolescents initiating end-stage renal disease care in the United States of America. Pediatr Nephrol 2006 Jul;21(7):1020-1026. [doi: 10.1007/s00467-006-0059-9] [Medline: 16773416]
24. Ferris ME, Miles JA, Seamon ML. Adolescents and young adults with chronic or end-stage kidney disease. Blood Purif 2016;41(1-3):205-210. [doi: 10.1159/000441317] [Medline: 26765837]
25. Li X, Wang L, Zhou H, Xu H. Association between triglyceride-glucose index and chronic kidney disease: results from NHANES 1999-2020. Int Urol Nephrol 2024 Nov;56(11):3605-3616. [doi: 10.1007/s11255-024-04103-8] [Medline: 38856937]
26. Lee HJ, Lee YH, Kim JS, et al. Triglyceride-glucose index and risk of renal function decline and death-censored renal allograft loss in kidney transplant recipients. Kidney Res Clin Pract 2025 Nov;44(6):974-983. [doi: 10.23876/j.krcp.25.095] [Medline: 41224321]
27. Fritz J, Brozek W, Concin H, et al. The association of excess body weight with risk of ESKD is mediated through insulin resistance, hypertension, and hyperuricemia. J Am Soc Nephrol 2022 Jul;33(7):1377-1389. [doi: 10.1681/ASN.2021091263] [Medline: 35500938]
28. Ko YE, Kim HW, Park JT, et al. Triglyceride-glucose index is an independent predictor of coronary artery calcification progression in patients with chronic kidney disease. Kidney Res Clin Pract 2024 May;43(3):381-390. [doi: 10.23876/j.krcp.23.264] [Medline: 38738273]
29. Bornfeldt KE, Tabas I. Insulin resistance, hyperglycemia, and atherosclerosis. Cell Metab 2011 Nov 2;14(5):575-585. [doi: 10.1016/j.cmet.2011.07.015] [Medline: 22055501]
30. Reardon CA, Lingaraju A, Schoenfelt KQ, et al. Obesity and insulin resistance promote atherosclerosis through an IFNγ-regulated macrophage protein network. Cell Rep 2018 Jun 5;23(10):3021-3030. [doi: 10.1016/j.celrep.2018.05.010] [Medline: 29874587]
31. Jia G, Aroor AR, DeMarco VG, Martinez-Lemus LA, Meininger GA, Sowers JR. Vascular stiffness in insulin resistance and obesity. Front Physiol 2015;6:231. [doi: 10.3389/fphys.2015.00231] [Medline: 26321962]
32. Duan S, Zhou M, Lu F, et al. Triglyceride-glucose index is associated with the risk of chronic kidney disease progression in type 2 diabetes. Endocrine 2023 Jul;81(1):77-89. [doi: 10.1007/s12020-023-03357-z] [Medline: 37004636]
33. Pan Y, Zhong S, Zhou K, et al. Association between diabetes complications and the triglyceride-glucose index in hospitalized patients with type 2 diabetes. J Diabetes Res 2021;2021:8757996. [doi: 10.1155/2021/8757996] [Medline: 34671683]
34. Liu L, Xia R, Song X, et al. Association between the triglyceride–glucose index and diabetic nephropathy in patients with type 2 diabetes: a cross‐sectional study. J Diabetes Investig 2021 Apr;12(4):557-565. [doi: 10.1111/jdi.13371] [Medline: 33319507]

XSL•FO
RenderX

35. Dalle Grave R, Calugi S, Centis E, Marzocchi R, El Ghoch M, Marchesini G. Lifestyle modification in the management of the metabolic syndrome: achievements and challenges. Diabetes Metab Syndr Obes 2010 Nov 2;3:373-385. [doi: 10.2147/DMSOTT.S13860] [Medline: 21437107]
36. Zhu Q, Chen Y, Cai X, et al. The non-linear relationship between triglyceride-glucose index and risk of chronic kidney disease in hypertensive patients with abnormal glucose metabolism: a cohort study. Front Med 2022;9:1018083. [doi: 10.3389/fmed.2022.1018083]

Abbreviations

aHR: adjusted hazard ratio
CKD: chronic kidney disease
DBP: diastolic blood pressure
eGFR: estimated glomerular filtration rate
ESKD: end-stage kidney disease
HDL-C: high-density lipoprotein-cholesterol
HR: hazard ratio
***ICD-10*:** *International Statistical Classification of Diseases and Related Health Problems, 10th Revision*
MetS: metabolic syndrome
NHIS: National Health Insurance Service
SBP: systolic blood pressure
TyG: triglyceride-glucose

Edited by Z Wu; submitted 11.Mar.2025; peer-reviewed by R Oh, WY Park; revised version received 02.Jan.2026; accepted 20.Apr.2026; published 03.Jun.2026.

Please cite as:
Kim JS, Lee YH, Chung S, Hwang HS, Han K, Yoon HE
Triglyceride-Glucose Index and Risk of End-Stage Kidney Disease in Young Adults: National Population-Based Cohort Study
JMIR Public Health Surveill 2026;12:e73750
URL: https://publichealth.jmir.org/2026/1/e73750
doi:10.2196/73750

© Jin Sug Kim, Yu Ho Lee, Sungjin Chung, Hyeon Seok Hwang, Kyungdo Han, Hye Eun Yoon. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 3.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Alcohol Consumption and Age-Specific Risk of Esophageal Cancer: Prospective Cohort Study

Yongbo Yang[1,2], MD; Xing Xing[3], MPH; Zhen Qin[3], BMed; Chunyang Han[4], BMed; He Zhu[3,5,6], PhD

[1]First Department of Thoracic Surgery, Peking University Cancer Hospital and Institute, Beijing, China
[2]Key Laboratory of Carcinogenesis and Translational Research, Ministry of Education, Peking University Cancer Hospital and Institute, Beijing, China
[3]School of Public Health, Peking University, 38 Xueyuan Rd, Haidian District, Beijing, China
[4]The First Clinical School, Huazhong University of Science and Technology, Wuhan, Hubei, China
[5]National Health Commission Key Laboratory of Health System Reform and Governance (Peking University), Peking University, Beijing, China
[6]Beijing Institute for Health Development, Peking University, Beijing, China

Corresponding Author:
He Zhu, PhD
School of Public Health, Peking University, 38 Xueyuan Rd, Haidian District, Beijing, China

Abstract

Background: China accounts for more than 40% of new global cases and deaths from esophageal cancer, and has a relatively high rate of past-year alcohol use, reaching up to 27%. The incidence and risk factors of esophageal cancer exhibit marked age-related variation; however, the impact of alcohol consumption on the risk of esophageal cancer across different age groups remains poorly understood.

Objective: This study aimed to investigate the age-specific associations between alcohol consumption and the risk of esophageal cancer.

Methods: Data were obtained from the China Kadoorie Biobank, a large-scale, nationwide prospective cohort study. The final analysis included 489,664 adult participants aged 30 to 79 years enrolled at baseline between 2004 and 2008. Age-specific incidence rates per 100,000 person-years of esophageal cancer were calculated. Cox proportional hazards models were used to assess the associations between alcohol consumption and the risk of esophageal cancer, stratified by age group.

Results: In the study cohort, incidence rates of esophageal cancer consistently increased with higher frequency of alcohol consumption among all age groups younger than 70 years. The association exhibited an age-specific pattern: within each age group, the rate ratio for weekly versus never drinkers was highest in the youngest group (8.31 for the 30‐49 year age group) and declined with age; across age groups, among weekly drinkers, the incidence rate increased sharply from the 30‐49 year to the 50‐54 year age group (rate ratio 2.81). Within each age group, compared with never drinkers, the adjusted hazard ratios for weekly drinkers decreased progressively with age group, from 4.06 (95% CI 2.73‐6.02) in adults aged 30-49 years to 3.17 (95% CI 2.34-4.30), 2.88 (95% CI 2.20-3.77), 2.36 (95% CI 1.78-3.14), 2.25 (95% CI 1.69-3.00), and 1.61 (95% CI 1.15-2.25) in those aged 50‐54, 55‐59, 60‐64, 65‐69, and 70‐74 years, respectively.

Conclusions: Our findings highlight a potential age-modifying effect of alcohol consumption on the risk of esophageal cancer, with a strong relative risk observed in young adults and a marked acceleration during the transition to midlife. This underscores the need to develop age-specific public health strategies to reduce alcohol use and to strengthen screening and intervention efforts to reduce the burden of esophageal cancer.

(JMIR Public Health Surveill 2026;12:e92949) doi:10.2196/92949

KEYWORDS

esophageal cancer; age; alcohol; China Kadoorie Biobank; cohort study; age-specific association

Introduction

Esophageal cancer is a highly aggressive malignancy characterized by a poor prognosis, reduced quality of life, and low survival rates, representing a substantial global health burden [1]. According to the Global Cancer Observatory estimates, esophageal cancer was the 11th most commonly diagnosed cancer, with 2.6% of all new cancer cases and the seventh leading cause of cancer-related deaths, accounting for 4.6% of all cancer deaths worldwide in 2022 [2]. Notably, in 2022, China accounted for 43.8% of new global cases and 42.1% of deaths from esophageal cancer [3], and an age-standardized incidence rate of 8.32 per 100,000 population and an age-standardized mortality rate of 6.68 per 100,000 population were higher than the global average, ranking as the seventh most prevalent cancer and the fifth leading cause of

XSL•FO
RenderX

cancer-related death within the country [4]. Furthermore, most patients are diagnosed at advanced stages, and esophageal cancer is associated with dysphagia and weight loss, reduced functional status, increased comorbidities, higher risk of treatment-related toxicities from aggressive multimodal therapies, and increased mortality risk [1,5,6]. Given this substantial clinical and economic burden, there is a clear need to better understand the underlying risk factors to inform early detection and targeted intervention efforts.

Esophageal cancer exhibits pronounced age-related characteristics, and the risk increases significantly with age, with a particularly high prevalence among older adults [1,7]. However, recent estimates indicate a concerning increase in incidence rates among younger populations [8]. National and global data indicate that incidence rates generally peak between the ages of 70 and 85 years [5,9-11]. An analysis of Chinese national data from 2005 to 2015 found that age is an independent risk factor for esophageal cancer; the risk of developing the disease increases significantly, with an odds ratio of 1.18 for each additional year of age and an odds ratio of 1.62 for every additional 5 years [12]. Evidence suggests that the impact of certain risk factors may vary across different age groups; however, these variations across the age spectrum remain poorly understood. For example, a study using large-scale pooled data from 8 population-based case-control studies within the international Barrett's and Esophageal Adenocarcinoma Consortium (BEACON) revealed that smoking and reflux were positively associated with an increased risk of esophageal adenocarcinoma (EAC) in all age groups, but obesity and recurrent gastroesophageal reflux were significantly stronger risk factors for individuals diagnosed before the age of 50 compared with older age groups [8]. There is a clear need to investigate age-specific risk factors for esophageal cancer to better inform targeted prevention strategies for differential susceptibility.

Alcohol consumption is a major modifiable risk factor for esophageal cancer, particularly for esophageal squamous cell carcinoma (ESCC), which is the predominant histological subtype accounting for approximately 90% of cases in China [13]. This differs from the Western population, where EAC is more common and exhibits a distinct etiological profile [2,14]. Moreover, alcohol consumption warrants particular attention due to its high prevalence in the Chinese population, and the recent national data reported that 27.6% of individuals aged 15 years and older consumed alcohol in the past year in China, with this rate reaching a striking 44.5% among men [15]. However, the impact of alcohol on esophageal cancer risk across the lifespan remains poorly characterized, particularly in Chinese populations. One study, largely derived from Western studies, suggested that alcohol shows little to no association with EAC risk among individuals aged <50 years and those in the 50 - 59 and 60 - 69 year age groups [8]. In contrast, research on ESCC demonstrated a monotonic increase in risk associated with younger age at drinking initiation, prolonged duration, and higher intensity of alcohol intake [16]. Furthermore, the age-specific etiological role of alcohol across the lifespan remains poorly characterized, largely due to historical sample size limitations. Because early-onset esophageal malignancies are relatively rare, previous studies assessing risk factors have often been constrained by small sample sizes, precluding detailed and statistically robust investigations of alcohol consumption across multiple distinct age groups [8]. Therefore, there is a clear need to explore large-scale, population-based data to systematically examine the age-specific associations between alcohol consumption and esophageal cancer risk to add evidence for the Chinese population.

Therefore, this study sought to comprehensively investigate the associations between alcohol consumption and the risk of esophageal cancer across various age groups using a cohort of 0.5 million individuals in China. The findings are expected to provide robust evidence to inform the development of more precise, efficient, and cost-effective risk-stratified screening strategies to prevent and intervene against esophageal cancer in China.

Methods

Data Source and Study Cohort

The data were derived from the China Kadoorie Biobank (CKB) resource (request numbers DAR-2025 - 00208 and DAR-2025 - 00316). The CKB is a large-scale, nationwide prospective cohort study designed to investigate the genetic and environmental determinants of major chronic diseases in China, established across 10 geographically diverse research sites, including 5 urban and 5 rural areas [17-19]. The selection of 10 sites considered the prevalence of chronic diseases, exposure profiles, economic development, stability of the local population, and local capacity to capture diverse disease patterns and risk exposures. At baseline, more than 512,000 adults aged 30 to 79 years were enrolled between 2004 and 2008, and demographic characteristics, socioeconomic status, lifestyle habits (including smoking, alcohol consumption, diet, and physical activity), and medical history were collected via an interviewer-administered electronic questionnaire. At the follow-up, the vital status and incident disease events of the cohort participants were continuously monitored through established linkages with regional death registries, disease surveillance systems, and national health insurance claim databases. Publication of results does not require or imply approval by the membership of the CKB Collaborative Group.

In this study, the data, including a total of 512,724 participants, were obtained from the CKB, and we excluded participants with a prior diagnosis of any cancer at baseline (n=2578, 0.5%), those with missing family history information (n=20,480, 3.99%), and those with missing BMI data (n=2, 0%). After these exclusions, a total of 489,664 (95.5%) participants were included in the final analysis. To examine age-related differences, all participants were categorized based on their baseline age into the following groups: 30 - 49 years (n=223,074, 45.55%), 50 - 54 years (n=84,708, 17.3%), 55 - 59 years (n=66,865, 13.66%), 60 - 64 years (n=46,663, 9.53%), 65 - 69 years (n=38,344, 7.83%), 70 - 74 years (n=26,862, 5.49%), and 75 - 79 years (n=3148, 0.64%). The 5-year interval was selected because it allows for a more granular exploration of age-specific associations between alcohol consumption and esophageal cancer risk, captures potential nonlinear patterns, and improves

XSL•FO
RenderX

direct comparisons with previous studies that widely used similar age-specific groupings [10,20]. Additionally, we grouped ages 30 - 49 years into one age group mainly due to the relatively low incidence of esophageal cancer among individuals younger than 50 years in the CKB data, thereby avoiding sparse data and limited statistical power and ensuring sufficient case numbers and statistical robustness for the regression models examining alcohol-related risks.

Study Variables

Esophageal Cancer

The primary outcome of this analysis was esophageal cancer, defined according to the *International Classification of Diseases, Tenth Revision* (*ICD-10*) code C15. Cases of esophageal cancer were identified through linkage with multiple sources by the CKB, including regional cancer registries, mortality databases, and the national health insurance system. To ensure comprehensive case identification and minimize loss to follow-up, all disease outcomes were coded by trained research staff, and annual active surveillance was performed using residential records, health insurance data, and administrative databases, with additional verification through direct contact with participants or their relatives when necessary [18].

Alcohol Consumption at Baseline

Information on alcohol consumption habits was collected from CKB participants using a standardized questionnaire during the baseline survey. Participants were asked about their drinking frequency over the past 12 months, which was categorized as follows: (1) never drinking (ie, never or almost never), (2) only occasionally (ie, only occasionally or only at certain seasons), (3) every month but less than weekly, and (4) at least once a week or weekly drinking [21]. This variable was used as the exposure in this study. Furthermore, participants who reported drinking at least once a week were additionally asked about their age at initiation of alcohol consumption. This was further categorized into 5 groups: <15, 15 - 24, 25 - 34, 35 - 44, and ≥45 years to examine the associations between the age of alcohol initiation and the risk of esophageal cancer among weekly drinkers.

Covariates at Baseline

To reduce the possible confounding influences related to the estimated association, covariates were selected mainly in accordance with the risk factors identified in previous studies [13,16,22-24] and the availability of the CKB database at baseline, including (1) demographic variables: age group, sex (male or female), education level (uneducated, primary school, middle school, high school, technical school or college, or university), household income (≤¥4999, ¥5000 - ¥19,999, or ≥¥20,000; an exchange rate of ¥1=US $0.15 is applicable), area (rural or urban), and marital status (unmarried, married, or divorced or widowed); (2) lifestyle factors: physical activity (low, medium, or high), BMI (normal, underweight, overweight, or obese), smoking status (did not smoke, only occasionally, on most days, or daily), and family history of cancer (no or yes). In detail, BMI was calculated using the standard formula: weight (kg)/height (m)^2, derived from objective measurements of weight and height. Body weight and height were measured by trained staff using calibrated instruments. Physical activity was defined using metabolic equivalent task hours per day (MET-h/d) spent on work, transportation, housework, and nonsedentary recreation, and participants were then categorized into low, medium, or high groups based on the 25th and 75th percentiles.

Statistical Analysis

First, we described demographic characteristics and lifestyle factors stratified by alcohol consumption subgroups. Categorical variables were presented as percentages and compared using chi-square tests. Second, we calculated age-specific incidence rates (per 100,000 person-years) of esophageal cancer by baseline age group and alcohol consumption subgroup. For each age stratum, person-years were accrued from the baseline date until the date of esophageal cancer diagnosis, loss to follow-up, or December 31, 2018, whichever occurred first, stratified by alcohol consumption subgroup. Rates were then computed within each stratum and alcohol consumption category. Third, Cox proportional hazards models were used to evaluate the association between alcohol consumption and the risk of esophageal cancer in the total sample and across different age groups, after adjusting for sex, education level, household income, area, marital status, smoking status, physical activity, BMI, and family history of cancer, and across sex and age groups after adjusting for these covariates. For further analysis among weekly drinkers, we used Cox models to examine the association between age at alcohol initiation and esophageal cancer risk, adjusting for the same set of covariates. The proportional hazards assumption was tested using Schoenfeld residuals. Results are presented as hazard ratios with corresponding 95% CIs. Finally, to address potential selection bias arising from our exclusion criteria, we conducted a sensitivity analysis using inverse probability weighting. We modeled the probability of inclusion using logistic regression based on demographic characteristics and applied these probabilities as weights in our Cox proportional hazards models. All statistical analyses were performed using R (version 4.4.2; R Foundation for Statistical Computing).

Ethical Considerations

Ethical approvals for the CKB were obtained from Oxford Tropical Research Ethics Committee (025 - 04) at the University of Oxford, Oxford, United Kingdom, and the Ethical Review Committee from the Chinese Center for Disease Control and Prevention (005/2004), Beijing, China. All participants provided written informed consent upon recruitment into the CKB. This study, which used CKB data to examine the association between alcohol consumption and the risk of esophageal cancer, was approved by the institutional review board of Peking University, Beijing, China (IRB00001052-25128).

Results

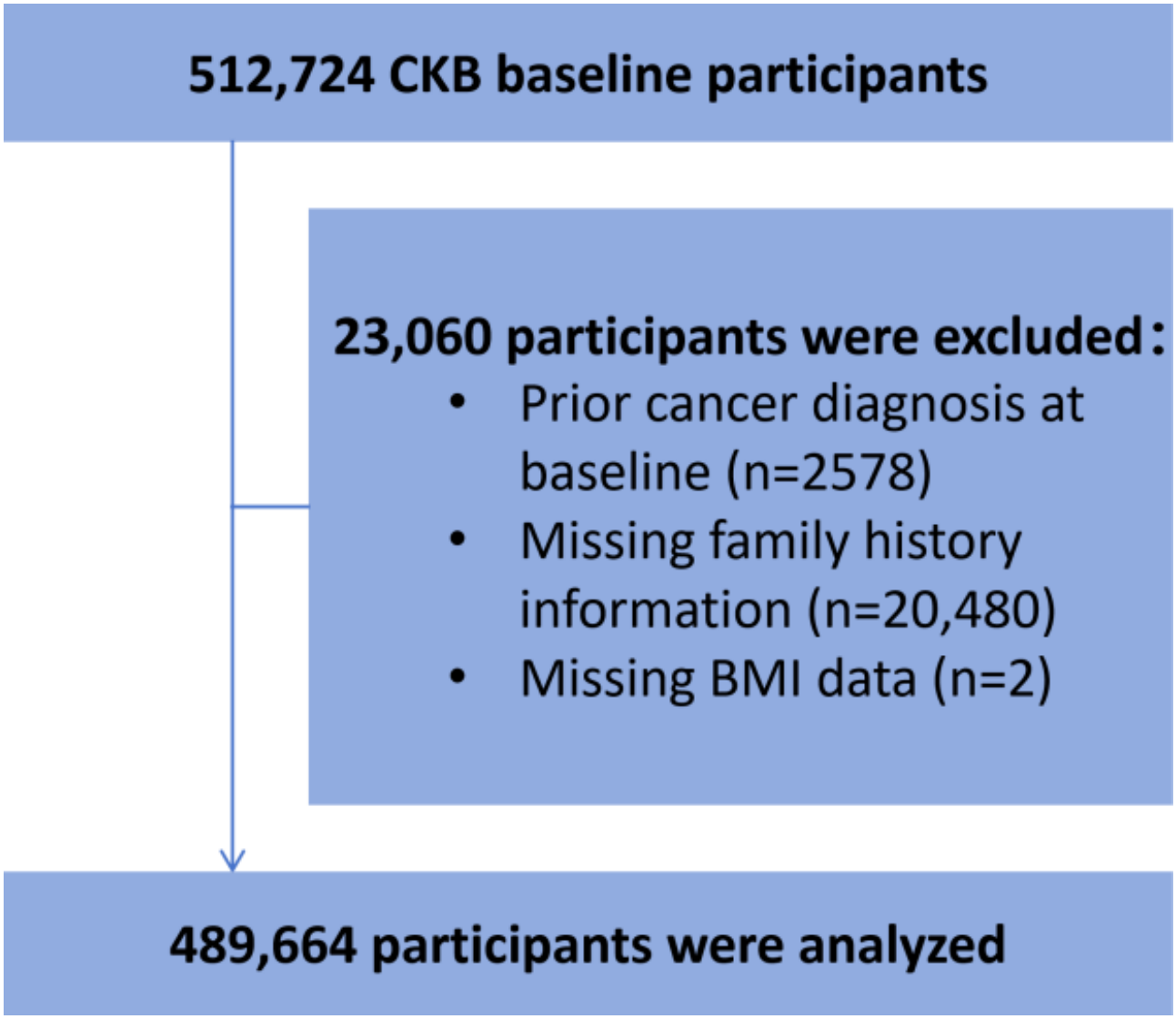
Baseline Characteristics of the Study Cohort

Overall, of the 489,664 participants (Figure 1), the study cohort was predominantly female (n=291,555, 59.54%), aged 30 to 49 years (n=223,074, 45.56%), had primary school education (n=156,853, 32.03%), were married (n=448,075, 91.51%), resided in rural areas (n=273,376, 55.83%), and reported a

XSL•FO
RenderX

medium income level (n=232,737, 47.53%) at baseline (Table 1).

Figure 1. Flow diagram of the study cohort selection. CKB: China Kadoorie Biobank.



Most participants had a normal BMI (n=255,030, 52.08%), did not smoke (n=347,209, 70.91%), engaged in high levels of physical activity (n=167,952, 34.30%), and had no family history of cancer (n=403,815, 82.47%).

Regarding alcohol consumption at baseline, of the total sample (n=489,664), 233,038 (47.59%) reported never drinking, 165,137 (33.72%) drank only occasionally, 18,908 (3.86%) drank every month but less than weekly, and 72,581 (14.82%) drank at least once a week. Participants aged 30 to 49 years accounted for more than half of the monthly and weekly drinkers. All subgroups differed significantly across alcohol consumption categories (P<.001).

XSL•FO
RenderX

Table . Characteristics of the study cohort stratified by alcohol consumption.

Variables	Overall (n=489,664), n (%)	Never (n=233,038), n (%)	Only occasionally (n=165,137), n (%)	Every month but less than weekly (n=18,908), n (%)	At least once a week (n=72,581), n (%)	*P* value
Age group (years)						<.001
30 - 49	223,074 (45.56)	95,708 (41.07)	82,713 (50.09)	10,917 (57.74)	33,736 (46.48)	
50 - 54	84,708 (17.30)	39,648 (17.01)	28,474 (17.24)	3108 (16.44)	13,478 (18.57)	
55 - 59	66,865 (13.66)	34,026 (14.60)	20,985 (12.71)	2065 (10.92)	9789 (13.49)	
60 - 64	46,663 (9.53)	25,041 (10.75)	13,866 (8.40)	1223 (6.47)	6533 (9.00)	
65 - 69	38,344 (7.83)	21,455 (9.21)	10,832 (6.56)	885 (4.68)	5172 (7.13)	
70 - 74	26,862 (5.49)	15,343 (6.58)	7366 (4.46)	642 (3.40)	3511 (4.84)	
75 - 79	3148 (0.64)	1817 (0.78)	901 (0.55)	68 (0.36)	362 (0.50)	
Sex						<.001
Male	198,109 (40.46)	46,647 (20.02)	70,203 (42.51)	14,654 (77.50)	66,605 (91.77)	
Female	291,555 (59.54)	186,391 (79.98)	94,934 (57.49)	4254 (22.50)	5976 (8.23)	
Education						<.001
Uneducated	89,927 (18.37)	65,710 (28.20)	17,038 (10.32)	1053 (5.57)	6126 (8.44)	
Primary school	156,853 (32.03)	81,850 (35.12)	47,640 (28.85)	4374 (23.13)	22,989 (31.67)	
Middle school	140,126 (28.62)	55,404 (23.77)	53,630 (32.48)	6618 (35.00)	24,474 (33.72)	
High school	74,528 (15.22)	23,709 (10.17)	33,064 (20.02)	4580 (24.22)	13,175 (18.15)	
Technical school or college	17,248 (3.52)	3955 (1.70)	8506 (5.15)	1306 (6.91)	3481 (4.80)	
University	10,982 (2.24)	2410 (1.03)	5259 (3.18)	977 (5.17)	2336 (3.22)	
Income (¥)[a]						<.001
≤4999	45,913 (9.38)	24,438 (10.49)	15,064 (9.12)	1057 (5.59)	5354 (7.38)	
5000 - 19,999	232,737 (47.53)	108,756 (46.67)	85,129 (51.55)	8722 (46.13)	30,130 (41.51)	
≥20,000	211,014 (43.09)	99,844 (42.84)	64,944 (39.33)	9129 (48.28)	37,097 (51.11)	
Area						<.001
Rural	273,376 (55.83)	137,325 (58.93)	89,727 (54.33)	10,020 (52.99)	36,304 (50.02)	
Urban	216,288 (44.17)	95,713 (41.07)	75,410 (45.67)	8888 (47.01)	36,277 (49.98)	
Marital status						<.001
Unmarried	554 (0.11)	211 (0.09)	199 (0.12)	37 (0.20)	107 (0.15)	
Married	448,075 (91.51)	209,335 (89.83)	152,550 (92.38)	17,770 (93.98)	68,420 (94.27)	
Divorced or widowed	41,035 (8.38)	23,492 (10.08)	12,388 (7.50)	1101 (5.82)	4054 (5.59)	
BMI						<.001
Normal	255,030 (52.08)	123,177 (52.86)	84,375 (51.09)	9572 (50.62)	37,906 (52.23)	
Underweight	21,528 (4.40)	12,850 (5.51)	5656 (3.43)	528 (2.79)	2494 (3.44)	
Overweight	162,015 (33.09)	73,401 (31.50)	56,865 (34.44)	6801 (35.97)	24,948 (34.37)	
Obese	51,091 (10.43)	23,610 (10.13)	18,241 (11.05)	2007 (10.61)	7233 (9.97)	
Smoking						<.001
Did not smoke	347,209 (70.91)	202,886 (87.06)	115,472 (69.92)	7992 (42.27)	20,859 (28.74)	
Only occasionally	21,637 (4.42)	3612 (1.55)	10,474 (6.34)	2035 (10.76)	5516 (7.60)	
On most days	3580 (0.73)	712 (0.31)	1358 (0.82)	410 (2.17)	1100 (1.52)	

XSL•FO
RenderX

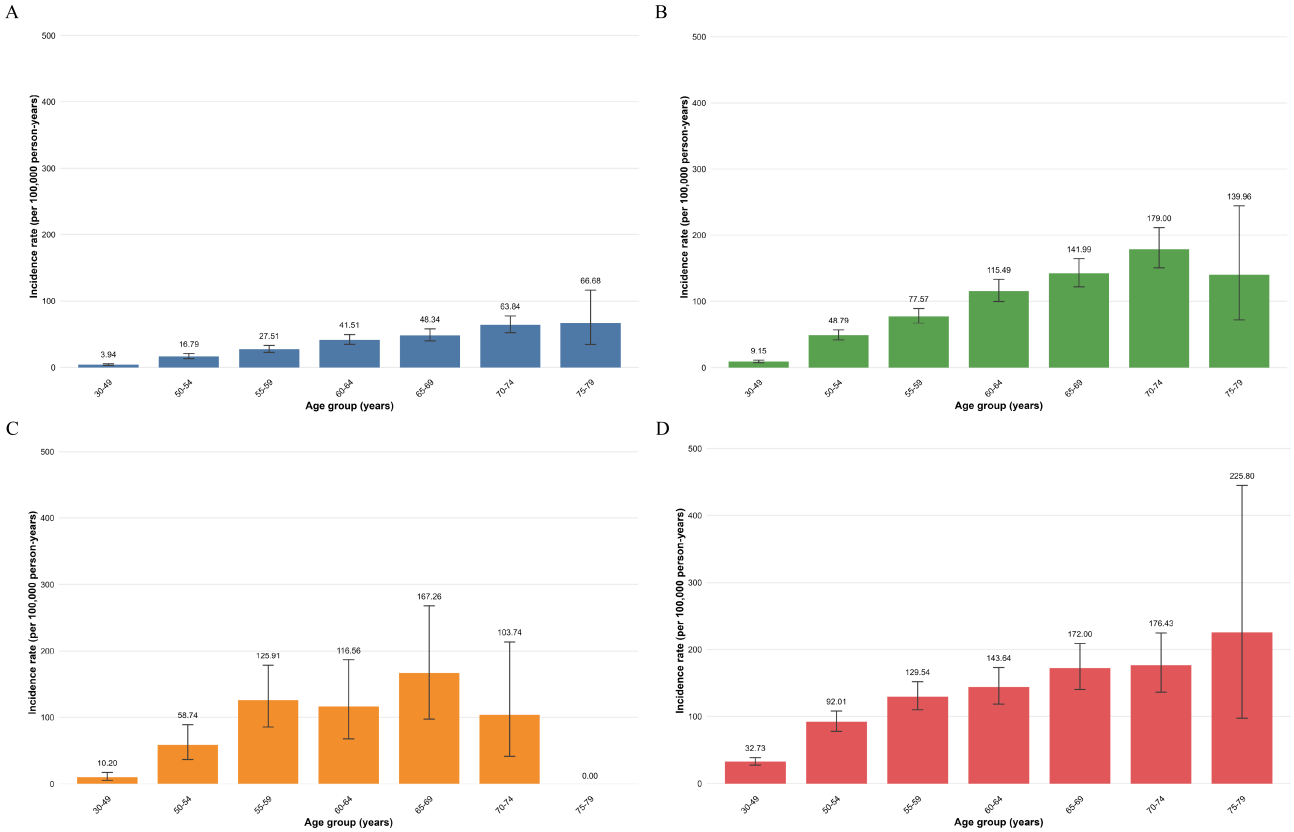
Variables	Overall (n=489,664), n (%)	Never (n=233,038), n (%)	Only occasionally (n=165,137), n (%)	Every month but less than weekly (n=18,908), n (%)	At least once a week (n=72,581), n (%)	*P* value
Daily	117,238 (23.94)	25,828 (11.08)	37,833 (22.91)	8471 (44.80)	45,106 (62.15)	
Physical activity						<.001
Low	160,732 (32.82)	81,026 (34.77)	54,537 (33.03)	5252 (27.78)	19,917 (27.44)	
Medium	160,980 (32.88)	73,300 (31.45)	57,726 (34.96)	6203 (32.81)	23,751 (32.72)	
High	167,952 (34.30)	78,712 (33.78)	52,874 (32.02)	7453 (39.42)	28,913 (39.84)	
Family history of cancer						<.001
No	403,815 (82.47)	197,866 (84.91)	132,350 (80.15)	15,052 (79.61)	58,547 (80.66)	
Yes	85,849 (17.53)	35,172 (15.09)	32,787 (19.85)	3856 (20.39)	14,034 (19.34)	

[a]An exchange rate of ¥1=US $0.15 is applicable.

Incidence Rates of Esophageal Cancer Stratified by Age Group and Alcohol Consumption at Baseline

Within each age stratum, incidence rates consistently increased with higher frequency of alcohol consumption among age groups younger than 70 years (Figure 2). The incidence rate ratio comparing weekly drinkers to never drinkers was highest in the youngest age group and progressively declined with advancing age. Specifically, the rate ratio was 8.31 in the 30 - 49 year age group, decreasing to 5.48 in the 50 - 54 year age group, 4.71 in the 55 - 59 year age group, 3.46 in the 60 - 64 year age group, 3.56 in the 65 - 69 year age group, 2.76 in the 70 - 74 year age group, and 3.39 in the 75 - 79 year age group.

Figure 2. Incidence rates per 100,000 person-years of esophageal cancer stratified by age group and alcohol consumption: (A) never, (B) only occasionally, (C) every month but less than weekly, and (D) at least once a week.



Across age groups, among weekly drinkers, the rate ratio comparing successive age groups was 2.81 for ages 50 - 54 vs 30 - 49 years (92.01/32.73), 1.41 for ages 55 - 59 vs 50 - 54 years (129.54/92.01), 1.11 for ages 60 - 64 vs 55 - 59 years (143.64/129.54), 1.20 for ages 65 - 69 vs 60 - 64 years (172.00/143.64), 1.03 for ages 70 - 74 vs 65 - 69 years (176.43/172.00), and 1.28 for ages 75 - 79 vs 70 - 74 years (225.80/176.43). Because of the small number of participants in the 75 - 79 year age group, these findings should be considered as preliminary results.

XSL•FO
RenderX

Alcohol Consumption and the Risk of Esophageal Cancer in the Overall Study Cohort

In the overall study cohort, after adjusting for covariates, all levels of alcohol consumption were significantly associated with an increased risk of esophageal cancer (all $P<.001$; Table 2). Compared with never drinkers, the adjusted hazard ratios (aHRs) were 2.02 (95% CI 1.82 - 2.24) for participants who drank only occasionally, 1.75 (95% CI 1.41 - 2.18) for those who drank monthly but less than weekly, and 2.64 (95% CI 2.34 - 2.98) for weekly drinkers.

XSL•FO
RenderX

Table . Adjusted hazard ratios (HRs) for the association between alcohol consumption and the risk of esophageal cancer in the overall study cohort.

Variables	aHR (95% CI)	*P* value
Alcohol consumption (reference: never drinking)		
Only occasionally	2.02 (1.82 - 2.24)	<.001
Every month but less than weekly	1.75 (1.41 - 2.18)	<.001
At least once a week	2.64 (2.34 - 2.98)	<.001
Sex (reference: male)		
Female	0.22 (0.13 - 0.36)	<.001
Age group (y; reference: 30 - 49 y)		
50 - 54	2.40 (1.89 - 3.06)	<.001
55 - 59	3.10 (2.42 - 3.99)	<.001
60 - 64	3.31 (2.52 - 4.36)	<.001
65 - 69	3.94 (2.93 - 5.28)	<.001
70 - 74	3.93 (2.78 - 5.56)	<.001
75 - 79	5.46 (2.49 - 11.84)	<.001
Education (reference: uneducated)		
Primary school	1.04 (0.83 - 1.31)	.73
Middle school	0.79 (0.60 - 1.04)	.09
High school	0.64 (0.44 - 0.91)	.01
Technical school or college	0.54 (0.28 - 1.01)	.06
University	0.39 (0.16 - 0.98)	.04
Marital status (reference: unmarried)		
Married	0.71 (0.38 - 1.32)	.27
Divorced or widowed	0.73 (0.39 - 1.37)	.33
Household income (¥; reference: ≤¥4999)[a]		
5000 - 19,999	0.55 (0.45 - 0.67)	<.001
≥20,000	0.35 (0.28 - 0.44)	<.001
Area (reference: rural)		
Urban	0.58 (0.48 - 0.70)	<.001
Smoking (reference: did not smoke)		
Only occasionally	0.90 (0.61 - 1.33)	.58
On most days	1.47 (0.77 - 2.79)	.24
Daily	1.59 (1.31 - 1.93)	<.001
BMI (reference: normal BMI)		
Underweight	1.25 (1.07 - 1.46)	.008
Overweight	0.88 (0.80 - 0.97)	.009
Obese	0.89 (0.76 - 1.13)	.13
Physical activity (reference: low physical activity)		
Medium	0.85 (0.70 - 1.02)	.08
High	0.77 (0.63 - 0.93)	.01
Family history of cancer (reference: no family history of cancer)		
Yes	1.54 (1.29 - 1.83)	<.001

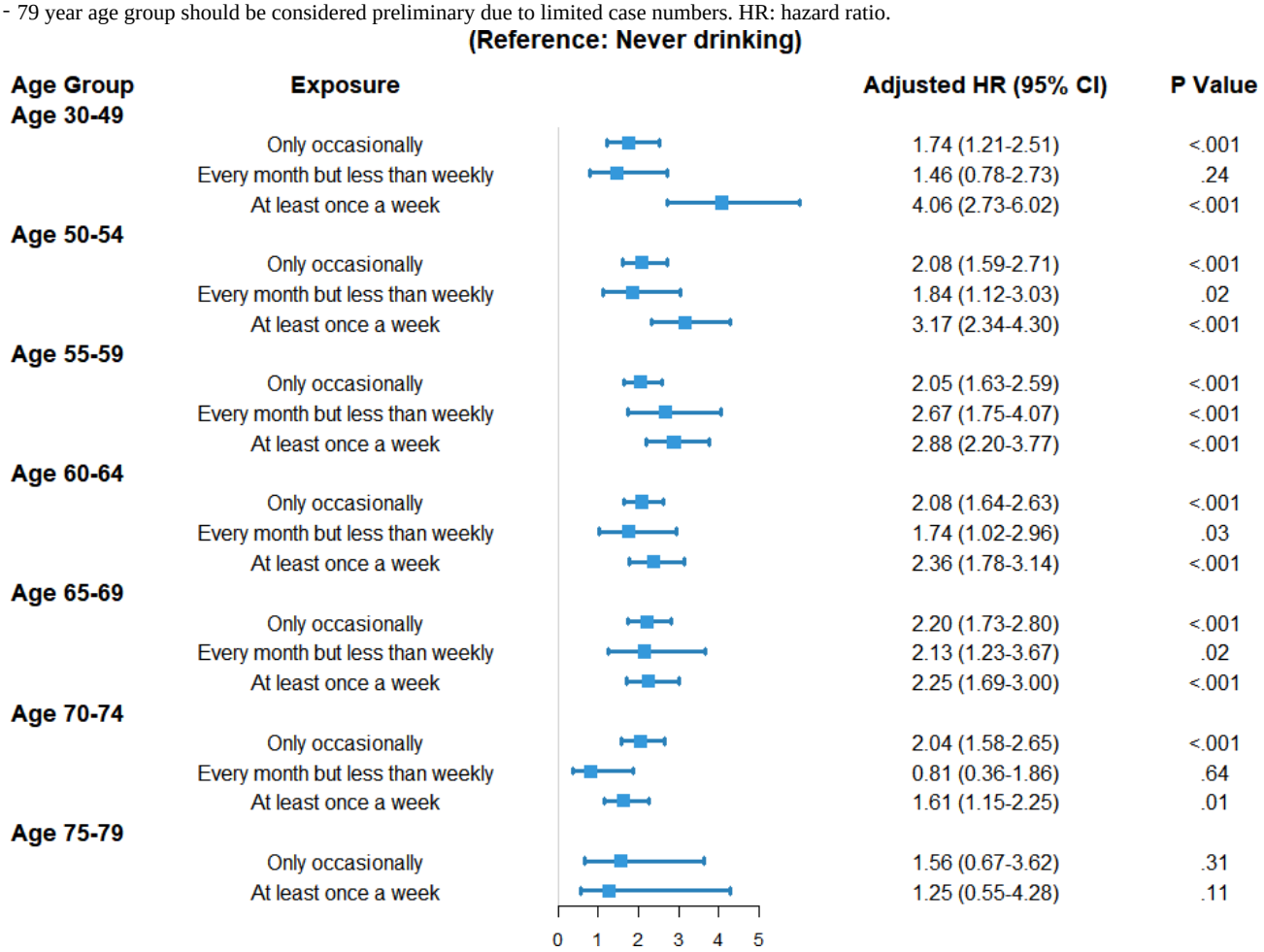
[a]An exchange rate of ¥1=US $0.15 is applicable.

XSL•FO
RenderX

Alcohol Consumption and the Risk of Esophageal Cancer Stratified by Age Group

Stratified analyses indicated a significant association between weekly drinking and the risk of esophageal cancer among all age groups except the 75 - 79 year age group, with a markedly stronger association observed in younger individuals (Figure 3; Figure S1 and Table S1 in Multimedia Appendix 1).

Figure 3. Forest plot of the associations between alcohol consumption and the risk of esophageal cancer, stratified by age group. Estimates for the 75 - 79 year age group should be considered preliminary due to limited case numbers. HR: hazard ratio.



Among participants aged 30 - 49 years, weekly drinkers had an aHR of 4.06 (95% CI 2.73 - 6.02) compared with never drinkers, and the aHRs for weekly drinkers were 3.17, 2.88, 2.36, 2.25, and 1.61 for the age groups 50 - 54, 55 - 59, 60 - 64, 65 - 69, and 70 - 74 years, respectively. Significant positive associations for occasional drinking were observed across all age groups from 30 - 49 to 70 - 74 years, with aHRs ranging from 1.74 to 2.20. For participants who drank every month but less than weekly, significant associations were found primarily in middle-aged and older groups (50 - 54, 55 - 59, 60 - 64, and 65 - 69 years), with aHRs ranging from 1.74 to 2.67, although estimates were less stable.

In sex-stratified analyses, the association between alcohol consumption and esophageal cancer risk was stronger and more consistent in men than in women (Figure 4). Among men, weekly drinkers had significantly higher risks across all age groups except for those aged 75-79 years compared with never drinkers, with aHRs ranging from 1.51 to 3.59 (all $P<.05$). Occasional drinking also consistently increased the risk in men. In women, the association was less consistent, with wide CIs and mostly nonsignificant results, likely due to low drinking prevalence and fewer female drinkers, which limited statistical power.

Additionally, among participants who reported drinking at least once a week, initiating alcohol consumption at ages 15-24 years was associated with an increased risk of esophageal cancer compared with initiating alcohol consumption after age 45 years (aHR 1.28, 95% CI 1.03-1.64; Figure S2 in Multimedia Appendix 1).

XSL•FO
RenderX

Figure 4. Forest plot of the associations between alcohol consumption and the risk of esophageal cancer, stratified by sex and age group: (A) men and (B) women. Estimates for the 75-79 year age group and for women should be considered preliminary due to limited case numbers. HR: hazard ratio.

XSL•FO
RenderX

Sensitivity Analysis

The results from the weighted models were consistent with our primary findings, suggesting that the exclusion criteria did not introduce substantial selection bias or alter our main conclusions. Detailed results are presented in Tables S2-S4 in Multimedia Appendix 1.

Discussion

Using data from CKB, one of the world's largest prospective cohorts, we comprehensively evaluated the age-specific associations between alcohol consumption and the risk of developing esophageal cancer. This study had 3 key findings. First, the significant positive association between drinking frequency and the risk of esophageal cancer was consistently observed across age groups. Second, alcohol exposure exhibited a potential age-modifying effect, and the highest relative risk between alcohol consumption and the risk of esophageal cancer was observed in younger adults aged 30‑49 years (aHR 4.06 for weekly vs never drinkers) and gradually declined with increasing age. Finally, among weekly drinkers, a marked increase in incidence from the 30‑49 to the 50‑54 year age group (rate ratio 2.81) highlighted a potential window for targeted prevention.

Our findings reinforce the association between alcohol consumption and the risk of esophageal cancer and further add evidence regarding age- and frequency-specific incidence rates and hazard ratios. The positive associations observed both in the overall cohort and across age groups are consistent with previous studies. For example, data from the Global Burden of Disease (GBD) 2019 study demonstrated that the rates of deaths specifically attributable to alcohol-induced esophageal cancer increased sequentially with age [25]. Notably, the prevalence and high-risk patterns of alcohol use remain a severe public health concern in China across all age groups, serving as an important driver of the associated disease burden. According to 2024 national data for individuals aged 15 years and older, the prevalence of alcohol use is 20.3% in the past 30 days and 27.6% in the past 12 months, with higher rates observed among men and young to middle-aged adults [15]. More than 40% of current drinkers engage in heavy episodic drinking, and a predominant preference was for high-alcohol spirits [15]. These deeply entrenched, high-risk drinking behaviors, combined with our findings of pronounced age-specific vulnerability, underscore a critical need for targeted public health interventions. More efforts are urgently needed to support the implementation of comprehensive primary prevention strategies, including stricter alcohol regulations, increased taxation, and widespread educational campaigns advocating alcohol reduction or cessation [25]. Future studies should aim to refine risk-stratified, age-targeted, and subtype-specific screening strategies to improve the effectiveness of esophageal cancer prevention in China.

Our findings indicate the risk of alcohol consumption on esophageal cancer, particularly among young heavy drinkers. Evidence regarding the increased risk of esophageal cancer associated with alcohol consumption among young people remains limited. The overall prevalence of alcohol consumption peaks among young to middle-aged adults aged 25 to 44 years, reaching a past-year prevalence of 32.5% and a past-month prevalence of 23.2% [15]. A study conducted in a high-risk area in Taixing, China, found that compared with never drinkers, individuals who started drinking between ages 21 and 28 years and those who started before age 21 years had a significantly increased risk of ESCC by 174% and 168%, respectively [16]. Notably, this age-related risk gradient exhibited a clear monotonic increasing trend—the younger the age at which drinking began, the greater the risk increase [16]. A cohort study from Shanghai, China, found that the average age at which esophageal cancer patients started regular drinking (26.0 years) was significantly younger than that of the noncase group (29.8 years) [26]. However, these studies primarily originated from specific high-risk areas for esophageal cancer in China, limiting their generalizability. Analyses based on GBD 2019 data indicated that risk factors such as smokeless tobacco use, low fruit intake, and low vegetable intake contributed more significantly to esophageal cancer risk among younger populations [27]. Therefore, despite the relatively low incidence rate of esophageal cancer among young people, the occurrence of early-onset cancer and other adverse health outcomes necessitates the development of targeted prevention and intervention strategies for adolescents and young adults to effectively reduce alcohol consumption and mitigate the rising future burden of early-onset esophageal cancer.

Our findings reveal a sharp increase in the incidence of esophageal cancer in weekly drinkers between the 30‑49 and 50‑54 year age groups. While global and national epidemiological data indicate that the absolute incidence burden of esophageal cancer generally peaks in much older populations, typically those aged 65 years and older [14,28,29], our results suggest that regular alcohol consumption might accelerate disease onset during midlife. This distinct risk escalation highlights a crucial window of opportunity for targeted clinical interventions. Currently, the National Health Commission of China has launched the 2022 Guideline for Esophageal Cancer Screening, Early Diagnosis, and Early Treatment. This guideline specifies high-risk groups for esophageal cancer, including individuals older than 45 years who live in high-incidence areas, have a family history among first-degree relatives (eg, parents, children, or siblings), exhibit poor lifestyle habits (eg, hot foods, high-salt diets, preserved foods, smoking, and heavy alcohol use), or have preexisting conditions (eg, chronic esophagitis, Barrett esophagus, and a history of precancerous lesions) [30]. The high-incidence areas in China were also attributed to environmental, lifestyle, or socioeconomic factors associated with esophageal cancer risk with geographic clustering [11,13,29]. Similarly, the American College of Gastroenterology and the British Society of Gastroenterology guidelines recommend periodic endoscopy for individuals older than 50 years with additional risk factors [31,32], and these guidelines are primarily focused on EAC. Beyond geographic, familial risk, and preexisting conditions, the findings reemphasize the necessity to aggressively target regular drinkers prior to and during this midlife transition to reduce the risk.

The results of this study should be interpreted in light of several limitations. First, alcohol exposure data were self-reported, and

XSL•FO
RenderX

heavy drinking may still be underreported. The CKB implemented multiple strategies to ensure data quality during collection. For example, field interviewers were required to have a medical background and fieldwork experience and to have received standardized training [19]. Second, alcohol consumption data were collected only at baseline, and potential changes during the long-term follow-up may have influenced the risk estimates. However, evidence indicates that lifestyle factors remained relatively stable over an extended period for most CKB participants, suggesting that these potential biases remain within acceptable ranges [18]. Third, the types of esophageal cancer, such as ESCC and EAC, cannot be defined by *ICD-10* codes in the CKB data. Finally, despite our efforts to adjust for various known and suspected confounders, unmeasured or poorly measured factors (eg, dietary risk factors or potential carcinogens) may have an effect.

In conclusion, this study used the large-scale CKB data to provide evidence on the age-modifying associations between alcohol consumption and esophageal cancer risk in China. The findings of a strong relative risk observed in young adults and a marked acceleration during the transition to midlife underscore the importance of developing age-specific public health strategies to mitigate the burden of esophageal cancer, particularly focusing on early alcohol intervention in young populations and enhanced screening in the middle-aged and older adult population.

Acknowledgments

The most important acknowledgment was directed to the participants in the study and the members of the survey teams in each of the 10 regional centers, as well as to the project development and management teams based in Beijing, Oxford, and the 10 regional centers.

Funding

Funding for the development and maintenance of the China Kadoorie Biobank resource was received from the Ministry of Science and Technology of the People's Republic of China, the National Natural Science Foundation of China, the National Health Commission of the People's Republic of China, the Kadoorie Charitable Foundation, the UK Medical Research Council, Cancer Research UK, and the Wellcome Trust. Unless specified elsewhere, these funders did not support the research required for the preparation of this paper.

Data Availability

The data were obtained from the China Kadoorie Biobank (CKB; request numbers DAR-2025-00208 and DAR-2025-00316). The data are available upon application and were accessed in accordance with the CKB data management guidelines.

Authors' Contributions

YY and HZ contributed to the study design, literature selection, data interpretation, and manuscript drafting. XX and ZQ conducted data analyses, interpreted the data, and revised the manuscript. CH conducted literature searches and revised the manuscript. All authors approved the final draft of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary figures and tables, including incidence rates of esophageal cancer stratified by age group and alcohol consumption, associations between age of alcohol initiation and esophageal cancer risk, and weighted hazard ratio analyses.
[DOCX File, 92 KB - publichealth_v12i1e92949_app1.docx]

References

1. Enzinger PC, Mayer RJ. Esophageal cancer. N Engl J Med 2003 Dec 4;349(23):2241-2252. [doi: 10.1056/NEJMra035010] [Medline: 14657432]
2. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74(3):229-263. [doi: 10.3322/caac.21834] [Medline: 38572751]
3. Jiao Y, Wang T, Fu L, et al. Trends, patterns, and risk factors of esophageal cancer mortality in China, 2008-2021: a National Mortality Surveillance System data analysis. J Adv Res 2026 Feb;80:825-834. [doi: 10.1016/j.jare.2025.05.021] [Medline: 40381911]
4. Han B, Zheng R, Zeng H, et al. Cancer incidence and mortality in China, 2022. J Natl Cancer Cent 2024;4(1):47-53. [doi: 10.1016/j.jncc.2024.01.006] [Medline: 39036382]

XSL•FO
RenderX

5. Zhu H, Ma X, Ye T, et al. Esophageal cancer in China: practice and research in the new era. Int J Cancer 2023 May 1;152(9):1741-1751. [doi: 10.1002/ijc.34301] [Medline: 36151861]
6. Farrow NE, Raman V, Jawitz OK, et al. Impact of age on surgical outcomes for locally advanced esophageal cancer. Ann Thorac Surg 2021 Mar;111(3):996-1003. [doi: 10.1016/j.athoracsur.2020.06.055] [Medline: 32853569]
7. Wang S, Zheng R, Arnold M, et al. Global and national trends in the age-specific sex ratio of esophageal cancer and gastric cancer by subtype. Int J Cancer 2022 Nov 1;151(9):1447-1461. [doi: 10.1002/ijc.34158] [Medline: 35678331]
8. Drahos J, Xiao Q, Risch HA, et al. Age-specific risk factor profiles of adenocarcinomas of the esophagus: a pooled analysis from the international BEACON consortium. Int J Cancer 2016 Jan 1;138(1):55-64. [doi: 10.1002/ijc.29688] [Medline: 26175109]
9. Guo D, Jin J, Li D, He Y, Lin Y. Analysis of the incidence and mortality trends of esophageal cancer in cancer registry areas of China and Japan. Int J Cancer 2024 Oct 15;155(8):1376-1386. [doi: 10.1002/ijc.35003] [Medline: 38771567]
10. Jiang Q, Shu Y, Jiang Z, et al. Burdens of stomach and esophageal cancer from 1990 to 2019 and projection to 2030 in China: findings from the 2019 Global Burden of Disease Study. J Glob Health 2024 Jan 5;14:04025. [doi: 10.7189/jogh.14.04025] [Medline: 38180951]
11. Zhang JG, Xu HF, Chen Q, et al. Time-trend of the incidence and mortality of esophageal cancer from 2010 to 2018 and its statistics in 2018 in Henan, China. Ann Transl Med 2022 Aug;10(16):899. [doi: 10.21037/atm-22-4027] [Medline: 36111000]
12. Li B, Liu Y, Peng J, Sun C, Rang W. Trends of esophageal cancer incidence and mortality and its influencing factors in China. Risk Manag Healthc Policy 2021;14:4809-4821. [doi: 10.2147/RMHP.S312790] [Medline: 34876863]
13. Zhao YX, Zhao HP, Zhao MY, et al. Latest insights into the global epidemiological features, screening, early diagnosis and prognosis prediction of esophageal squamous cell carcinoma. World J Gastroenterol 2024 May 28;30(20):2638-2656. [doi: 10.3748/wjg.v30.i20.2638] [Medline: 38855150]
14. Chen R, Zheng R, Zhang S, et al. Patterns and trends in esophageal cancer incidence and mortality in China: an analysis based on cancer registry data. J Natl Cancer Cent 2023;3(1):21-27. [doi: 10.1016/j.jncc.2023.01.002] [Medline: 39036314]
15. Ji N, Wang Y, Sang Y, Zeng X, Liu Y, Liu S. Alcohol use prevalence among Chinese people aged 15 years and above - China, 2024. China CDC Wkly 2026 Jan 9;8(2):35-42. [doi: 10.46234/ccdcw2026.006] [Medline: 41551334]
16. Yang X, Chen X, Zhuang M, et al. Smoking and alcohol drinking in relation to the risk of esophageal squamous cell carcinoma: a population-based case-control study in China. Sci Rep 2017;7(1):17249. [doi: 10.1038/s41598-017-17617-2]
17. Chen Z, Lee L, Chen J, et al. Cohort profile: the Kadoorie Study of Chronic Disease in China (KSCDC). Int J Epidemiol 2005 Dec;34(6):1243-1249. [doi: 10.1093/ije/dyi174] [Medline: 16131516]
18. Chen Z, Chen J, Collins R, et al. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. Int J Epidemiol 2011 Dec;40(6):1652-1666. [doi: 10.1093/ije/dyr120] [Medline: 22158673]
19. Li LM, Lv J, Guo Y, et al. The China Kadoorie Biobank: related methodology and baseline characteristics of the participants [Article in Chinese]. Zhonghua Liu Xing Bing Xue Za Zhi 2012 Mar;33(3):249-255. [Medline: 22613372]
20. GBD 2019 Lip, Oral, and Pharyngeal Cancer Collaborators, Cunha AR, Compton K, et al. The global, regional, and national burden of adult lip, oral, and pharyngeal cancer in 204 countries and territories: a systematic analysis for the Global Burden of Disease Study 2019. JAMA Oncol 2023 Oct 1;9(10):1401-1416. [doi: 10.1001/jamaoncol.2023.2960] [Medline: 37676656]
21. Millwood IY, Li L, Smith M, et al. Alcohol consumption in 0.5 million people from 10 diverse regions of China: prevalence, patterns and socio-demographic and health-related correlates. Int J Epidemiol 2013 Jun;42(3):816-827. [doi: 10.1093/ije/dyt078] [Medline: 23918852]
22. Huang FL, Yu SJ. Esophageal cancer: risk factors, genetic association, and treatment. Asian J Surg 2018 May;41(3):210-215. [doi: 10.1016/j.asjsur.2016.10.005] [Medline: 27986415]
23. Cao Y, Willett WC, Rimm EB, Stampfer MJ, Giovannucci EL. Light to moderate intake of alcohol, drinking patterns, and risk of cancer: results from two prospective US cohort studies. BMJ 2015 Aug 18;351:h4238. [doi: 10.1136/bmj.h4238] [Medline: 26286216]
24. Li P, Jing J, Liu W, Wang J, Qi X, Zhang G. Spatiotemporal patterns of esophageal cancer burden attributable to behavioral, metabolic, and dietary risk factors from 1990 to 2019: longitudinal observational study. JMIR Public Health Surveill 2023 Oct 6;9:e46051. [doi: 10.2196/46051] [Medline: 37801354]
25. Yan D, Wei L, Liang H, Zhang J, Shi D, Wang Z. Burden of oesophageal cancer attributed to alcohol use in 204 countries and territories, 1990-2019: results from the Global Burden of Disease Study 2019. BMJ Open 2025 Feb 13;15(2):e086343. [doi: 10.1136/bmjopen-2024-086343] [Medline: 39947819]
26. Fan Y, Yuan JM, Wang R, Gao YT, Yu MC. Alcohol, tobacco, and diet in relation to esophageal cancer: the Shanghai Cohort Study. Nutr Cancer 2008;60(3):354-363. [doi: 10.1080/01635580701883011] [Medline: 18444169]
27. Ilic I, Zivanovic Macuzic I, Ravic-Nikolic A, Ilic M, Milicic V. Global burden of esophageal cancer and its risk factors: a systematic analysis of the Global Burden of Disease Study 2019. Life (Basel) 2024;15(1):24. [doi: 10.3390/life15010024]
28. Li S, Chen H, Man J, et al. Changing trends in the disease burden of esophageal cancer in China from 1990 to 2017 and its predicted level in 25 years. Cancer Med 2021 Mar;10(5):1889-1899. [doi: 10.1002/cam4.3775] [Medline: 33586344]

XSL•FO
RenderX

29. Kou K, Baade PD, Gatton M, et al. Individual- and area-level socioeconomic inequalities in esophageal cancer survival in Shandong Province, China: a multilevel analysis. Cancer Epidemiol Biomarkers Prev 2019 Sep;28(9):1427-1434. [doi: 10.1158/1055-9965.EPI-19-0203] [Medline: 31239265]
30. He J, Chen WQ, Li ZS, et al. China guideline for the screening, early detection and early treatment of esophageal cancer (2022, Beijing) [Article in Chinese]. Zhonghua Zhong Liu Za Zhi 2022 Jun 23;44(6):491-522. [doi: 10.3760/cma.j.cn112152-20220517-00348] [Medline: 35754225]
31. Vantanasiri K, Kamboj AK, Kisiel JB, Iyer PG. Advances in screening for Barrett esophagus and esophageal adenocarcinoma. Mayo Clin Proc 2024 Mar;99(3):459-473. [doi: 10.1016/j.mayocp.2023.07.014] [Medline: 38276943]
32. Fitzgerald RC, di Pietro M, Ragunath K, et al. British Society of Gastroenterology guidelines on the diagnosis and management of Barrett's oesophagus. Gut 2014 Jan;63(1):7-42. [doi: 10.1136/gutjnl-2013-305372] [Medline: 24165758]

Abbreviations

aHR: adjusted hazard ratio
BEACON: Barrett's and Esophageal Adenocarcinoma Consortium
CKB: China Kadoorie Biobank
EAC: esophageal adenocarcinoma
ESCC: esophageal squamous cell carcinoma
GBD: Global Burden of Disease
***ICD-10*:** *International Classification of Diseases, Tenth Revision*

Edited by A Mavragani, T Sanchez; submitted 05.Feb.2026; peer-reviewed by R Gongye, Y Wang; revised version received 27.Apr.2026; accepted 30.Apr.2026; published 03.Jun.2026.

Please cite as:
Yang Y, Xing X, Qin Z, Han C, Zhu H
Alcohol Consumption and Age-Specific Risk of Esophageal Cancer: Prospective Cohort Study
JMIR Public Health Surveill 2026;12:e92949
URL: https://publichealth.jmir.org/2026/1/e92949
doi:10.2196/92949

© Yongbo Yang, Xing Xing, Zhen Qin, Chunyang Han, He Zhu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 3.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Determining the Association Between Hearing Disability and Injury Risk in Older Adults Using Propensity Score Matching: Quasi-Experimental Study

Woo-Ri Lee[1], PhD; Sungyoun Chun[2], PhD; Youyoung An[3], MD; Hyun Seung Choi[3], MD

[1]Division of Cancer Control and Policy, National Cancer Control Institute, National Cancer Center, Goyang, South Korea

[2]Department of Research and Analysis, National Health Insurance Service Ilsan Hospital, Goyang, South Korea

[3]Department of Otorhinolaryngology, National Health Insurance Service Ilsan Hospital, Ilsan-ro 100, Goyang, South Korea

Corresponding Author:
Hyun Seung Choi, MD
Department of Otorhinolaryngology, National Health Insurance Service Ilsan Hospital, Ilsan-ro 100, Goyang, South Korea

Abstract

Background: With rapid population aging, both traumatic injuries and hearing disability have become increasingly prevalent among older adults. Hearing disability may increase vulnerability to injury through impaired balance, reduced environmental awareness, and functional limitations; however, longitudinal evidence examining the association between hearing disability and injury risk remains limited.

Objective: This study aimed to examine the association between hearing disability and the risk of injury among older adults using a quasi-experimental design with propensity score matching (PSM).

Methods: This population-based cohort study included individuals aged 60 years and older with hearing disabilities and a matched control group without disabilities using data from the National Health Insurance Service–Senior cohort from 2008 to 2019. Injury admission, defined using *International Classification of Diseases* codes S00 to S99, was the primary outcome. A quasi-experimental design was applied using PSM at a 1:3 ratio to balance baseline characteristics between the hearing-disabled and nondisabled groups. Cox proportional hazard regression models adjusted for all covariates were used to estimate hazard ratios. Sensitivity analyses were conducted according to disability severity and injury site.

Results: The total number of participants was 43,944, with 10,986 (25%) in the hearing-disabled group and 32,958 (75%) in the nondisabled group, thus confirming a 1:3 matching ratio. The PSM results showed that the standardized mean difference values for all covariates were below the absolute value of 0.1, thus indicating that PSM was successfully performed. The incidence of injury admissions was higher in the hearing-disabled group (1567/10,986, 14.3% of patients) than in the nondisabled group (3966/32,958, 12%), and this difference was statistically significant ($P<.001$). During the follow-up period, older adults with hearing disability had a significantly higher risk of injury admission compared with those without hearing disability (hazard ratio=1.21, 95% CI 1.14 - 1.28; $P<.001$). The association was stronger among individuals with more severe hearing disability and varied by injury site.

Conclusions: Hearing disability in older adults is independently associated with increased injury admission risk, with greater severity conferring a higher risk and variation by injury site. Interventions such as hearing aid provision, targeted traffic safety measures, and enhanced community and family support are warranted to mitigate this burden.

(JMIR Public Health Surveill 2026;12:e78826) doi:10.2196/78826

KEYWORDS

hearing disability; injury; fall; older patients; propensity score matching

Introduction

The proportion of older individuals is increasing worldwide. By 2050, individuals aged 65 years and older are projected to account for approximately 15.9% of the total population [1]. In South Korea, the proportion of older adults has exceeded 20% as of 2024, thus officially classifying the country as a superaged society [2]. As a consequence of this demographic shift, traumatic injury in older adults has emerged as a significant public health concern [3]. Geriatric injuries are characterized by distinct epidemiological patterns in which ground-level falls and motor vehicle crashes are the predominant mechanisms of injury, frequently resulting in predictable injury types such as fractures (particularly of the hip and spine), head trauma, and thoracic injuries among older adults [4]. A study that analyzed the trends in trauma incidence using the National Trauma Data Bank (United States) and TraumaRegister DGU (Germany)

XSL•FO
RenderX

reported an increasing trend in geriatric injury cases in both datasets [5]. According to data from the Korea Centers for Disease Control and Prevention, approximately 320,000 older adults sustained injuries in 2024, which accounted for 38.4% of all injury cases in the country [6]. Prior studies have identified traffic accidents and falls as the leading causes of injury among older adults [3,7-9], with geriatric injury being one of the major contributors to preventable mortality [10].

In parallel with the aging population, the prevalence of hearing disability is also increasing worldwide [11]. In South Korea, approximately 24.9% of registered individuals with disabilities aged 65 years and older had a hearing disability as of 2023, and this number is expected to continue growing [12]. Hearing disability in older adults extends beyond auditory difficulties and negatively affects physical and social functioning [11]. From a functional perspective, hearing impairment may influence balance control and situational awareness, which are essential for injury prevention in daily life. One study demonstrated that individuals with hearing loss experience reduced awareness of environmental sounds and warning signals while walking, thereby increasing the risk of falls [12]. Additionally, the social and psychological consequences of hearing loss, including isolation and depression, are associated with decreased physical activity, which may further impair balance and gait function [13]. Current studies have reported that consistent hearing aid use is associated with a lower prevalence and risk of falls among older adults, suggesting a potential link between hearing-related functional limitations and injury risk [14]. Furthermore, longitudinal evidence indicates that hearing loss is associated with an increased risk of falls and fall-related injuries in older populations [15].

However, the number of studies directly examining the association between hearing disability and injury in older adults remains limited. Most studies have used cross-sectional designs or relatively small sample sizes, thus restricting the ability to establish causal inferences [16-18]. Furthermore, many studies have had limited ability to adequately control for baseline differences between individuals with and without hearing disability. Therefore, this study aimed to examine the association between hearing disability and the occurrence of injury among older adults, and we hypothesized that older adults with hearing disability would have a higher risk of injury than those without hearing disability.

Methods

Study Data

We used the National Health Insurance Service (NHIS)–Senior cohort data provided by the NHIS between 2002 and 2019. South Korea operates a single-payer national health insurance (NHI) system; the NHIS, as the sole insurer, provides NHI coverage for 97% of the population and Medical Aid program coverage for the remaining 3%, thus ensuring universal health care coverage [19]. As NHIS enrollees, all citizens pay monthly insurance premiums based on their income and assets. When receiving medical services, patients pay a portion of their total medical expenses as out-of-pocket payments, and the remaining costs are covered by the NHIS as copayments to medical institutions. To receive copayments, medical institutions must submit reimbursement claims to the NHIS. Given that all medical institutions in South Korea must submit claims to the NHIS, the NHIS database comprehensively stores and manages all medical service use records [20,21]. Therefore, NHIS data can be considered nationally representative. The NHIS-Senior cohort was established using a stratified random sampling method. In 2008, the older adult population aged 60 to 80 years was stratified by age, sex, health insurance premium, and region, and 8% of the individuals were randomly selected from each stratum [22]. Accordingly, in this study, older adults were defined as individuals aged 60 years or older in accordance with the sampling framework of the NHIS-Senior cohort.

Variables

Outcome Measures

The main outcome measure of this study was injury admission. Injury admissions were recorded between 2008 and 2019. Injury admission was defined based on the 10th revision of the *International Classification of Diseases*, where hospitalization with principal diagnosis codes S00 to S99 was classified as an injury admission.

Independent Variable

The key independent variable in this study was hearing disability status. A diagnosis of hearing disability is defined in the Medical Service Act in South Korea and requires an evaluation by an otolaryngologist at a medical institution equipped with a soundproof booth [23]. Hearing disability is classified based on pure-tone audiometry thresholds at 4 frequencies (0.5, 1, 2, and 4 kHz); hearing loss of 80 dB or higher in one ear and 40 dB or higher in the other ear qualifies as hearing disability, and thresholds of 80 dB or higher in both ears indicate severe hearing disability [23]. The severity of disability in the Korea National Disability Registration System (KNDRS) is classified into 6 grades; for hearing disability, grades 2 and 3 indicate severe impairment, and grades 4 to 6 indicate mild impairment [23,24].

Since 1989, South Korea has used the KNDRS to register individuals with disabilities [23]. The NHIS database has been linked to the KNDRS to provide disability status information. Therefore, in this study, hearing disability status was identified using linked data [23].

Covariates

The following variables were used as covariates: sex (male and female), age group (60s, 70s, and >80 years), income level (quintile), type of health care insurance (Medical Aid, NHI self-employed, and NHI employee), region (Seoul, Gyeonggi, metropolitan, and rural), Charlson Comorbidity Index score (0, 1, 2, and ≥3), and index year. Age was included as a categorical variable to improve interpretability, account for potential nonlinear associations, and facilitate covariate balance in the propensity score matching (PSM).

Statistical Analysis

The analysis was conducted in 5 steps. First, PSM was performed to ensure homogeneity between the case and control groups. PSM is a quasi-experimental method that reduces selection bias and mimics a randomized clinical trial in

XSL•FO
RenderX

observational studies [25]. By using binomial logistic regression, PSM calculates the probability value corresponding to the independent variable. The computed probability value was referred to as the propensity score. Matching was performed using the greedy method at a 1:3 ratio with replacement, with a caliper of 0.1. Exact matching was not applied. All the covariates included in the study (sex, age, income, type of health care insurance, region, Charlson Comorbidity Index, and index year) were matched. After PSM, standardized mean difference values below an absolute value of 0.1 were considered to indicate homogeneity between the case and control groups. For binary variables, standardized mean difference values were calculated as the standardized differences in proportions between groups. Second, the chi-square test was used to examine the relationship between the participants' general characteristics and injury admission. Statistical significance was set at a value of P below .05. Third, survival periods and incidence rates were calculated according to the participants' general characteristics. Differences in survival periods were assessed using 2-tailed t tests and ANOVA, and statistical significance was set at a value of $P<.05$. The incidence rates were calculated as the number of events per 100,000 person-days. Fourth, regression analysis using the Cox proportional hazard model was conducted to evaluate the association between hearing disability and injury admission. The proportional hazard assumption was verified using Kaplan-Meier survival curves, and the differences in survival curves according to hearing disability status were tested using the log-rank test. Statistical significance for the log-rank test and Cox regression analysis was set at a value of $P<.05$. Finally, sensitivity analysis was performed to enhance the robustness of the results. Sensitivity analysis was conducted using two approaches: (1) subgroup analysis, in which hearing disability was classified into mild and severe and wherein each subgroup was compared with the control group; and (2) site-specific injury analysis, in which injury admissions were categorized based on the affected body region. All regression analyses were adjusted for all covariates used in PSM to improve robustness [25]. All analyses were performed using SAS (version 9.4; SAS Institute) and R software (version 3.6.3; R Foundation for Statistical Computing).

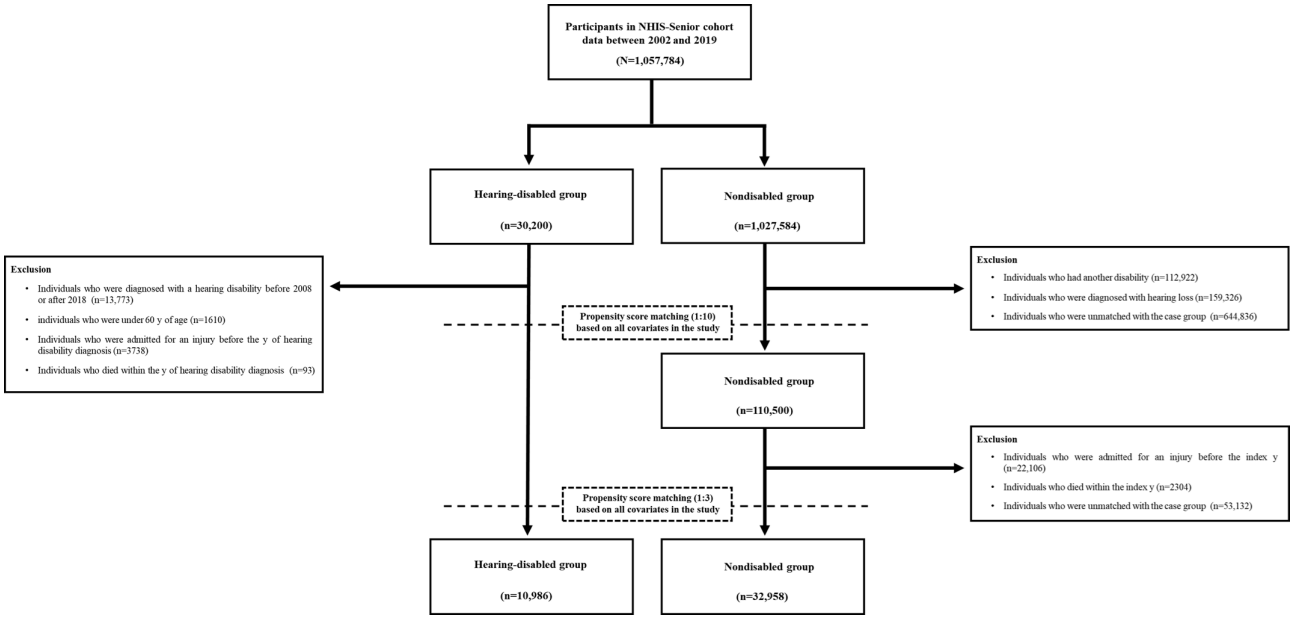
Ethical Considerations

Ethical approval for this study was waived by the institutional review board of the NHIS Ilsan Hospital, South Korea (NHIMC 2025-03-027), because the study used secondary data containing anonymized and encrypted personal information. The requirement for informed consent was also waived. The data were analyzed in a deidentified form to protect the privacy and confidentiality of individuals. As the study used anonymized secondary data, informed consent and participant compensation were not applicable.

Results

The NHIS-Senior cohort initially included 1,057,784 individuals. The hearing-disabled group (case group) and nondisabled group (control group) consisted of 30,200 and 1,027,584 individuals, respectively. The exclusion criteria were applied to define the final study population. In the nondisabled group, the following individuals were excluded: those with other disabilities (112,922/1,027,584, 11%), those with a history of hearing loss (159,326/1,027,584, 15.5%), those excluded from PSM for the index year matching (644,836/1,027,584, 62.8%), those with injury-related hospital admissions before the index year (22,106/1,027,584, 2.2%), those who died in the index year (2304/1,027,584, 0.2%), and those further excluded from PSM (53,132/1,027,584, 5.2%). In the hearing-disabled group, the following individuals were excluded: those diagnosed with hearing disability before 2008 or after 2018 (13,773/30,200, 45.6%), those younger than 60 years (1610/30,200, 5.3%), those with injury-related hospital admissions before hearing disability diagnosis (3738/30,200, 12.4%), and those who died within the year of hearing disability diagnosis (93/30,200, 0.3%). After applying the exclusion criteria, the final study population included 10,986 and 38,411 patients in the hearing-disabled and nondisabled group, respectively (Figure 1).

Figure 1. Participant selection flowchart. NHIS: National Health Insurance Service.



XSL•FO
RenderX

Multimedia Appendix 1 shows the general characteristics of the study population before and after PSM. After PSM, the total number of study participants was 43,944, with 10,986 (25%) in the hearing-disabled group and 32,958 (75%) in the nondisabled group, thus confirming a 1:3 matching ratio. The PSM results showed that the standardized mean difference values for all covariates were below the absolute value of 0.1, thus indicating that PSM was successfully performed.

Table 1 shows the relationships between the participants' general characteristics and admissions for injury. Of the 43,944 participants, 5533 (12.6%) experienced admissions for injuries. The incidence of injury admissions was higher in the hearing-disabled group (1567/10,986, 14.3% of the participants) than in the nondisabled group (3966/32,958, 12%), and this difference was statistically significant ($P<.001$).

XSL•FO
RenderX

Table . General characteristics of the study population.

Variables	Injury admission, n (%)		*P* value[a]
	Yes	No	
Total (n=43,944)	5533 (12.6)	38,411 (87.4)	__[b]
Hearing disability			<.001
Nondisabled (n=32,958)	3966 (12.0)	28,992 (88.0)	
Disabled (n=10,986)	1567 (14.3)	9419 (85.7)	
Severity of hearing disability			<.001
Nondisabled (n=32,750)	3932 (12.0)	28,818 (88.0)	
Mildly disabled (n=9923)	1340 (13.5)	8583 (86.5)	
Severely disabled (n=1271)	261 (20.5)	1010 (79.5)	
Sex			<.001
Male (n=24,348)	2354 (9.7)	21,994 (90.3)	
Female (n=19,596)	3179 (16.2)	16,417 (83.8)	
Age (years)			<.001
60s (n=11,355)	1363 (12.0)	9992 (88.0)	
70s (n=20,660)	2865 (13.9)	17,795 (86.1)	
Older than 80 (n=11,929)	1305 (10.9)	10,624 (89.1)	
Income			.20
Quintile 1 (lowest; n=10,565)	1397 (13.2)	9168 (86.8)	
Quintile 2 (n=4326)	531 (12.3)	3795 (87.7)	
Quintile 3 (n=5689)	713 (12.5)	4976 (87.5)	
Quintile 4 (n=8242)	1042 (12.6)	7200 (87.4)	
Quintile 5 (highest; 15,122)	1850 (12.2)	13,272 (87.8)	
Type of health care insurance			<.001
Medical Aid (n=3976)	601 (15.1)	3375 (84.9)	
NHI[c] self-employed (n=11,525)	1496 (13.0)	10,029 (87.0)	
NHI employee (n=28,443)	3436 (12.1)	25,007 (87.9)	
Region			<.001
Seoul (n=6837)	748 (10.9)	6089 (89.1)	
Gyeonggi (n=7296)	783 (10.7)	6513 (89.3)	
Metropolitan (n=11,238)	1185 (10.5)	10,053 (89.5)	
Rural (n=18,573)	2817 (15.2)	15,756 (84.8)	
CCI[d]			<.001
0 (n=17,663)	2234 (12.6)	15,429 (87.4)	
1 (n=9511)	1298 (13.6)	8213 (86.4)	
2 (n=6802)	804 (11.8)	5998 (88.2)	
≥3 (n=9968)	1197 (12.0)	8771 (88.0)	
Index year			<.001
2008 (n=4753)	1181 (24.8)	3572 (75.2)	
2009 (n=4899)	1143 (23.3)	3756 (76.7)	
2010 (n=4066)	829 (20.4)	3237 (79.6)	
2011 (n=1814)	323 (17.8)	1491 (82.2)	

XSL•FO
RenderX

Variables	Injury admission, n (%)		*P* value[a]
	Yes	No	
2012 (n=1246)	212 (17.0)	1034 (83.0)	
2013 (n=1248)	168 (13.5)	1080 (86.5)	
2014 (n=1227)	165 (13.4)	1062 (86.6)	
2015 (n=1564)	175 (11.2)	1389 (88.8)	
2016 (n=5699)	542 (9.5)	5157 (90.5)	
2017 (n=7477)	456 (6.1)	7021 (93.9)	
2018 (n=9951)	339 (3.4)	9612 (96.6)	

[a]*P* values were obtained using the chi-square test.

[b]Not applicable.

[c]NHI: national health insurance.

[d]CCI: Charlson Comorbidity Index.

Table 2 shows the analysis comparing injury admission incidence rates and survival times according to the participants' general characteristics. The overall incidence rate was 8.4 cases per 100,000 person-days, and the mean survival time was 1497 (SD 1230) days. The hearing-disabled group had an incidence rate of 9.7 cases per 100,000 person-days, which was higher than that in the nondisabled group (8.0 cases per 100,000 person-days). However, the difference in survival time between the 2 groups was not statistically significant.

XSL•FO
RenderX

Table . Incidence rate per 100,000 person-years and survival time.

Variables	Injury admission		*P* value
	Incidence rate	Survival time (days), mean (SD)	
Total	8.4	1497 (1230)	__[a]
Hearing disability			.051
Nondisabled	8.0	1503 (1236)	
Disabled	9.7	1477 (1212)	
Hearing disability severity			<.001
Nondisabled	8.0	1505 (1237)	
Mildly disabled	9.5	1423 (1190)	
Severely disabled	11.0	1864 (1292)	
Sex			.08
Male	6.4	1506 (1226)	
Female	10.9	1485 (1235)	
Age (y)			<.001
60s	5.8	2081 (1336)	
70s	9.0	1545 (1223)	
Older than 80	12.8	856 (737)	
Income			<.001
Quintile 1 (lowest)	9.0	1474 (1211)	
Quintile 2	8.0	1529 (1230)	
Quintile 3	8.1	1555 (1255)	
Quintile 4	8.0	1582 (1270)	
Quintile 5 (highest)	8.5	1435 (1208)	
Type of health care insurance			<.001
Medical Aid	11.0	1371 (1173)	
NHI[b] self-employed	8.3	1571 (1254)	
NHI employee	8.1	1484 (1226)	
Region			<.001
Seoul	6.8	1620 (1275)	
Gyeonggi	7.2	1499 (1240)	
Metropolitan	7.4	1426 (1222)	
Rural	10.2	1493 (1211)	
CCI[c]			<.001
0	7.3	1725 (1301)	
1	8.9	1527 (1218)	
2	8.7	1364 (1173)	
≥3	10.4	1153 (1046)	
Index year			<.001
2008	8.3	2982 (1300)	
2009	8.5	2752 (1176)	
2010	7.9	2572 (1020)	
2011	7.7	2327 (902)	

XSL•FO
RenderX

Variables	Injury admission		*P* value
	Incidence rate	Survival time (days), mean (SD)	
2012	8.2	2082 (765)	
2013	7.2	1876 (590)	
2014	8.4	1605 (459)	
2015	8.4	1325 (328)	
2016	9.5	1002 (235)	
2017	8.8	692 (125)	
2018	9.5	359 (39)	

[a]Not applicable.

[b]NHI: national health insurance.

[c]CCI: Charlson Comorbidity Index.

Table 3 shows the Cox regression analysis exploring the association between hearing disability and injury admission. The Kaplan-Meier survival curve demonstrated a statistically significant difference in the according to hearing disability status (Figure 2; $P<.001$). The Cox regression analysis indicated that the injury admission risk was higher in the hearing-disabled group than in the nondisabled group (hazard ratio [HR] 1.21, 95% CI 1.14 - 1.28).

XSL•FO
RenderX

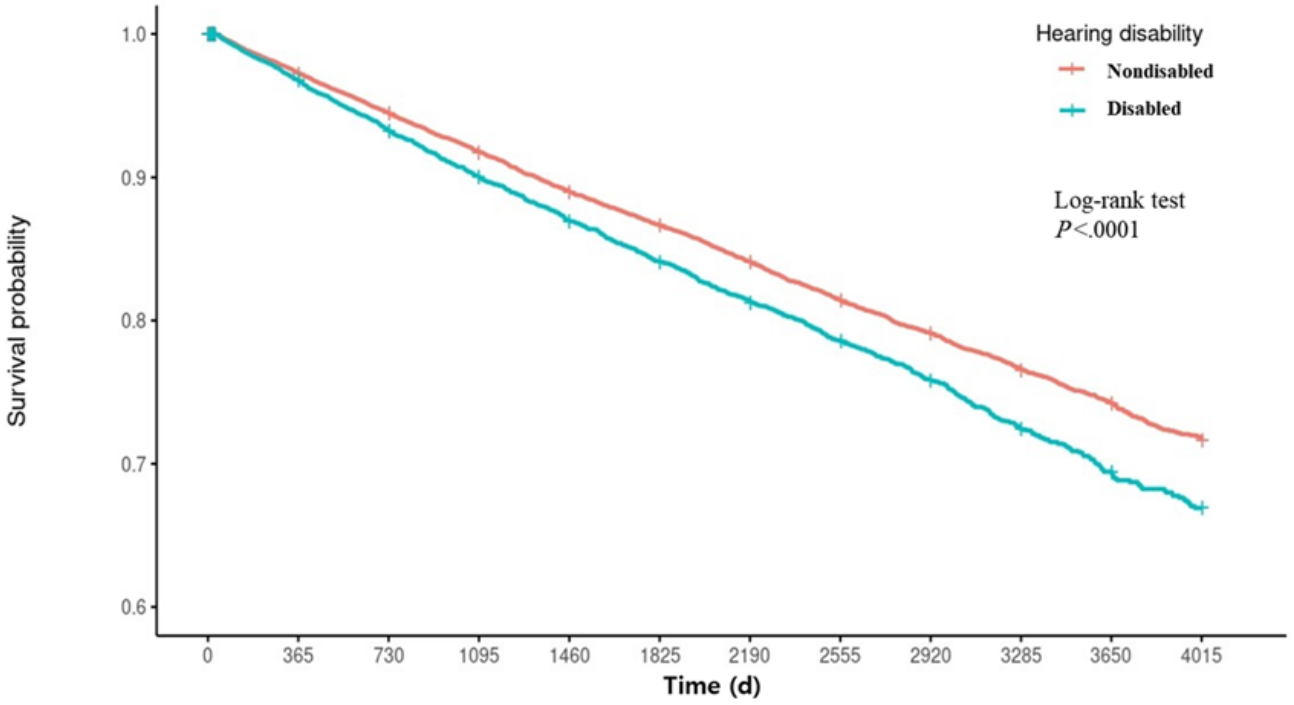
Table . Results of the multivariable Cox regression analysis on the association between hearing disability and injury admission.

Variables	Injury admission, HR[a] (95% CI)	*P* value
Hearing disability		
Nondisabled	1.00 (reference)	—[b]
Disabled	1.21 (1.14 - 1.28)	<.001
Sex		
Male	1.00 (reference)	—
Female	1.56 (1.48 - 1.65)	<.001
Age (years)		
60s	1.00 (reference)	—
70s	1.49 (1.39 - 1.59)	<.001
Older than 80	2.11 (1.93 - 2.29)	<.001
Income		
Quintile 1 (lowest)	1.00 (reference)	—
Quintile 2	1.05 (0.94 - 1.17)	.40
Quintile 3	1.03 (0.93 - 1.14)	.57
Quintile 4	1.04 (0.95 - 1.14)	.37
Quintile 5 (highest)	1.02 (0.94 - 1.11)	.70
Type of health care insurance		
Medical Aid	1.00 (reference)	—
NHI[c] self-employed	0.89 (0.79 - 0.99)	.04
NHI employee	0.87 (0.78 - 0.97)	.01
Region		
Seoul	1.00 (reference)	—
Gyeonggi	1.05 (0.95 - 1.17)	.30
Metropolitan	1.09 (0.99 - 1.20)	.06
Rural	1.42 (1.31 - 1.54)	<.001
CCI[d]		
0	1.00 (reference)	—
1	1.19 (1.11 - 1.28)	<.001
2	1.18 (1.09 - 1.28)	<.001
≥3	1.40 (1.30 - 1.50)	<.001
Index year	1.00 (0.99 - 1.01)	.79

[a]HR: hazard ratio.
[b]Not applicable.
[c]NHI: national health insurance.
[d]CCI: Charlson Comorbidity Index.

XSL•FO
RenderX

Figure 2. Kaplan-Meier survival curve according to hearing disability status.



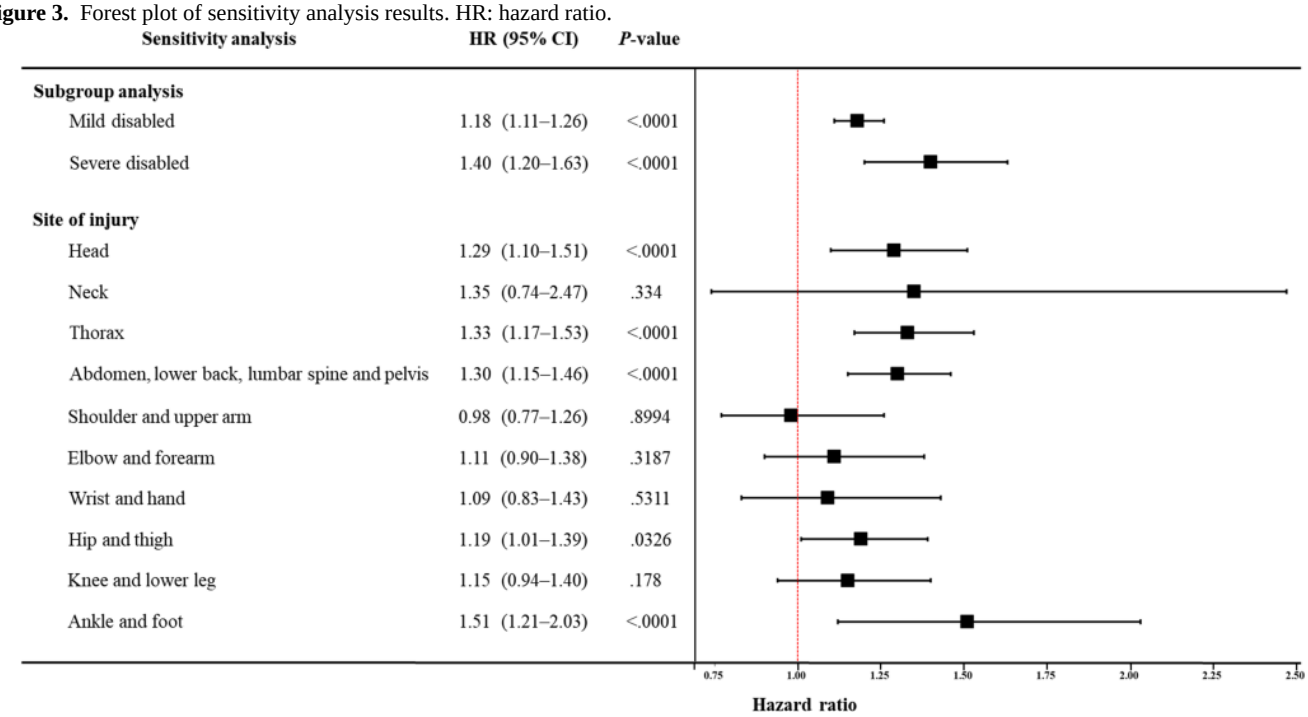
Number at risk

Nondisabled	32958	31982	23227	17091	12737	11096	9772	8497	7391	6088	3997	1869
Disabled	10986	10620	7725	5654	4173	3605	3146	2730	2339	1892	1209	555

A sensitivity analysis was conducted to enhance the robustness of the findings (Figure 3). The results showed that individuals with hearing disability had a higher injury admission risk regardless of severity (mild disability: HR 1.18, 95% CI 1.11 - 1.26; severe disability: HR 1.40, 95% CI 1.20 - 1.63). A further analysis of site-specific injury risk revealed statistically significant differences across multiple body regions (head: HR 1.29, 95% CI 1.10 - 1.51; thorax: HR 1.33, 95% CI 1.17 - 1.53; abdomen, lower back, lumbar spine, and pelvis: HR 1.30, 95% CI 1.15 - 1.46; hips and thighs: HR 1.19, 95% CI 1.01 - 1.39; ankles and feet: HR 1.51, 95% CI 1.21 - 2.03). These anatomical injury sites commonly correspond to clinically meaningful injury types in geriatric populations, such as head injuries, fractures of the hip or lower extremities, and other fall-related blunt injuries. Such injury patterns are consistent with mechanisms frequently observed among older adults, particularly falls and traffic-related trauma.

XSL•FO
RenderX

Figure 3. Forest plot of sensitivity analysis results. HR: hazard ratio.



Discussion

Principal Findings

This study aimed to explore the association between hearing disability and injury admissions among older adults. The analysis revealed that older adults with hearing disability were at a significantly higher risk of injury. Moreover, the association was more pronounced with greater disability severity, and the influence varied according to the injury site.

Interpretation

Mechanisms Linking Hearing Disability and Injury

Several mechanisms may explain the observed association between hearing disability and injury risk among older adults. Hearing disability has been associated with accelerated cognitive decline, including impairments in memory, attention, and executive function [26,27]. Such cognitive changes may reduce the ability to detect environmental hazards and appropriately respond to warning signals, thereby increasing susceptibility to falls and other injuries [12]. In addition, auditory deprivation may compromise balance control and spatial orientation, which are essential for safe mobility during daily activities, further increasing injury risk [13]. Taken together, these mechanisms suggest that addressing cognitive and balance-related impairments associated with hearing disability, as well as improving hearing levels, may represent an important component of injury prevention strategies for older adults.

Hearing Aid Use and Barriers to Adoption

Functional restoration of auditory capacity may help reduce injury risk among older adults with hearing disability. Hearing aid use has been shown to improve auditory perception as well as depressive symptoms, emotional stability, and social functioning [28-30]. A previous study further indicates that hearing aid use can improve social relationships and overall quality of life among older adults with age-related hearing loss, highlighting social engagement as an important pathway through which auditory rehabilitation may influence broader health outcomes [31]. By enhancing environmental awareness and reducing cognitive load, hearing aids may also decrease the risk of falls and related injuries [32]. Despite these benefits, approximately 30% of individuals with hearing disability in South Korea do not use hearing aids [33]. Barriers to hearing aid use may include financial burden, stigma associated with device use, discomfort, and limited awareness of potential benefits [31,34]. Addressing these barriers may be critical for reducing injury risk in this population. These findings suggest that promoting hearing aid adoption may represent a practical intervention to help reduce injury risk in this vulnerable population.

Psychosocial Pathways and Social Support

Beyond direct sensory and functional impairments, hearing disability in older adults is also associated with psychosocial challenges, including social isolation and strained interpersonal relationships [28,35-38]. Reduced social engagement may lead to lower levels of physical activity and decreased participation in daily routines, both of which are known risk factors for falls and injury. Previous studies have shown that stronger social relationships within families and communities are associated with lower rates of fall-related injuries among older adults [39,40]. These findings suggest that social support may play a protective role in mitigating injury risk among individuals with hearing disability. Accordingly, interventions that strengthen social support and promote social engagement may help reduce injury risk in older adults with hearing disability.

Injury Patterns and Traffic-Related Injuries

As the older adult population continues to grow, the incidence of traffic-related injuries among this population is also expected to increase [41]. Older adults are particularly vulnerable to

XSL•FO
RenderX

traffic-related injuries due to slower reaction times, sensory impairments, and a higher likelihood of being pedestrians. Previous studies have identified common injury sites in traffic accidents, including the head, thorax, abdomen, pelvis, and lower extremities [42,43]. Consistent with these findings, our study observed elevated injury risks in older adults with hearing impairment across multiple body regions, including the head, trunk, pelvis, and lower extremities. Furthermore, pedestrians with sensory disabilities, including hearing loss, have been shown to have a higher risk of traffic-related injury and mortality [44]. Reduced perception of environmental sounds and warning signals may further increase vulnerability to traffic accidents among individuals with hearing disability. Together, these findings underscore the importance of incorporating sensory impairment considerations into traffic safety policies and pedestrian protection strategies for older adults.

Implications for Health Care and Policy

These findings have important implications for health care systems and public health policy. First, routine hearing screening and early identification of hearing disability in older adults may facilitate the timely recognition of individuals at increased risk of injury. Second, integrating hearing assessment into fall prevention programs and geriatric care pathways could enable more comprehensive risk management and contribute to reductions in injury-related morbidity by addressing both sensory and functional vulnerabilities. Third, improving access to hearing aids and promoting their use through financial support and public awareness campaigns may serve as practical and effective strategies to mitigate injury risk at the individual level. Finally, beyond clinical and community-based interventions, system-level approaches are also warranted. Targeted traffic safety interventions—such as pedestrian-friendly road design and enhanced auditory warning systems—may further help protect older adults with hearing disability and reduce traffic-related injuries.

Limitations

This study has a few limitations. First, owing to the nature of the claims data, only primary disabilities could be identified, and information on multiple co-occurring disabilities was unavailable. Consequently, individuals with other severe types of disability and mild hearing impairment could not be classified as hearing disabled and were excluded from the analysis. Future studies should incorporate methods that detect and include participants with multiple disabilities. Second, we could not directly ascertain whether the injuries resulted from traffic accidents. Given that injury admissions were defined solely by *International Classification of Diseases, 10th Revision*, diagnosis codes, we were unable to distinguish traffic-related trauma from other causes. To address this limitation, we conducted subgroup analyses according to injury site and compared our results with the major injury locations reported in traffic accident studies. Third, some older adults may have experienced improvements in auditory function through hearing aid use during follow-up; however, based on the available data, we could not assess hearing aid adoption or auditory improvement. Despite these limitations, this study is meaningful because it confirmed the association between hearing disability status and injury incidence among older adults using a longitudinal design.

Conclusions

Our findings demonstrated a clear disparity in injury risk according to hearing disability status in older adults. To reduce these health disparities, it is essential to promote hearing aid use, implement targeted traffic safety measures, and strengthen social support at both the community and family levels.

Funding

This work was supported by the National Health Insurance Service Ilsan Hospital (grant NHIMC-2024-CR-064).

Data Availability

This study used National Health Insurance Service (NHIS)–Senior cohort data (NHIS-2025-04-2-049) collected by the NHIS. The data for this study are public secondary data, and they can be accessed through the following NHIS website by submitting an application form and paying a fee.

Conflicts of Interest

None declared.

Multimedia Appendix 1
General characteristics of the study population before and after propensity score matching.
[DOCX File, 48 KB - publichealth_v12i1e78826_app1.docx]

References

1. Gu D, Andreev K, Dupre ME. Major trends in population growth around the world. China CDC Wkly 2021 Jul 9;3(28):604-613. [doi: 10.46234/ccdcw2021.160] [Medline: 34594946]
2. Resident population in five-year age groups. Korean Statistical Information Service (KOSIS). 2025. URL: https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1B04005N&conn_path=I2&language=en [accessed 2025-03-14]
3. Kehoe A, Smith JE, Edwards A, Yates D, Lecky F. The changing face of major trauma in the UK. Emerg Med J 2015 Dec;32(12):911-915. [doi: 10.1136/emermed-2015-205265] [Medline: 26598629]

XSL•FO
RenderX

4. Welsch JM, Patin BB, Mohammad IY, Palladino-Davis AG, Davis CS. Geriatric injury: a contemporary review of patterns, prevention, and treatment. Curr Emerg Hosp Med Rep 2023;11(3):117-125. [doi: 10.1007/s40138-023-00270-7]
5. Jiang L, Zheng Z, Zhang M. The incidence of geriatric trauma is increasing and comparison of different scoring tools for the prediction of in-hospital mortality in geriatric trauma patients. World J Emerg Surg 2020 Oct 19;15(1):59. [doi: 10.1186/s13017-020-00340-1] [Medline: 33076958]
6. Injury with older ages. Korea Centers for Disease Control & Prevention (KCDC). 2024. URL: https://www.kdca.go.kr/injury/biz/injury/damgInfo/odsnDamgMain.do [accessed 2025-03-14]
7. Khorgami Z, Fleischer WJ, Chen YJA, Mushtaq N, Charles MS, Howard CA. Ten-year trends in traumatic injury mechanisms and outcomes: a trauma registry analysis. Am J Surg 2018 Apr;215(4):727-734. [doi: 10.1016/j.amjsurg.2018.01.008] [Medline: 29397887]
8. Benhamed A, Batomen B, Boucher V, et al. Epidemiology, injury pattern and outcome of older trauma patients: a 15-year study of level-I trauma centers. PLOS ONE 2023;18(1):e0280345. [doi: 10.1371/journal.pone.0280345] [Medline: 36716316]
9. Beck B, Cameron P, Lowthian J, Fitzgerald M, Judson R, Gabbe BJ. Major trauma in older persons. BJS Open 2018 Sep;2(5):310-318. [doi: 10.1002/bjs5.80] [Medline: 30263982]
10. Johnson NB, Hayes LD, Brown K, Hoo EC, Ethier KA, Centers for Disease Control and Prevention (CDC). CDC National Health Report: leading causes of morbidity and mortality and associated behavioral risk and protective factors--United States, 2005-2013. MMWR Suppl 2014 Oct 31;63(4):3-27. [Medline: 25356673]
11. World health organization (WHO). Deafness and hearing loss. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss [accessed 2025-03-14]
12. Jiam NTL, Li C, Agrawal Y. Hearing loss and falls: a systematic review and meta-analysis. Laryngoscope 2016 Nov;126(11):2587-2596. [doi: 10.1002/lary.25927] [Medline: 27010669]
13. Mick P, Kawachi I, Lin FR. The association between hearing loss and social isolation in older adults. Otolaryngol Head Neck Surg 2014 Mar;150(3):378-384. [doi: 10.1177/0194599813518021] [Medline: 24384545]
14. Campos L, Prochazka A, Anderson M, Kaizer A, Foster C, Hullar T. Consistent hearing aid use is associated with lower fall prevalence and risk in older adults with hearing loss. J American Geriatrics Society 2023 Oct;71(10):3163-3171. [doi: 10.1111/jgs.18461]
15. Sakurai R, Kawai H, Suzuki H, et al. Increased risk of falls in older adults with hearing loss and slow gait: results from the Otassha Study. Geroscience 2025 Apr;47(2):2235-2244. [doi: 10.1007/s11357-024-01412-9] [Medline: 39485656]
16. Gür K, Kılınç E, Günaydın E, Gülsün H. The epidemiology of injury among adolescents with hearing loss, health beliefs regarding injury and associated factors. Disabil Health J 2021 Apr;14(2):100994. [doi: 10.1016/j.dhjo.2020.100994] [Medline: 32958403]
17. Mick P, Foley D, Lin F, Pichora-Fuller MK. Hearing difficulty is associated with injuries requiring medical care. Ear Hear 2018;39(4):631-644. [doi: 10.1097/AUD.0000000000000535] [Medline: 29309398]
18. Jeong J, Lim H, Choi HS. Association between hearing loss and trauma based on population data of Korea. Acta Otolaryngol 2023 May;143(5):392-395. [doi: 10.1080/00016489.2023.2204116] [Medline: 37104855]
19. Population coverage. National Health Insurance Services (NHIS). 2025. URL: https://www.nhis.or.kr/english/wbheaa02400m01.do [accessed 2025-03-14]
20. Song YJ. The South Korean health care system. Jpn Med Assoc J 2009;52(3):206-209 [FREE Full text]
21. Health promotion. National Health Insurance Services (NHIS). 2025. URL: https://www.nhis.or.kr/english/wbheaa02700m01.do [accessed 2025-03-14]
22. Sample DB research introduction: elderly cohort database. National Health Insurance Sharing Service (NHIS). 2025. URL: https://nhiss.nhis.or.kr/en/a/a/100/lpaa100m01en.do [accessed 2025-03-14]
23. Kim M, Jung W, Kim SY, Park JH, Shin DW. The Korea National Disability Registration System. Epidemiol Health 2023;45:e2023053. [doi: 10.4178/epih.e2023053] [Medline: 37189275]
24. Jeon B, Noguchi H, Kwon S, Ito T, Tamiya N. Disability, poverty, and role of the basic livelihood security system on health services utilization among the elderly in South Korea. Soc Sci Med 2017 Apr;178:175-183. [doi: 10.1016/j.socscimed.2017.02.013] [Medline: 28237862]
25. Faries D, Zhang X, Kadziola Z, et al. Real World Health Care Data Analysis: Causal Methods and Implementation Using SAS: SAS Institute; 2020.
26. Lin FR, Yaffe K, Xia J, et al. Hearing loss and cognitive decline in older adults. JAMA Intern Med 2013 Feb 25;173(4):293-299. [doi: 10.1001/jamainternmed.2013.1868] [Medline: 23337978]
27. Lin FR, Ferrucci L, Metter EJ, An Y, Zonderman AB, Resnick SM. Hearing loss and cognition in the Baltimore longitudinal study of aging. Neuropsychology 2011 Nov;25(6):763-770. [doi: 10.1037/a0024238] [Medline: 21728425]
28. Boi R, Racca L, Cavallero A, et al. Hearing loss and depressive symptoms in elderly patients. Geriatr Gerontol Int 2012 Jul;12(3):440-445. [doi: 10.1111/j.1447-0594.2011.00789.x] [Medline: 22212622]
29. Jerger J, Chmiel R, Wilson N, Luchi R. Hearing impairment in older adults: new concepts. J Am Geriatr Soc 1995 Aug;43(8):928-935. [doi: 10.1111/j.1532-5415.1995.tb05539.x] [Medline: 7636104]
30. Yorgason JB, Piercy FP, Piercy SK. Acquired hearing impairment in older couple relationships: an exploration of couple resilience processes. J Aging Stud 2007 Aug;21(3):215-228. [doi: 10.1016/j.jaging.2006.10.002]

XSL•FO
RenderX

31. Lee WR, Chun S, Son K, Choi HS, An YY. Impact of hearing aid use on depression and quality of life in older patients with age-related hearing loss: a focus on social relationships. SSM Popul Health 2025 Dec;32:101861. [doi: 10.1016/j.ssmph.2025.101861] [Medline: 41050210]
32. Xu D, Newell MD, Francis AL. Fall-related injuries mediate the relationship between self-reported hearing loss and mortality in middle-aged and older adults. J Gerontol A Biol Sci Med Sci 2021 Aug 13;76(9):e213-e220. [doi: 10.1093/gerona/glab123] [Medline: 33929532]
33. Byun H, Kim EM, Kim I, Lee SH, Chung JH. The trend in adoption of hearing aids following changes in provision policy in South Korea. Sci Rep 2022 Aug 4;12(1):13389. [doi: 10.1038/s41598-022-17459-7] [Medline: 35927560]
34. Jenstad L, Moon J. Systematic review of barriers and facilitators to hearing aid uptake in older adults. Audiol Res 2011 May 10;1(1):e25. [doi: 10.4081/audiores.2011.e25] [Medline: 26557310]
35. Cosh S, Carriere I, Daien V, et al. The relationship between hearing loss in older adults and depression over 12 years: findings from the Three-City prospective cohort study. Int J Geriatr Psychiatry 2018 Dec;33(12):1654-1661. [doi: 10.1002/gps.4968] [Medline: 30209835]
36. Lawrence BJ, Jayakody DMP, Bennett RJ, Eikelboom RH, Gasson N, Friedland PL. Hearing loss and depression in older adults: a systematic review and meta-analysis. Gerontologist 2020 Apr 2;60(3):e137-e154. [doi: 10.1093/geront/gnz009] [Medline: 30835787]
37. Scarinci N, Worrall L, Hickson L. The effect of hearing impairment in older people on the spouse. Int J Audiol 2008 Mar;47(3):141-151. [doi: 10.1080/14992020701689696] [Medline: 18307094]
38. Wallhagen MI, Strawbridge WJ, Shema SJ, Kaplan GA. Impact of self-assessed hearing loss on a spouse: a longitudinal analysis of couples. J Gerontol B Psychol Sci Soc Sci 2004 May;59(3):S190-S196. [doi: 10.1093/geronb/59.3.s190] [Medline: 15118025]
39. Quach L, Dugan E, Gagnon D, Burr J. Social determinants of falls among older adults: the role of social support and depression. Gerontologist 2016 Nov;56(Suppl_3):366-366. [doi: 10.1093/geront/gnw162.1479]
40. Faulkner KA, Cauley JA, Zmuda JM, Griffin JM, Nevitt MC. Is social integration associated with the risk of falling in older community-dwelling women? J Gerontol A Biol Sci Med Sci 2003 Oct;58(10):M954-M959. [doi: 10.1093/gerona/58.10.m954] [Medline: 14570865]
41. Lee WY, Cameron PA, Bailey MJ. Road traffic injuries in the elderly. Emerg Med J 2006 Jan;23(1):42-46. [doi: 10.1136/emj.2005.023754] [Medline: 16381081]
42. Heinrich D, Holzmann C, Wagner A, et al. What are the differences in injury patterns of young and elderly traffic accident fatalities considering death on scene and death in hospital? Int J Legal Med 2017 Jul;131(4):1023-1037. [doi: 10.1007/s00414-017-1531-8] [Medline: 28180986]
43. Abou-Raya S, ElMeguid LA. Road traffic accidents and the elderly. Geriatr Gerontol Int 2009 Sep;9(3):290-297. [doi: 10.1111/j.1447-0594.2009.00535.x] [Medline: 19702940]
44. Schwartz N, Buliung R, Daniel A, Rothman L. Disability and pedestrian road traffic injury: A scoping review. Health Place 2022 Sep;77:102896. [doi: 10.1016/j.healthplace.2022.102896] [Medline: 36037674]

Abbreviations

HR: hazard ratio
KNDRS: Korea National Disability Registration System
NHI: national health insurance
NHIS: National Health Insurance Service
PSM: propensity score matching

Edited by A Mavragani, T Sanchez; submitted 10.Jun.2025; peer-reviewed by LF Belo, R Vesselinov; revised version received 31.Dec.2025; accepted 27.Feb.2026; published 27.Mar.2026.

Please cite as:
Lee WR, Chun S, An Y, Choi HS
Determining the Association Between Hearing Disability and Injury Risk in Older Adults Using Propensity Score Matching: Quasi-Experimental Study
JMIR Public Health Surveill 2026;12:e78826
URL: https://publichealth.jmir.org/2026/1/e78826
doi:10.2196/78826

© Woo-Ri Lee, Sungyoun Chun, Youyoung An, Hyun Seung Choi. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 27.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons

XSL•FO
RenderX

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

XSL•FO
RenderX

Association Between Low Chinese Visceral Adiposity Index Values and Chronic Lung Disease in Older Adults: Cohort Study

Xiaorong Zhu[1*], MBBS; Hui Tian[1*], MBBS; Zhi Liu[2], MBBS; Zhoubin Huang[2], MBBS; Miaomiao Yuan[3], MSc; Zhangshun Tu[3], MSc; Kaiwang Cui[1*], MSc; Jianping Liu[1,4*], MSc; Xiangwen Gong[1*], PhD

[1]Department of Respiratory and Critical Care Medicine, Ganzhou Key Laboratory of Respiratory Diseases, Ganzhou Institute of Respiratory Diseases, The Fifth People's Hospital of Ganzhou, Dongjiangyuan Road 666, Zhanggong District, Ganzhou, China

[2]Xingguo County Pulmonary Hospital, Ganzhou, China

[3]Gannan Medical University, Ganzhou, China

[4]Jiangxi Province Key Laboratory of Preventive Medicine, School of Public Health, Nanchang University, Nanchang, China

[*]these authors contributed equally

Corresponding Author:
Xiangwen Gong, PhD
Department of Respiratory and Critical Care Medicine, Ganzhou Key Laboratory of Respiratory Diseases, Ganzhou Institute of Respiratory Diseases, The Fifth People's Hospital of Ganzhou, Dongjiangyuan Road 666, Zhanggong District, Ganzhou, China

Abstract

Background: Visceral fat content, a key indicator of obesity, has been increasingly linked to chronic lung diseases (CLDs). However, the precise relationship between the Chinese visceral adiposity index (CVAI) and CLD remains unclear.

Objective: This study aimed to investigate the association between the CVAI and CLD risk.

Methods: Data from 6695 participants aged ≥45 years in the China Health and Retirement Longitudinal Study were analyzed. Restricted cubic spline models were applied to evaluate dose-response relationships between the CVAI and CLD across 3 population subgroups. Cox models were used to assess the hazard ratios and 95% CIs for the relationship between CVAI quintiles and CLD incidence in the significantly associated population.

Results: Restricted cubic spline analysis revealed a significant L-shaped relationship between the CVAI and CLD incidence in the older adult population (≥60 years; *P* value for overall=.01; *P* value for nonlinearity=.04), rather than in the middle-aged and older adult population (≥45 years; *P* value for overall=.12) or middle-aged population (45-59 years; *P* value for overall=.55). The inflection point of the CVAI for CLD risk was identified at 126 in the older adults. After multivariable adjustment, compared with participants in the fourth quintile (Q4), those in the first quintile (Q1) exhibited a significantly increased risk of CLD (hazard ratio=1.38, 95% CI 1.05 - 1.83). The subgroup analyses, which were stratified by sociodemographic factors and lifestyle, showed consistent results across most subgroups (*P* value for interaction>.05).

Conclusions: In adults aged ≥60 years, a lower CVAI was associated with a higher risk of CLD. No such association was observed in middle-aged adults.

(JMIR Public Health Surveill 2026;12:e78627) doi:10.2196/78627

KEYWORDS

chronic lung diseases; Chinese visceral adiposity index; China Health and Retirement Longitudinal Study; restricted cubic spline models; cohort study

Introduction

Chronic lung diseases (CLDs), which are recognized as a critical global public health challenge [1], demonstrate a pathological spectrum encompassing diverse respiratory disorders, including chronic bronchitis, emphysema, cor pulmonale, and asthma. Epidemiological evidence indicates a high prevalence of these conditions among middle-aged and older adult populations, with substantial impacts on patients' quality of life and considerable socioeconomic burdens. Current research has established tobacco smoke exposure, environmental pollution, and genetic predisposition as primary etiological factors [2], which collectively induce irreversible pulmonary functional impairment through sustained disruption of alveolar architecture and pulmonary tissue homeostasis.

In recent years, the interplay between metabolic dysregulation and chronic diseases has emerged as a pivotal research focus. Notably, aberrant visceral adiposity has been conclusively linked to type 2 diabetes mellitus, cardiovascular diseases, and metabolic syndrome [3,4]. While BMI remains widely used for obesity assessment, it fails to delineate the spatial distribution characteristics of adipose tissue [5]. The weight-adjusted waist

XSL•FO
RenderX

circumference index, calculated as waist circumference divided by the square root of body weight, improves abdominal adiposity evaluation but retains limitations in clinical utility due to its regional specificity [6]. In contrast, the Chinese visceral adiposity index (CVAI) integrates biochemical parameters such as triglycerides and high-density lipoprotein cholesterol (HDL-C) to establish a multidimensional metabolic profiling system. Emerging evidence demonstrates the CVAI's robust prognostic associations with hypertension [7], chronic kidney disease, metabolic syndrome [8], and diabetes mellitus [9], solidifying its role as a critical biomarker for quantifying visceral fat–related metabolic derangements.

Emerging evidence suggests that visceral adiposity may impair pulmonary function through mechanisms involving proinflammatory cytokine release and exacerbated oxidative stress [10]. Nevertheless, the relationship between the CVAI and chronic pulmonary diseases remains underexplored. Leveraging the large-scale dataset from the China Health and Retirement Longitudinal Study (CHARLS), this study aims to elucidate the epidemiological association between the CVAI and chronic pulmonary diseases, thereby providing mechanistic insights to inform the development of metabolism-targeted prevention and therapeutic strategies for respiratory disorders.

Methods

Ethical Considerations

The CHARLS survey was reviewed and approved by the institutional review board of Peking University (IRB00001052-11015). Before participation, all respondents were provided with a comprehensive introduction to the potential risks and benefits of the study. Thereafter, each participant provided written informed consent and repository consent, permitting the storage and secondary sharing of their data for research purposes. Participants' privacy and confidentiality were strictly protected: personal identifiers were removed, and data were stored on secure servers accessible only to authorized study personnel. Data sharing was conducted in accordance with applicable data protection regulations. Participants did not receive direct financial compensation; however, they were informed about the potential indirect benefits of participation, such as contributing to national health research and policy development.

Study Design and Study Population

The data for this study were obtained from the CHARLS, which was a national representative study. The details of study design and assessment protocols for the CHARLS have been clearly detailed in a previous study [11]. In summary, the CHARLS is the first high-quality microdatabase designed to accurately represent families and individuals aged ≥45 years in China. This database serves as a vital resource for examining the challenges associated with population aging in China. Between 2011 and 2012, a sample of over 17,000 individuals aged ≥45 years was recruited from 450 communities and administrative villages across 28 provinces nationwide, using a multistage probability sampling method. Since the initial recruitment, these participants have undergone biennial follow-up assessments.

Data from waves 1 to 5 (collected in 2011, 2013, 2015, 2017, and 2020, respectively) were included in this study. At baseline (wave 1), 17,596 participants were enrolled. Participants were excluded based on the following criteria: (1) age <45 years, (2) missing information on the CVAI, (3) missing covariate data, and (4) lost to follow-up during waves 2-5.

The start time was defined as the date of the 2011 CHARLS baseline interview. Event time was calculated in years from baseline to the earliest of either the first self-reported CLD or the wave 5 (2020) interview. Censoring was applied to participants who were lost to follow-up at any point during the follow-up period (waves 2 - 5) or remained free of CLD at study end.

Assessment of the CVAI

The CVAI is a reliable and applicable index for evaluation of visceral fat dysfunction in Chinese, which has been used in many studies. It was calculated with sex-specific formulas as follows [12]:

Females: $CVAI = -187.32 + 1.71 \times age + 4.23 \times BMI + 1.12 \times WC + 39.76 \times Log_{10}TG - 11.66 \times HDL\text{-}C$

Males: $CVAI = -267.93 + 0.68 \times age + 0.03 \times BMI + 4.00 \times WC + 22.00 \times Log_{10}TG - 16.32 \times HDL\text{-}C$

Assessment of CLD

CLD was assessed by asking participants whether they have been diagnosed with CLD by a doctor. They were classified as patients with CLD if the answer was positive. In contrast, they were not classified as patients with CLD if the answer was negative.

Assessment of Covariates

Covariates include the following: demographic data (age, sex, education, area of residence, and marital status), lifestyles (smoking and alcohol consumption), and chronic disease. Age was divided into 2 categories: middle-aged (45 - 59 years) and older (≥60 years). Education was categorized into 3 groups: illiterate, primary school and below, or middle school and above. Residence was grouped into urban or village. Marital status was divided into 2 categories: married and others (including unmarried, separated, divorced, and widowed). Smoking and alcohol consumption were categorized into 3 groups: never, former, and current. Chronic diseases were assessed in a similar manner to CLD and grouped into yes or no categories.

Statistical Analysis

Boruta is an all-relevant feature selection wrapper algorithm that identifies which features are important and which are not. We used it to determine the important features associated with incident CLD [13].

Restricted cubic spline (RCS) analyses were performed with the "rcssci" R package (version 0.4.0; R Foundation for Statistical Computing) [14] to assess the dose-response associations of the CVAI with CLD in three populations: (1) middle-aged and older adults (≥45 years), (2) middle-aged adults (45 - 59 years), and (3) older adults (≥60 years). A 2-piecewise Cox model was applied when a cutoff point was identified in

XSL•FO
RenderX

the RCS analysis. The number of knots was automatically selected by the "rcssci" R package through minimization of the Akaike information criterion. Three RCS models ultimately used 3 knots placed at the 10th, 50th, and 90th percentiles, ensuring a parsimonious yet flexible fit to any nonlinear relationship.

Subsequently, participants exhibiting a significant association between the CVAI and CLD were grouped by CVAI quintiles. Differences in CVAI quintile groups were tested using 1-way ANOVA for normally distributed data, the Kruskal-Wallis test for nonnormally distributed data, and the χ^2 test for categorical data. Cox models were used to assess the hazard ratios (HRs) and 95% CIs for the relationship between CVAI quintiles and CLD incidence. Model 1 was the crude model. In model 2, we adjusted for age, sex, education, marital status, residence, smoking, and alcohol consumption. In model 3, we further adjusted for several chronic diseases (hypertension, dyslipidemia, diabetes, heart attack, stroke, stomach disease, and memory-related disease). The proportional hazards assumptions for all Cox models were assessed using Schoenfeld residuals and were not violated (all *P* values >.05).

Finally, subgroup analyses stratified by sex (male or female), education (illiterate or literate), marital status (married or other), residence (village or urban), smoking status (no or yes), and alcohol consumption (no or yes) were conducted. Additionally, the interactions of the CVAI with the stratified factors were tested using the likelihood ratio test by comparing models with and without the interaction term.

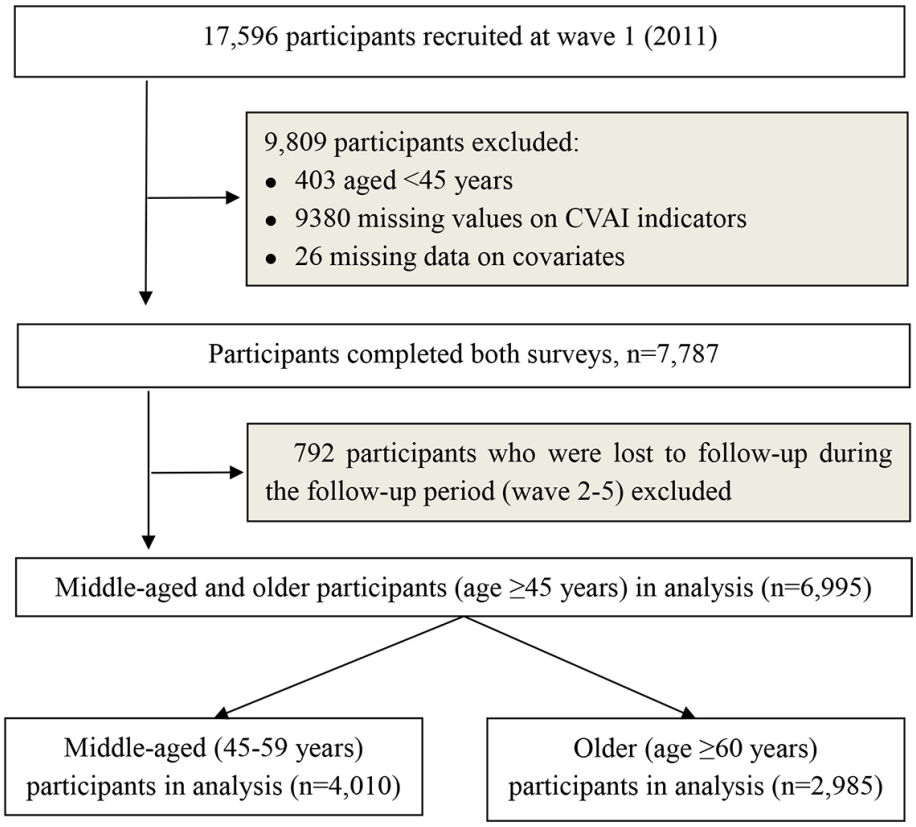
The Stata 15 SE (StataCorp) and R (version 4.4.1; R Foundation for Statistical Computing) were used for analyses, and statistical significance was defined as 2-tailed *P* values <.05. *P* values reported throughout the study are defined as follows: *P* value for overall indicates the overall significance of RCS model, *P* value for nonlinearity assesses whether the association exhibits a non-linear relationship in RCS analyses, and *P* value for interaction evaluates the significance of interaction terms between covariates and the main exposure.

Results

Study Population

A total of 6695 participants aged ≥45 years with complete data were included in the analysis. Of these, 4010 were middle-aged (45–59 years old) and 2685 were older (≥60 years old). The process of participant selection, including all exclusions at each step, is illustrated in Figure 1.

Figure 1. Flowchart of study participants. CVAI: Chinese visceral adiposity index.

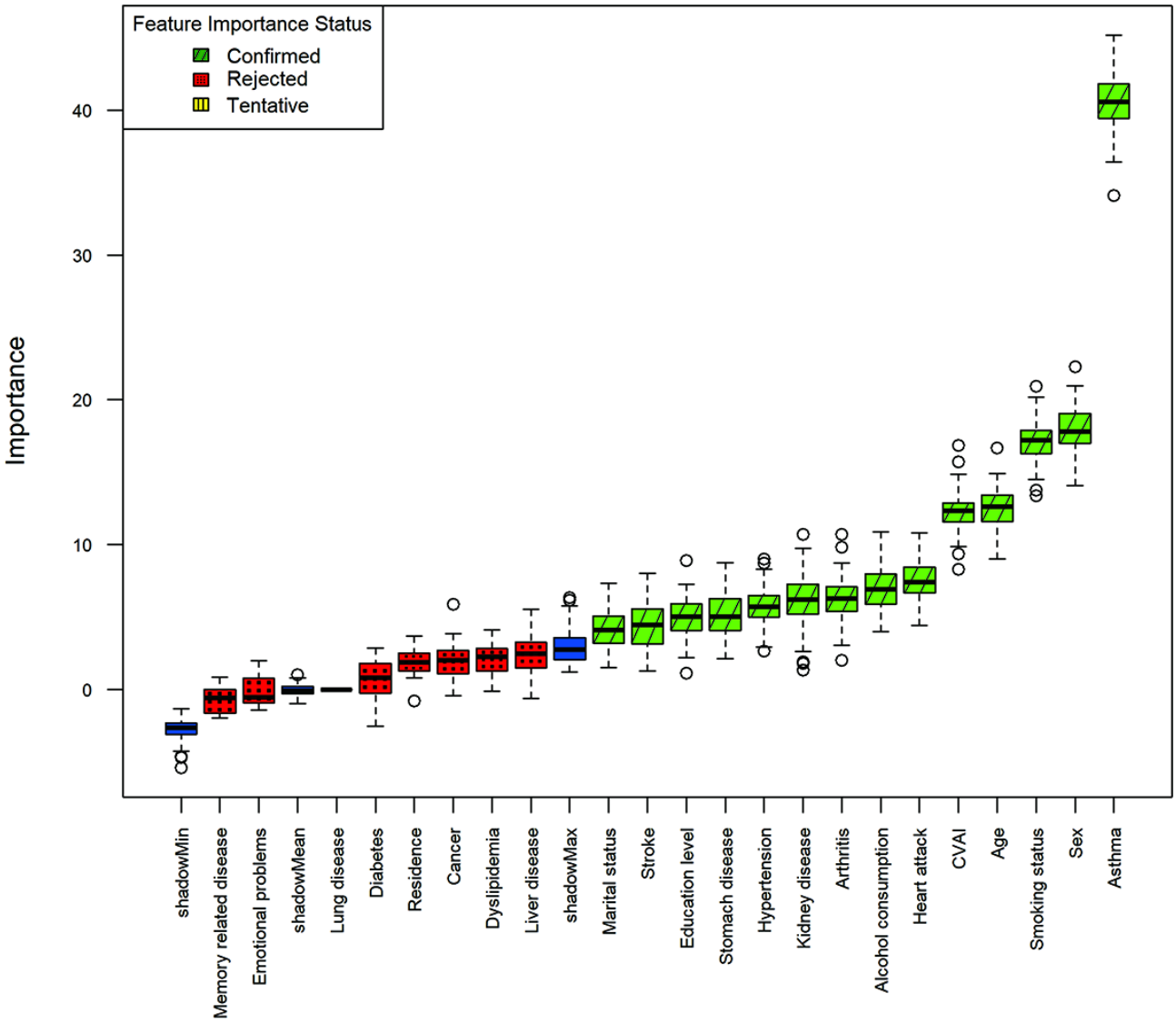


Feature Selection

Figure 2 shows 14 features strongly associated with incident CLD identified by the Boruta algorithm. The 14 confirmed variables are the CVAI, age, sex, education level, marital status, smoking status, alcohol consumption, hypertension, heart attack, stroke, kidney disease, stomach disease, arthritis, and asthma. Residence, though falling below the shadow-feature threshold, was retained on the basis of prior evidence and clinical experience [15,16].

XSL•FO
RenderX

Figure 2. Feature selection for chronic lung disease incidence using the Boruta algorithm (ie, Boruta feature importance). ShadowMin, shadowMax, and shadowMean indicate the minimum, maximum, and mean importance values of shadow (randomized) features used internally by the Boruta algorithm. CVAI: Chinese visceral adiposity index.



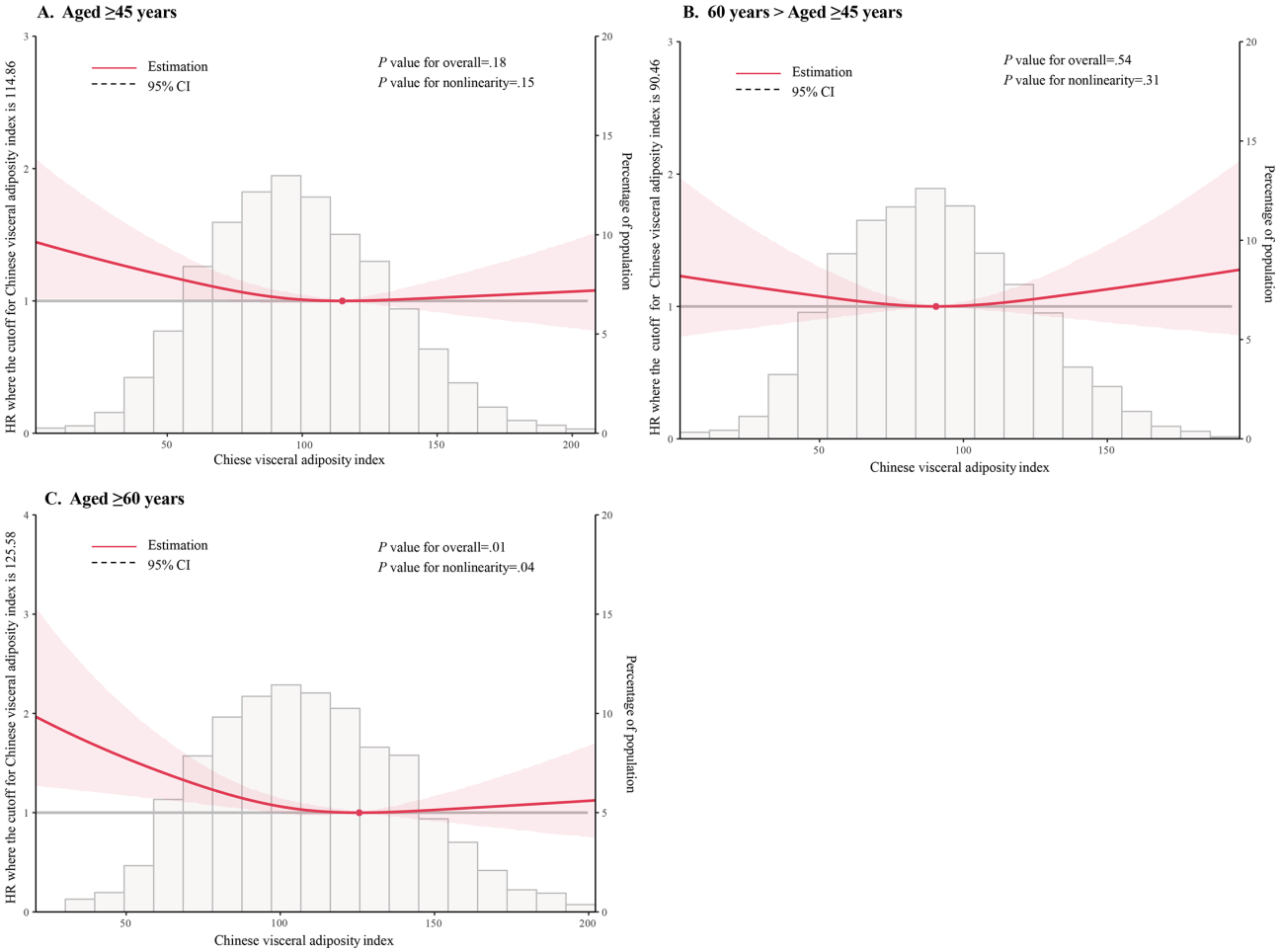
Dose-Response Relationship

Figure 3 shows the results of the RCS analyses examining the association between the CVAI and CLD across different age groups. The results suggested a significant L-shaped relationship between the CVAI and CLD in older adults (aged ≥60 years; *P* value for overall=.01; *P* value for nonlinearity=.04). The cutoff point for the CVAI was 126. In a multiadjusted 2-piecewise Cox model, when the CVAI was below 126, each 1-point decrease in the CVAI was associated with a 0.6 % increase in CLD risk (HR=1.006, 95% CI 1.002-1.011; *P*=.007). When the CVAI was above 126, CLD risk showed a nonsignificant increasing trend with each 1-point increase in the CVAI (HR=1.001, 95% CI 0.992-1.010; *P*=.86).

However, the CVAI was not associated with CLD in either the middle-aged and older adult population (aged ≥45 years; *P* value for overall=.12) or the middle-aged population (aged 45 - 59 years; *P* value for overall=.55).

XSL•FO
RenderX

Figure 3. Dose-response relationship between the Chinese visceral adiposity index (CVAI) and chronic lung disease. Spline analyses were adjusted for age, sex, education, marital status, residence, smoking, alcohol consumption, hypertension, heart attack, stroke, kidney disease, stomach disease, arthritis, and asthma. HR: hazard ratio.



Baseline Characteristics of the Older Adult Population

According to the L-shaped association between the CVAI and CLD in the older adult population, the CVAI was categorized in quintiles: Q1 (CVAI<80), Q2 (80≤CVAI<98), Q3 (98≤CVAI<116), Q4 (116≤CVAI<136), and Q5 (136≤CVAI). Table 1 presents the characteristics of the older adult population. Participants with a higher CVAI were more likely to be older, female, married, living in urban areas, non-smokers, and non-drinkers, and to have lower educational attainment, hypertension, a history of heart attack, stroke, and stomach disease.

XSL•FO
RenderX

Table . Characteristics of older adult participants (aged ≥60 years) by quintiles of the Chinese visceral adiposity index (N=2985).

Characteristics	Chinese visceral adiposity index[a]					*P* value
	Q1 (n=597)	Q2 (n=597)	Q3 (n=597)	Q4 (n=597)	Q5 (n=597)	
Age (years), mean (SD)	64.75 (4.51)	66.39 (5.34)	66.95 (5.87)	67.30 (5.89)	68.38 (6.36)	<.001
Sex, n (%)						<.001
Male	332 (55.61)	340 (56.95)	277 (46.40)	259 (43.38)	239 (40.03)	
Female	265 (44.39)	257 (43.05)	320 (53.60)	338 (56.62)	358 (59.97)	
Education level, n (%)						<.001
Illiterate	199 (33.33)	201 (33.67)	233 (39.03)	210 (35.18)	241 (40.37)	
Primary school and below	302 (50.59)	308 (51.59)	260 (43.55)	276 (46.23)	231 (38.69)	
Middle school and above	96 (16.08)	88 (14.74)	104 (17.42)	111 (18.59)	125 (20.94)	
Marital status, n (%)						.008
Married	88 (14.74)	98 (16.42)	111 (18.59)	127 (21.27)	128 (21.44)	
Other	509 (85.26)	499 (83.58)	486 (81.41)	470 (78.73)	469 (78.56)	
Residence, n (%)						<.001
Village	467 (78.22)	439 (73.53)	403 (67.50)	365 (61.14)	331 (55.44)	
Urban	130 (21.78)	158 (26.47)	194 (32.50)	232 (38.86)	266 (44.56)	
Smoking status, n (%)						<.001
Never smoker	291 (48.74)	305 (51.09)	371 (62.14)	376 (62.98)	398 (66.67)	
Former smoker	49 (8.21)	56 (9.38)	55 (9.21)	59 (9.88)	73 (12.23)	
Current smoker	257 (43.05)	236 (39.53)	171 (28.64)	162 (27.14)	126 (21.11)	
Alcohol consumption, n (%)						<.001
Never drinker	311 (52.09)	316 (52.93)	347 (58.12)	365 (61.14)	391 (65.49)	
Former drinker	49 (8.21)	62 (10.39)	63 (10.55)	57 (9.55)	63 (10.55)	
Current drinker	237 (39.70)	219 (36.68)	187 (31.32)	175 (29.31)	143 (23.95)	
Hypertension, n (%)	84 (14.07)	123 (20.60)	168 (28.14)	209 (35.01)	313 (52.43)	<.001
Heart attack, n (%)	48 (8.05)	54 (9.05)	73 (12.23)	92 (15.41)	124 (20.77)	<.001
Stroke, n (%)	14 (2.35)	4 (0.67)	14 (2.35)	16 (2.68)	31 (5.19)	<.001
Kidney disease, n (%)	21 (3.52)	23 (3.85)	29 (4.86)	31 (5.19)	34 (5.70)	.34
Stomach disease, n (%)	164 (27.47)	135 (22.61)	142 (23.79)	123 (20.60)	108 (18.09)	.002
Arthritis, n (%)	222 (37.25)	218 (36.52)	206 (34.51)	211 (35.34)	223 (37.35)	.82
Asthma, n (%)	8 (1.34)	11 (1.84)	12 (2.01)	12 (2.01)	18 (3.02)	.35

[a]Q1-Q5 represent the quintiles of Chinese Visceral Adiposity Index (CVAI) in the older population: Q1 (lowest quintile, CVAI<80), Q2 (80≤CVAI<98), Q3 (98≤CVAI<116), Q4 (116≤CVAI<136), and Q5 (highest quintile, CVAI≥136).

CVAI and CLD in the Older Adult Population

During follow-up (mean 8.42, SD 1.73 years; total 25,132 person-years), 527 participants developed incident CLD, yielding an incidence density of 20.97 per 1000 person-years.

The association of the CVAI and CLD in the older adult population is shown in Table 2. We selected Q4 as the reference group because the cutoff point (126) falls within the range of the Q4 group (116≤CVAI<136). After adjusting multiple covariates, the HR of participants in the Q1 group was 1.38 (95% CI 1.05-1.83) compared to the Q4 group. However, no significant difference was observed between the Q4 group and either the Q2, Q3, and Q5 groups. Kaplan-Meier plots by CVAI quintiles also showed the lowest CVAI group to have the highest CLD incident risk (Multimedia Appendix 1).

XSL•FO
RenderX

Table . Hazard ratios (HRs) and 95% CIs for the association between the Chinese visceral adiposity index (CVAI) and chronic lung disease in older adults (N=2985).[a]

CVAI	Cases, n (%)	Model 1[b]		Model 2[c]		Model 3[d]	
		HR (95% CI)	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value
Q1	127 (21.27)	1.42 (1.04-1.77)	.01	1.35 (1.03-1.78)	.03	1.38 (1.05-1.83)	.02
Q2	104 (17.42)	1.16 (0.87-1.54)	.30	1.11 (0.84-1.48)	.45	1.14 (0.86-1.52)	.36
Q3	103 (17.25)	1.13 (0.85-1.50)	.39	1.13 (0.85-1.50)	.40	1.13 (0.85-1.49)	.40
Q4	91 (15.24)	1 (reference)	__[e]	1 (reference)	—	1 (reference)	—
Q5	102 (17.09)	1.11 (0.84-1.47)	.46	1.14 (0.86-1.51)	.37	1.07 (0.80-1.42)	.65

[a]Q1-Q5 represent the quintiles of CVAI in the older population: Q1 (lowest quintile, CVAI<80), Q2 (80≤CVAI<98), Q3 (98≤CVAI<116), Q4 (116≤CVAI<136), and Q5 (highest quintile, CVAI≥136).

[b]Model 1 was unadjusted.

[c]Model 2 was adjusted for age, sex, education, marital status, residence, smoking, and alcohol consumption.

[d]Model 3 was adjusted for age, sex, education, marital status, residence, smoking, alcohol consumption, hypertension, heart attack, stroke, kidney disease, stomach disease, arthritis, and asthma.

[e]Not applicable.

Subgroup Analysis

Subgroup analyses stratified by sociodemographic factors and lifestyle were performed after full adjustment (Table 3). The results showed that the relationship between the CVAI and CLD was consistent across most subgroups (*P* value for interaction>.05). A significant interaction for education and the CVAI was observed (*P* value for interaction=.03). Among the illiterate older adult population, compared with the Q4 group, the HRs were 2.03 (1.21-3.46) and 1.70 (1.07-2.76) in the Q1 and Q5 groups, respectively. Conversely, the CVAI was not related with CLD in the literate older adult population.

XSL•FO
RenderX

Table . The subgroup analysis for the association between the Chinese visceral adiposity index (CVAI) and chronic lung disease in older adults (N=2984).

Subgroup	CVAI[a,b]					*P* value for interaction
	Q1, HR[c] (95% CI)	Q2, HR (95% CI)	Q3, HR (95% CI)	Q4, HR (95% CI)	Q5, HR (95% CI)	
Sex						.83
Male	1.28 (0.87-1.88)	1.09 (0.74-1.59)	1.10 (0.74-1.63)	1 (reference)	0.87 (0.57-1.34)	
Female	1.52 (0.98-2.30)	1.26 (0.82-1.93)	1.20 (0.80-1.81)	1 (reference)	1.26 (0.85-1.86)	
Education						.03
Illiterate	2.03 (1.21-3.46)	1.16 (0.87-2.18)	1.72 (1.04-2.87)	1 (reference)	1.70 (1.07-2.76)	
literate	1.13 (0.81-1.58)	1.12 (0.80-1.56)	0.90 (0.63-1.28)	1 (reference)	0.84 (0.59-1.21)	
Marital status						.25
Married	1.35 (0.72-2.55)	1.13 (0.59-2.17)	1.09 (0.53-1.86)	1 (reference)	1.56 (0.88-2.73)	
Other	1.38 (0.99-1.89)	1.14 (0.83-1.57)	1.17 (0.85-1.60)	1 (reference)	0.91 (0.65-1.28)	
Residence						.94
Village	1.43 (1.02-2.00)	1.12 (0.79-1.59)	1.15 (0.81-1.63)	1 (reference)	1.11 (0.77-1.61)	
Urban	1.31 (0.77-2.22)	1.24 (0.75-2.07)	1.17 (0.72-1.90)	1 (reference)	1.03 (0.62-1.55)	
Smoking status						.98
No	1.28 (0.89-1.85)	1.09 (0.76-1.58)	1.14 (0.81-1.62)	1 (reference)	1.09 (0.78-1.54)	
Yes	1.68 (1.06-2.66)	1.31 (0.81-2.10)	1.17 (0.71-1.93)	1 (reference)	1.04 (0.61-1.78)	
Alcohol consumption						.52
No	1.50 (1.06-2.13)	1.17 (0.81-1.67)	1.03 (0.72-1.47)	1 (reference)	1.11 (0.79-1.57)	
Yes	1.20 (0.74-1.93)	1.06 (0.66-1.72)	1.29 (0.80-2.08)	1 (reference)	0.91 (0.53-1.57)	

[a]Model was adjusted for age, sex, education, marital status, residence, smoking, alcohol consumption, hypertension, heart attack, stroke, kidney disease, stomach disease, arthritis, and asthma.

[b]Q1-Q5 represent the quintiles of CVAI in the older population: Q1 (lowest quintile, CVAI<80), Q2 (80≤CVAI<98), Q3 (98≤CVAI<116), Q4 (116≤CVAI<136), and Q5 (highest quintile, CVAI≥136).

[c]HR: hazard ratio.

Discussion

Principal Findings

This study systematically investigated the relationship between the CVAI and CLD in a cohort from the CHARLS. The results demonstrated a significant L-shaped relationship between the CVAI and CLD, with a low CVAI being negatively correlated with the risk of CLDs. This finding not only improves the understanding of the association between visceral fat and respiratory diseases but also provides a quantifiable biomarker reference value for the risk stratification of CLD based on body fat distribution in clinical practice.

In the study, we revealed that participants with lower visceral adiposity exhibited an elevated risk of chronic pulmonary diseases, particularly pronounced in older adult populations. In cardiovascular research, the functional heterogeneity of adipose tissue across anatomical depots contributes to differential cardiovascular risks, with visceral fat being mechanistically implicated in metabolic syndrome and cardiovascular disease pathogenesis through multiple pathways [17]. Consequently, clinical obesity metrics have evolved beyond weight-centric indices such as BMI to incorporate visceral adiposity surrogates. The CVAI—integrating BMI, waist circumference, triglycerides, and HDL-C—has been validated as a robust predictor of diabetes, cardiovascular disease, hypertension, and metabolic syndrome in Chinese cohorts [18-21]. However, no prior studies have explored the CVAI–chronic pulmonary disease nexus. Intriguingly, a study using Asia-Pacific BMI criteria demonstrated an inverse correlation between BMI and lung function [22], whereas analyses applying World Health Organization (WHO) BMI thresholds revealed U-shaped associations for forced expiratory volume in one second and diffusing capacity for carbon monoxide. This discrepancy underscores the critical influence of region-specific obesity classification criteria on pulmonary function assessments. Our CVAI-based analysis identified an L-shaped relationship between visceral adiposity and chronic pulmonary disease incidence in Chinese adults. While existing literature predominantly reports linear inverse correlations between obesity severity and lung function decline [23,24], most studies contrast pulmonary metrics between overweight or obese and normal-weight groups [25,26]. Notably, the current evidence base overlooks pulmonary dysfunction risks associated with hypoadiposity. Clinically, malnutrition-induced hypoadiposity warrants attention, as it impairs respiratory muscle architecture

XSL•FO
RenderX

and contractility, heightening fatigue susceptibility and ventilatory dysfunction [27-30]. Concurrent protein deficiency compromises immunoglobulin and complement production, exacerbating immunosuppression and infection susceptibility, thereby aggravating respiratory impairment in underweight populations [29-31]. These mechanisms may underlie the observed association between a low CVAI and heightened chronic pulmonary disease risk.

Excessive visceral adiposity has been implicated in pulmonary function impairment and elevated chronic pulmonary disease incidence. While our findings suggested an increasing trend in CLD risk when the CVAI was above 126, this trend was not statistically significant (HR=1.001, 95% CI 0.992-1.010, *P*=.86). Nevertheless, the obesity–lung function relationship remains multifaceted. Most studies corroborate an inverse association between obesity and pulmonary function, though paradoxically, some report attenuated lung function decline—for example, slower forced expiratory volume in one second reduction—in obese individuals [32]. Notably, overweight or obesity has demonstrated a protective effect against mortality in patients with chronic obstructive pulmonary disease [33], aligning with the "obesity paradox," wherein excess adiposity correlates with improved clinical outcomes in chronic disease populations [34]. Adiposity may serve as a nutritional reserve, particularly critical for older, frail, or comorbid patients. Overweight or obese patients with chronic obstructive pulmonary disease exhibit higher fat-free mass and respiratory muscle mass compared to underweight or normal-weight counterparts—a physiological advantage in compensating for elevated airway resistance and airflow obstruction [35]. These factors may explain the survival benefits observed in obese patients with chronic respiratory conditions [36]. Thus, the CVAI–chronic pulmonary disease association warrants further investigation, particularly through disease-specific analyses or site-specific assessments of visceral fat distribution. Future studies should disentangle the dualistic effects of adiposity-metabolic detriment versus cardiopulmonary reserve to reconcile these seemingly contradictory epidemiological observations.

In the subgroup analysis, higher educational attainment appeared to buffer the effect of an unfavorable CVAI on CLD risk. The reason may be that individuals with higher education usually possess better health literacy and self-management abilities, enabling them to offset the negative effects of unreasonable visceral adipose tissue through active lifestyles and regular health check-ups.

Inevitably, there are some limitations in our study. First, data on chronic pulmonary disease were collected via self-reported questionnaires, introducing potential recall bias. Older adults are more likely to forget previous diagnoses, and CLD is a condition with a low diagnosis rate in China. Thus, the true association between a low CVAI and CLD may be underestimated. Future studies should link electronic medical records, conduct pulmonary function tests, or perform chest imaging verification to reduce recall bias. Second, some other confounders, including physical activity, diet quality, or detailed socioeconomic position, were not adjusted for. In theory, higher levels of these factors could attenuate the association between the CVAI and CLD. Finally, the study population was restricted to individuals aged >45 years, limiting generalizability to younger adult demographics.

Conclusions

In conclusion, this study is the first to demonstrate an L-shaped association between the CVAI and CLD in the older adult population, with a low CVAI linked to a higher risk of CLD, but not in middle-aged individuals. These findings, while preliminary, highlight the need for validation through large-scale, multicenter prospective studies to elucidate age-stratified mechanisms linking visceral adiposity dynamics to respiratory health outcomes.

Funding

XG received grants from the 2023-2025 First Batch of Provincial Clinical Key Specialty Construction Projects in Jiangxi (Gan Wei Yi Zi [2023] No. 82) and the Gannan Soviet Area Talent Support Program (Gan Shi Cai Zi [2021] No. 1).

Data Availability

The datasets used and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

XZ and HT conceived the study and were responsible for formal analysis, data curation, and drafting the manuscript. ZL, ZH, MY, and ZT contributed to writing (review and editing), as well as formal analysis and data curation. KC, JL, and XG contributed to writing (review and editing) and data curation. All authors approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Kaplan-Meier plots.
[PNG File, 44 KB - publichealth_v12i1e78627_app1.png]

References

XSL•FO
RenderX

1. Hu Z, Tian Y, Song X, Zeng F, Yang A. Associations between sarcopenia with asthmatic prevalence, lung function and comorbidity. BMC Geriatr 2022 Aug 24;22(1):703. [doi: 10.1186/s12877-022-03394-9] [Medline: 36002808]
2. López-Campos JL, Tan W, Soriano JB. Global burden of COPD. Respirology 2016 Jan;21(1):14-23. [doi: 10.1111/resp.12660] [Medline: 26494423]
3. Blüher M. Obesity: global epidemiology and pathogenesis. Nat Rev Endocrinol 2019 May;15(5):288-298. [doi: 10.1038/s41574-019-0176-8] [Medline: 30814686]
4. Sánchez-Bacaicoa C, Santano-Mogena E, Rico-Martín S, et al. Association between asymptomatic hyperuricemia with adiposity indices: a cross-sectional study in a Spanish population. Nutrients 2023 Nov 16;15(22):4798. [doi: 10.3390/nu15224798] [Medline: 38004193]
5. Bray GA, Heisel WE, Afshin A, et al. The science of obesity management: an Endocrine Society scientific statement. Endocr Rev 2018 Apr 1;39(2):79-132. [doi: 10.1210/er.2017-00253] [Medline: 29518206]
6. Silveira EA, Castro MCR, Rezende ATO, et al. Body composition assessment in individuals with class II/III obesity: a narrative review. BMC Nutr 2024 Oct 22;10(1):142. [doi: 10.1186/s40795-024-00913-2] [Medline: 39438968]
7. Zhang Z, Shi D, Zhang Q, et al. Visceral adiposity index (VAI), a powerful predictor of incident hypertension in prehypertensives. Intern Emerg Med 2018 Jun;13(4):509-516. [doi: 10.1007/s11739-018-1836-8] [Medline: 29569088]
8. Lazzer S, D'Alleva M, Isola M, et al. Cardiometabolic index (CMI) and visceral adiposity index (VAI) highlight a higher risk of metabolic syndrome in women with severe obesity. J Clin Med 2023 Apr 23;12(9):3055. [doi: 10.3390/jcm12093055] [Medline: 37176497]
9. Ahn N, Baumeister SE, Amann U, et al. Visceral adiposity index (VAI), lipid accumulation product (LAP), and product of triglycerides and glucose (TyG) to discriminate prediabetes and diabetes. Sci Rep 2019 Jul 4;9(1):9693. [doi: 10.1038/s41598-019-46187-8] [Medline: 31273286]
10. Palma G, Sorice GP, Genchi VA, et al. Adipose tissue inflammation and pulmonary dysfunction in obesity. Int J Mol Sci 2022 Jul 1;23(13):7349. [doi: 10.3390/ijms23137349] [Medline: 35806353]
11. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). Int J Epidemiol 2014 Feb;43(1):61-68. [doi: 10.1093/ije/dys203] [Medline: 23243115]
12. Xia MF, Chen Y, Lin HD, et al. A indicator of visceral adipose dysfunction to evaluate metabolic health in adult Chinese. Sci Rep 2016 Dec 1;6:38214. [doi: 10.1038/srep38214] [Medline: 27905531]
13. Kursa MB, Rudnicki WR. Feature selection with Boruta package. J Stat Softw 2010;36(11):1-13. [doi: 10.18637/jss.v036.i11]
14. Rcssci: visualization of restricted cubic splines. CRAN. URL: https://cran.r-project.org/web/packages/rcssci/index.html [accessed 2026-03-15]
15. Pan XF, Wang L, Pan A. Epidemiology and determinants of obesity in China. Lancet Diabetes Endocrinol 2021 Jun;9(6):373-392. [doi: 10.1016/S2213-8587(21)00045-0] [Medline: 34022156]
16. Chen H, Liu X, Gao X, et al. Epidemiological evidence relating risk factors to chronic obstructive pulmonary disease in China: a systematic review and meta-analysis. PLoS One 2021;16(12):e0261692. [doi: 10.1371/journal.pone.0261692]
17. Neeland IJ, Ross R, Després JP, et al. Visceral and ectopic fat, atherosclerosis, and cardiometabolic disease: a position statement. Lancet Diabetes Endocrinol 2019 Sep;7(9):715-725. [doi: 10.1016/S2213-8587(19)30084-1] [Medline: 31301983]
18. Pan L, Xu Q, Liu J, et al. Dose-response relationship between Chinese visceral adiposity index and type 2 diabetes mellitus among middle-aged and elderly Chinese. Front Endocrinol (Lausanne) 2022;13:959860. [doi: 10.3389/fendo.2022.959860] [Medline: 36277708]
19. Ren Y, Hu Q, Li Z, Zhang X, Yang L, Kong L. Dose-response association between Chinese visceral adiposity index and cardiovascular disease: a national prospective cohort study. Front Endocrinol (Lausanne) 2024;15:1284144. [doi: 10.3389/fendo.2024.1284144] [Medline: 38699393]
20. Duan Y, Zhang W, Li Z, et al. Predictive ability of obesity- and lipid-related indicators for metabolic syndrome in relatively healthy Chinese adults. Front Endocrinol (Lausanne) 2022;13:1016581. [doi: 10.3389/fendo.2022.1016581] [Medline: 36465613]
21. Han M, Qie R, Li Q, et al. Chinese visceral adiposity index, a novel indicator of visceral obesity for assessing the risk of incident hypertension in a prospective cohort study. Br J Nutr 2021 Aug 28;126(4):612-620. [doi: 10.1017/S0007114520004298] [Medline: 33143773]
22. Lim JU, Lee JH, Kim JS, et al. Comparison of World Health Organization and Asia-Pacific body mass index classifications in COPD patients. Int J Chron Obstruct Pulmon Dis 2017;12:2465-2475. [doi: 10.2147/COPD.S141295] [Medline: 28860741]
23. He S, Yang J, Li X, Gu H, Su Q, Qin L. Visceral adiposity index is associated with lung function impairment: a population-based study. Respir Res 2021 Jan 6;22(1):2. [doi: 10.1186/s12931-020-01599-3] [Medline: 33407481]
24. Forno E, Han YY, Mullen J, Celedón JC. Overweight, obesity, and lung function in children and adults-a meta-analysis. J Allergy Clin Immunol Pract 2018;6(2):570-581. [doi: 10.1016/j.jaip.2017.07.010] [Medline: 28967546]
25. Klepaker G, Svendsen MV, Hertel JK, et al. Influence of obesity on work ability, respiratory symptoms, and lung function in adults with asthma. Respiration 2019;98(6):473-481. [doi: 10.1159/000502154] [Medline: 31461714]
26. Xing X, Hu L, Guo Y, et al. Interactions between ambient air pollution and obesity on lung function in children: the Seven Northeastern Chinese Cities (SNEC) Study. Sci Total Environ 2020 Jan 10;699:134397. [doi: 10.1016/j.scitotenv.2019.134397] [Medline: 31677469]

XSL•FO
RenderX

27. Li L, Zhong X, Zheng A, et al. Prevalence and risk factors of chronic obstructive pulmonary disease in Kashi region, Northwestern China. Int J Chron Obstruct Pulmon Dis 2021;16:655-663. [doi: 10.2147/COPD.S289620] [Medline: 33758502]
28. Benz E, Trajanoska K, Lahousse L, et al. Sarcopenia in COPD: a systematic review and meta-analysis. Eur Respir Rev 2019 Nov 13;28(154):190049. [doi: 10.1183/16000617.0049-2019] [Medline: 31722892]
29. Lee SJ, Kim SW, Kong KA, Ryu YJ, Lee JH, Chang JH. Risk factors for chronic obstructive pulmonary disease among never-smokers in Korea. Int J Chron Obstruct Pulmon Dis 2015;10:497-506. [doi: 10.2147/COPD.S77662] [Medline: 25784796]
30. Li J, Zhu L, Wei Y, et al. Association between adiposity measures and COPD risk in Chinese adults. Eur Respir J 2020 Apr;55(4):1901899. [doi: 10.1183/13993003.01899-2019] [Medline: 31980495]
31. Dobner J, Kaser S. Body mass index and the risk of infection - from underweight to obesity. Clin Microbiol Infect 2018 Jan;24(1):24-28. [doi: 10.1016/j.cmi.2017.02.013] [Medline: 28232162]
32. Zhang X, Chen H, Gu K, Chen J, Jiang X. Association of body mass index with risk of chronic obstructive pulmonary disease: a systematic review and meta-analysis. COPD 2021 Feb;18(1):101-113. [doi: 10.1080/15412555.2021.1884213] [Medline: 33590791]
33. Chailleux E, Laaban JP, Veale D. Prognostic value of nutritional depletion in patients with COPD treated by long-term oxygen therapy: data from the ANTADIR observatory. Chest 2003 May;123(5):1460-1466. [doi: 10.1378/chest.123.5.1460] [Medline: 12740261]
34. Spelta F, Fratta Pasini AM, Cazzoletti L, Ferrari M. Body weight and mortality in COPD: focus on the obesity paradox. Eat Weight Disord 2018 Feb;23(1):15-22. [doi: 10.1007/s40519-017-0456-z] [Medline: 29110280]
35. Gea J, Pascual S, Casadevall C, Orozco-Levi M, Barreiro E. Muscle dysfunction in chronic obstructive pulmonary disease: update on causes and biological findings. J Thorac Dis 2015 Oct;7(10):E418-E438. [doi: 10.3978/j.issn.2072-1439.2015.08.04] [Medline: 26623119]
36. Keller K, Hobohm L, Münzel T, et al. Survival benefit of obese patients with pulmonary embolism. Mayo Clin Proc 2019 Oct;94(10):1960-1973. [doi: 10.1016/j.mayocp.2019.04.035] [Medline: 31585580]

Abbreviations

CHARLS: China Health and Retirement Longitudinal Study
CLD: chronic lung disease
CVAI: Chinese visceral adiposity index
HDL-C: high-density lipoprotein cholesterol
HR: hazard ratio
RCS: restricted cubic spline
WHO: World Health Organization

Edited by Z Wu; submitted 06.Jun.2025; peer-reviewed by F Adjei, L Huang, Y Zhang; revised version received 16.Dec.2025; accepted 20.Apr.2026; published 20.May.2026.

Please cite as:
Zhu X, Tian H, Liu Z, Huang Z, Yuan M, Tu Z, Cui K, Liu J, Gong X
Association Between Low Chinese Visceral Adiposity Index Values and Chronic Lung Disease in Older Adults: Cohort Study
JMIR Public Health Surveill 2026;12:e78627
URL: https://publichealth.jmir.org/2026/1/e78627
doi:10.2196/78627

© Xiaorong Zhu, Hui Tian, Zhi Liu, Zhoubin Huang, Miaomiao Yuan, Zhangshun Tu, Kaiwang Cui, Jianping Liu, Xiangwen Gong. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 20.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Postpandemic Change in Demographic and Clinical Features of Patients With Omicron Who Were Hospitalized: Territory-Wide Retrospective Repeated Cross-Sectional Study in Hong Kong

Christie J Y Ching[1], MPH; Sunny C L Chan[1], MStat; Teddy T L Lee[1], MPH; Hugo H H Pui[1], BNurs; Bosco K H Leung[1]; Man Sing Wong[2], PhD; Tafu Yamamoto[3], MBChB; Chak Kwan Tong[4], MBBS; Cantian Wang[5], BMed; Timothy H Rainer[1], MD; Abraham K C Wai[1,5,6], MD

[1]Department of Emergency Medicine, School of Clinical Medicine, University of Hong Kong, G06, G/F, University of Hong Kong the Hong Kong Jockey Club Building for Interdisciplinary Research, 5 Sassoon Road, Pokfulam, Hong Kong, China

[2]Department of Land Surveying and Geo-informatics, Faculty of Construction and Environment, The Hong Kong Polytechnic University, Hong Kong, China

[3]Department of Accident & Emergency, Yan Chai Hospital, Hong Kong, China

[4]Intensive Care Unit, Tuen Mun Hospital, Hong Kong, China

[5]Department of Accident & Emergency, University of Hong Kong - Shenzhen Hospital, Shenzhen, China

[6]Department of Accident & Emergency, Queen Mary Hospital, Hong Kong, China

Corresponding Author:
Abraham K C Wai, MD
Department of Emergency Medicine, School of Clinical Medicine, University of Hong Kong, G06, G/F, University of Hong Kong the Hong Kong Jockey Club Building for Interdisciplinary Research, 5 Sassoon Road, Pokfulam, Hong Kong, China

Abstract

Background: The Omicron variant of SARS-CoV-2 underwent several mutations since it was first identified in November 2021, with a large outbreak in Hong Kong in early 2022. Yet, local cases of Omicron infections persist, even though the COVID-19 pandemic ended in May 2023.

Objective: This study aims to describe the changes in demographic and clinical characteristics of patients infected with COVID-19 across different Omicron waves in Hong Kong and determine whether the changes continued into the postpandemic period.

Methods: This retrospective repeated cross-sectional study collected data on patients infected with COVID-19 admitted to public hospitals in Hong Kong between May 1, 2022, and May 31, 2024. These data were later categorized into 3 periods based on the Omicron strain. A subsequent age-stratified descriptive analysis was conducted on each characteristic to identify any significant differences across the periods.

Results: First, the case fatality ratio significantly lowered among those older than 85 years (1.5% proportion decrease, period 1: 11.6%, period 2: 10.1%, effect size: 0.02; $P<.001$). Second, most patients were Chinese (≥68.7% per age group and period), and females were predominantly infected for those aged older than 85 years (≥56.9% per period). Third, the Charlson Comorbidity Index scores in most age groups showed a predominant proportion of infected individuals with 0 scores (more than 70% per period). Fourth, most cases were from slightly disadvantaged populations in Hong Kong (≥30.5% per age group per period). Fifth, clinical management of Omicron hospitalizations showed lowered length of hospital stays among adults and older individuals (≥1 d decrease between periods 1 and 3, per age group), as well as increased administration of bronchodilators.

Conclusions: Despite the decreasing incidence of Omicron cases admitted to public hospitals in Hong Kong, the increasing case fatality ratio with age suggests that long-term surveillance of COVID-19 should be maintained to prepare for potential mutations and outbreaks.

(JMIR Public Health Surveill 2026;12:e75635) doi:10.2196/75635

KEYWORDS

emergency department; COVID-19; infectious disease epidemiology; Omicron; SARS-CoV-2

Introduction

The Omicron outbreak in early 2022 resulted in a crude population mortality rate of 37.7 per million, overwhelming the Hong Kong health system [1,2]. On May 5, 2023, the World Health Organization (WHO) officially declared an end to the pandemic [3]. Nevertheless, reports of COVID-19 outbreaks persisted in Hong Kong and globally [4,5]. By May 2024, the

XSL•FO
RenderX

Communicable Disease Branch of the Center for Health Protection in Hong Kong reported that JN.1, the local dominant SARS-CoV-2 variant, did not cause more severe disease compared to XBB and its descendant lineages, the previous dominant strain [6].

Postpandemic surveillance is crucial for understanding the ongoing impact of COVID-19 and improving future pandemic preparedness. Current literature on the post–COVID-19 era primarily focuses on COVID-19 (long COVID-19) and delayed health care for patients with noncommunicable diseases during the pandemic. Long COVID-19 is a multisystemic condition manifesting as new onset cardiovascular disease, thrombotic disease, cerebrovascular disease, myalgic encephalomyelitis or chronic fatigue syndrome, type 2 diabetes, or postural orthostatic tachycardia syndrome [7]. Delayed health care for patients with noncommunicable diseases has resulted in complications and multimorbidity, contributing to increased health care costs [8,9].

An updated characterization of hospitalized patients with COVID-19 can provide a clearer clinical picture after the COVID-19 pandemic. Previous studies have identified several factors associated with worse prognosis in hospitalized patients with COVID-19, including blood biomarkers [10-13], hypertensive and diabetic medications [14,15], older age [16,17], males [12,13,17], hypertension and diabetes comorbidities [18,19], frailty [20,21], length of hospital stay [22], and lower socioeconomic background [23,24]. A longitudinal investigation into these parameters may enhance our understanding of the ongoing impact of COVID-19.

The aim of this study is to describe the changes in demographic and clinical characteristics of hospitalized patients infected with Omicron across different waves in Hong Kong, and to determine whether these changes continued into the postpandemic period. By conducting these comparisons, we provide insights into the evolving clinical profile of COVID-19 to inform the development of future public health strategies.

Methods

Participants and Context

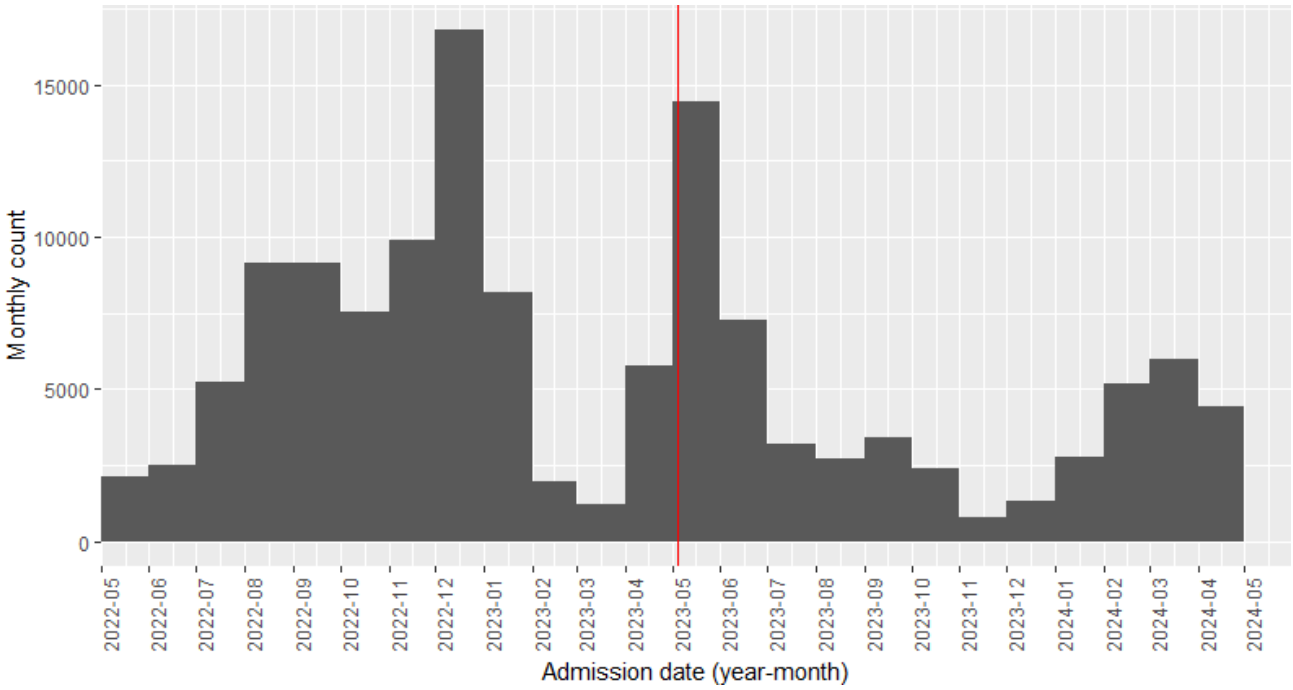
This study used a territory-wide, retrospective, repeated, cross-sectional study design to collect electronic clinical data from May 1, 2022, to May 31, 2024, sourced from the Clinical Data Analysis and Reporting System, managed by the Hong Kong Hospital Authority [25]. The Clinical Data Analysis and Reporting System encompasses data from 18 public hospitals in Hong Kong and has been validated in previous studies for its reliability [26]. Also, the patients included in this study were identified through confirmatory laboratory test results (reverse transcription polymerase chain reaction, multiplex polymerase chain reaction, and polymerase chain reaction) conducted in hospitals or government public health laboratories during the study period.

The study period was divided into 3 groups according to the epidemic waves illustrated in Figure 1:

1. Period 1: between May 1, 2022, and February 28, 2023.
2. Period 2: between March 1, 2023, and November 30, 2023.
3. Period 3: between December 1, 2023, and May 31, 2024.

Based on WHO's declaration of an end to the pandemic on May 5, 2023, [3], changes that continued into the postpandemic period were identified by differences in characteristics between periods 2 and 3. In addition, this study adopted episode count as the primary unit of measurement. An episode was defined as a single hospital admission to discharge occurrence, regardless of readmission of unique patients.

Figure 1. A histogram plot showing the monthly episodes of patients infected with Omicron hospitalized in Hong Kong between 2022 and 2024 in this repeated cross-sectional study (total episodes=136,544). The red line indicates the end of the pandemic.



XSL•FO
RenderX

Defining Demographic Characteristics

This study categorized ages into 5 groups: 0 - 17, 18 - 64, 65 - 74, 75 - 84, and ≥85 years. Socioeconomic Deprivation Index (SDI) was defined using the deprivation score in each tertiary planning unit in Hong Kong [27]. The score was calculated based on the 2021 population census, considering factors such as marital status, school attendance, working population, monthly domestic household income, household size, and tenure of accommodation [28]. Charlson Comorbidity Index (CCI) score was calculated using *International Statistical Classification of Diseases and Related Health Problems, 10th Revision* (*ICD-10*) codes [29], and categorized into 4 groups: 0, 1, 2 - 3, and ≥4. Specific comorbidities were identified, including unspecified essential hypertension (*ICD-10*: I10), type 2 diabetes mellitus (*ICD-10*: E11.0-E11.9), and unspecified hyperlipidemia (*ICD-10*: E78.5). Frailty-related episodes were identified based on co-occurring *ICD-10* diagnoses related to frailty markers, validated in a previous study [30]. The case fatality ratio (CFR) was defined, according to the WHO definition, as the proportion of COVID-19–related deaths per confirmed case of COVID-19 [31].

Defining Clinical Characteristics

All in-patient drugs administered during each episode were categorized according to Table S2 in Multimedia Appendix 1. For the biomarkers of interest, the earliest laboratory requests made during each episode were obtained to reflect the biomarker value closest to the date of COVID-19 diagnosis. For biomarkers with less than 25% missing data, multiple imputation using k-nearest neighbors with Gower Distance was applied [32]. Also, patients with more than 2 hospital admissions were used to measure readmission.

Descriptive Analysis

Age-stratified descriptive analyses were used to characterize the changes in demographic and clinical characteristics across distinct pandemic periods. Categorical variables (≥2 hospital readmissions, sex, comorbidities, race, CCI score, frailty-related episodes, Social Deprivation Index, and drug administration) were presented as frequency (n, %). Continuous variables (length of stay and blood biomarkers) were reported as median (IQR).

To account for varying period durations, the measurements were standardized: categorical variables as episodes per month and continuous variables as mean values per month.

Continuous variables across periods were compared using the Kruskal-Wallis test; categorical variables were compared using the proportions test (Fisher exact test was used if the proportions test assumptions were violated). Post hoc pairwise comparisons of significant differences were performed (proportions test or Fisher exact test for categorical, Dunn test for continuous variables), with Bonferroni correction for multiple testing.

The effect sizes for significant variables were measured using Cramér V for categorical and Eta squared test for continuous variables. All *P* values were 2-sided; statistical significance was defined as $P<.05$. Analyses were conducted with R (version 4.2.2; R Core Team) [33].

Ethical Considerations

This study received ethical approval from The University of Hong Kong Institutional Review Board (UW 20 - 112). Informed consent was waived by the institutional review board as all the patients' data were collected anonymously. The privacy and confidentiality of human subjects were protected by maintaining anonymity, not collecting personal data, and ensuring that data were securely stored by the principal investigator and deleted after the storage period. This study was conducted in accordance with the Declaration of Helsinki, and no compensation was provided for the human subjects involved.

Results

Hospitalization and Epidemic Trends

Between May 2022 and May 2024, Figure 1 and Table 1 illustrated the 3 distinct epidemic curves corresponding to periods 1-3, each exhibiting progressively lower peaks. This trend was further highlighted in Tables 2-6, which showed depleting total number of hospitalized Omicron cases across the study period among all age groups: 0 - 17 years (75.2% decrease, Period 1: 8399, Period 3: 2081); 18 - 64 years (71.6% decrease, Period 1: 15,274, Period 3: 4340); 65 - 74 years (69.9% decrease, Period 1: 13,549, Period 3: 4072); 75 - 84 years (67.5% decrease, Period 1: 15,622, Period 3: 5076); and older than 85 years (62.8% decrease, Period 1: 19,717, Period 3: 7337). The difference between periods 1 and 3 diminished as the age group got older. Moreover, the rate of hospital readmissions per month decreased across all ages: 0 - 17 years (14.8 episodes per month decrease, Period 1: 41.6 episodes per month, Period 3: 26.8 episodes per month); 18 - 64 years (225.8 episodes per month decrease, Period 1: 296.0 episodes per month, Period 3: 70.2 episodes per month); 65 - 74 years (249.8 episodes per month decrease, Period 1: 346.8 episodes per month, Period 3: 97 episodes per month); 75 - 84 years (332 episodes per month decrease, Period 1: 484.5 episodes per month, Period 3: 152.5 episodes per month); and older than 85 years (480.8 episodes per month decrease, Period 1: 754 episodes per month, Period 3: 273.2 episodes per month).

XSL•FO
RenderX

Table . Frequency of hospitalized patients infected with Omicron in Hong Kong from this repeated cross-sectional study (2022-2024) (total episodes, N=136,544).

Date	Values, n (%)
May 1, 2022	2076 (1.5)
June 1, 2022	2447 (1.8)
July 1, 2022	5202 (3.8)
January 8, 2022	8929 (6.5)
September 1, 2022	9314 (6.8)
October 1, 2022	7483 (5.5)
November 1, 2022	9735 (7.1)
December 1, 2022	16,695 (12.2)
January 1, 2023	8492 (6.2)
February 1, 2023	2098 (1.5)
March 1, 2023	1165 (0.9)
April 1, 2023	5487 (4)
May 1, 2023	14,395 (10.5)
June 1, 2023	7487 (5.5)
July 1, 2023	3276 (2.4)
August 1, 2023	2654 (1.9)
January 9, 2023	3379 (2.5)
October 1, 2023	2484 (1.8)
November 1, 2023	840 (0.6)
December 1, 2023	1246 (0.9)
January 1, 2024	2757 (2)
February 1, 2024	5039 (3.7)
March 1, 2024	6077 (4.5)
April 1, 2024	4407 (3.2)
May 1, 2024	3380 (2.5)

XSL•FO
RenderX

Table . Characteristics of patients infected with Omicron in Hong Kong from this repeated cross-sectional study (2022 - 2024), stratified by ages 0 - 17 years (total episodes, N=14,556).

Characteristics	Period 1[a]		Period 2[b]		Period 3[c]		*P* value	Effect size
	Observed	Adjusted	Observed	Adjusted	Observed	Adjusted		
Case fatality ratio	0.1	__[d]	0.1	—	0	—	—	—
≥2 hospital readmissions, n (%)	416 (5)	41.6 (0.5)	250 (6.1)	27.8 (0.68)	161 (7.7)	26.8 (1.28)	<.001[e]	0.89[f]
Sex, n (%)								
Male	4689 (55.8)	468.9 (5.58)	2167 (53.2)	240.8 (5.91)	1128 (54.2)	188 (9.03)	.02[e]	0.02[f]
Female	3710 (44.2)	371 (4.42)	1909 (46.8)	212.1 (5.2)	953 (45.8)	158.8 (7.63)	.02[e]	0.02[f]
Comorbidities, n (%)								
Essential primary hypertension	1 (0)	0.1 (0)	0 (0)	0 (0)	0 (0)	0 (0)	.37[g]	—
Type 2 diabetes mellitus	3 (0)	0.3 (0)	2 (0)	0.2 (0)	0 (0)	0 (0)	.69[g]	—
Hyperlipidemia (unspecified)	1 (0)	0.1 (0)	0 (0)	0 (0)	0 (0)	0 (0)	>.99[g]	—
Race, n (%)								
Chinese	6483 (77.2)	648.3 (7.72)	3035 (74.5)	337.2 (8.28)	1429 (68.7)	238.2 (11.45)	<.001[e]	0.07[f]
Non-Chinese	1916 (22.8)	191.6 (2.28)	1041 (25.5)	115.7 (2.83)	652 (31.3)	108.7 (5.21)	<.001[e]	0.07[f]
Length of hospital stay (days), median (IQR)	2 (2)	3.19 (0.77)	2 (2)	2.79 (0.83)	2 (2)	2.56 (0.54)	<.001[h]	0.009[i]
Charlson Comorbidity Index score, n (%)								
0	8236 (98.1)	823.6 (9.81)	3962 (97.2)	440.2 (10.8)	2023 (97.6)	337.2 (16.26)	.003[e]	0.03[f]
1	69 (0.8)	6.9 (0.08)	47 (1.2)	5.2 (0.13)	18 (0.9)	3 (0.15)	.18[e]	—
2 - 3	87 (1)	8.7 (0.1)	63 (1.5)	7 (0.17)	28 (1.4)	4.7 (0.23)	.04[e]	0.02[f]
≥4	5 (0.1)	0.5 (0.01)	4 (0.1)	0.4 (0.01)	3 (0.1)	0.5 (0.02)	.36[g]	—
Frailty-related episodes, n (%)	58 (0.7)	5.8 (0.07)	24 (0.6)	2.7 (0.07)	8 (0.4)	1.3 (0.07)	.26[e]	—
Social deprivation index, n (%)								
1 (least disadvantaged)	1146 (14)	114.6 (1.4)	620 (15.3)	68.9 (1.7)	309 (14.9)	51.5 (2.48)	.04[e]	0.02[f]
2 (slightly disadvantaged)	2935 (35.7)	293.5 (3.57)	1477 (36.4)	164.1 (4.04)	750 (36.1)	125 (6.02)	.31[e]	—
3 (moderately disadvantaged)	2439 (29.7)	243.9 (2.97)	1070 (26.4)	118.9 (2.93)	544 (26.2)	90.7 (4.37)	<.001[e]	0.03[f]
4 (most disadvantaged)	1695 (20.6)	169.5 (2.06)	891 (22)	99 (2.44)	474 (22.8)	79 (3.8)	.01[e]	0.02[f]
Drug administration, n (%)								

XSL•FO
RenderX

Characteristics	Period 1[a]		Period 2[b]		Period 3[c]		*P* value	Effect size
	Observed	Adjusted	Observed	Adjusted	Observed	Adjusted		
Angiotensin-converting enzyme inhibitors	23 (0.3)	2.3 (0.03)	12 (0.3)	1.3 (0.03)	7 (0.3)	1.2 (0.05)	.89[e]	—
Antidiabetics	29 (0.3)	2.9 (0.03)	16 (0.4)	1.8 (0.04)	8 (0.4)	1.3 (0.07)	.91[e]	—
Antiplatelets and anticoagulants	60 (0.7)	6 (0.07)	48 (1.2)	5.3 (0.13)	25 (1.2)	4.2 (0.2)	.01[e]	0.02[f]
Beta blockers	26 (0.3)	2.6 (0.03)	5 (0.1)	0.6 (0.01)	1 (0)	0.2 (0.0)	.02[g]	0.02[f]
Bronchodilators	243 (2.9)	24.3 (0.29)	220 (5.4)	24.4 (0.6)	148 (7.1)	24.7 (1.18)	<.001[e]	0.08[f]
Calcium channel blocker	14 (0.2)	1.4 (0.02)	17 (0.4)	1.9 (0.04)	2 (0.1)	0.3 (0.02)	.01[g]	0.02[f]
Diuretics	49 (0.6)	4.9 (0.06)	42 (1)	4.7 (0.11)	20 (1)	3.3 (0.17)	.01[e]	0.02[f]
Inhaled corticosteroids	174 (2.1)	17.4 (0.21)	114 (2.8)	12.7 (0.31)	87 (4.2)	14.5 (0.7)	<.001[e]	0.04[f]
Rheumatoid drugs	21 (0.3)	2.1 (0.03)	17 (0.4)	1.9 (0.04)	5 (0.2)	0.8 (0.03)	.24[e]	—
Statins	1 (0)	0.1 (0)	3 (0.1)	0.3 (0.01)	0 (0)	0 (0)	.17[g]	—
Systemic corticosteroids	715 (8.5)	71.5 (0.85)	400 (9.8)	44.4 (1.09)	189 (9.1)	31.5 (1.52)	.06[e]	—
Blood biomarkers, median (IQR)								
Albumin (g/L)	38.82 (3.98)	40.4 (0.3)	38.82 (3.18)	39.8 (0.1)	38.82 (3.58)	39.9 (0.2)	<.001[h]	0.003[i]
Neutrophil ($\times 10^9$/L)	4.48 (1.27)	4.4 (0.2)	4.48 (1.46)	4.6 (0.1)	4.48 (1.75)	4.8 (0.1)	<.001[h]	0.0009[i]
Bilirubin (μmol/L)	9.30 (14.20)	12.4 (0.6)	9.03 (14.10)	13 (0.6)	8.80 (14.10)	12.6 (0.3)	.67[h]	—
Lymphocyte ($\times 10^9$/L)	0.86 (0.89)	1.6 (0.2)	1 (1.44)	1.9 (0.2)	1.40 (2.14)	2.3 (0.04)	<.001[h]	0.02[i]
Platelet ($\times 10^9$/L)	189 (127.50)	216 (7.5)	211 (155.25)	234.6 (7.5)	232 (182)	252.2 (3.2)	<.001[h]	0.01[i]

[a]Episodes=8399 and head count=8179.
[b]Episodes=4076 and head count=3940.
[c]Episodes=2081 and head count=1995.
[d]Not available.
[e]Proportions test.
[f]Cramér V test.
[g]Fisher exact test.
[h]Kruskal-Wallis test.
[i]Eta squared test.

XSL•FO
RenderX

Table . Characteristics of patients infected with Omicron in Hong Kong from this repeated cross-sectional study (2022 - 2024), stratified by ages 18 - 64 years (total episodes, N=28,183).

Characteristics	Period 1[a]		Period 2[b]		Period 3[c]		*P* value	Effect size
	Observed	Adjusted	Observed	Adjusted	Observed	Adjusted		
Case fatality ratio	2.7	__[d]	2.9	—	2.6	—	—	—
≥2 hospital readmissions, n (%)	2960 (19.4)	296 (1.94)	936 (10.9)	104 (1.21)	421 (9.7)	70.2 (1.62)	<.001[e]	0.72[f]
Sex, n (%)								
Male	7523 (49.3)	752.3 (4.93)	3984 (46.5)	442.7 (5.17)	2053 (47.3)	342.2 (7.88)	<.001[e]	0.02[f]
Female	7751 (50.7)	775.1 (5.07)	4585 (53.5)	509.4 (5.94)	2287 (52.7)	381.2 (8.78)	<.001[e]	0.02[f]
Comorbidities, n (%)								
Essential primary hypertension	432 (2.8)	43.2 (0.28)	286 (3.3)	31.8 (0.37)	161 (3.7)	26.8 (0.62)	.005[e]	0.02[f]
Type 2 diabetes mellitus	465 (3)	46.5 (0.3)	290 (3.4)	32.2 (0.38)	122 (2.8)	20.3 (0.47)	.16[e]	—
Hyperlipidemia (unspecified)	173 (1.1)	17.3 (0.11)	110 (1.3)	12.2 (0.14)	47 (1.1)	7.8 (0.18)	.49[e]	—
Race, n (%)								
Chinese	13,868 (90.8)	1386.8 (9.08)	7801 (91)	866.8 (10.1)	3928 (90.5)	654.7 (15.08)	.60[e]	—
Non-Chinese	1406 (9.2)	140.6 (0.92)	768 (9)	85.3 (1)	412 (9.5)	68.7 (1.58)	.60[e]	—
Length of hospital stay (days), median (IQR)	4 (6)	11 (2.7)	3 (5)	10 (3)	3 (4)	6.9 (2.3)	<.001[g]	0.007[h]
Charlson Comorbidity Index score, n (%)								
0	11,950 (78.5)	1195 (7.85)	6657 (78)	739.7 (8.67)	3248 (76.5)	541.3 (12.75)	<.001[e]	0.03[f]
1	1160 (7.6)	116 (0.76)	744 (8.7)	82.7 (0.97)	405 (9.5)	67.5 (1.58)	<.001[e]	0.02[f]
2 - 3	1644 (10.8)	164.4 (1.08)	848 (9.9)	94.2 (1.1)	444 (10.5)	74 (1.75)	.10[e]	—
≥4	471 (3.1)	47.1 (0.31)	287 (3.4)	31.9 (0.38)	150 (3.5)	25 (0.58)	.34[e]	—
Frailty-related episodes, n (%)	443 (2.9)	44.3 (0.29)	272 (3.2)	30.2 (0.36)	124 (2.9)	20.7 (0.48)	.43[e]	—
Social deprivation index, n (%)								
1 (least disadvantaged)	2148 (14.3)	214.8 (1.43)	1118 (13.1)	124.2 (1.46)	573 (13.3)	95.5 (2.22)	.06[e]	—
2 (slightly disadvantaged)	5241 (34.9)	524.1 (3.49)	3042 (35.7)	338 (3.97)	1541 (35.7)	256.8 (5.95)	.11[e]	—
3 (moderately disadvantaged)	4343 (28.9)	434.3 (2.89)	2449 (28.7)	272.1 (3.19)	1222 (28.3)	203.7 (4.72)	.88[e]	—
4 (most disadvantaged)	3283 (21.9)	328.3 (2.19)	1915 (22.5)	212.8 (2.5)	977 (22.7)	162.8 (3.78)	.18[e]	—
Drug administration, n (%)								

XSL•FO
RenderX

Characteristics	Period 1[a]		Period 2[b]		Period 3[c]		*P* value	Effect size
	Observed	Adjusted	Observed	Adjusted	Observed	Adjusted		
Angiotensin-converting enzyme inhibitors	2496 (16.3)	249.6 (1.63)	1559 (18.2)	173.2 (2.02)	802 (18.5)	133.7 (3.08)	<.001[e]	0.02[f]
Antidiabetics	2627 (17.2)	262.7 (1.72)	1682 (19.6)	186.9 (2.18)	822 (18.9)	137 (3.15)	<.001[e]	0.03[f]
Antiplatelets and anticoagulants	4311 (28.2)	431.1 (2.82)	2526 (29.5)	280.7 (3.28)	1275 (29.4)	212.5 (4.9)	.08[e]	—
Beta blockers	2046 (13.4)	204.6 (1.34)	332 (3.9)	36.9 (0.43)	40 (0.9)	6.7 (0.15)	<.001[e]	0.52[f]
Bronchodilators	1017 (6.7)	101.7 (0.67)	779 (9.1)	86.6 (1.01)	530 (12.2)	88.3 (2.03)	<.001[e]	0.07[f]
Calcium channel blocker	3406 (22.3)	340.6 (2.23)	2060 (24)	228.9 (2.67)	1079 (24.9)	179.8 (4.15)	<.001[e]	0.02[f]
Diuretics	1760 (11.5)	176 (1.15)	1074 (12.5)	119.3 (1.39)	613 (14.1)	102.2 (2.35)	<.001[e]	0.03[f]
Inhaled corticosteroids	783 (5.1)	78.3 (0.51)	479 (5.6)	53.2 (0.62)	279 (6.4)	46.5 (1.07)	<.001[e]	0.02[f]
Rheumatoid drugs	881 (5.8)	88.1 (0.58)	704 (8.2)	78.2 (0.91)	336 (7.7)	56 (1.28)	<.001[e]	0.04[f]
Statins	3038 (19.9)	303.8 (1.99)	2026 (23.6)	225.1 (2.62)	1039 (23.9)	173.2 (3.98)	<.001[e]	0.05[f]
Systemic corticosteroids	2867 (18.8)	286.7 (1.88)	2330 (27.2)	258.9 (3.02)	1113 (25.6)	185.5 (4.27)	<.001[e]	0.09[f]
Blood biomarkers, median (IQR)								
Albumin (g/L)	38.70 (7.10)	37.3 (0.9)	38 (7.50)	36.9 (0.3)	37.90 (8)	36.7 (0.4)	<.001[g]	0.002[h]
Neutrophil ($\times 10^9$/L)	4.48 (3.13)	5.6 (0.4)	4.90 (3.58)	5.8 (0.1)	5.12 (3.91)	6.4 (0.2)	<.001[g]	0.005[h]
Bilirubin (μmol/L)	8 (7.50)	11.6 (0.7)	7.90 (6.60)	10.8 (0.5)	8 (6.70)	11 (0.7)	<.001[g]	0.001[h]
Lymphocyte ($\times 10^9$/L)	1 (0.89)	1.5 (0.06)	0.96 (0.87)	1.6 (0.1)	1.04 (0.96)	1.7 (0.2)	<.001[g]	0.002[h]
Platelet ($\times 10^9$/L)	208 (109)	221.5 (6)	206 (104)	220.5 (4.6)	217 (115)	227.8 (4.7)	<.001[g]	0.001[h]

[a]Episodes=15,274 and head count=13,626.

[b]Episodes=8569 and head count=8063.

[c]Episodes=4340 and head count=4107.

[d]Not available.

[e]Proportions test.

[f]Cramér V test.

[g]Kruskal-Wallis test.

[h]Eta squared test.

XSL•FO
RenderX

Table . Characteristics of patients infected with Omicron in Hong Kong from this repeated cross-sectional study (2022 - 2024), stratified by ages 65 - 74 years (total episodes, N=25,045).

Characteristics	Period 1[a]		Period 2[b]		Period 3[c]		*P* value	Effect size
	Observed	Adjusted	Observed	Adjusted	Observed	Adjusted		
Case fatality ratio	4.9	__[d]	5.2	—	5.5	—	—	—
≥2 hospital readmissions, n (%)	3468 (25.8)	346.8 (2.58)	1074 (14.5)	119.3 (1.61)	582 (14.3)	97 (2.38)	<.001[e]	0.64[f]
Sex, n (%)								
Male	8190 (60.9)	819 (6.09)	4287 (57.7)	476.3 (6.41)	2445 (60)	407.5 (10)	<.001[e]	0.02[f]
Female	5269 (39.1)	526.9 (3.91)	3137 (42.3)	348.6 (4.7)	1627 (40)	271.2 (6.67)	<.001[e]	0.02[f]
Comorbidities, n (%)								
Essential primary hypertension	662 (4.9)	66.2 (0.49)	445 (6)	49.4 (0.67)	227 (5.6)	37.8 (0.93)	.002[e]	0.02[f]
Type 2 diabetes mellitus	677 (5)	67.7 (0.5)	466 (6.3)	51.8 (0.7)	191 (4.7)	31.8 (0.78)	<.001[e]	0.08[f]
Hyperlipidemia (unspecified)	294 (2.2)	29.4 (0.22)	215 (2.9)	23.9 (0.32)	106 (2.6)	17.7 (0.43)	.004[e]	0.02[f]
Race, n (%)								
Chinese	12,876 (95.7)	1287.6 (9.57)	7124 (96)	791.6 (10.67)	3894 (95.6)	649 (15.93)	.006[e]	0.02[f]
Non-Chinese	583 (4.3)	58.3 (0.43)	300 (4)	33.3 (0.44)	178 (4.4)	29.7 (0.73)	.006[e]	0.02[f]
Length of hospital stay (days), median (IQR)	6 (7)	10.4 (1.3)	4 (7)	9.1 (1.8)	4 (6)	7.8 (2.3)	<.001[g]	0.01[h]
Charlson Comorbidity Index score, n (%)								
0	9740 (72.5)	974 (7.25)	5286 (71.3)	587.3 (7.92)	2845 (71.2)	474.2 (11.87)	.04[e]	0.01[f]
1	1706 (12.7)	170.6 (1.27)	1072 (14.5)	119.1 (1.61)	551 (13.8)	91.8 (2.3)	<.001[e]	0.02[f]
2 - 3	1516 (11.3)	151.6 (1.13)	769 (10.4)	85.4 (1.16)	432 (10.8)	72 (1.8)	.16[e]	—
≥4	479 (3.6)	47.9 (0.36)	282 (3.8)	31.3 (0.42)	169 (4.2)	28.2 (0.7)	.17[e]	—
Frailty-related episodes, n (%)	674 (5)	67.4 (0.5)	490 (6.6)	54.4 (0.73)	218 (5.4)	36.3 (0.9)	<.001[e]	0.03[f]
Social deprivation index, n (%)								
1 (least disadvantaged)	1788 (13.3)	178.8 (1.33)	996 (13.5)	110.7 (1.5)	544 (13.4)	90.7 (2.23)	.90[e]	—
2 (slightly disadvantaged)	4879 (36.4)	487.9 (3.64)	2636 (35.6)	292.9 (3.96)	1465 (36.1)	244.2 (6.02)	.76[e]	—
3 (moderately disadvantaged)	3901 (29.1)	390.1 (2.91)	2119 (28.6)	235.4 (3.18)	1184 (29.2)	197.3 (4.87)	.83[e]	—
4 (most disadvantaged)	2826 (21.1)	282.6 (2.11)	1649 (22.3)	183.2 (2.48)	868 (21.4)	144.7 (3.57)	.07[e]	—
Drug administration, n (%)								

XSL•FO
RenderX

Characteristics	Period 1[a]		Period 2[b]		Period 3[c]		*P* value	Effect size
	Observed	Adjusted	Observed	Adjusted	Observed	Adjusted		
Angiotensin-converting enzyme inhibitors	4496 (33.4)	449.6 (3.34)	2545 (34.3)	282.8 (3.81)	1342 (33)	223.7 (5.5)	.20[e]	—
Antidiabetics	4454 (33.1)	445.4 (3.31)	2631 (35.4)	292.3 (3.93)	1412 (34.7)	235.3 (5.78)	<.001[e]	0.02[f]
Antiplatelets and anticoagulants	6940 (51.6)	694 (5.16)	3873 (52.2)	430.3 (5.8)	2115 (51.9)	352.5 (8.65)	.38[e]	—
Beta blockers	3311 (24.6)	331.1 (2.46)	569 (7.7)	63.2 (0.86)	52 (1.3)	8.7 (0.22)	<.001[e]	0.27[f]
Bronchodilators	1807 (13.4)	180.7 (1.34)	1117 (15)	124.1 (1.67)	813 (20)	135.5 (3.33)	<.001[e]	0.07[f]
Calcium channel blocker	5867 (43.6)	586.7 (4.36)	3240 (43.6)	360 (4.84)	1727 (42.4)	287.8 (7.07)	.44[e]	—
Diuretics	2412 (17.9)	241.2 (1.79)	1326 (17.9)	147.3 (1.99)	912 (22.4)	152 (3.73)	<.001[e]	0.04[f]
Inhaled corticosteroids	1013 (7.5)	101.3 (0.75)	540 (7.3)	60 (0.81)	359 (8.8)	59.8 (1.47)	.007[e]	0.02[f]
Rheumatoid drugs	1358 (10.1)	135.8 (1.01)	936 (12.6)	104 (1.4)	441 (10.8)	73.5 (1.8)	<.001[e]	0.04[f]
Statins	5901 (43.8)	590.1 (4.38)	3701 (49.9)	411.2 (5.54)	1993 (48.9)	332.2 (8.15)	<.001[e]	0.06[f]
Systemic corticosteroids	3665 (27.2)	366.5 (2.72)	2622 (35.3)	291.3 (3.92)	1491 (36.6)	248.5 (6.1)	<.001[e]	0.09[f]
Blood biomarkers, median (IQR)								
Albumin (g/L)	36.80 (7.90)	35.8 (1.3)	36.30 (7.50)	35.4 (0.3)	35.60 (8.20)	34.6 (0.3)	<.001[g]	0.004[h]
Neutrophil ($\times 10^9$/L)	4.62 (3.57)	5.7 (0.5)	5.18 (3.69)	6.1 (0.2)	5.86 (4.47)	6.9 (0.08)	<.001[g]	0.02[h]
Bilirubin (μmol/L)	8.80 (7.20)	11.7 (0.7)	9 (7)	11.5 (0.6)	9 (7.20)	12.5 (0.5)	.16[g]	—
Lymphocyte ($\times 10^9$/L)	0.99 (0.74)	1.4 (0.07)	0.90 (0.79)	1.5 (0.2)	1 (0.83)	1.5 (0.09)	<.001[g]	0.001[h]
Platelet ($\times 10^9$/L)	201 (104.15)	215.9 (7.9)	195 (103)	209 (4.2)	206 (108)	222.8 (2.8)	<.001[g]	0.002[h]

[a]Episodes=13,549 and head count=11,517.

[b]Episodes=7424 and head count=6848.

[c]Episodes=4072 and head count=3760.

[d]Not available.

[e]Proportions test.

[f]Cramér V test.

[g]Kruskal-Wallis test.

[h]Eta squared test.

XSL•FO
RenderX

Table . Characteristics of patients infected with Omicron in Hong Kong from this repeated cross-sectional study (2022 - 2024), stratified by ages 75 - 84 years (total episodes, N=29,743).

Characteristics	Period 1[a]		Period 2[b]		Period 3[c]		*P* value	Effect size
	Observed	Adjusted	Observed	Adjusted	Observed	Adjusted		
Case fatality ratio	7.5	__[d]	7.3	—	7.5	—	—	—
≥2 hospital readmissions, n (%)	4845 (31)	484.5 (3.1)	1453 (16.1)	161.4 (1.79)	915 (18)	152.5 (3)	<.001[e]	0.57[e]
Sex, n (%)								
Male	9265 (59.3)	926.5 (5.93)	5280 (58.4)	586.7 (6.49)	2890 (58.7)	481.7 (9.78)	.01[f]	0.02[e]
Female	6357 (40.7)	635.7 (4.07)	3765 (41.6)	418.3 (4.62)	2096 (41.3)	349.3 (6.88)	.01[f]	0.02[e]
Comorbidities, n (%)								
Essential primary hypertension	958 (6.1)	95.8 (0.61)	625 (6.9)	69.4 (0.77)	320 (6.3)	53.3 (1.05)	.05[f]	—
Type 2 diabetes mellitus	884 (5.7)	88.4 (0.57)	573 (6.3)	63.7 (0.7)	269 (5.3)	44.8 (0.88)	<.001[f]	0.01[e]
Hyperlipidemia (unspecified)	372 (2.4)	37.2 (0.24)	293 (3.2)	32.6 (0.36)	156 (3.1)	26 (0.52)	<.001[f]	0.02[e]
Race, n (%)								
Chinese	15,178 (97.2)	1517.8 (9.72)	8752 (96.8)	972.4 (10.8)	4902 (96.6)	817 (16.1)	.06[f]	—
Non-Chinese	444 (2.8)	44.4 (0.28)	293 (3.2)	32.6 (0.36)	174 (3.4)	29 (0.57)	.06[f]	—
Length of hospital stay (days), median (IQR)	6 (8)	10.5 (0.4)	5 (6)	9.1 (1.1)	4 (6)	7.9 (2.2)	<.001[g]	0.01[h]
Charlson Comorbidity Index score, n (%)								
0	11,354 (72.7)	1135.4 (7.27)	6639 (73.5)	737.7 (8.17)	3558 (71.3)	593 (11.88)	<.001[f]	0.02[e]
1	2327 (14.9)	232.7 (1.49)	1401 (15.5)	155.7 (1.72)	815 (16.3)	135.8 (2.72)	.11[f]	—
2 - 3	1560 (10)	156 (1.0)	774 (8.6)	86 (0.96)	482 (9.7)	80.3 (1.62)	.001[f]	0.02[e]
≥4	372 (2.4)	37.2 (0.24)	223 (2.5)	24.8 (0.28)	136 (2.7)	22.7 (0.45)	.49[f]	—
Frailty-related episodes, n (%)	1306 (8.4)	130.6 (0.84)	750 (8.3)	83.3 (0.92)	420 (8.3)	70 (1.38)	.97[f]	—
Social deprivation index, n (%)								
1 (least disadvantaged)	2187 (14)	218.7 (1.4)	1315 (14.6)	146.1 (1.62)	671 (13.3)	111.8 (2.22)	.09[f]	—
2 (slightly disadvantaged)	5247 (33.7)	524.7 (3.37)	3039 (33.6)	337.7 (3.73)	1831 (36.2)	305.2 (6.03)	.003[f]	0.02[e]
3 (moderately disadvantaged)	4560 (29.3)	456 (2.93)	2503 (27.7)	278.1 (3.08)	1337 (26.4)	222.8 (4.4)	<.001[f]	0.02[e]
4 (most disadvantaged)	3589 (23)	358.9 (2.3)	2177 (24.1)	241.9 (2.68)	1219 (24.1)	203.2 (4.02)	.09[f]	—
Drug administration, n (%)								

XSL•FO
RenderX

Characteristics	Period 1[a]		Period 2[b]		Period 3[c]		*P* value	Effect size
	Observed	Adjusted	Observed	Adjusted	Observed	Adjusted		
Angiotensin-converting enzyme inhibitors	6016 (38.5)	601.6 (3.85)	3593 (39.7)	399.2 (4.41)	2044 (40.3)	340.7 (6.72)	.04[f]	0.01[e]
Antidiabetics	5439 (34.8)	543.9 (3.48)	3268 (36.1)	363.1 (4.01)	1844 (36.3)	307.3 (6.05)	.04[f]	0.01[e]
Antiplatelets and anticoagulants	9859 (63.1)	985.9 (6.31)	5696 (63)	632.9 (7)	3171 (62.5)	528.5 (10.42)	.71[f]	—
Beta blockers	4319 (27.6)	431.9 (2.76)	735 (8.1)	81.7 (0.9)	85 (1.7)	14.2 (0.28)	<.001[f]	0.29[e]
Bronchodilators	2745 (17.6)	274.5 (1.76)	1891 (20.9)	210.1 (2.32)	1260 (24.8)	210 (4.13)	<.001[f]	0.07[e]
Calcium channel blocker	8185 (54.2)	818.5 (5.42)	4613 (51)	512.6 (5.67)	2589 (51)	431.5 (8.5)	.06[f]	—
Diuretics	3316 (21.2)	331.6 (2.12)	2011 (22.2)	223.4 (2.47)	1315 (25.9)	219.2 (4.32)	<.001[f]	0.04[e]
Inhaled corticosteroids	1296 (8.3)	129.6 (0.83)	813 (9)	90.3 (1)	503 (9.9)	83.8 (1.65)	.001[f]	0.02[e]
Rheumatoid drugs	1894 (12.1)	189.4 (1.21)	1234 (13.6)	137.1 (1.51)	649 (12.8)	108.2 (2.13)	.003[f]	0.02[e]
Statins	7839 (50.2)	783.9 (5.02)	5044 (55.8)	560.4 (6.2)	2876 (56.7)	479.3 (9.45)	<.001[f]	0.06[e]
Systemic corticosteroids	4881 (31.2)	488.1 (3.12)	3460 (38.3)	384.4 (4.26)	1907 (37.6)	317.8 (6.27)	<.001[f]	0.07[e]
Blood biomarkers, median (IQR)								
Albumin (g/L)	35.20 (8.02)	34.6 (1.1)	35 (7.90)	34.4 (0.4)	34.50 (8.50)	33.8 (0.4)	<.001[g]	0.002[h]
Neutrophil ($\times 10^9$/L)	4.83 (3.70)	5.9 (0.5)	5.28 (3.77)	6.2 (0.2)	5.70 (4.24)	6.7 (0.2)	<.001[g]	0.01[h]
Bilirubin (μmol/L)	9 (7.10)	11.3 (0.5)	9.20 (7.20)	12.1 (0.4)	9 (6.80)	11.3 (0.3)	.03[g]	0.002[h]
Lymphocyte ($\times 10^9$/L)	0.98 (0.72)	1.4 (0.1)	0.90 (0.77)	1.7 (0.3)	1(0.84)	1.7 (0.2)	<.001[g]	0.002[h]
Platelet ($\times 10^9$/L)	194 (102)	210.9 (7.6)	189 (96)	203.7 (3.9)	200 (108)	217.1 (3.2)	<.001[g]	0.002[h]

[a]Episodes=15,622 and head count=12,935.

[b]Episodes=9045 and head count=8266.

[c]Episodes=5076 and head count=4572.

[d]Not available.

[e]Cramér V test.

[f]Proportions test.

[g]Kruskal-Wallis test.

[h]Eta squared test.

XSL•FO
RenderX

Table . Characteristics of patients infected with Omicron in Hong Kong from this repeated cross-sectional study (2022 - 2024), stratified by ages older than 85 years (total episodes, N=39,107).

Characteristics	Period 1[a]		Period 2[b]		Period 3[c]		*P* value	Effect size
	Observed	Adjusted	Observed	Adjusted	Observed	Adjusted		
Case fatality ratio	11.6	__[d]	10.1	—	10.1	—	—	—
≥2 hospital readmissions, n (%)	7540 (38.2)	754 (3.82)	2626 (21.8)	291.8 (2.42)	1639 (22.3)	273.2 (3.72)	<.001[e]	0.48[f]
Sex, n (%)								
Male	8444 (42.8)	844.4 (4.28)	5152 (42.7)	572.4 (4.74)	3161 (43.1)	526.8 (7.18)	.89[e]	—
Female	11,273 (57.2)	1127.3 (5.72)	6901 (57.3)	766.8 (6.37)	4176 (56.9)	696 (9.48)	.89[e]	—
Comorbidities, n (%)								
Essential primary hypertension	1369 (6.9)	136.9 (0.69)	910 (7.5)	101.1 (0.83)	509 (6.9)	84.8 (1.15)	.10[e]	—
Type 2 diabetes mellitus	933 (4.7)	93.3 (0.47)	550 (4.6)	61.1 (0.51)	264 (3.6)	44 (0.6)	<.001[e]	0.02[f]
Hyperlipidemia (unspecified)	418 (2.1)	41.8 (0.21)	299 (2.5)	33.2 (0.28)	192 (2.6)	32 (0.43)	.02[e]	0.01[f]
Race, n (%)								
Chinese	19,356 (98.2)	1935.6 (9.82)	11,832 (98.2)	1314.7 (10.91)	7193 (98)	1198.8 (16.3)	.76[e]	—
Non-Chinese	361 (1.8)	36.1 (0.18)	221 (1.8)	24.6 (0.2)	144 (2)	24 (0.33)	.76[e]	—
Length of hospital stay (days), median (IQR)	7 (9)	10.1 (0.8)	5 (6)	8.7 (1.1)	5 (5)	7.6 (1.7)	<.001[g]	0.02[h]
Charlson Comorbidity Index score, n (%)								
0	14,820 (75.2)	1482 (7.52)	9364 (77.7)	1040.4 (8.63)	5576 (77.3)	929.3 (12.88)	<.001[e]	0.02[f]
1	2952 (15)	295.2 (1.5)	1696 (14.1)	188.4 (1.57)	1014 (14.1)	169 (2.35)	.02[e]	0.01[f]
2 - 3	1617 (8.2)	161.7 (0.82)	810 (6.7)	90 (0.74)	529 (7.3)	88.2 (1.22)	<.001[e]	0.02[f]
≥4	310 (1.6)	31 (0.16)	178 (1.5)	19.8 (0.17)	98 (1.4)	16.3 (0.23)	.35[e]	—
Frailty-related episodes, n (%)	1903 (9.7)	190.3 (0.97)	1095 (9.1)	121.7 (1.01)	633 (8.6)	105.5 (1.43)	.02[e]	0.01[f]
Social deprivation index								
1 (least disadvantaged)	2715 (13.8)	271.5 (1.38)	1717 (14.3)	190.8 (1.59)	956 (13)	159.3 (2.17)	.06[e]	—
2 (slightly disadvantaged)	6294 (32)	629.4 (3.2)	3671 (30.5)	407.9 (3.39)	2327 (31.8)	387.8 (5.3)	.02[e]	0.01[f]
3 (moderately disadvantaged)	6016 (30.6)	601.6 (3.06)	3656 (30.4)	406.2 (3.38)	2184 (29.8)	364 (4.97)	.50[e]	—
4 (most disadvantaged)	4666 (23.7)	466.6 (2.37)	3001 (24.9)	333.4 (2.77)	1860 (25.4)	310 (4.23)	.004[e]	0.02[f]
Drug administration, n (%)								

XSL•FO
RenderX

Characteristics	Period 1[a]		Period 2[b]		Period 3[c]		*P* value	Effect size
	Observed	Adjusted	Observed	Adjusted	Observed	Adjusted		
Angiotensin-converting enzyme inhibitors	6981 (35.4)	698.1 (3.54)	4419 (36.7)	491 (4.08)	2656 (36.2)	442.7 (6.03)	.07[e]	—
Antidiabetics	4840 (24.5)	484 (2.45)	3140 (26.1)	348.9 (2.9)	1932 (26.3)	322 (4.38)	.001[e]	0.02[f]
Antiplatelets and anticoagulants	13,461 (68.3)	1346.1 (6.83)	8039 (66.7)	893.2 (7.41)	4851 (66.1)	808.5 (11.02)	<.001[e]	0.02[f]
Beta blockers	4823 (24.5)	482.3 (2.45)	945 (7.8)	105 (0.87)	104 (1.4)	17.3 (0.23)	<.001[e]	0.27[f]
Bronchodilators	3790 (19.2)	379 (1.92)	2698 (22.4)	299.8 (2.49)	1934 (26.4)	322.3 (4.4)	<.001[e]	0.09[f]
Calcium channel blocker	10,906 (55.3)	1090.6 (5.53)	6713 (55.7)	745.9 (6.19)	4024 (54.8)	670.7 (9.13)	.51[e]	—
Diuretics	4996 (25.3)	499.6 (2.53)	2968 (24.6)	329.8 (2.73)	2144 (29.2)	357.3 (4.87)	<.001[e]	0.04[f]
Inhaled corticosteroids	1442 (7.3)	144.2 (0.73)	918 (7.6)	102 (0.84)	615 (8.4)	102.5 (1.4)	.01[e]	0.01[f]
Rheumatoid drugs	2472 (12.5)	247.2 (1.25)	1596 (13.2)	177.3 (1.47)	972 (13.2)	162 (2.2)	.11[e]	—
Statins	8400 (42.6)	840 (4.26)	5783 (48)	642.6 (5.33)	3565 (48.6)	594.2 (8.1)	<.001[e]	0.06[f]
Systemic corticosteroids	6769 (34.3)	676.9 (3.43)	4951 (41.1)	550.1 (4.57)	2882 (39.3)	480.3 (6.55)	<.001[e]	0.06[f]
Blood biomarkers, median (IQR)								
Albumin (g/L)	33.30 (8.10)	32.8 (1.1)	33.70 (7.80)	33.1 (0.3)	32.90 (8.20)	32.3 (0.3)	<.001[g]	0.002[h]
Neutrophil ($\times 10^9$/L)	5 (3.85)	6.1 (0.6)	5.44 (4)	6.5 (0.2)	5.92 (4.37)	7 (0.2)	<.001[g]	0.01[h]
Bilirubin (μmol/L)	9 (7.40)	11.3 (0.4)	9 (7)	11 (0.3)	9 (7.30)	11.1 (0.3)	.01[g]	0.0002[h]
Lymphocyte ($\times 10^9$/L)	1 (0.75)	1.4 (0.05)	0.99 (0.80)	1.6 (0.04)	1.03 (0.87)	1.7 (0.1)	<.001[g]	0.001[h]
Platelet ($\times 10^9$/L)	193 (103)	210.1 (9)	187 (99)	204.6 (4.5)	197 (104)	213.6 (2.5)	<.001[g]	0.001[h]

[a]Episodes=19,717 and head count=15,529.
[b]Episodes=12,053 and head count=10,639.
[c]Episodes=7337 and head count=6456.
[d]Not available.
[e]Proportions test.
[f]Cramér V test.
[g]Kruskal-Wallis test.
[h]Eta squared test.

Demographic Profile of Hospitalized Patients Across Periods

Tables 2-6 and Table S1 in Multimedia Appendix 1 demonstrated demographic shifts across periods 1 to 3 by age subgroups. First, the CFR increased in each period as the age groups got older (Figure S1 in Multimedia Appendix 1) and lowered in patients older than 85 years (1.5% difference, period 1: 11.6%, period 3: 10.1%, *P*<.001, effect size: 0.02). Second, the monthly episodic rate of males and females hospitalized for COVID-19 infection decreased across periods for ages 0 - 84. Females were predominantly accounted for infections in patients

XSL•FO
RenderX

18 - 64, aged older than 85, whereas more males were infected in age groups 0 - 17, 64 - 74, and 75 - 84 (Figure S2 in Multimedia Appendix 1).

Third, the proportion of individuals of the Chinese race dominated across all age groups, becoming more prominent as age increased (Figure S4 in Multimedia Appendix 1). Only those aged 0 - 17 reported decreased proportions of Chinese race (8.5% difference, period 1: 77.2%, period 3: 68.7%, *P*<.001, effect size: 0.07). Also reflected in the adjusted measurements (410.1 episodes per month decrease, period 1: 648.3 episodes per month, period 2: 238.2 episodes per month).

Fourth, the Charlson Comorbidity Index scores in most age groups showed a predominant proportion of infected individuals with scores of 0 (more than 70% per period; Figure S5 in Multimedia Appendix 1), with notable decrease across the study period in age groups 0 - 17 (0.5% difference, period 1: 98.1%, period 3: 97.6%, *P*=.003, effect size: 0.03; 486.4 episodes per month decrease, period 1: 823.6 episodes per month, period 3: 337.2 episodes per month), 18 - 64 (2% difference, period 1: 78.5%, period 3: 76.5%, *P*<.001, effect size: 0.03; 653.7 episodes per month decrease, period 1: 1195 episodes per month, period 3: 541.3 episodes per month), 65 - 74 (1.3% difference, period 1: 72.5%, period 3: 71.2%, *P*=.04, effect size: 0.01; 499.8 episodes per month decrease, period 1: 974 episodes per month, period 3: 474.2 episodes per month), 75 - 84 (1.4% difference, period 1: 72.7%, period 3: 71.3%, *P*<.001, effect size: 0.02; 332 episodes per month decrease, period 1: 484.5 episodes per month, period 3: 152.5 episodes per month). In contrast, infected patients older than 85 showed an increase (2.1% difference, period 1: 75.2%, period 3: 77.3%, *P*<.001, effect size: 0.02), but the adjusted measurements showed otherwise (552.7 episodes per month decrease, period 1: 1482 episodes per month, period 3: 929.3 episodes per month). Although the proportion of infected patients with comorbid essential primary hypertension, type 2 diabetes mellitus, and hyperlipidemia (unspecified) remained below 10% and decreased in rate across periods and age groups (Figure S3 in Multimedia Appendix 1). Additionally, less than 10% of episodes per period and age group were related to frailty, although this proportion increased with age (Figure S8 in Multimedia Appendix 1). The rates decreased across periods for each age group.

Fifth, the distribution of the social deprivation index of infected individuals centered around slightly disadvantaged populations in Hong Kong (Figure S6 in Multimedia Appendix 1). This demographic showed higher proportions across periods among age groups 75 - 84 (2.5% difference, period 1: 33.7%, period 3: 36.2%, *P*=.003, effect size: 0.02) and lower proportions in those older than 85 years (0.2% difference, period 1: 32%, period 3: 31.8%, *P*=.02, effect size: 0.01). Although both age groups decreased in rate across period (age group 75 - 84: 219.5 episodes per month decrease, period 1: 524.7 episodes per month, period 3: 305.2 episodes per month; age group older than 85 years: 241.6 episodes per month decrease, period 1: 629.4 episodes per month, period 3: 387.8 episodes per month).

Clinical Changes in Hospitalized Patients Across Periods

The findings from Tables 2-6 and Table S1 in Multimedia Appendix 1 also showed changes in the clinical management of hospitalized patients infected with Omicron across the study period when stratified by age groups. First, the length of hospital stay decreased across periods for age groups 18 - 64 (1 d decrease, period 1: 4 d, period 3: 3 d, *P*<.001, effect size: 0.007; 4.1 d per month decrease, period 1: 11 d per month, period 3: 6.9 d per month), 65 - 74 (2 d decrease, period 1: 6 d, period 3: 4 d, *P*<.001, effect size: 0.01; 2.6 d per month decrease, period 1: 10.4 d per month, period 3: 7.8 d per month), 75 - 84 (2 d decrease, period 1: 6 d, period 3: 4 d, *P*<.001, effect size: 0.01; 2.6 d per month decrease, period 1: 10.5 d per month, period 3: 7.9 d per month), and older than 85 (2 days decrease, period 1: 7 d, period 3: 5 d, *P*<.001, effect size: 0.02; 2.5 d per month decrease, period 1: 10.1 d per month, period 3: 7.6 d per month). Additionally, a longer length of stay was observed as the demographic aged (Figure S9 in Multimedia Appendix 1).

Second, calcium channel blockers, statins, antiplatelets, and anticoagulants were the most frequently administered to patients aged 18 and older (Figure S7 in Multimedia Appendix 1), but the rate of administration reduced over time. Moreover, patients 0 - 17 years received 4.2% more bronchodilators over the course of the study period (period 1: 2.9%, period 3: 7.1%, *P*<.001, effect size: 0.08) and slightly increased rate of administration (0.4 episodes per month, period 1: 24.7 episodes per month decrease, period 3: 24.3 episodes per month); 5.5% more bronchodilators in 18 - 64 years (period 1: 6.7%, period 3: 12.2%, *P*<.001, effect size: 0.07) and decreased rate of administration (13.4 episodes per month decrease, period 1: 101.7 episodes per month, period 3: 88.3 episodes per month); 6.6% more bronchodilators in 65 - 74 years (period 1: 13.4%, period 3: 20%, *P*<.001, effect size: 0.07) and decreasing rate of administration across periods (45.2 episodes per month decrease, period 1: 180.7 episodes per month, period 3: 135.5 episodes per month); 7.2% more bronchodilators in 75 - 84 years (period 1: 17.6%, period 3: 24.8%, *P*<.001, effect size: 0.07) and decreasing rate of administration (64.5 episodes per month decrease, period 1: 274.5 episodes per month, period 3: 210 episodes per month); and 7.2% more bronchodilators in older than 85 (period 1:19.2%, period 3: 26.4%, *P*<.001, effect size: 0.09) and decreasing rate of administration (56.7 episodes per month decrease, period 1: 379 episodes per month, period 3: 322.3 episodes per month). Meanwhile, systemic corticosteroids were administered the most frequently in the 18 - 64 age group (6.8% difference, period 1: 18.8%, period 3: 25.6%, *P*<.001, effect size: 0.09) with decreasing rate of administration (101.2 episodes per month decrease, period 1: 286.7 episodes per month, period 3: 185.5 episodes per month); and the 65 - 74 age group (9.4% difference, period 1: 27.2%, period 3: 36.6%, *P*<.001, effect size: 0.09) with decreasing rate of administration (118 episodes per month, period 1: 366.5 episodes per month, period 3: 248.5 episodes per month). The highest effect size was observed with the decline in the proportion of patients administered beta blockers over time (effect size in aged 0 - 17: 0.52, 64 - 74: 0.27, 75 - 84: 0.29, older than 85: 0.27).

XSL•FO
RenderX

Third, all the blood biomarkers exhibited minimal changes across periods in each age stratification (Figure S10 in Multimedia Appendix 1). Median albumin levels lowered, and median neutrophil levels elevated as age groups got older, especially in patients older than 85 years (albumin: 32.9 to 38.82g/L and 37.3 to 36.7g/L per month; neutrophil: 4.48 to 5.92×10^9/L and 5.6 to 6.4×10^9/L per month). Meanwhile, median bilirubin, lymphocyte, and platelet levels fluctuated across age groups.

Discussion

Principal Findings

To our knowledge, this is among the first retrospective, repeated cross-sectional studies to review changes in demographic and clinical characteristics of all patients infected with Omicron admitted to public hospitals in Hong Kong since the Omicron outbreak, including postpandemic periods. Additionally, this study expanded upon existing literature by providing insight into all age groups and incorporating social deprivation index scores.

Trends in Hospitalizations and Case Fatality Rate

Although the epidemic curves demonstrated a decrease in the number of hospitalized Omicron cases, this study observed an increase in the CFR with advancing age, which was consistent with reports from other local and international studies [34,35]. In contrast, the CFR started to decline among patients aged older than 85 years, a trend not observed in previous studies. This finding suggested there were age-related disparities among Omicron infections. Furthermore, the rate of hospital readmission decreased across periods for all ages, which contradicted the findings from a study in the UK [36]. They suggested that recent vaccination reduced the risk of reinfection, and subsequent reinfections demonstrated lower severity [36], but further research is needed to ascertain this in our population group.

Comorbidity Differences

The shift in Omicron infections to patients with more than 1 CCI over time aligned with local and international studies conducted in a postpandemic setting. They further attributed this shift to prolonged viral shedding from SARS-CoV-2 rebound [37-40], but subsequent research is needed to ascertain this hypothesis in our study population. Namely, our findings indicated that the proportion of patients with hypertension and hyperlipidemia comorbidities was low across periods, even though past literature suggested that existing low-grade chronic systemic inflammatory diseases can complicate clinical management of infected patients and increase the risk of poorer outcomes [41].

Gender Differences

Our results also indicated a predominance of male patients in most age groups (0 - 17 y, 64 - 74 y, and 75 - 84 y) throughout the study period, an observation that was well established during the pandemic [42]. Several hypotheses were proposed: one was the lower uptake of the second dose of COVID-19 vaccines among males [43], while another hypothesized a link to cardiovascular factors [43].

Socioeconomic Demographic Shift

Most infections in our study population occurred among individuals from slightly disadvantaged backgrounds, even though the rate decreased across periods. Few studies explored the impact of the Omicron outbreak on patients from varied socioeconomic statuses. A study conducted during the early phase of COVID-19 in Hong Kong reported an association between socioeconomic disadvantage and a broader spread of infections [44], and was particularly related to essential activities such as living and working. International studies similarly found an association with lower socioeconomic status, hypothesizing that vaccine coverage might be an underlying reason [45,46].

Changes in In-Hospital Drug Administrations

Systemic corticosteroids became an increasingly common and effective treatment for COVID-19, as demonstrated by the rising proportion and reduced rate of administration among patients aged 18 years to 74 years and across periods in our findings, as well as in other studies [47,48]. One study suggested that while such treatments have anti-inflammatory effects, this could induce immunosuppression, potentially delaying viral clearance and increasing the risk of poorer outcomes [49]. This led to differing clinical management recommendations, with the WHO advising against corticosteroid treatments in severe cases [50], while local and Chinese studies recommended low-dose, short-course treatments for severe cases [51,52].

Also, cardiovascular treatments, including calcium channel blockers, statins, antiplatelets, and anticoagulants, were the most frequently administered, with a reduced rate of administration among adults and older individuals in this study. Global studies established a relationship between COVID-19 and cardiovascular complications [53,54]. In particular, studies from China reported a high prevalence of myocardial injury among infected patients with poor prognosis [55]. To mitigate cardiovascular complications, subsequent studies demonstrated improved prognosis in patients with COVID-19 administered with calcium channel blockers [56,57], statins [58], anticoagulants [59], and antiplatelet agents [60]. Therefore, our findings may reflect the common clinical treatment for Omicron cases.

Additionally, we observed an increased use of bronchodilators and a decreasing rate of administration over time. This might have reflected evolving clinical management strategies rather than increased severity, as by 2022, over 90% of the Hong Kong population was vaccinated, and the CFR in our study remained largely static across all age groups [61]. However, further research was required to clarify these observations.

Changes in Biomarker Indicators

Recent studies aligned with our findings and further demonstrated an association with COVID-19-related mortality. For example, one study found that elevated neutrophil levels were associated with increased COVID-19 mortality [62], while another study from China showed that hypoalbuminemia (albumin <35g/L) increased the odds of mortality [63].

XSL•FO
RenderX

Therefore, further research can help better understand the clinical management of infected patients, particularly in those aged over 85 years.

Age Disparities in Omicron Cases

Older age saw a higher case-mortality ratio, likely from the higher CCI scores and frailty also observed in other studies [21,64], making this demographic more susceptible to infection. This was possibly exacerbated by the excess mortality among the older individuals during the Omicron outbreak in Hong Kong [65], the largest local outbreak to date. Such age disparities were reported to persist in the postpandemic setting [65], although differences between age groups appeared to have narrowed. However, the EuCARE-HOSPITALISED international study observed opposite trends [66]. This may be attributed to Hong Kong's low vaccine uptake during the Omicron outbreak, but further research was needed to confirm this hypothesis.

Limitations

Although this study collected a large sample size from all the public hospitals in Hong Kong, several limitations remained. First, our sample is not representative of all cases of COVID-19 infection in Hong Kong as it did not include unreported and asymptomatic cases, data from private hospitals, or non-Hong Kong residents. Second, our SDI score was based on calculations from 2008. Such definitions may be outdated and differ from the current situation in Hong Kong. The SDI distribution may also be skewed towards more disadvantaged populations with the exclusion of private hospital data. So, the respective findings from Tables 2-6 should be interpreted with caution. Third, the list of frailty-related diagnoses was defined by clinicians from an acute hospital trust in England, which may not be representative of frail patients in Hong Kong. Therefore, our findings should be interpreted with caution. Fourth, the patient's vaccination status was not available, so we cannot understand the impact of vaccination on incidence.

Conclusions

Our study provides an updated descriptive overview of postpandemic Omicron hospitalizations in Hong Kong. The findings highlight the need for age-specific interventions, particularly among older individuals. Further research is essential to understand the effectiveness of the vaccine booster dose in a postpandemic setting. All to improve pandemic preparedness and to develop more effective public health strategies in the ongoing fight against COVID-19.

Acknowledgments

Generative artificial intelligence was not used in any portion of the article writing.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Data Availability

The dataset generated or analyzed during this study is available from the corresponding author (AW) on reasonable request.

Authors' Contributions

AKCW and CJYC contributed to the conception and design of the study. CJYC, HHHP, and BKHL contributed to data collection. CJYC, TTLL, and SCLC carried out the data analysis. MSW, TY, CKT, CW, and THR contributed to the revision and read and approved the submitted version of the article.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Drugs and post hoc analysis characterizing patients infected with Omicron.
[DOCX File, 200 KB - publichealth_v12i1e75635_app1.docx]

References

1. Wong SC, Au AKW, Lo JYC, et al. Evolution and control of COVID-19 epidemic in Hong Kong. Viruses 2022 Nov 14;14(11):2519. [doi: 10.3390/v14112519] [Medline: 36423128]
2. Xiong X, Wai AKC, Wong JYH, et al. Impact of varying wave periods of COVID-19 on in-hospital mortality and length of stay for admission through emergency department: a territory-wide observational cohort study. Influenza Other Respir Viruses 2022 Mar;16(2):193-203. [doi: 10.1111/irv.12919] [Medline: 34643047]
3. WHO chief declares end to COVID-19 as a global health emergency. Unied Nations. URL: https://news.un.org/en/story/2023/05/1136367#:~:text=The%20head%20of%20the%20UN,no%20longer%20a%20global%20threat [accessed 2026-01-08]

XSL•FO
RenderX

4. COVID-19 & Flu Express. Centre for Health Protection. URL: https://www.chp.gov.hk/en/resources/29/100148.html [accessed 2026-01-08]
5. COVID-19 epidemiological update. World Health Organization. URL: https://www.who.int/publications/m/item/covid-19-epidemiological-update-edition-169 [accessed 2026-01-08]
6. COVID-19 & flu express. : Centre for Health Protection URL: https://www.chp.gov.hk/files/pdf/covid_flux_week17_2_5_2024_eng.pdf [accessed 2026-01-08]
7. Davis HE, McCorkell L, Vogel JM, Topol EJ. Long COVID: major findings, mechanisms and recommendations. Nat Rev Microbiol 2023 Mar;21(3):133-146. [doi: 10.1038/s41579-022-00846-2] [Medline: 36639608]
8. Hu Z, Youn HM, Jin Q, et al. The negative spillover impact of the COVID-19 pandemic on Hong Kong diabetes patients' mortality and cardiovascular disease. Br J Gen Pract 2024 Jun 20;74(suppl 1):bjgp24X738045. [doi: 10.3399/bjgp24X738045] [Medline: 38902054]
9. Mak IL, Wan EYF, Wong TKT, et al. The spill-over impact of the novel Coronavirus-19 pandemic on medical care and disease outcomes in non-communicable diseases: a narrative review. Public Health Rev 2022;43:1604121. [doi: 10.3389/phrs.2022.1604121] [Medline: 35574567]
10. Cekic D, Issever K, Genc AC, Yaylaci S, Genc AB, Tamer A. Association of c-reactive protein/albumin, procalcitonin/albumin, platelet/lymphocyte, and lymphocyte/monocyte ratio with mortality in hospitalised COVID-19 patients. J Coll Physicians Surg Pak 2022 Sep;32(9):1191-1195. [doi: 10.29271/jcpsp.2022.09.1191] [Medline: 36089719]
11. Samprathi M, Jayashree M. Biomarkers in COVID-19: an up-to-date review. Front Pediatr 2020;8:607647. [doi: 10.3389/fped.2020.607647] [Medline: 33859967]
12. Sargin Altunok E, Satici C, Dinc V, et al. Comparison of demographic and clinical characteristics of hospitalized COVID-19 patients with severe/critical illness in the first wave versus the second wave. J Med Virol 2022 Jan;94(1):291-297. [doi: 10.1002/jmv.27319] [Medline: 34491575]
13. Tsui ELH, Lui CSM, Woo PPS, et al. Development of a data-driven COVID-19 prognostication tool to inform triage and step-down care for hospitalised patients in Hong Kong: a population-based cohort study. BMC Med Inform Decis Mak 2020 Dec 7;20(1):323. [doi: 10.1186/s12911-020-01338-0] [Medline: 33287804]
14. Kastora S, Patel M, Carter B, Delibegovic M, Myint PK. Impact of diabetes on COVID-19 mortality and hospital outcomes from a global perspective: an umbrella systematic review and meta-analysis. Endocrinol Diabetes Metab 2022 May;5(3):e00338. [doi: 10.1002/edm2.338] [Medline: 35441801]
15. Khairy Y, Naghibi D, Moosavi A, Sardareh M, Azami-Aghdash S. Prevalence of hypertension and associated risks in hospitalized patients with COVID-19: a meta-analysis of meta-analyses with 1468 studies and 1,281,510 patients. Syst Rev 2022 Nov 17;11(1):242. [doi: 10.1186/s13643-022-02111-2] [Medline: 36397129]
16. Yang B, Tsang TK, Gao H, et al. Universal community nucleic acid testing for coronavirus disease 2019 (COVID-19) in Hong Kong reveals insights into transmission dynamics: a cross-sectional and modeling study. Clin Infect Dis 2022 Aug 24;75(1):e216-e223. [doi: 10.1093/cid/ciab925] [Medline: 34718464]
17. Wong MCS, Ng RWY, Chong KC, et al. Stringent containment measures without complete city lockdown to achieve low incidence and mortality across two waves of COVID-19 in Hong Kong. BMJ Glob Health 2020 Oct;5(10):e003573. [doi: 10.1136/bmjgh-2020-003573] [Medline: 33028700]
18. Swamy S, Koch CA, Hannah-Shmouni F, Schiffrin EL, Klubo-Gwiezdzinska J, Gubbi S. Hypertension and COVID-19: updates from the era of vaccines and variants. J Clin Transl Endocrinol 2022 Mar;27(100285):100285. [doi: 10.1016/j.jcte.2021.100285] [Medline: 34900602]
19. Vasbinder A, Anderson E, Shadid H, et al. Inflammation, hyperglycemia, and adverse outcomes in individuals with diabetes mellitus hospitalized for COVID-19. Diabetes Care 2022 Mar 1;45(3):692-700. [doi: 10.2337/dc21-2102] [Medline: 35045184]
20. Verduri A, Short R, Carter B, et al. Comparison between first and second wave of COVID-19 outbreak in older people: the COPE multicentre European observational cohort study. Eur J Public Health 2022 Oct 3;32(5):807-812. [doi: 10.1093/eurpub/ckac108] [Medline: 35997587]
21. Tam E, Kwan YK, Ng YY, Yam PW. Clinical course and mortality in older patients with COVID-19: a cluster-based study in Hong Kong. Hong Kong Med J 2022 Jun;28(3):215-222. [doi: 10.12809/hkmj219367] [Medline: 35686472]
22. Kwan TH, Wong NS, Yeoh EK, Lee SS. Shifts of SARS-CoV-2 exposure settings in the transmission clusters of 2 epidemic waves in Hong Kong. Int J Environ Health Res 2023 Sep;33(9):911-923. [doi: 10.1080/09603123.2022.2064438] [Medline: 35437073]
23. Liao Q, Dong M, Yuan J, et al. Assessing community vulnerability over 3 waves of COVID-19 pandemic, Hong Kong, China. Emerg Infect Dis 2021 Jul;27(7):1935-1939. [doi: 10.3201/eid2707.204076] [Medline: 34152957]
24. Liu Y, Wang Z, Rader B, et al. Associations between changes in population mobility in response to the COVID-19 pandemic and socioeconomic factors at the city level in China and country level worldwide: a retrospective, observational study. Lancet Digit Health 2021 Jun;3(6):e349-e359. [doi: 10.1016/S2589-7500(21)00059-5] [Medline: 34045001]
25. Kwok TC, Cheng IT, Chung AP, Leung JK, Fung MT, Cheung NT. MEDINFO. 2010 URL: http://139.91.210.27/CBML/PROCEEDINGS/2010_MEDINFO/browse/pdfs/poster%20I6a.pdf [accessed 2026-01-30]

XSL•FO
RenderX

26. Sing CW, Woo YC, Lee ACH, et al. Validity of major osteoporotic fracture diagnosis codes in the clinical data analysis and reporting system in Hong Kong. Pharmacoepidemiol Drug Saf 2017 Aug;26(8):973-976. [doi: 10.1002/pds.4208] [Medline: 28371079]
27. Wong CM, Ou CQ, Chan KP, et al. The effects of air pollution on mortality in socially deprived urban areas in Hong Kong, China. Environ Health Perspect 2008 Sep;116(9):1189-1194. [doi: 10.1289/ehp.10850] [Medline: 18795162]
28. 2021 Population Census - Older Persons in Hong Kong. 2021 Hong Kong Population Census. URL: https://www.census2021.gov.hk/doc/pub/21C_Articles_Older_Persons.pdf [accessed 2026-01-09]
29. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care 2005 Nov;43(11):1130-1139. [doi: 10.1097/01.mlr.0000182534.19832.83] [Medline: 16224307]
30. Gilbert T, Neuburger J, Kraindler J, et al. Development and validation of a Hospital Frailty Risk Score focusing on older people in acute care settings using electronic hospital records: an observational study. Lancet 2018 May 5;391(10132):1775-1782. [doi: 10.1016/S0140-6736(18)30668-8] [Medline: 29706364]
31. Estimating mortality from COVID-19. World Health Organization. URL: https://www.who.int/news-room/commentaries/detail/estimating-mortality-from-covid-19 [accessed 2026-01-09]
32. Kenyhercz MW, Passalacqua NV. Missing data imputation methods and their performance with biodistance analyses. Biological Distance 2016:181-194. [doi: 10.1016/B978-0-12-801966-5.00009-3]
33. The R project for statistical computing. R. URL: https://www.R-project.org [accessed 2026-01-09]
34. Miyashita K, Hozumi H, Furuhashi K, et al. Outcomes of COVID-19 in the Omicron-predominant wave: large-scale real-world data analysis with a comparison to influenza. Pneumonia (Nathan) 2025 Feb 5;17(1):3. [doi: 10.1186/s41479-025-00158-y] [Medline: 39905566]
35. Smith DJ, Hakim AJ, Leung GM, et al. COVID-19 mortality and vaccine coverage - Hong Kong special administrative region, China, January 6, 2022-March 21, 2022. MMWR Morb Mortal Wkly Rep 2022 Apr 15;71(15):545-548. [doi: 10.15585/mmwr.mm7115e1] [Medline: 35421076]
36. Wei J, Stoesser N, Matthews PC, et al. Risk of SARS-CoV-2 reinfection during multiple Omicron variant waves in the UK general population. Nat Commun 2024 Feb 2;15(1):1008. [doi: 10.1038/s41467-024-44973-1] [Medline: 38307854]
37. Lin Y, Wu P, Tsang TK, et al. Viral kinetics of SARS-CoV-2 following onset of COVID-19 in symptomatic patients infected with the ancestral strain and Omicron BA.2 in Hong Kong: a retrospective observational study. Lancet Microbe 2023 Sep;4(9):e722-e731. [doi: 10.1016/S2666-5247(23)00146-5] [Medline: 37659420]
38. Chong KC, Wei Y, Jia KM, et al. SARS-CoV-2 rebound and post-acute mortality and hospitalization among patients admitted with COVID-19: cohort study. Nat Commun 2025 Jul 28;16(1):6924. [doi: 10.1038/s41467-025-61737-7] [Medline: 40721474]
39. Zuo W, He D, Liang C, et al. The persistence of SARS-CoV-2 in tissues and its association with long COVID symptoms: a cross-sectional cohort study in China. Lancet Infect Dis 2024 Aug;24(8):845-855. [doi: 10.1016/S1473-3099(24)00171-3] [Medline: 38663423]
40. Ghafari M, Hall M, Golubchik T, et al. Prevalence of persistent SARS-CoV-2 in a large community surveillance study. Nature New Biol 2024 Feb;626(8001):1094-1101. [doi: 10.1038/s41586-024-07029-4] [Medline: 38383783]
41. Buicu AL, Cernea S, Benedek I, Buicu CF, Benedek T. Systemic inflammation and COVID-19 mortality in patients with major noncommunicable diseases: chronic coronary syndromes, diabetes and obesity. J Clin Med 2021 Apr 7;10(8):1545. [doi: 10.3390/jcm10081545] [Medline: 33916917]
42. Fabião J, Sassi B, Pedrollo EF, et al. Why do men have worse COVID-19-related outcomes? A systematic review and meta-analysis with sex adjusted for age. Braz J Med Biol Res 2022;55:e11711. [doi: 10.1590/1414-431X2021e11711] [Medline: 35195196]
43. Law AHT, Presanis AM, Cheung JK, et al. Estimating changes in life expectancy in Hong Kong during the COVID-19 pandemic: a longitudinal ecological study. Lancet Reg Health West Pac 2025 May;58:101571. [doi: 10.1016/j.lanwpc.2025.101571] [Medline: 40470520]
44. Chung GKK, Chan SM, Chan YH, et al. Socioeconomic patterns of COVID-19 clusters in low-incidence city, Hong Kong. Emerg Infect Dis 2021 Nov;27(11):2874-2877. [doi: 10.3201/eid2711.204840] [Medline: 34469286]
45. Guerrero-Araya E, Ravello C, Rosemblatt M, Perez-Acle T. Socioeconomic status correlates with COVID-19 vaccination coverage among primary and secondary students in the most populated city of Chile. Sci Rep 2025 Jan 9;15(1):1509. [doi: 10.1038/s41598-024-84260-z] [Medline: 39789033]
46. Bergen N, Kirkby K, Fuertes CV, et al. Global state of education-related inequality in COVID-19 vaccine coverage, structural barriers, vaccine hesitancy, and vaccine refusal: findings from the Global COVID-19 Trends and Impact Survey. Lancet Glob Health 2023 Feb;11(2):e207-e217. [doi: 10.1016/S2214-109X(22)00520-4] [Medline: 36565702]
47. RECOVERY Collaborative Group, Horby P, Lim WS, et al. Dexamethasone in hospitalized patients with Covid-19. N Engl J Med 2021 Feb 25;384(8):693-704. [doi: 10.1056/NEJMoa2021436] [Medline: 32678530]
48. Corral-Gudino L, Bahamonde A, Arnaiz-Revillas F, et al. Methylprednisolone in adults hospitalized with COVID-19 pneumonia: an open-label randomized trial (GLUCOCOVID). Wien Klin Wochenschr 2021 Apr;133(7-8):303-311. [doi: 10.1007/s00508-020-01805-8] [Medline: 33534047]

XSL•FO
RenderX

49. Zhuang R, Xia H, Xu L, et al. Corticosteroids for hospitalized patients with severe/critical COVID-19: a retrospective study in Chongqing, China. Sci Rep 2024 Oct 16;14(1):24317. [doi: 10.1038/s41598-024-75926-9] [Medline: 39414922]
50. Mehta N, Mazer-Amirshahi M, Alkindi N, Pourmand A. Pharmacotherapy in COVID-19; a narrative review for emergency providers. Am J Emerg Med 2020 Jul;38(7):1488-1493. [doi: 10.1016/j.ajem.2020.04.035] [Medline: 32336586]
51. Shang L, Zhao J, Hu Y, Du R, Cao B. On the use of corticosteroids for 2019-nCoV pneumonia. Lancet 2020 Feb 29;395(10225):683-684. [doi: 10.1016/S0140-6736(20)30361-5] [Medline: 32122468]
52. Yam LYC, Lau ACW, Lai FYL, et al. Corticosteroid treatment of severe acute respiratory syndrome in Hong Kong. J Infect 2007 Jan;54(1):28-39. [doi: 10.1016/j.jinf.2006.01.005] [Medline: 16542729]
53. Clerkin KJ, Fried JA, Raikhelkar J, et al. COVID-19 and cardiovascular disease. Circulation 2020 May 19;141(20):1648-1655. [doi: 10.1161/CIRCULATIONAHA.120.046941] [Medline: 32200663]
54. Hoffmann M, Kleine-Weber H, Schroeder S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. Cell 2020 Apr 16;181(2):271-280. [doi: 10.1016/j.cell.2020.02.052] [Medline: 32142651]
55. He W, Xu K, Ni L, et al. Myocardial injury and related mortality in hospitalized patients with COVID-19 during the Omicron pandemic: new perspectives and insights. Virol Sin 2023 Dec;38(6):940-950. [doi: 10.1016/j.virs.2023.10.005] [Medline: 37839550]
56. Kow CS, Ramachandram DS, Hasan SS. Clinical outcomes of hypertensive patients with COVID-19 receiving calcium channel blockers: a systematic review and meta-analysis. Hypertens Res 2022 Feb;45(2):360-363. [doi: 10.1038/s41440-021-00786-z] [Medline: 34754084]
57. Peng C, Wang H, Guo YF, et al. Calcium channel blockers improve prognosis of patients with coronavirus disease 2019 and hypertension. Chin Med J (Engl) 2021 Jun 16;134(13):1602-1609. [doi: 10.1097/CM9.0000000000001479] [Medline: 34133354]
58. Golomb BA, Han JH, Langsjoen PH, Dinkeloo E, Zemljic-Harpf AE. Statin use in relation to COVID-19 and other respiratory infections: muscle and other considerations. J Clin Med 2023 Jul 13;12(14):4659. [doi: 10.3390/jcm12144659] [Medline: 37510774]
59. Flumignan RL, Civile VT, Tinôco JDS, et al. Anticoagulants for people hospitalised with COVID-19. Cochrane Database Syst Rev 2022 Mar 4;3(3):CD013739. [doi: 10.1002/14651858.CD013739.pub2] [Medline: 35244208]
60. Sokolski M, Reszka K, Adamik B, et al. Antiplatelet therapy prior to COVID-19 infection impacts on patients mortality: a propensity score-matched cohort study. Sci Rep 2024 Feb 28;14(1):4832. [doi: 10.1038/s41598-024-55407-9] [Medline: 38413716]
61. Wei Y, Jia KM, Zhao S, et al. Estimation of vaccine effectiveness of coronaVac and BNT162b2 against severe outcomes over time among patients with SARS-CoV-2 Omicron. JAMA Netw Open 2023 Feb 1;6(2):e2254777. [doi: 10.1001/jamanetworkopen.2022.54777] [Medline: 36735253]
62. Tizazu AM, Gize A, Ali S. Age influences blood cell-based immune deregulation antibody response and unfavorable clinical outcomes in COVID-19 patients. Sci Rep 2025 May 20;15(1):17431. [doi: 10.1038/s41598-025-95722-3] [Medline: 40394243]
63. Huang J, Cheng A, Kumar R, et al. Hypoalbuminemia predicts the outcome of COVID-19 independent of age and co-morbidity. J Med Virol 2020 Oct;92(10):2152-2158. [doi: 10.1002/jmv.26003] [Medline: 32406952]
64. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet 2020 Mar 28;395(10229):1054-1062. [doi: 10.1016/S0140-6736(20)30566-3] [Medline: 32171076]
65. Li K, Wei Y, Hung CT, et al. Post-pandemic excess mortality of COVID-19 in Hong Kong: a retrospective study. Lancet Reg Health West Pac 2025 May;58:101554. [doi: 10.1016/j.lanwpc.2025.101554] [Medline: 40336577]
66. Hedberg P, Parczewski M, Serwin K, et al. In-hospital mortality during the wild-type, alpha, delta, and omicron SARS-CoV-2 waves: a multinational cohort study in the EuCARE project. Lancet Reg Health Eur 2024 Mar;38:100855. [doi: 10.1016/j.lanepe.2024.100855] [Medline: 38476753]

Abbreviations

CCI: Charlson Comorbidity Index
CFR: case fatality ratio
***ICD-10*:** *International Statistical Classification of Diseases and Related Health Problems 10th Revision*
SDI: Social Deprivation Index
WHO: World Health Organization

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 08.Apr.2025; peer-reviewed by S Lau, S Jonathan; revised version received 27.Nov.2025; accepted 02.Dec.2025; published 09.Feb.2026.

Please cite as:
Ching CJY, Chan SCL, Lee TTL, Pui HHH, Leung BKH, Wong MS, Yamamoto T, Tong CK, Wang C, Rainer TH, Wai AKC
Postpandemic Change in Demographic and Clinical Features of Patients With Omicron Who Were Hospitalized: Territory-Wide Retrospective Repeated Cross-Sectional Study in Hong Kong
JMIR Public Health Surveill 2026;12:e75635
URL: https://publichealth.jmir.org/2026/1/e75635
doi:10.2196/75635

© Christie J Y Ching, Sunny C L Chan, Teddy T L Lee, Hugo H H Pui, Bosco K H Leung, Man Sing Wong, Tafu Yamamoto, Chak Kwan Tong, Cantian Wang, Timothy H Rainer, Abraham K C Wai. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 9.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Developmental Trajectories of Positive Expectancies of Cannabis Use Effects Among Early Adolescents: Longitudinal Observational Study Using Latent Class Growth Analysis

Weisiyu Abraham Qin[1], MPH, MHA, PhD; Dong-Chul Seo[2], PhD; Wura Jacobs[2], PhD; Sijia Huang[3], PhD; Kit K Elam[2], PhD

[1]Center for Tobacco Control Research and Education, Cardiovascular Research Institute, University of California, San Francisco, 530 Parnassus Ave., Ste 366, San Francisco, CA, United States

[2]Department of Applied Health Science, School of Public Health, Indiana University Bloomington, Bloomington, IN, United States

[3]Department of Applied Psychology in Education and Research Methodology, School of Education, Indiana University Bloomington, Bloomington, IN, United States

Corresponding Author:
Weisiyu Abraham Qin, MPH, MHA, PhD
Center for Tobacco Control Research and Education, Cardiovascular Research Institute, University of California, San Francisco, 530 Parnassus Ave., Ste 366, San Francisco, CA, United States

Abstract

Background: Positive expectancies of cannabis use effects, which are the beliefs about the anticipated positive effects of cannabis, are robust cognitive precursors of adolescent cannabis initiation and escalation. However, little is known about how sociodemographic, familial, and psychopathological factors predict positive expectancies of cannabis use effects or how these expectancies evolve across early adolescence.

Objective: This study aimed to identify distinct developmental trajectories of positive expectancies of cannabis use effects among early adolescents, as well as the longitudinal effects of familial factors on positive expectancies of cannabis use effects over time.

Methods: This study used latent class growth analysis with 3 waves of longitudinal data from the Adolescent Brain Cognitive Development Study (ABCD Study) to identify distinct trajectories of positive expectancies of cannabis use effects among a large, demographically diverse cohort of early adolescents (aged 10‑13 years). Multinomial logistic regression was used to examine whether baseline sociodemographic and policy-level factors were associated with class membership. Time-varying effects of familial factors (ie, parental monitoring, family cannabis use rules, and family conflict) and adolescents' psychopathology were examined within and across trajectory classes using class-specific and common effects models.

Results: Four distinct trajectories of positive expectancies of cannabis use effects emerged with different profiles: moderate-increasing (3118/7409, 42.1%), high-increasing (2111/7409, 28.5%), low-increasing (1496/7409, 20.2%), and high-decreasing (684/7409, 9.2%) trajectories. Parental monitoring and strict family cannabis use rules consistently predicted lower positive expectancies of cannabis use effects, particularly in the moderate- and high-increasing groups, while family conflict emerged as a robust risk factor. Psychopathological symptoms became increasingly predictive of positive expectancies of cannabis use effects at later ages, suggesting a developmental shift in vulnerability.

Conclusions: The development of positive expectancies of cannabis use effects in early adolescence is heterogeneous and shaped by the interplay among sociodemographic, familial, and psychopathological factors. These findings highlight the critical window for early, family-based prevention and underscore the importance of tailoring intervention strategies to specific developmental and risk profiles.

(JMIR Public Health Surveill 2026;12:e85652) doi:10.2196/85652

KEYWORDS

positive cannabis use expectancy; latent class growth analysis; family dynamics; early adolescents; parental monitoring; family cannabis use rules; family conflict

XSL•FO
RenderX

Introduction

Background

Adolescent cannabis use is a significant public health concern in the United States. Despite its federally illegal status [1], it is estimated that 11.2% (2.9 million) of US adolescents (aged 12 - 17 years) used cannabis during the past 12 months [2]. Prior research has shown that early initiation, frequent use, and escalating cannabis use during adolescence are associated with a range of adverse developmental outcomes, including academic underachievement, impaired social functioning, increased risks for depression and suicidality, elevated likelihood of developing substance use disorders, and poorer psychosocial and occupational functioning in later adulthood [3-6]. Understanding cognitive antecedents of cannabis use, particularly positive expectancies of cannabis use effects, is critical for effective prevention.

Substance use expectancies are beliefs about the anticipated effects of using a particular substance, which can serve as critical proximal cognitive mechanisms determining whether an individual will initiate the use of a substance or continue substance use later in life [7-12]. Furthermore, substance use expectancies serve as a core construct in various psychological theories explaining substance use behavior [13], including social learning theory [14-16]; expectancy theory [17]; and plans, responses, impulses, motives, and evaluations (PRIME) theory [18]. Social learning theory emphasizes that substance-related cognitions are acquired through observational learning, modeling, and reinforcement in salient social contexts, such as the family. Expectancy theory and PRIME theory expand on this by conceptualizing that expectancies form as part of a broader evaluative cognitive network that guides motivation, decision-making, and dynamic behavioral choices, which precedes and organizes actual substance use behavior.

Guided primarily by social learning theory, this study focuses on examining how early adolescents, who are particularly sensitive to familial cues, are likely to form positive expectancies of cannabis use effects in response to familial factors. In this context, defining trajectories of positive expectancies of cannabis use effects and identifying family factors (eg, household rules, parental monitoring, and family conflict) that predict membership in different trajectories of positive expectancies of cannabis use effects are essential for informing early interventions and refining theoretical models of cannabis use during early adolescence.

Positive Expectancies of Cannabis Use Effects

Positive expectancies of cannabis use effects include anticipated feelings of relaxation, enhanced creativity, and social connection when using cannabis [19,20], and have been consistently identified as key cognitive drivers of cannabis use behaviors [21-28]. Adolescents who hold more positive beliefs about the anticipated effects of cannabis use are significantly more likely to initiate cannabis use at an earlier age and engage in sustained and/or escalating use over time, even after controlling for other known established risk factors [29,30]. This underscores the unique etiological role of positive expectancies of cannabis use effects in shaping the developmental trajectories of future cannabis initiation. More importantly, positive expectancies of cannabis use effects are modifiable, making them compelling targets for early interventions, before beliefs become firmly entrenched.

Despite growing concerns surrounding adolescent cannabis use and the need for prevention, research has largely focused on behaviors emerging in late adolescence, often neglecting early adolescence (ages 10 - 14 years), a critical period when expectancies develop before direct cannabis experimentation [31,32]. During this period, there is rapid cognitive, emotional, and social development, and environmental influences play formative roles in shaping substance-related expectancies. Among these, family factors, such as parental monitoring, household rules, and family conflict, are particularly influential as they structure adolescents' early views of substance use [33,34].

Family Influences on Positive Expectancies of Cannabis Use Effects

Family rules regarding substance use, parental monitoring, and family conflict all have robust influences on shaping adolescents' substance-related expectancies but have yet to be examined relative to positive expectancies of cannabis use effects. Substantial research has demonstrated that parental alcohol and tobacco rules influence alcohol and tobacco expectancies and subsequent use [35,36]. Empirical studies focusing on cannabis use have shown that clear, well-defined family rules on cannabis use serve as protective factors for cannabis use, whereas the absence or ambiguity of such rules is linked to increased cannabis use [37]. However, research has yet to examine the influence of family rules on positive expectancies of cannabis use effects. It may be that parents who clearly communicate the risks of substance use and enforce explicit household rules indirectly cultivate lower positive expectancies of cannabis use effects in their children, whereas permissive or neutral parental attitudes on substance use may promote more favorable expectancies about cannabis effects. Notably, findings from broader literature on substance use expectancies may not fully extend to positive expectancies of cannabis use effects. The distinct social, legal, and perceived medicinal aspects of cannabis use may lead adolescents to form unique expectancies that differ from those observed for alcohol or tobacco.

Parental monitoring, defined as active supervision and awareness of adolescents' activities [38], represents another critical protective factor across various domains of adolescent risk behavior. Parental monitoring has been consistently associated with reduced alcohol, cannabis, and nicotine use across diverse demographic groups [39-42]. In addition to deterring actual use behaviors, higher levels of parental monitoring are associated with a lower intention to initiate substance use [43]. Given the demonstrated impact on behavioral intentions and decision-making, higher levels of parental monitoring may also reduce adolescents' positive expectancies of cannabis use effects, although direct empirical evidence remains limited.

Family conflict has been robustly associated with an increased risk of substance use and more favorable expectancies of alcohol use [44-47]. Mechanistically, conflict may undermine parental

XSL•FO
RenderX

authority, increase psychological distress, and elicit maladaptive coping strategies, thereby engendering positive attitudes toward substance use [48,49]. While most research has centered on alcohol, the underlying mechanisms are likely applicable to cannabis, warranting extension of these findings to positive expectancies of cannabis use effects. Thus, family rules, parental monitoring, and family conflict represent key proximal determinants of the formation and trajectory of positive expectancies of cannabis use effects. Understanding their dynamics provides important leverage points for targeted interventions that may disrupt adolescent cannabis risk trajectories.

Of note, previous studies on substance use expectancies have relied on conventional growth models to examine developmental trajectories [11,50]. These variable-centered approaches assume that all individuals within a population follow a single average growth trajectory and posit that covariates influencing growth factors affect all individuals uniformly [51]. These variable-centered approaches overlook the possibility of distinct subgroups with divergent developmental pathways, limiting the ability to capture the complexity of early adolescent development [52]. To address this limitation, we used latent class growth analysis (LCGA), a person-centered alternative that classifies individuals into distinct subgroups following similar trajectories, thereby capturing unobserved variation in adolescent development [53,54].

Our Study

In this study, we used LCGA to identify distinct developmental trajectories of positive expectancies of cannabis use effects and examine how parental monitoring, family cannabis use rules, and family conflict are associated with trajectories of positive expectancies of cannabis use effects both within and across trajectory classes while adjusting for demographic characteristics. By integrating a person-centered, longitudinal approach, this study seeks to advance our understanding of how familial factors shape the formation and progression of positive expectancies of cannabis use effects during this critical developmental period.

Methods

Data and Study Sample

The data were drawn from the Adolescent Brain Cognitive Development Study (ABCD Study), the largest ongoing longitudinal investigation of development and health among early adolescents in the United States. Funded by the National Institutes of Health (NIH) and conducted across 21 research sites using a rigorous multistage sampling design, the ABCD Study provides a unique opportunity to understand the factors shaping adolescent development, substance use behaviors, and mental health outcomes [55,56]. Recruitment was carried out between 2016 and 2018 using a systematic school-based sampling approach designed to approximate the demographic composition of the national population of 9- and 10-year-old children [57]. Schools were selected through probability sampling methods stratified by geographic region, race and ethnicity distributions, and socioeconomic characteristics. ABCD Study teams coordinated with school administrators to distribute study information, conduct on-site presentations, and invite families to participate [57]. Additional information regarding study design, methodology, and data accessibility can be found on the study website [58]. The present analysis included data from the 1-year (mean age 10-11 years), 2-year (mean age 11-12 years), and 3-year (mean age 12-13 years) follow-up waves. The 1-year follow-up was designated as baseline (T1), with subsequent waves designated as time 2 (T2) and time 3 (T3) for this study.

Participants who had valid data on the outcome variables at T1 were included, yielding an initial analytic sample of 8841 participants. Between T1 and T2, 418 participants were lost to follow-up, and an additional 606 participants were lost between T2 and T3. Furthermore, 408 cases were excluded due to missing poststratification weights, which are required for population-representative estimates. These exclusions led to a final analytic sample of 7409 participants.

Ethical Considerations

This is a secondary analysis of data collected by the ABCD Study. The ABCD Study was approved by the central Institutional Review Board (IRB) of the University of California, San Diego (IRB# 160091) and by the IRB at each of the 21 participating research sites [59,60]. Written informed consent was obtained from all parents or legal guardians prior to data collection [59,60]. This analysis used the deidentified, publicly available ABCD Study dataset obtained through the NIH Data Hub, and it was deemed exempt from human subject review by the investigators' IRB (Indiana University Bloomington; 2008226356). Participants and families were compensated for the time spent participating in the study, with amounts varying by data collection site.

Measurements

Outcome Variables

Positive Expectancies of Cannabis Use Effects

Positive expectancies of cannabis use effects were assessed using youth self-report on the Marijuana Effect Expectancy Questionnaire-Brief (MEEQ-B) [25]. The items assessed the degree to which adolescents believe that (1) "marijuana helps a person relax and feel less tense," (2) "marijuana helps people get along better with others or feel more romantic," and (3) "marijuana enhances creativity or alters perceptions." The MEEQ-B has been validated among adolescents and young adults, effectively capturing beliefs about the effects of cannabis [61]. Youth responded to 3 positive expectancy questions on a 5-point Likert scale, with higher summed scores (range 3 - 15) reflecting stronger positive expectancies. Internal consistency (Cronbach α) for the positive expectancies of cannabis use effects scale indicated good reliability (α=.77 [T1], .80 [T2], and .83 [T3]).

Predictor Variables

Family Cannabis Use Rules

Aligning with previous studies [62-64], cannabis use rules were measured using parental report on the following question: "What are the family rules about using marijuana for your

XSL•FO
RenderX

son/daughter?" [65-68] Responses were dichotomized as "strict rules" ("not allowed to use marijuana under any circumstances") versus "lenient/no rules," which included all other responses (ie, "not allowed to use marijuana in the home but no rules outside the home," "allowed to use marijuana in the home with permission," "allowed to use marijuana in the home whenever desired," "no rules set about marijuana use," and "have not yet made rules about my child's marijuana use"). Given our study's focus on cannabis and the high correlation of family rules regarding cannabis, alcohol, and nicotine use (r>0.70), alcohol- and nicotine-specific rules were excluded from the present analyses to minimize multicollinearity and improve model interpretability.

Parental Monitoring

Parental monitoring was assessed using youth self-report on 4 items measured on a 5-point Likert scale (1=Never, 2=Almost never, 3=Sometimes, 4=Often, and 5=Always or almost always), with higher mean scores (range 1 - 5) indicating greater parental knowledge, involvement, oversight, and communication [69]. The four items were as follows: (1) "How often do your parents/guardians know where you are?" (2) "If you are at home while your parents or guardians are away, how often do you know how to contact them?" (3) "How often do you talk to your parents or guardians about your plans for the following day, such as school activities or other engagements?" and (4) "How many times do you and your parents/guardians eat dinner together?" This measure reflects the widely used conceptualization of parental monitoring in adolescent development research [40,70-72].

Family Conflict

Consistent with previous studies using the ABCD Study dataset for developmental research [47,72-74], family conflict was assessed using youth self-report on 9 items from the Family Conflict Subscale of the ABCD Study Parent Family Environment Scale, adapted from the PhenX toolkit [75]. Items were coded as True=1 and False=0, with reverse coding applied to positively worded items. The following items were included: "We fight a lot in our family" (1=True), "Family members sometimes get so angry they throw things" (1=True), "Family members often criticize each other" (1=True), and "Family members sometimes hit each other" (1=True). Reverse-coded items included statements such as "Family members rarely become openly angry" (1=False), "Family members hardly ever lose their tempers" (1=False), "If there's a disagreement in our family, we try hard to smooth things over and keep the peace" (1=False), and "In our family, we believe you don't ever get anywhere by raising your voice" (1=False). The items were averaged together, with higher mean scores indicating greater conflict (range 0 - 9). Internal consistency values (Cronbach α) for this study were .67 (T1), .64 (T2), and .68 (T3).

Covariates

Psychopathology

Consistent with previous studies [76-79], youth psychopathology was assessed with parent-reported standardized total scores from the Child Behavior Checklist (CBCL) [80]. This questionnaire comprises 112 items rated on a 3-point Likert scale (0=Not at all true, 1=Somewhat true, and 2=Very true). The total t scores were adjusted for age and sex norms derived from population studies, ensuring comparability across participants [79]. Higher t scores reflect more severe psychopathological problems (range 24 - 88). Cronbach α was .95 for each time point.

Demographic Covariates

Demographic covariates included participant age (in years), biological sex assigned at birth (male/female), parent-reported race and ethnicity, and parental highest education and household income [40,55,73,74,81]. Following the ABCD Study–provided race-ethnicity variable and established frameworks developed by sociocultural literature using the ABCD Study [73,82], parent-reported youth race and ethnicity were categorized as Hispanic, non-Hispanic White, non-Hispanic Black, non-Hispanic Asian, and non-Hispanic other/mixed race (including youth whose parents selected multiple racial categories or "Other race") [83]. Parental education was dichotomized as high school or less versus some college or higher, and household income was dichotomized as less than US $75,000 versus US $75,000 or higher [84-86].

State Recreational Cannabis Legalization Status

State recreational cannabis legalization status was coded as legal (Yes) or not legal (No) by the ABCD Study administration based on the participant's state of residence at baseline in the ABCD Study (approximately 1 year before study T1). Because the dataset does not include time-varying recreational cannabis use policy indicators, this baseline measure served as a proxy for legalization status at T1 (study reference time point).

Statistical Analysis

For the descriptive analysis, unweighted frequencies and weighted proportions were assessed for categorical variables, and weighted means with SDs were calculated for continuous variables at each time point. Differences across time points were evaluated using weighted chi-square tests for categorical variables and weighted ANOVA for continuous variables. Prior to modeling the latent growth models, a bivariate correlation matrix was examined to assess multicollinearity between predictors.

A series of latent growth models was fitted to examine developmental trajectories of positive expectancies of cannabis use effects. Unconditional latent growth curve models (LGCMs) were first examined to assess within-person change and determine whether sufficient heterogeneity existed to justify latent class modeling [50,87,88]. Both linear and quadratic LGCMs were tested using maximum likelihood estimation with robust SEs.

Subsequently, LCGA models were used to identify distinct subgroups of adolescents with similar trajectories of positive expectancies of cannabis use effects. Consistent with standard practice, all the LCGA models were specified with intercept and slope variances fixed to zero and residual variances constrained to equality across time points [54]. Unconditional LCGA models with 1 to 7 classes were evaluated to assess the optimal number of trajectory classes. Model fit was assessed

XSL•FO
RenderX

using multiple criteria: Akaike information criterion (AIC), Bayesian information criterion (BIC), sample size–adjusted Bayesian information criterion (aBIC), entropy, and Lo-Mendell-Rubin adjusted likelihood ratio test (LMR-aLRT) [89-91]. The optimal model was defined as having the lowest information criterion values, significant LMR-aLRT, entropy ≥0.80, and no class size smaller than 5% of the total sample, which was considered statistically unstable [52].

After determining the best-fitting model, R3STEP (auxiliary procedure that implements the 3-step method for adding predictors of latent class membership specified by Mplus) was applied to examine associations between class membership and time-invariant covariates (ie, age, biological sex, race/ethnicity, parental education, recreational cannabis legal status, and total family income). This procedure accounts for the uncertainty in class assignments by incorporating posterior probabilities into auxiliary multinomial logistic regressions. This approach improves estimation accuracy and protects against biased parameter estimates [92,93]. Race/ethnicity was specified as a nominal variable in Mplus, which dummy-coded the variable using non-Hispanic White as the reference category. Mplus then reported the overall omnibus effect of the race/ethnicity block rather than separate coefficients for each category unless individual dummy-coded contrasts produced statistically separable estimates across class comparisons. The full set of dummy contrasts nevertheless contributed internally to the estimation of the classification error–adjusted multinomial logistic model.

To further explore predictive associations with time-varying variables (ie, family cannabis rules, parental monitoring, family conflict, and psychopathology), 2 complementary models were estimated. A class-specific effects model was used to assess the different effects of time-varying predictors across latent classes without interfering with the predefined trajectory classes. This approach revealed heterogeneity in the associations between time-varying predictors and outcome variables across developmental trajectories. As a sensitivity analysis, a common effects model was used to estimate population-average associations between the time-varying predictors and outcome variables under the assumption that the effects of the time-varying predictors on the outcome variables are homogeneous across all latent classes. This model provides insights to understand the general exposure effects that are consistent across subpopulations.

All LCGA models were estimated using Mplus 8.11 [94]. To ensure model stability and reduce the risk of convergence on the local maxima, a multistage estimation procedure was used. Each model was initialized with 1000 random sets of starting values, from which the 250 best-fitting solutions were retained for final optimization. To further verify solution stability, log-likelihood values were required to replicate across 20 iterations. Likelihood ratio tests (eg, LMR-aLRT) were conducted with an additional set of 1000 random start replications, with 200 used for preliminary evaluation and 500 selected for final optimization, repeated 100 times, to ensure the reliability of model comparison results. Missing data of predictors were imputed using the nonparametric random forest–based approach, which has been shown to perform well in retaining nonlinear relationships and interactions among variables in mixed-type datasets [95]. Descriptive statistics were conducted with R 4.5.0 (via RStudio, Posit). The reporting of this study followed the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines (Checklist 1) [96].

Results

Descriptive Statistics

Table 1 presents the descriptive statistics of the study sample across the 3 time points. The positive expectancies of cannabis use effects score measured among the participants demonstrated an increasing trend (T1: mean 6.41, SD 2.85; T2: mean 7.20, SD 2.93; T3: mean 7.96, SD 2.99; $P<.001$). Similarly, the proportion of strict family cannabis use rules increased over time from 97.9% (T1) to 98.9% (T3) ($P<.001$). Moreover, family conflict scores increased from 1.90 (SD 1.88) to 2.09 (SD 1.96) ($P<.001$). Parental monitoring scores showed a slight decrease across waves (T1: mean 4.51, SD 0.45; T3: mean 4.38, SD 0.50; $P<.001$).

XSL•FO
RenderX

Table . Descriptive statistics of participants from time 1 to time 3 in the Adolescent Brain Cognitive Development Study (ABCD Study) (N=7409).

Variable[a]	Time 1 (age 10-11 years)[b]	Time 2 (age 11-12 years)[b]	Time 3 (age 12-13 years)[b]	*P* value[c]
Positive expectancies of cannabis use effects (range 3-15), mean (SD)	6.41 (2.85)	7.20 (2.93)	7.96 (2.99)	<.001
Biological sex, n (%)				—
Male	4040 (49.0)	—[d]	—	
Female	3369 (51.0)	—	—	
Race/ethnicity, n (%)				—
NH[e] White	4126 (20.7)	—	—	
NH Black	918 (37.6)	—	—	
Hispanic	1421 (28.9)	—	—	
NH Asian	144 (0.4)	—	—	
NH others	800 (12.3)	—	—	
Parental education, n (%)				—
High school or less	1163 (21.5)	—	—	
Some college or higher	6221 (78.5)	—	—	
Recreational cannabis legal status[f], n (%)				—
No	5114 (72.8)	—	—	
Yes	1973 (27.2)	—	—	
Total family income, n (%)				—
Less than US $75,000	2647 (74.4)	—	—	
US $75,000 or higher	4269 (25.6)	—	—	
Age (range 7 - 16 years), mean (SD)	10.55 (0.64)	11.56 (0.71)	12.51 (0.68)	<.001
Standardized psychopathology *t* score (range 24-88), mean (SD)	45.67 (11.06)	45.11 (11.13)	45.01 (11.25)	<.001
Family conflict score (range 0 - 9), mean (SD)	1.90 (1.88)	1.90 (1.84)	2.09 (1.96)	<.001
Parental monitoring score (range 1 - 5), mean (SD)	4.51 (0.45)	4.50 (0.46)	4.38 (0.50)	<.001
Family cannabis use rules, n (%)				<.001
Lenient/no rules	1685 (2.3)	1331 (1.8)	1010 (1.1)	
Strict rules	5712 (97.9)	6033 (98.2)	6207 (98.9)	

[a]Except for the baseline sociodemographic characteristics (biological sex, race/ethnicity, parental education, and recreational cannabis legal status), all other variables were measured repeatedly from time 1 (T1) to time 3 (T3).

[b]Values represent unweighted frequencies and weighted proportions for categorical variables, and weighted means with SDs for continuous variables. Frequencies may not sum to the total sample size due to missing data.

[c]*P* values were generated using weighted ANOVA for continuous variables and weighted chi-square tests for categorical variables to test differences across waves.

[d]Not applicable.

[e]NH: non-Hispanic.

[f]The cannabis recreational legal status was determined based on the participant's state of residence at the time of their baseline interview in the ABCD Study, which is approximately 1 year prior to T1 in this study.

XSL•FO
RenderX

Unconditional LCGA Model Statistics Regarding Positive Expectancies of Cannabis Use Effects

Table 2 presents the latent class model fit comparisons of the optimal class solutions, ranging from 1 class to 7 classes. Table 3 presents the sizes of the individual classes. While models with a greater number of classes (5-class to 7-class trajectory solutions) were explored, they yielded subgroups with minimal representation (ie, group size <5% of the total sample), raising concerns about model overfitting and limited interpretability.

Table . Latent class model fit comparisons for unconditional latent class growth analysis models regarding positive expectancies of cannabis use effects.

Trajectory (model)	Log likelihood	BIC[a]	aBIC[b]	AIC[c]	LMR-aLRT[d] *P* value	BLRT[e] *P* value	Entropy[f]	Minimal class membership[g] (%)
1 class	−55350.179	110744.910	110729.021	110710.358	__[h]	—	—	—
2 classes	−53156.244	106383.722	106358.349	106328.488	<.001	<.001	0.734	38.9
3 classes	−52609.740	105317.495	105282.540	105241.480	<.001	<.001	0.776	16.9
4 classes[i]	−52116.646	104358.038	104313.549	104261.292	<.001	<.001	0.827	9.2
5 classes	−51746.873	103645.224	103591.202	103527.746	<.001	<.001	0.841	4.4
6 classes	−51583.031	103344.270	103280.714	103206.061	<.001	<.001	0.838	2.9
7 classes	−51396.675	102998.291	102925.202	102839.351	<.001	<.001	0.866	0.4

[a]BIC: Bayesian information criterion.

[b]aBIC: sample size–adjusted Bayesian information criterion.

[c]AIC: Akaike information criterion.

[d]LMR-aLRT: Lo-Mendell-Rubin adjusted likelihood ratio test; *P* value for k-1 refers to a significant improvement in model fit between the class (k) and the preceding class (k-1), which compares whether a profile solution with k profiles fits significantly better than a profile.

[e]BLRT: parametric bootstrapped likelihood ratio test, which is similar to the LMR-aLRT; *P* value refers to a significant improvement in model fit between the class (k) and the preceding class (k-1).

[f]Entropy indicates classification accuracy, with a higher value indicating better classification (range 0-1).

[g]Minimal class membership represents the proportion of participants in the latent class with the smallest membership.

[h]Not applicable.

[i]Selected model.

Table . Sizes of the classes (N=7409).

Trajectory (model)	Class[a]						
	Class 1, n (%)	Class 2, n (%)	Class 3, n (%)	Class 4, n (%)	Class 5, n (%)	Class 6, n (%)	Class 7, n (%)
1 class	7409 (100.0)	__[b]	—	—	—	—	—
2 classes	4530 (61.1)	2879 (38.9)	—	—	—	—	—
3 classes	4125 (55.7)	2035 (27.5)	1249 (16.9)	—	—	—	—
4 classes[c]	2111 (28.5)	684 (9.2)	1496 (20.2)	3118 (42.1)	—	—	—
5 classes	1897 (25.6)	328 (4.4)	3054 (41.2)	1470 (19.8)	660 (8.9)	—	—
6 classes	1864 (25.2)	674 (9.1)	502 (6.8)	1375 (18.6)	2774 (37.4)	220 (2.9)	—
7 classes	1247 (16.8)	31 (0.4)	1439 (19.4)	219 (2.9)	1433 (19.3)	2398 (32.4)	642 (8.7)

[a]Each cell displays the frequency and corresponding proportion of individuals within each latent class. Frequencies represent the unweighted counts, while proportions are calculated relative to the total number within each class or group.

[b]Not applicable.

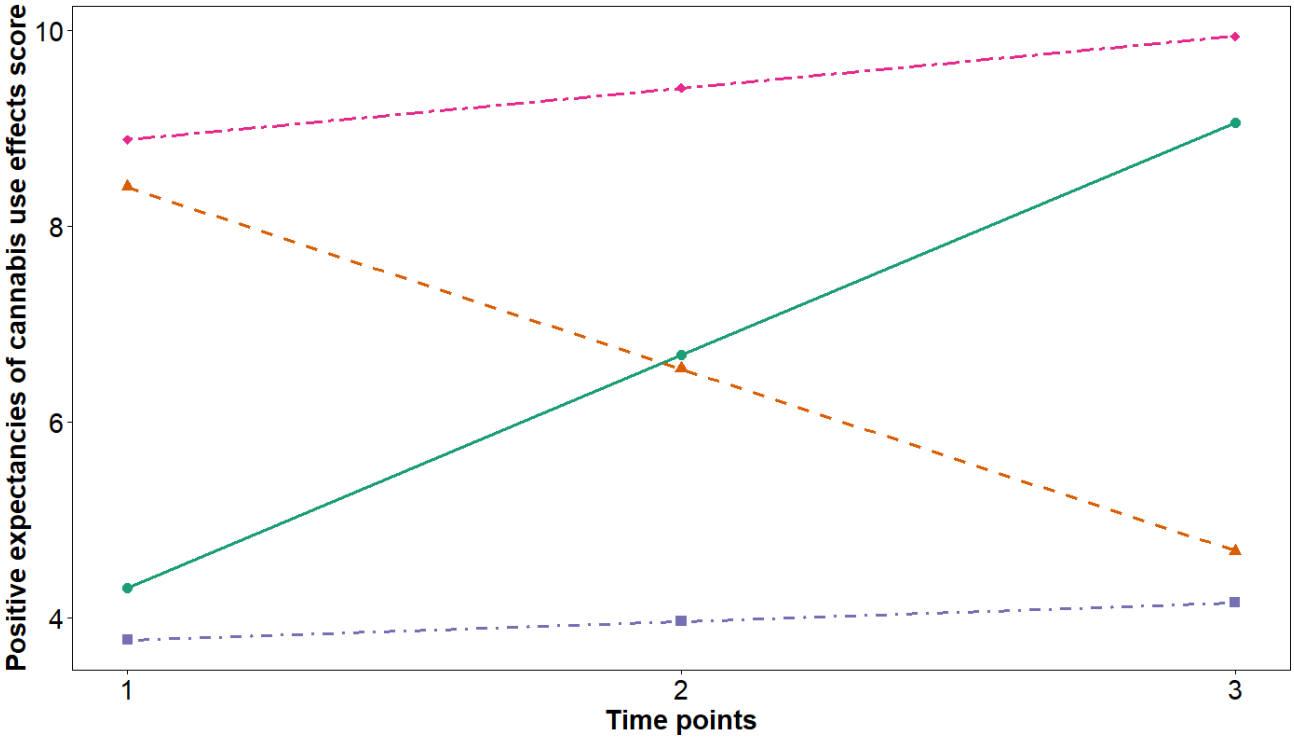
[c]Selected model.

The 4-class model was selected as the optimal solution based on both statistical fitness and conceptual interpretability. This model demonstrated comparatively lower values for the BIC, aBIC, and AIC and higher entropy compared with the 3-class model, indicating improved classification precision. Compared with the 5-class model, it maintained a balanced class distribution, with each subgroup exceeding the recommended 5% minimum threshold (the smallest group being class 2, with 684 cases or 9.2% of the total sample). Therefore, the 4-class solution balanced parsimony with meaningful subgroup differentiation, which avoided interpretive challenges posed by extremely small latent classes observed in higher-order models.

XSL•FO
RenderX

Figure 1 visualizes the 4-class trajectories. The largest subgroup followed a moderate-increasing trajectory that was characterized by a high baseline level but a moderate increasing trend over time (class 4: moderate-increasing class; n=3118, 42.1% of the sample). The second most prevalent trajectory followed a high-increasing trajectory that was characterized by a moderate baseline level and a steep increase across the study period (class 1: high-increasing class; n=2111, 28.5% of the sample). The third most prevalent trajectory followed a low-increasing trajectory that was characterized by a low baseline level with a slight increase (class 3: low-increasing class; n=1496, 20.2% of the sample). The smallest group followed a high-decreasing trajectory that was characterized by a high baseline level that declined sharply (class 2: high-decreasing class; n=684, 9.2% of the sample). Parameter estimates and detailed trajectory features are reported in Table S3 in Multimedia Appendix 1.

Figure 1. The 4-class developmental trajectories of positive expectancies of cannabis use effects at 3 time points. The y-axis represents the mean positive expectancies of cannabis use effects score, and the x-axis represents the 3 examined time points. The 4 trajectories represent latent classes identified through latent class growth modeling: class 1 is plotted with a solid green line and filled circle markers (2111/7409, 28.5%), class 2 is plotted with an orange dashed line and filled triangle markers (684/7409, 9.2%), class 3 is plotted with a purple dot-dash line and filled square markers (1496/7409, 20.2%), and class 4 is plotted with a pink 2-dash line and filled diamond markers (3118/7409, 42.1%).



Associations Between Baseline Time Invariant Variables and Latent Class Membership

Table 4 presents the results of multinomial logistic regression with the low-increasing class (class 3) serving as the reference category. Compared with this group, youth in the high- and moderate-increasing classes were older (adjusted odds ratio [aOR] 1.13, 95% CI 1.01 - 1.25; *P*=.04 and aOR 1.45, 95% CI 1.31 - 1.60; *P*<.001, respectively). Those in the moderate-increasing class were also more likely to reside in states with legalized recreational cannabis use and be from higher-income families. Multinomial logistic regression findings with other groups as reference categories are reported in Tables S5 - S7 in Multimedia Appendix 1.

XSL•FO
RenderX

Table . Multinomial logistic regression predicting latent class membership (reference class: class 3).

Class[a] and variable[b]	aOR[c,d] (95% CI)	*P* value
Class 1		
Biological sex	0.94 (0.82 - 1.08)	.39
Race/ethnicity	1.03 (0.98 - 1.08)	.27
Parental education	1.06 (0.88 - 1.28)	.53
Recreational cannabis legal status	1.13 (0.96 - 1.32)	.16
Age	1.13 (1.01 - 1.25)	.04[e]
Total family income	1.00 (0.87 - 1.16)	.97
Class 2		
Biological sex	1.05 (0.88 - 1.27)	.59
Race/ethnicity	0.97 (0.90 - 1.04)	.40
Parental education	0.85 (0.67 - 1.09)	.17
Recreational cannabis legal status	1.06 (0.85 - 1.32)	.60
Age	0.99 (0.85 - 1.14)	.88
Total family income	1.09 (0.90 - 1.32)	.40
Class 4		
Biological sex	1.01 (0.89 - 1.15)	.84
Race/ethnicity	1.02 (0.97 - 1.07)	.38
Parental education	0.99 (0.83 - 1.17)	.88
Recreational cannabis legal status	1.28 (1.11 - 1.49)	.003[f]
Age	1.45 (1.31 - 1.60)	<.001[g]
Total family income	1.25 (1.09 - 1.42)	.004[f]

[a]The reference category for this model is class 3 (low-increasing), which represents 1496 participants (20.2%).

[b]For covariates, the reference groups are as follows: female for biological sex, non-Hispanic White for race/ethnicity, high school education or less for parental education, non-legalized status for recreational cannabis legal status, and total family income below US $75,000 for family income. Race/ethnicity was specified as a 5-category nominal covariate and was dummy-coded internally by Mplus. Mplus reports a single omnibus effect representing the overall effect of this multicategory covariate.

[c]aOR: adjusted odds ratio.

[d]Reported odds ratios represent the relative odds of belonging to each latent class versus the reference class for each covariate category.

[e]$P<.05$.

[f]$P<.01$.

[g]$P<.001$.

Dynamic Associations of Family Environment and Psychopathology With the Trajectories of Positive Expectancies of Cannabis Use Effects Within Each Latent Trajectory Class

Table 5 presents the results from the class-specific effects model estimating the time-varying associations of familial and psychopathological predictors with positive expectancies of cannabis use effects across the 4 identified latent trajectory classes. Each class was modeled independently to capture heterogeneity in relation to the predictors over time. Distinct time-varying familial and psychopathological predictors emerged, underscoring differential developmental processes. In the high-increasing class, lower parental monitoring predicted greater expectancy growth at both T1 (β=−0.152, SE=0.072; *P*=.04) and T2 (β=−0.477, SE=0.122; *P*<.001), while increased family conflict at T3 (β=0.071, SE=0.019; *P*<.001) predicted elevated positive expectancies of cannabis use effects. In the high-decreasing class, only family conflict at T2 (β=0.124, SE=0.063; *P*=.047) was a significant risk factor, possibly reflecting transient reinforcement of positive expectancies before decline. The low-increasing class exhibited no significant associations across time points, though family conflict at T3 approached significance (β=0.038, SE=0.021; *P*=.07). In contrast, the moderate-increasing class (the largest group) showed the most consistent effects: less strict family cannabis use rules were significantly associated with expectancy increases at T1 (β=−0.171, SE=0.063; *P*=.006) and T2 (β=−0.212, SE=0.091; *P*=.02), lower parental monitoring was significant at both T2 (β=−0.275, SE=0.088; *P*=.002) and T3 (β=−0.307, SE=0.066; *P*<.001), and family conflict was the most robust predictor from T1 to T3 (*P*<.005).

XSL•FO
RenderX

Table . Class-specific estimates of time-varying predictors of positive expectancies of cannabis use effects across 3 time points.

Class, time, and variable[a]	β	SE	z-statistics	*P* value
Class 1 (high-increasing trajectory)				
Time 1				
Family cannabis use rules	−0.079	0.071	−1.108	.27
Parental monitoring	−0.152	0.072	−2.111	.04[b]
Psychopathology *t* score	−0.004	0.003	−1.293	.20
Family conflict	−0.016	0.017	−0.959	.34
Time 2				
Family cannabis use rules	−0.226	0.149	−1.519	.13
Parental monitoring	−0.477	0.122	−3.922	<.001[c]
Psychopathology *t* score	−0.008	0.006	−1.312	.19
Family conflict	0.019	0.033	0.587	.56
Time 3				
Family cannabis use rules	0.023	0.099	0.229	.82
Parental monitoring	−0.123	0.077	−1.603	.11
Psychopathology *t* score	0.005	0.004	1.527	.13
Family conflict	0.071	0.019	3.687	<.001[c]
Class 2 (high-decreasing trajectory)				
Time 1				
Family cannabis use rules	−0.143	0.141	−1.009	.31
Parental monitoring	0.048	0.139	0.349	.73
Psychopathology *t* score	−0.004	0.006	−0.765	.44
Family conflict	−0.007	0.037	−0.195	.85
Time 2				
Family cannabis use rules	0.150	0.275	0.544	.59
Parental monitoring	−0.369	0.267	−1.381	.17
Psychopathology *t* score	−0.006	0.010	−0.620	.54
Family conflict	0.124	0.063	1.982	.047[b]
Time 3				
Family cannabis use rules	−0.123	0.169	−0.729	.47
Parental monitoring	−0.135	0.129	−1.044	.30
Psychopathology *t* score	−0.010	0.006	−1.825	.07
Family conflict	0.050	0.033	1.500	.13
Class 3 (low-increasing trajectory)				
Time 1				
Family cannabis use rules	−0.065	0.073	−0.886	.38
Parental monitoring	−0.063	0.064	−0.978	.33
Psychopathology *t* score	0.004	0.003	1.408	.16
Family conflict	0.013	0.017	0.801	.42
Time 2				
Family cannabis use rules	0.099	0.150	0.661	.51
Parental monitoring	−0.143	0.119	−1.204	.23

XSL•FO
RenderX

Class, time, and variable[a]	β	SE	z-statistics	*P* value
Psychopathology *t* score	0.002	0.005	0.410	.68
Family conflict	0.040	0.036	1.111	.27
Time 3				
Family cannabis use rules	0.074	0.107	0.697	.49
Parental monitoring	−0.045	0.069	−0.645	.52
Psychopathology *t* score	−0.002	0.003	−0.814	.42
Family conflict	0.038	0.021	1.805	.07
Class 4 (moderate-increasing trajectory)				
Time 1				
Family cannabis use rules	−0.171	0.063	−2.733	.006[d]
Parental monitoring	−0.089	0.068	−1.317	.19
Psychopathology *t* score	0.004	0.003	1.555	.12
Family conflict	0.046	0.016	2.859	.004[d]
Time 2				
Family cannabis use rules	−0.212	0.091	−2.317	.02[b]
Parental monitoring	−0.275	0.088	−3.105	.002[d]
Psychopathology *t* score	0.012	0.004	2.908	.004[d]
Family conflict	0.079	0.022	3.689	<.001[c]
Time 3				
Family cannabis use rules	0.031	0.078	0.402	.69
Parental monitoring	−0.307	0.066	−4.654	<.001[c]
Psychopathology *t* score	0.010	0.003	3.085	.002[d]
Family conflict	0.074	0.016	4.661	<.001[c]

[a]The outcome variable is the individual's level of positive expectancies of cannabis use effects at each time point (time 1, time 2, and time 3) within each latent class. All covariates are repeated measures within respondents from time 1 to time 3. This model does not predict the growth trajectory, but instead, it estimates how time-varying predictors are associated with variation in the positive expectancies of cannabis use effects score over time within each trajectory class.

[b]*P*<.05.

[c]*P*<.001.

[d]*P*<.01.

Shared Associations of Family Environment and Psychopathology With Positive Expectancies of Cannabis Use Effects Across Latent Trajectory Classes

Results from the common effects model are presented in Table S8 in Multimedia Appendix 1, where time-varying familial and psychopathological predictors were constrained to have equal influence across all latent trajectory classes. Strict family cannabis use rules were significantly associated with lower positive expectancies of cannabis use effects at T1 only (β=−0.130, SE=0.040; *P*=.001). However, parental monitoring remained a significant predictor for positive expectancies of cannabis use effects across all 3 time points (T1: β=−0.089, SE=0.040; *P*=.03; T2: β=−0.315, SE=0.061; *P*<.001; T3: β=−0.185, SE=0.040; *P*<.001). Family conflict was a consistent and robust risk factor, with its influence increasing from T2 to T3 (*P*<.001).

Discussion

Heterogeneous Trajectories of Positive Expectancies of Cannabis Use Effects in Early Adolescence

To the best of our knowledge, this is the first study to use a large-scale longitudinal dataset to examine the developmental trajectories of positive expectancies of cannabis use effects in early adolescents, using a person-centered analytic framework. By modeling the development of positive expectancies of cannabis use effects across early adolescence, this study offers novel insights into the dynamic, heterogeneous nature of the formation of positive expectancies of cannabis use effects during this sensitive developmental phase. It identified the following

XSL•FO
RenderX

4 distinct trajectories of positive expectancies of cannabis use effects: high-increasing, high-decreasing, low-increasing, and moderate-increasing trajectories. These trajectories highlight the substantial variability in both the baseline levels and patterns of change in positive expectancies of cannabis use effects, underscoring early adolescence as an important period for tailoring interventions to prevent cannabis use.

Although weighted descriptive statistics across the sample indicated relatively modest population-level changes over the 3 waves, this pattern is expected given the narrow but critical developmental window of early adolescence (approximately ages 11 - 13 years) represented in our sample from the ABCD Study cohort. During this period, many psychosocial and contextual characteristics exhibit relative stability at the population level [97], yet substantial *within-person* variability persists in cognitive-affective processes such as substance-related expectancies, social-emotional development, and dynamic familial factors [7,98-100]. LCGA, a person-centered approach, is uniquely suited to capturing individual-level heterogeneity because it identifies subgroups of youth who share similar developmental trajectories even when the overall mean trend appears relatively flat. Accordingly, the 4 trajectory classes identified in this study represent distinct and meaningful expectancy development over time rather than simple cross-sectional differences based on average levels [52,101]. In addition, although the trajectory classes are derived from repeated measures of positive expectancies of cannabis use effects, the LCGA analytic framework used in this study helps minimize potential bias arising from factors, such as parental monitoring and family conflict, which may influence both the development of positive expectancies of cannabis use effects and the predictors included in the R3STEP model. LCGA does not stratify individuals on a single observed positive expectancies of cannabis use effects score; instead, it forms subgroups based on model-estimated posterior probabilities that reflect the overall pattern of trajectories. Moreover, R3STEP estimates covariate associations only after class formation and incorporates adjustment for classification uncertainty, reducing the bias that can arise when treating uncertain class assignments as if they are certain [92]. Within this framework, associations between predictors and class membership represent correlational patterns among latent developmental pathways rather than artifacts of the analytic strategy. Thus, the heterogeneity observed across classes reflects meaningful differences in expectancy development over time, highlighting the advantage of mixture modeling for uncovering nuanced developmental processes that would remain obscure in traditional variable-centered analyses.

Using this approach, the largest trajectory class identified in this study was the moderate-increasing trajectory, which was characterized by relatively high initial levels of positive expectancies of cannabis use effects that increased steadily with time, suggesting an active expectancy formation phase. This pattern may reflect normative developmental processes in early adolescence, where adolescents increasingly seek self-identity and autonomy, and become increasingly susceptible to substance use opportunities [102]. Multivariable comparisons revealed that youth in this trajectory were more likely to be older, from high-income families, and living in states with recreational cannabis legalization. Notably, the consistently elevated and gradually intensifying positive expectancies of cannabis use effects observed in this group are concerning, as the findings point to a sustained expectancy formation process that may heighten the risk for future initiation and persistent use. Youth in this trajectory may be actively shaping their cognitive belief around cannabis use prior to engaging in cannabis use, which may be reinforced by their developmental maturity (older age) and the legal status of recreational cannabis use in their environment. These findings suggest that prevention efforts should extend beyond traditionally high-risk youth to include those on seemingly normative developmental pathways who may nonetheless be building pro-cannabis expectancies that increase long-term vulnerability. The other 3 trajectories provide further insights into heterogeneity in the development of positive expectancies of cannabis use effects, highlighting the substantial variability in both the onset and developmental course of positive expectancies of cannabis use effects and pointing to multiple pathways of risk and resilience in early cannabis-related cognitions.

Familial Protective and Risk Factors

Findings from the common effects model showed that stricter family cannabis use rules and higher levels of parental monitoring were protective against positive expectancies of cannabis use effects during early adolescence. These effects were the strongest at earlier ages, particularly at 10 - 11 years (T1) and 11 - 12 years (T2). Parental monitoring demonstrated a consistent inverse association with positive expectancies of cannabis use effects across all 3 time points, with the strongest effect at ages 11 - 12 years. In contrast, cannabis-specific rules were significant only at T1, with diminished predictive value at later time points. These protective effects align with the findings of a large body of developmental research emphasizing the critical role of structured and engaged parenting in deterring adolescent substance-related cognitions and behaviors [43,103].

The early and pronounced influence of family cannabis use rules underscores their importance in shaping adolescents' cognitive attitudes regarding substance use when they are most embedded within the family context and are more receptive to parental expectations and boundaries [104]. These rules function as clear behavioral norms, potentially counterbalancing early exposure to peer influences and emerging social scripts around cannabis. Although the direct statistical effect weakened by ages 11 - 12 years, such rules may establish enduring internalized norms that persist even when external risks are present. Prior research suggests that early parental rule-setting exerts long-term influence on substance-related decision-making. For example, the authoritative parenting style is characterized by setting limits and is linked to lower substance use and less positive attitudes toward drugs throughout adolescence [105-107].

Parental monitoring demonstrated a more enduring and stable protective effect across the developmental period studied. Unlike rule-setting, monitoring reflects an ongoing dynamic engagement with the daily lives of adolescents, which provides not only behavioral oversight but also emotional attunement

XSL•FO
RenderX

and accountability [108]. This form of proactive parenting has been consistently shown to reduce adolescents' opportunities to engage in risk behaviors and to shape substance-related cognitions in a protective direction [65,108]. The heightened impact observed at ages 11-12 years may indicate a critical developmental "sweet spot," when adolescents begin to seek autonomy but remain highly responsive to external regulation and support. These findings emphasize the importance of initiating family-based prevention efforts during early adolescence, leveraging this window to reinforce cognitive resistance to substance use before peer norms and societal influences exert stronger effects.

Family conflict emerged as a robust risk factor for elevated positive expectancies of cannabis use effects at ages 12-13 years (T3). In addition, class-specific models showed that this effect was the strongest in the moderate- and high-increasing classes, suggesting that sustained family tensions may accelerate the formation of positive expectancies of cannabis use effects, particularly when not offset by protective parenting practices. This finding is consistent with the findings of a substantial body of literature linking family dysfunction to increased vulnerability to substance use [109,110]. Chronic conflict may erode emotional regulation capacities and model maladaptive coping strategies, thereby reinforcing cannabis as a perceived tool for managing stress [111]. These findings highlight the dual importance of reinforcing protective parenting practices and reducing family conflict during this critical developmental period.

Class-Specific Nuances

Unlike the common effects model, which assumes that predictors operate uniformly across all trajectories, the class-specific model allows the effects of parental and familial factors to vary by trajectory and captures heterogeneity in developmental processes. The findings revealed notable differences in the influences of familial and psychopathological factors across the 4 trajectories of positive expectancies of cannabis use effects.

Parental monitoring was pronounced among adolescents who followed trajectories marked by moderate- and high-increasing risk, where higher monitoring during early and mid-adolescence (ages 10 - 12 years) was associated with significantly lower positive expectancies of cannabis use effects. Additionally, stricter family cannabis use rules were associated with lower positive expectancies of cannabis use effects in the moderate-increasing group at ages 10-11 and 11-12 years. These effects were most pronounced prior to the age of 13 years, underscoring a sensitive developmental window when parental guidance may shape substance-related cognitions before peer norms and autonomy-seeking dominate [112,113]. It is also possible that elevated parental monitoring reflects a reactive process, wherein parents increase oversight in response to perceiving their child's heightened risk or early signs of problematic behaviors. In this interpretation, monitoring may serve as a preventive and responsive strategy, suggesting that parents who recognize vulnerability may intensify supervision as a preemptive measure against further risk escalation.

However, these protective effects were less evident or absent in the high-decreasing and low-increasing trajectories. In the high-decreasing group, early high levels of positive expectancies of cannabis use effects declined over time, and neither parental monitoring nor rule-setting was significantly associated with these shifts. This may suggest that reductions in positive expectancies of cannabis use effects were driven by other factors, such as experiential disconfirmation or broader contextual influences, and further research is warranted for this group of youth. In the low-increasing group, neither monitoring nor rule-setting showed significant associations, though family conflict emerged as a marginal risk factor at later ages. The absence of early effects in this group could reflect other protective dispositional factors (eg, low sensation-seeking) or structural buffers (eg, strong school engagement) not captured in our analysis.

Family conflict emerged as the most robust and class-differentiating risk factor across trajectories, particularly for those at elevated or increasing risk. Adolescents in the moderate- and high-increasing groups exhibited significantly elevated positive expectancies of cannabis use effects in association with greater family conflict, especially by ages 12-13 years. This pattern suggests that interpersonal stress within the home may amplify the development of positive expectancies of cannabis use effects during a period of heightened social and emotional reactivity. In line with developmental cascade models [114], chronic exposure to family conflict may erode previously adaptive cognitive trajectories and accelerate the adoption of risk-promoting beliefs. Conflict was also a significant predictor in the high-decreasing group at mid-stage (ages 11-12 years), suggesting a contemporaneous, level-shifting effect of conflict rather than a change in growth rate. In contrast, conflict showed no impact among adolescents in the low-increasing class, potentially reflecting greater resilience or the presence of unmeasured compensatory mechanisms, such as school connectedness and temperamentally based self-regulation.

Psychopathological symptoms played a more nuanced, temporally specific role. Across most trajectories, they were not significant predictors in early adolescence but became increasingly relevant by ages 11-12 and 12-13 years in the moderate-increasing class. This shift may reflect the growing salience of emotional distress in early adolescence, when academic, social, and identity-related demands intensify, and cannabis may be perceived as a coping tool.

Collectively, several effects were close to statistical significance based on the *P* value. It is likely that we detected small effects because of our large sample size. Therefore, the findings warrant replication.

Limitations

Several limitations should be considered when interpreting the findings of this study. First, the analysis was limited to 3 waves of data collected during early adolescence, which restricts the ability to capture the full developmental trajectory of positive expectancies of cannabis use effects into middle and late adolescence. Given that substance-related attitudes and behaviors often intensify during these later periods, future research with a longer follow-up is needed to determine whether the identified trajectories persist, shift, or predict distal actual cannabis use behavior. Second, the measure of recreational

XSL•FO
RenderX

cannabis legalization status was based on state-level policy 1 year prior to this study's baseline (T1) assessment. While useful as a contextual marker, this static measure may not fully reflect the evolving influence of recreational legalization over time, particularly as policy implementation and social norms continue to change. Third, this study used a 3-step LCGA in which covariates and predictors were deliberately excluded from the trajectory formation process and incorporated only in the R3STEP multinomial logistic regression and subsequent class-specific and common effects models. While this approach preserves the integrity of class estimation, unmeasured or imperfectly measured factors may still be associated with both the trajectory of positive expectancies of cannabis use effects and class membership, introducing the possibility of residual confounding. Fourth, the study was unable to account for a range of other social and contextual influences that are becoming increasingly relevant to adolescent development and substance-related behaviors. Given emerging evidence that cannabis-related content on social media can shape adolescents' attitudes, expectancies, and perceived norms, future studies should integrate time-matched assessments of digital media exposure to provide a more comprehensive understanding of expectancy formation.

Public Health Implications

This study fills a critical gap in the literature by identifying 4 distinct developmental trajectories of positive expectancies of cannabis use effects among early adolescents, underscoring the need for prevention strategies that extend beyond universal, one-size-fits-all models. While universal prevention remains important, interventions must be tailored to the heterogeneous developmental pathways identified in this study.

In addition to these practical implications, the observed trajectory patterns and associated family predictors contribute to the development of the current theory. The identification of distinct developmental trajectories of positive expectancies of cannabis use effects suggests that expectancy formation during early adolescence may be heterogeneous rather than uniformly increasing. The findings that stricter household rules and higher parental monitoring were associated with membership in lower or declining classes of positive expectancies of cannabis use effects and that greater family conflict was associated with higher-risk classes of positive expectancies of cannabis use effects are consistent with the social learning theory that emphasizes the role of family environments in shaping evaluative beliefs and related motivational states. Together, these findings suggest that theoretical models of adolescent cannabis use may benefit from incorporating heterogeneity in the development of positive expectancies of cannabis use effects and accounting for the structuring influence of family dynamics during this critical developmental period.

Parental monitoring and clear cannabis use rules were most protective in early adolescence, particularly between the ages of 10 and 12 years, when parental influence remains salient. Intervening during this sensitive period may delay or suppress the rise in positive expectancies of cannabis use effects before peer norms and autonomy-seeking behaviors exert greater influences. These practices are especially critical for adolescents in moderate- and high-increasing trajectories, where sustained parental engagement may disrupt escalation in expectancies. In contrast, family conflict emerged as a robust risk factor, particularly in the later stages of early adolescence. Chronic conflict may amplify cognitive vulnerability, destabilize otherwise adaptive trajectories, and accelerate the internalization of risk-promoting beliefs. Prevention programs that educate parents about conflict management and communication skills could therefore provide meaningful protection. Psychopathology also became increasingly salient by ages 11‑13 years, reinforcing the importance of integrated prevention that includes mental health screening and timely intervention. Addressing emotional distress and teaching adaptive coping strategies may reduce the perceived utility of cannabis for managing stress.

Collectively, these findings highlight the value of early, sustained, and nuanced family involvement across developmental stages. Cannabis use prevention programs should also focus on enhancing parental self-efficacy by providing practical tools for effective communication, conflict resolution, and the implementation of developmentally appropriate cannabis use rules [115]. Early adolescence is a critical period during which parental authority still has a dominant influence, particularly for adolescents not yet embedded in high-risk trajectories. For these adolescents, clear and consistent rule-setting within a supportive context may help prevent escalation in expectancies and delay susceptibility.

Acknowledgments

Data used in the preparation of this article were obtained from the Adolescent Brain Cognitive Development Study (ABCD Study) [58], held in the National Institute of Mental Health Data Archive (NDA). This is a multisite, longitudinal study designed to recruit more than 10,000 children aged 9-10 years and follow them over 10 years into early adulthood. The ABCD Study is supported by the National Institutes of Health (NIH) and additional federal partners under award numbers U01DA041048, U01DA050989, U01DA051016, U01DA041022, U01DA051018, U01DA051037, U01DA050987, U01DA041174, U01DA041106, U01DA041117, U01DA041028, U01DA041134, U01DA050988, U01DA051039, U01DA041156, U01DA041025, U01DA041120, U01DA051038, U01DA041148, U01DA041093, U01DA041089, U24DA041123, and U24DA041147. A full list of supporters is available on the ABCD Study website [116]. A list of participating sites and a complete list of the study investigators can be found on the ABCD Study website [117]. ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in the analysis or writing of this report. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators. The ABCD data repository grows and changes over time. The ABCD data used in this report came from the ABCD 5.1 release [118].

XSL•FO
RenderX

Funding

This research did not receive funding from any agencies in the public, commercial, or not-for-profit sectors.

Data Availability

Data for this study were derived from the Adolescent Brain Cognitive Development Study (ABCD Study) 5.1 release [118]. Deidentified ABCD Study data are available to qualified investigators through the National Institutes of Health Brain Development Cohorts (NBDC) Data Hub. Information on data access procedures is provided on the ABCD Study data sharing page [119]. The raw data of this study are available on the National Institute of Mental Health Data Archive (NDA) website [120]. Instructions on how to create an NDA study are available on the NDA website [121].

Authors' Contributions

Conceptualization: WAQ, KKE
Data curation: WAQ, SH
Formal analysis: WAQ
Methodology: WAQ, DCS, WJ, SH, KKE
Project administration: WAQ, KKE
Resources: KKE
Software: WAQ, KKE
Supervision: WJ, SH, KKE
Validation: DCS, WJ, SH, KKE
Writing – original draft: WAQ, WJ, KKE
Writing – review & editing: WAQ, DCS, WJ, SH, KKE

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional data to support the findings of the study.
[DOCX File, 70 KB - publichealth_v12i1e85652_app1.docx]

Checklist 1
STROBE checklist.
[PDF File, 164 KB - publichealth_v12i1e85652_app2.pdf]

References

1. Sacco LN, Lampe JR, Sheikh HZ. The federal status of marijuana and the policy gap with states. : Library of Congress; 2024 URL: https://www.congress.gov/crs-product/IF12270 [accessed 2025-12-30]
2. Key substance use and mental health indicators in the United States: results from the 2023 national survey on drug use and health. : Substance Abuse and Mental Health Services Administration; 2024 URL: https://www.drugsandalcohol.ie/41611/ [accessed 2025-12-30]
3. Ahuja M, Awasthi M, Gim S, et al. Early age of cannabis initiation and its association with suicidal behaviors. Subst Abuse 2022;16:11782218221116731. [doi: 10.1177/11782218221116731] [Medline: 35966616]
4. Blest-Hopley G, O'Neill A, Wilson R, Giampietro V, Bhattacharyya S. Disrupted parahippocampal and midbrain function underlie slower verbal learning in adolescent-onset regular cannabis use. Psychopharmacology (Berl) 2021 May;238(5):1315-1331. [doi: 10.1007/s00213-019-05407-9] [Medline: 31814047]
5. Volkow ND, Swanson JM, Evins AE, et al. Effects of cannabis use on human behavior, including cognition, motivation, and psychosis: a review. JAMA Psychiatry 2016 Mar;73(3):292-297. [doi: 10.1001/jamapsychiatry.2015.3278] [Medline: 26842658]
6. Guttmannova K, Kosterman R, White HR, et al. The association between regular marijuana use and adult mental health outcomes. Drug Alcohol Depend 2017 Oct 1;179:109-116. [doi: 10.1016/j.drugalcdep.2017.06.016] [Medline: 28763778]
7. Montes KS, Witkiewitz K, Pearson MR, Leventhal AM. Alcohol, tobacco, and marijuana expectancies as predictors of substance use initiation in adolescence: a longitudinal examination. Psychol Addict Behav 2019 Feb;33(1):26-34. [doi: 10.1037/adb0000422] [Medline: 30407027]
8. Brown SA. Drug effect expectancies and addictive behavior change. Exp Clin Psychopharmacol 1993;1(1-4):55-67. [doi: 10.1037//1064-1297.1.1-4.55]
9. Patrick ME, Wray-Lake L, Finlay AK, Maggs JL. The long arm of expectancies: adolescent alcohol expectancies predict adult alcohol use. Alcohol Alcohol 2010;45(1):17-24. [doi: 10.1093/alcalc/agp066] [Medline: 19808940]

XSL•FO
RenderX

10. Stewart SA, Copeland AL, Cherry KE. Risk factors for substance use across the lifespan. J Genet Psychol 2023;184(2):145-162. [doi: 10.1080/00221325.2022.2130025] [Medline: 36300651]
11. Smit K, Voogt C, Hiemstra M, Kleinjan M, Otten R, Kuntsche E. Development of alcohol expectancies and early alcohol use in children and adolescents: a systematic review. Clin Psychol Rev 2018 Mar;60:136-146. [doi: 10.1016/j.cpr.2018.02.002] [Medline: 29449029]
12. Kouimtsidis C, Stahl D, West R, Drummond C. How important are positive and negative outcome expectancies in the treatment of addiction: a narrative review of the literature. Drugs Alcohol Today 2014 Aug 26;14(3):137-149. [doi: 10.1108/DAT-11-2013-0051]
13. Fulton HG, Krank MD, Stewart SH. Outcome expectancy liking: a self-generated, self-coded measure predicts adolescent substance use trajectories. Psychol Addict Behav 2012 Dec;26(4):870-879. [doi: 10.1037/a0030354] [Medline: 23088405]
14. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [doi: 10.1037//0033-295x.84.2.191] [Medline: 847061]
15. Bandura A. Social cognitive theory: an agentic perspective. Annu Rev Psychol 2001;52(1):1-26. [doi: 10.1146/annurev.psych.52.1.1] [Medline: 11148297]
16. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory: Pearson; 1986.
17. Marlatt GA. A cognitive-behavioral model of the relapse process. NIDA Res Monogr 1979 Jun(25):191-200. [doi: 10.1037/e497382006-015] [Medline: 117369]
18. West R. Theory of Addiction: Wiley–Blackwell; 2006.
19. Jones CB, Hill ML, Pardini DA, Meier MH. Prevalence and correlates of vaping cannabis in a sample of young adults. Psychol Addict Behav 2016 Dec;30(8):915-921. [doi: 10.1037/adb0000217] [Medline: 27631612]
20. Connor JP, Gullo MJ, Feeney GFX, Young RM. Validation of the Cannabis Expectancy Questionnaire (CEQ) in adult cannabis users in treatment. Drug Alcohol Depend 2011 Jun 1;115(3):167-174. [doi: 10.1016/j.drugalcdep.2010.10.025] [Medline: 21163593]
21. Boden MT, McKay JR, Long WR, Bonn-Miller MO. The effects of cannabis use expectancies on self-initiated cannabis cessation. Addiction 2013 Sep;108(9):1649-1657. [doi: 10.1111/add.12233] [Medline: 23627879]
22. Farris SG, Zvolensky MJ, Boden MT, Bonn-Miller MO. Cannabis use expectancies mediate the relation between depressive symptoms and cannabis use among cannabis-dependent veterans. J Addict Med 2014;8(2):130-136. [doi: 10.1097/ADM.0000000000000010] [Medline: 24365805]
23. Bolts OL, Prince MA, Noel NE. Expectancies that predict cannabis initiation in response to legalization. Subst Use Misuse 2023 Jun 7;58(7):939-946. [doi: 10.1080/10826084.2023.2198596]
24. Anthenien AM, Prince MA, Wallace G, Jenzer T, Neighbors C. Cannabis outcome expectancies, cannabis use motives, and cannabis use among a small sample of frequent using adults. Cannabis 2021;4(1):69-84. [doi: 10.26828/cannabis/2021.01.005] [Medline: 37287995]
25. Torrealday O, Stein LAR, Barnett N, et al. Validation of the marijuana effect expectancy questionnaire-brief. J Child Adolesc Subst Abuse 2008 Aug 20;17(4):1-17. [doi: 10.1080/15470650802231861]
26. Ellickson PL, Tucker JS, Klein DJ, Saner H. Antecedents and outcomes of marijuana use initiation during adolescence. Prev Med 2004 Nov;39(5):976-984. [doi: 10.1016/j.ypmed.2004.04.013] [Medline: 15475032]
27. Flory K, Lynam D, Milich R, Leukefeld C, Clayton R. Early adolescent through young adult alcohol and marijuana use trajectories: early predictors, young adult outcomes, and predictive utility. Dev Psychopathol 2004;16(1):193-213. [doi: 10.1017/s0954579404044475] [Medline: 15115071]
28. Willner P. A view through the gateway: expectancies as a possible pathway from alcohol to cannabis. Addiction 2001 May;96(5):691-703. [doi: 10.1046/j.1360-0443.2001.9656915.x] [Medline: 11331028]
29. Korn L, Haynie DL, Luk JW, Sita K, Simons-Morton BG. Attitudes, subjective norms, and perceived behavioral control associated with age of first use of cannabis among adolescents. J Sch Health 2021 Jan;91(1):50-58. [doi: 10.1111/josh.12977] [Medline: 33216363]
30. Malmberg M, Overbeek G, Vermulst AA, Monshouwer K, Vollebergh WAM, Engels R. The theory of planned behavior: precursors of marijuana use in early adolescence? Drug Alcohol Depend 2012 Jun 1;123(1-3):22-28. [doi: 10.1016/j.drugalcdep.2011.10.011] [Medline: 22056217]
31. Blum RW, Astone NM, Decker MR, Mouli VC. A conceptual framework for early adolescence: a platform for research. Int J Adolesc Med Health 2014;26(3):321-331. [doi: 10.1515/ijamh-2013-0327] [Medline: 24486726]
32. Lane C, Brundage CL, Kreinin T. Why we must invest in early adolescence: early intervention, lasting impact. J Adolesc Health 2017 Oct;61(4 Suppl):S10-S11. [doi: 10.1016/j.jadohealth.2017.07.011] [Medline: 28915985]
33. Christiansen BA, Goldman MS. Alcohol-related expectancies versus demographic/background variables in the prediction of adolescent drinking. J Consult Clin Psychol 1983 Apr;51(2):249-257. [doi: 10.1037//0022-006x.51.2.249] [Medline: 6841769]
34. Shin Y, Miller-Day M. A longitudinal study of parental anti-substance-use socialization for early adolescents' substance-use behaviors. Commun Monogr 2017;84(3):277-297. [doi: 10.1080/03637751.2017.1300821] [Medline: 30595618]
35. Chassin L, Presson C, Rose J, Sherman SJ, Prost J. Parental smoking cessation and adolescent smoking. J Pediatr Psychol 2002 Sep;27(6):485-496. [doi: 10.1093/jpepsy/27.6.485] [Medline: 12177249]

XSL•FO
RenderX

36. Ennett ST, Foshee VA, Bauman KE, et al. The social ecology of adolescent alcohol misuse. Child Dev 2008;79(6):1777-1791. [doi: 10.1111/j.1467-8624.2008.01225.x] [Medline: 19037949]
37. Heerde JA, Bailey JA, Toumbourou JW, Catalano RF. Longitudinal associations between the adolescent family environment and young adult substance use in Australia and the United States. Front Psychiatry 2019;10:821. [doi: 10.3389/fpsyt.2019.00821] [Medline: 31780972]
38. Stattin H, Kerr M. Parental monitoring: a reinterpretation. Child Dev 2000;71(4):1072-1085. [doi: 10.1111/1467-8624.00210] [Medline: 11016567]
39. Marmorstein NR. Investigating associations between caffeinated beverage consumption and later alcohol consumption among early adolescents. Addict Behav 2019 Mar;90:362-368. [doi: 10.1016/j.addbeh.2018.11.033] [Medline: 30522076]
40. Pelham WE, Tapert SF, Gonzalez MR, et al. Parental knowledge/monitoring and adolescent substance use: a causal relationship? Health Psychol 2023 Dec;42(12):913-923. [doi: 10.1037/hea0001245] [Medline: 36355697]
41. Rusby JC, Light JM, Crowley R, Westling E. Influence of parent-youth relationship, parental monitoring, and parent substance use on adolescent substance use onset. J Fam Psychol 2018 Apr;32(3):310-320. [doi: 10.1037/fam0000350] [Medline: 29300096]
42. Dittus PJ, Li J, Verlenden JV, et al. Parental monitoring and risk behaviors and experiences among high school students — youth risk behavior survey, United States, 2021. MMWR Suppl 2021;72(1):37-44. [doi: 10.15585/mmwr.su7201a5]
43. Ryan SM, Jorm AF, Lubman DI. Parenting factors associated with reduced adolescent alcohol use: a systematic review of longitudinal studies. Aust N Z J Psychiatry 2010 Sep;44(9):774-783. [doi: 10.1080/00048674.2010.501759] [Medline: 20815663]
44. Bray JH, Adams GJ, Getz JG, Baer PE. Developmental, family, and ethnic influences on adolescent alcohol usage: a growth curve approach. J Fam Psychol 2001 Jun;15(2):301-314. [doi: 10.1037//0893-3200.15.2.301] [Medline: 11458635]
45. Chaplin TM, Sinha R, Simmons JA, et al. Parent-adolescent conflict interactions and adolescent alcohol use. Addict Behav 2012 May;37(5):605-612. [doi: 10.1016/j.addbeh.2012.01.004] [Medline: 22341765]
46. Hung CC, Yen LL, Wu WC. Association of parents' alcohol use and family interaction with the initiation of alcohol use by sixth graders: a preliminary study in Taiwan. BMC Public Health 2009 Jun 4;9:1-9. [doi: 10.1186/1471-2458-9-172] [Medline: 19497107]
47. Trevino AD, Jamil B, Su J, Aliev F, Elam KK, Lemery-Chalfant K. Alcohol use disorder polygenic risk scores and trajectories of early adolescent externalizing behaviors: examining the role of parenting and family conflict in the racially/ethnically diverse ABCD sample. Behav Genet 2024 Jan;54(1):101-118. [doi: 10.1007/s10519-023-10155-w] [Medline: 37792148]
48. Johnson V, Pandina RJ. Effects of the family environment on adolescent substance use, delinquency, and coping styles. Am J Drug Alcohol Abuse 1991;17(1):71-88. [doi: 10.3109/00952999108992811] [Medline: 2038985]
49. Skeer M, McCormick MC, Normand SLT, Buka SL, Gilman SE. A prospective study of familial conflict, psychological stress, and the development of substance use disorders in adolescence. Drug Alcohol Depend 2009 Sep 1;104(1-2):65-72. [doi: 10.1016/j.drugalcdep.2009.03.017] [Medline: 19446409]
50. Bollen KA, Curran PJ. Latent Curve Models: A Structural Equation Perspective: John Wiley & Sons; 2006. [doi: 10.1002/0471746096]
51. Curran PJ, Obeidat K, Losardo D. Twelve frequently asked questions about growth curve modeling. J Cogn Dev 2010;11(2):121-136. [doi: 10.1080/15248371003699969] [Medline: 21743795]
52. Jung T, Wickrama KAS. An introduction to latent class growth analysis and growth mixture modeling. Social & Personality Psych 2008 Jan;2(1):302-317. [doi: 10.1111/j.1751-9004.2007.00054.x]
53. Nagin DS, Land KC. Age, criminal careers, and population heterogeneity: specification and estimation of a nonparametric, mixed Poisson model. Criminology 1993 Aug;31(3):327-362. [doi: 10.1111/j.1745-9125.1993.tb01133.x]
54. Muthén B, Muthén LK. Integrating person-centered and variable-centered analyses: growth mixture modeling with latent trajectory classes. Alcohol Clin Exp Res 2000 Jun;24(6):882-891. [Medline: 10888079]
55. Barch DM, Albaugh MD, Avenevoli S, et al. Demographic, physical and mental health assessments in the adolescent brain and cognitive development study: rationale and description. Dev Cogn Neurosci 2018 Aug;32:55-66. [doi: 10.1016/j.dcn.2017.10.010] [Medline: 29113758]
56. Volkow ND, Koob GF, Croyle RT, et al. The conception of the ABCD study: from substance use to a broad NIH collaboration. Dev Cogn Neurosci 2018 Aug;32:4-7. [doi: 10.1016/j.dcn.2017.10.002] [Medline: 29051027]
57. Garavan H, Bartsch H, Conway K, et al. Recruiting the ABCD sample: design considerations and procedures. Dev Cogn Neurosci 2018 Aug;32:16-22. [doi: 10.1016/j.dcn.2018.04.004] [Medline: 29703560]
58. About the study. ABCD Study. URL: https://abcdstudy.org/about [accessed 2025-12-30]
59. Auchter AM, Hernandez Mejia M, Heyser CJ, et al. A description of the ABCD organizational structure and communication framework. Dev Cogn Neurosci 2018 Aug;32:8-15. [doi: 10.1016/j.dcn.2018.04.003] [Medline: 29706313]
60. Owens MM, Allgaier N, Hahn S, et al. Multimethod investigation of the neurobiological basis of ADHD symptomatology in children aged 9-10: baseline data from the ABCD study. Transl Psychiatry 2021 Jan 18;11(1):64. [doi: 10.1038/s41398-020-01192-8] [Medline: 33462190]

XSL•FO
RenderX

61. Kristjansson SD, Agrawal A, Lynskey MT, Chassin LA. Marijuana expectancies and relationships with adolescent and adult marijuana use. Drug Alcohol Depend 2012 Nov 1;126(1-2):102-110. [doi: 10.1016/j.drugalcdep.2012.04.024] [Medline: 22682980]
62. Wang Y, Zhang Y, Zhao Z, et al. Multiple discrimination and substance use intention in late childhood: findings from the adolescent brain cognitive development (ABCD) study. J Adolesc Health 2024 Jun;74(6):1217-1224. [doi: 10.1016/j.jadohealth.2024.01.028] [Medline: 38483374]
63. Bhatia D, Lewis B, Farrior H, Moore A, Nixon SJ. Substance familiarity in middle childhood and adolescent substance use. Drug Alcohol Depend 2023 Sep 1;250:110892. [doi: 10.1016/j.drugalcdep.2023.110892] [Medline: 37473699]
64. Martz ME, Heitzeg MM, Lisdahl KM, et al. Individual-, peer-, and parent-level substance use-related factors among 9- and 10-year-olds from the ABCD Study: prevalence rates and sociodemographic differences. Drug Alcohol Depend Rep 2022 Jun;3:100037. [doi: 10.1016/j.dadr.2022.100037] [Medline: 35445220]
65. Dishion TJ, Nelson SE, Kavanagh K. The family check-up with high-risk young adolescents: preventing early-onset substance use by parent monitoring. Behav Ther 2003;34(4):553-571. [doi: 10.1016/S0005-7894(03)80035-7]
66. Dishion TJ, Kavanagh K. Intervening in Adolescent Problem Behavior: A Family-Centered Approach: Guilford Press; 2003.
67. Jackson KM, Barnett NP, Colby SM, Rogers ML. The prospective association between sipping alcohol by the sixth grade and later substance use. J Stud Alcohol Drugs 2015 Mar;76(2):212-221. [doi: 10.15288/jsad.2015.76.212] [Medline: 25785796]
68. Jackson KM, Roberts ME, Colby SM, Barnett NP, Abar CC, Merrill JE. Willingness to drink as a function of peer offers and peer norms in early adolescence. J Stud Alcohol Drugs 2014 May;75(3):404-414. [doi: 10.15288/jsad.2014.75.404] [Medline: 24766752]
69. Karoly HC, Callahan T, Schmiege SJ, Ewing SWF. Evaluating the hispanic paradox in the context of adolescent risky sexual behavior: the role of parent monitoring. J Pediatr Psychol 2016 May;41(4):429-440. [doi: 10.1093/jpepsy/jsv039] [Medline: 25972373]
70. Handschuh C, Mokkink LB, Smaldone A. Perceived parental monitoring: a systematic review of monitoring instruments. J Nurs Meas 2020 Dec 1;28(3):E253-E292. [doi: 10.1891/JNM-D-19-00045] [Medline: 33199481]
71. Racz SJ, McMahon RJ. The relationship between parental knowledge and monitoring and child and adolescent conduct problems: a 10-year update. Clin Child Fam Psychol Rev 2011 Dec;14(4):377-398. [doi: 10.1007/s10567-011-0099-y] [Medline: 22086648]
72. Gong W, Rolls ET, Du J, Feng J, Cheng W. Brain structure is linked to the association between family environment and behavioral problems in children in the ABCD study. Nat Commun 2021 Jun 18;12(1):3769. [doi: 10.1038/s41467-021-23994-0] [Medline: 34145259]
73. Sanchez M, Gonzalez MR, Fernandez A, Barton A, Diaz V, Wang W. Sociocultural influences on alcohol expectancies in early adolescence: findings from the ABCD study. Health Psychol 2023 Dec;42(12):842-855. [doi: 10.1037/hea0001290] [Medline: 37227824]
74. Bristol SC, Johnson ME, Thompson WK, et al. Prospective associations of family conflict with alcohol expectancies in the adolescent brain cognitive development study: effects of race and ethnicity. Front Psychiatry 2024;15:1250351. [doi: 10.3389/fpsyt.2024.1250351] [Medline: 38550535]
75. Moos RH, Moos BS. Family Environment Scale Manual: Development, Applications, Research: Consulting Psychologists Press; 1994.
76. Jia L, Wei Z, Liu B, et al. The developmental and dynamic relationship between psychopathology and psychotic-like experiences trajectories in children and adolescents. J Affect Disord 2025 Jan 15;369:95-102. [doi: 10.1016/j.jad.2024.09.100] [Medline: 39306006]
77. Roffman JL, Sipahi ED, Dowling KF, et al. Association of adverse prenatal exposure burden with child psychopathology in the Adolescent Brain Cognitive Development (ABCD) Study. PLoS ONE 2021;16(4):e0250235. [doi: 10.1371/journal.pone.0250235] [Medline: 33909652]
78. Dooley N, Clarke M, Cotter D, Cannon M. Birth weight and childhood psychopathology in the ABCD cohort: association is strongest for attention problems and is moderated by sex. Res Child Adolesc Psychopathol 2022 May;50(5):563-575. [doi: 10.1007/s10802-021-00859-0] [Medline: 35072847]
79. Isaiah A, Uddin S, Ernst T, Cloak C, Li D, Chang L. Cognitive and behavioral outcomes of snoring among adolescents. JAMA Netw Open 2024 Nov 4;7(11):e2444057. [doi: 10.1001/jamanetworkopen.2024.44057] [Medline: 39514229]
80. Achenbach TM, Rescorla LA. Manual for ASEBA School-Age Forms & Profiles: University of Vermont, Research Center for Children, Youth & Families; 2001.
81. Barch DM, Albaugh MD, Baskin-Sommers A, et al. Demographic and mental health assessments in the adolescent brain and cognitive development study: updates and age-related trajectories. Dev Cogn Neurosci 2021 Dec;52:101031. [doi: 10.1016/j.dcn.2021.101031] [Medline: 34742018]
82. Karcher NR, Klaunig MJ, Elsayed NM, Taylor RL, Jay SY, Schiffman J. Understanding associations between race/ethnicity, experiences of discrimination, and psychotic-like experiences in middle childhood. Journal of the American Academy of Child & Adolescent Psychiatry 2022 Oct;61(10):1262-1272. [doi: 10.1016/j.jaac.2022.03.025]

XSL•FO
RenderX

83. Borrell LN, Elhawary JR, Fuentes-Afflick E, et al. Race and genetic ancestry in medicine - a time for reckoning with racism. N Engl J Med 2021 Feb 4;384(5):474-480. [doi: 10.1056/NEJMms2029562] [Medline: 33406325]
84. Dubow EF, Boxer P, Huesmann LR. Long-term effects of parents' education on children's educational and occupational success: mediation by family interactions, child aggression, and teenage aspirations. Merrill Palmer Q (Wayne State Univ Press) 2009 Jul;55(3):224-249. [doi: 10.1353/mpq.0.0030] [Medline: 20390050]
85. Davis-Kean PE, Tang S, Waters NE. Parent education attainment and parenting. In: Bornstein MH, editor. Handbook of Parenting: Routledge; 2019:400-420. [doi: 10.4324/9780429401459-12]
86. Waters NE, Ahmed SF, Tang S, Morrison FJ, Davis-Kean PE. Pathways from socioeconomic status to early academic achievement: the role of specific executive functions. Early Child Res Q 2021;54:321-331. [doi: 10.1016/j.ecresq.2020.09.008] [Medline: 33519062]
87. Collins LM, Horn JL. Best Methods for the Analysis of Change: Recent Advances, Unanswered Questions, Future Directions: American Psychological Association; 1991. [doi: 10.1037/10099-000]
88. Ferrer E, Ghisletta P. Methodological and analytical issues in the psychology of aging. In: Schaie KW, Willis SL, editors. Handbook of the Psychology of Aging: Elsevier; 2011:25-39. [doi: 10.1016/B978-0-12-380882-0.00002-4]
89. Grimm KJ, Ram N, Estabrook R. Growth Modeling: Structural Equation and Multilevel Modeling Approaches: Guilford Publications; 2016.
90. Nylund KL, Asparouhov T, Muthén BO. Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. Struct Equ Modeling 2007 Oct 23;14(4):535-569. [doi: 10.1080/10705510701575396]
91. Ram N, Grimm KJ. Growth mixture modeling: a method for identifying differences in longitudinal change among unobserved groups. Int J Behav Dev 2009;33(6):565-576. [doi: 10.1177/0165025409343765] [Medline: 23885133]
92. Asparouhov T, Muthén B. Auxiliary variables in mixture modeling: three-step approaches using M plus. Struct Equ Modeling 2014 Jul 3;21(3):329-341. [doi: 10.1080/10705511.2014.915181]
93. Vermunt JK. Latent class modeling with covariates: two improved three-step approaches. Polit anal 2010;18(4):450-469. [doi: 10.1093/pan/mpq025]
94. Muthén B, Muthén L. Mplus: a general latent variable modeling program. : Statmodel; 2019 URL: https://www.statmodel.com/download/Mplus-A%20General%20Latent%20Variable%20Modeling%20Program.pdf [accessed 2025-12-30]
95. Stekhoven DJ, Bühlmann P. MissForest--non-parametric missing value imputation for mixed-type data. Bioinformatics 2012 Jan 1;28(1):112-118. [doi: 10.1093/bioinformatics/btr597] [Medline: 22039212]
96. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Int J Surg 2014 Dec;12(12):1495-1499. [doi: 10.1016/j.ijsu.2014.07.013] [Medline: 25046131]
97. Steinberg LD. Age of Opportunity: Lessons from the New Science of Adolescence: Houghton Mifflin Harcourt; 2014.
98. D'Amico EJ, Rodriguez A, Tucker JS, et al. Early and late adolescent factors that predict co-use of cannabis with alcohol and tobacco in young adulthood. Prev Sci 2020 May;21(4):530-544. [doi: 10.1007/s11121-020-01086-7] [Medline: 31960260]
99. Henry KL, Oetting ER, Slater MD. The role of attachment to family, school, and peers in adolescents' use of alcohol: a longitudinal study of within-person and between-persons effects. J Couns Psychol 2009 Oct 1;56(4):564-572. [doi: 10.1037/a0017041] [Medline: 20228876]
100. Brumback T, Thompson W, Cummins K, Brown S, Tapert S. Psychosocial predictors of substance use in adolescents and young adults: longitudinal risk and protective factors. Addict Behav 2021 Oct;121:106985. [doi: 10.1016/j.addbeh.2021.106985] [Medline: 34087768]
101. Masyn KE. Latent class analysis and finite mixture modeling. In: Little TD, editor. The Oxford Handbook of Quantitative Methods: Statistical Analysis: Oxford University Press; 2013:551-611. [doi: 10.1093/oxfordhb/9780199934898.013.0025]
102. Steinberg L. Cognitive and affective development in adolescence. Trends Cogn Sci (Regul Ed) 2005 Feb;9(2):69-74. [doi: 10.1016/j.tics.2004.12.005]
103. Simons-Morton B, Chen RS. Over time relationships between early adolescent and peer substance use. Addict Behav 2006 Jul;31(7):1211-1223. [doi: 10.1016/j.addbeh.2005.09.006] [Medline: 16229958]
104. Steinberg L, Morris AS. Adolescent development. Annu Rev Psychol 2001;52(1):83-110. [doi: 10.1146/annurev.psych.52.1.83] [Medline: 11148300]
105. Baumrind D. The influence of parenting style on adolescent competence and substance use. J Early Adolesc 1991 Feb;11(1):56-95. [doi: 10.1177/0272431691111004]
106. Donaldson CD, Handren LM, Crano WD. The enduring impact of parents' monitoring, warmth, expectancies, and alcohol use on their children's future binge drinking and arrests: a longitudinal analysis. Prev Sci 2016 Jul;17(5):606-614. [doi: 10.1007/s11121-016-0656-1] [Medline: 27178008]
107. Roche KM, Ahmed S, Blum RW. Enduring consequences of parenting for risk behaviors from adolescence into early adulthood. Soc Sci Med 2008 May;66(9):2023-2034. [doi: 10.1016/j.socscimed.2008.01.009]
108. Yang H, Stanton B, Li X, Cottrel L, Galbraith J, Kaljee L. Dynamic association between parental monitoring and communication and adolescent risk involvement among African-American adolescents. J Natl Med Assoc 2007 May;99(5):517-524. [Medline: 17534009]

XSL•FO
RenderX

109. Dougherty DM, Lake SL, Mathias CW, et al. Behavioral impulsivity and risk-taking trajectories across early adolescence in youths with and without family histories of alcohol and other drug use disorders. Alcohol Clin Exp Res 2015 Aug;39(8):1501-1509. [doi: 10.1111/acer.12787] [Medline: 26173617]
110. Latendresse SJ, Rose RJ, Viken RJ, Pulkkinen L, Kaprio J, Dick DM. Parenting mechanisms in links between parents' and adolescents' alcohol use behaviors. Alcohol Clin Exp Res 2008 Feb;32(2):322-330. [doi: 10.1111/j.1530-0277.2007.00583.x] [Medline: 18162066]
111. Repetti RL, Taylor SE, Seeman TE. Risky families: family social environments and the mental and physical health of offspring. Psychol Bull 2002 Mar;128(2):330-366. [Medline: 11931522]
112. Bronfenbrenner U. The Ecology of Human Development: Experiments by Nature and Design: Harvard University Press; 1981.
113. Hirschi T. Causes of Delinquency: Routledge; 2017. [doi: 10.4324/9781315081649]
114. Masten AS, Cicchetti D. Developmental cascades. Dev Psychopathol 2010 Aug;22(3):491-495. [doi: 10.1017/S0954579410000222] [Medline: 20576173]
115. Kokotovič KO, Pšunder M, Kirbiš A. Cannabis use and parenting practices among young people: the impact of parenting styles, parental cannabis-specific rules, and parental cannabis use. Int J Environ Res Public Health 2022 Jul 1;19(13):8080. [doi: 10.3390/ijerph19138080] [Medline: 35805740]
116. Federal partners. ABCD Study. URL: https://abcdstudy.org/federal-partners.html [accessed 2025-12-30]
117. Complete ABCD roster. ABCD Study. URL: https://abcdstudy.org/wp-content/uploads/2019/04/Consortium_Members.pdf [accessed 2025-12-30]
118. Adolescent Brain Cognitive Development Study (ABCD) - annual release 5.1 #2313. NIMH Data Archive. URL: https://nda.nih.gov/study.html?id=2313 [accessed 2025-12-30]
119. Data sharing. ABCD Study. URL: https://abcdstudy.org/scientists/data-sharing/ [accessed 2025-12-30]
120. Raw data. NIMH Data Archive. URL: https://nda.nih.gov/edit_collection.html?id=2573 [accessed 2025-12-30]
121. Creating an NDA study. NIMH Data Archive. URL: https://nda.nih.gov/nda/tutorials/creating_an_nda_study [accessed 2025-12-30]

Abbreviations

ABCD Study: Adolescent Brain Cognitive Development Study
aBIC: sample size–adjusted Bayesian information criterion
AIC: Akaike information criterion
aOR: adjusted odds ratio
BIC: Bayesian information criterion
IRB: Institutional Review Board
LCGA: latent class growth analysis
LGCM: latent growth curve model
LMR-aLRT: Lo-Mendell-Rubin adjusted likelihood ratio test
MEEQ-B: Marijuana Effect Expectancy Questionnaire-Brief
NIH: National Institutes of Health
PRIME: plans, responses, impulses, motives, and evaluations

Edited by A Mavragani, T Sanchez; submitted 14.Oct.2025; peer-reviewed by MI Gallegos, VT Le; revised version received 25.Nov.2025; accepted 17.Dec.2025; published 09.Jan.2026.

Please cite as:
Qin WA, Seo DC, Jacobs W, Huang S, Elam KK
Developmental Trajectories of Positive Expectancies of Cannabis Use Effects Among Early Adolescents: Longitudinal Observational Study Using Latent Class Growth Analysis
JMIR Public Health Surveill 2026;12:e85652
URL: https://publichealth.jmir.org/2026/1/e85652
doi:10.2196/85652

© Weisiyu Abraham Qin, Dong-Chul Seo, Wura Jacobs, Sijia Huang, Kit K Elam. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 9.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly

XSL•FO
RenderX

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

XSL•FO
RenderX

Original Paper

Longitudinal Analysis of Variations in Daily Step Counts and Long-Term Implications of COVID-19 Waves and Restriction Phases in Qatar's Step Into Health Program: Mixed Methods Study

Lina Majed[1], PhD; Suzan Sayegh[2,3], MPH; Feriel Dalansi[4,5], MSc; Abdulla Saeed Al-Mohannadi[6], MPH; Marco Cardinale[4,7,8], PhD; Abdulaziz Farooq[4], PhD

[1]College of Health and Life Sciences, Hamad bin Khalifa University, Doha, Qatar

[2]Faculty of Health Sciences, Modern University for Business and Science, Beirut, Lebanon

[3]Shift Wellness Institute, Vancouver, WA, United States

[4]Research and Scientific Support, Aspetar, Orthopaedic and Sports Medicine Hospital, Doha, Qatar

[5]Department of Orthopedic Surgery and Sports Medicine, Amsterdam Movement Sciences, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands

[6]World Innovation Summit for Health (WISH), Qatar Foundation, Doha, Qatar

[7]Institute of Sport Exercise and Health, University College London, London, United Kingdom

[8]Department of Sport Exercise and Rehabilitation, Northumbria University, Newcastle Upon Tyne, United Kingdom

Corresponding Author:
Lina Majed, PhD
College of Health and Life Sciences
Hamad bin Khalifa University
Education City, Ar Rayyan, Penrose House (LAS Building)
Doha, P.O. Box 5825
Qatar
Phone: 974 66999851
Email: limajed@hbku.edu.qa

Abstract

Background: Public health restrictive measures adopted during the COVID-19 pandemic led to significant changes in lifestyles. Global declines in physical activity (PA) and increases in sedentary behavior were noted. These trends were observed within different regions of the world, pointing toward potential long-term implications for PA behaviors.

Objective: This mixed methods study aims to assess variations in daily step counts in Qatar using device-driven data throughout all 3 COVID-19 waves (February 2020 to February 2023) compared with a full pre–COVID-19 year. In-depth interviews were further conducted with randomly selected participants to gain insights into determinants, perceptions, and barriers of PA during the pandemic.

Methods: A total of 362 participants (60/362, 16.6% female) from the Step Into Health community-based program reported daily step counts using pedometers (170/362, 47%) or a mobile phone app (192/362, 53%). Linear mixed models examined changes in daily step counts across 19 phases of implementation and lifting of restrictions. Overall, 9 participants also completed semistructured interviews that were analyzed thematically and phenomenologically. Triangulation of quantitative and qualitative data was applied to interpret convergences and divergences between device-measured activity patterns and lived experiences.

Results: Significant declines in daily step counts (ie, from 689 to 1013 steps) were observed at the onset of each wave ($P<.001$), were especially marked at wave 2, and were followed by a recovery of step count following the lifting of restrictions at each wave (ie, increase of 609 to 1147 steps). Different patterns of change in step count emerged within sex ($P=.03$), age ($P=.03$), and BMI ($P=.01$) groups, where larger variations were seen among male individuals, pedometer users, and normal-weight participants. Qualitative themes (ie, disrupted routines, reliance on home-based exercise, and media influence) contextualized these patterns and explained subgroup differences.

Conclusions: The largest drops in daily step count coincided with increased case severity and Ramadan. Integration of quantitative and qualitative findings showed that declines in activity were shaped not only by restrictions but also by fear, motivation, and

XSL•FO
RenderX

contextual factors. These results underscore the importance of designing interventions that encourage outdoor activity and provide reliable social media–based guidance during public health crises.

(JMIR Public Health Surveill 2026;12:e76860) doi:10.2196/76860

KEYWORDS

community-based intervention; coronavirus disease; COVID-19; interviews; lockdown; longitudinal data; mixed methods; mobile app; pandemic; pedometer; step count

Introduction

COVID-19 was initially detected in December 2019 in Wuhan, China. This disease is caused by SARS-CoV-2 [1]. On March 11, 2020, the World Health Organization declared COVID-19 a global pandemic [2]. The first confirmed case of COVID-19 in Qatar was reported on February 27, 2020 [3]. Since then, progressive restrictions have been implemented in Qatar and worldwide as preventive measures to reduce the spread of the virus. These measures targeted various sectors, including business and leisure, transportation, outdoor and professional sports, education, health services, and social gatherings.

Physical activity (PA) is a critical determinant of health, contributing to the prevention of chronic diseases and the promotion of overall well-being [4]. According to the World Health Organization, adults aged 18-64 years are recommended to engage in at least 150-300 minutes per week of moderate-intensity PA, or a minimum of 75-150 minutes per week of vigorous-intensity PA, or a combination of both for substantial health benefits [5]. However, fulfilling these recommendations became particularly challenging during COVID-19 restrictions, although many countries have published targeted PA guidelines for times of pandemics, outbreaks, and home quarantine (eg, Qatar National PA Guidelines [6]). A global decline in total PA levels was observed [7-11] including vigorous-, moderate-, and light-intensity PA [8,9,12-14]. This was accompanied by a significant increase in total sedentary time [15,16].

Adequate and consistent PA has been suggested to enhance immune responses against respiratory viral infections, emphasizing its relevance during the pandemic [17]. Reductions in PA, along with increased sedentary behavior, can lead to adverse health outcomes such as diminished physical fitness and an impaired immune system response, negatively affecting quality of life [4].

In Qatar, a study examining PA levels before and during the first lockdown phase in 2020 revealed significant reductions in walking and both vigorous and moderate PA levels and increases in sedentary behavior [18]. A general decline in life satisfaction was also noted and reflected the profound physical and mental impact of the imposed restrictions [18]. Another survey-based study on Qatar's adult population confirmed increases in sedentary time and decreases in exercise time, which were associated with gains in body weight [19]. Both studies presented a few limitations, such as the reliance on subjective self-reported data and the lack of comparison with other periods of reduced PA in Qatar (eg, summer months or Ramadan). To address the latter, a recent study explored daily variations in step count before, during, and after the first COVID-19 wave in Qatar, covering a period of approximately 1 year [20]. Authors showed that the significant declines in daily step count following restrictions persisted after the lifting of restrictive measures, affecting mostly those who were physically active at baseline (ie, men, normal weight, and older individuals) [20]. The COVID-19 restrictions, which aimed to curb the spread of the virus, have continued disrupting normal life patterns in Qatar until November 2022. However, none of the studies from Qatar accounted for the second and third COVID-19 waves that resulted in potential long-term implications to PA. Such longitudinal data could provide better insights into the dynamics of PA behaviors in times of extended public health safety measures, particularly in regions where there is a paucity of data on PA changes during pandemics.

Lockdowns in the Arabian Gulf Peninsula and the Middle East and North Africa regions inadvertently decreased opportunities for regular PA and exacerbated sedentary lifestyles [21,22]. Although restrictions have been lifted, the long-term effects of the COVID-19 pandemic on physical and mental health are still emerging [23,24]. While much has been done on the immediate impacts, long-term behavioral shifts present a critical gap, particularly in regions like the Gulf Peninsula region, with its unique environmental and cultural characteristics [25].

Given the challenges and potential biases present in assessing PA through self-reported data and the rather short-term nature of most studies in the region, this study seeks to explore the impact of all 3 waves of COVID-19–related restrictions on device-measured daily step counts in Qatar. Understanding these disruptions over a period of 3 years is vital for shaping public health strategies that promote active lifestyles and address the ongoing post–COVID-19 challenges [26]. Using a mixed method approach, this study further aims to assess associated barriers, behavior changes, attitudes, and perceptions during the pandemic, providing valuable insights into how restrictions have influenced PA patterns in Qatar and potentially the region.

Methods

Ethical Considerations

All participants in the "Step Into Health" program (SIH) provided a disclaimer at enrollment permitting the use of their step count data for research and program evaluation. Only limited, deidentified data were extracted from the SIH database for this study. For the qualitative part, interview participants provided separate written informed consent, including permission to audio-record the interview and to use anonymized quotations in publications. No compensation or incentives were provided to participants for their involvement in this research.

XSL•FO
RenderX

Privacy and confidentiality were strictly maintained in accordance with institutional review board approval (E202104021).

Study Design and Population

This is a retrospective longitudinal cohort study, which aimed to explore the impact of all waves of COVID-19–related restrictions in Qatar on PA levels using mixed methods. Participants included in this research were members of the SIH, a pedometer-based community intervention designed to encourage residents of Qatar to walk 10,000 steps per day [20,27-30]. SIH is the largest PA community-based intervention in Qatar and is still ongoing since 2012. The program was initially publicized nationally through outreach advertisement campaigns targeting different settings (workplaces, campuses, and malls). Registration to the SIH was done on a voluntary basis through a dedicated online platform [20,27-30]. Upon registration, participants received a free pocket-sized pedometer (Omron HJ-324U, Omron Healthcare Co, Ltd) or a mobile phone app (Android and iOS operating systems; SIH; version 3.0.3; Aspire Zone), both of which were linked to a web database that recorded step counts [20,27-30].

Male and female participants of all nationalities (as registered in the SIH database) were included in the analysis. From the initial pool of participants that were actively recording their step count during the studied time frame (January 2019 to November 2022), a total of 3562 participants were retained after excluding those aged 18 years or younger or those presenting invalid step counts (ie, missing data or less than 500 or more than 60,000 steps per day). A further filtering step excluded participants presenting less than 3 activity days per week [31] or 5 days per month [32] over the entire studied period. After applying the last step, 362 participants were retained for the quantitative analysis.

Qualitative Research

From the retained sample, 9 participants were invited to participate in in-depth interviews. A random selection process was applied (ie, every 40th individual in the SIH database) to reduce researcher selection bias and give all participants an equal chance of recruitment. Although not purposive, this process resulted in a sample with diversity in sex, age, and activity behaviors, which allowed a range of perspectives to be captured. Data collection continued until thematic saturation was observed, with no new themes and codes emerging after the ninth interview. At that point, the information became repetitive across participants, indicating that the sample size was sufficient to capture the diversity of perspectives. Participants were contacted individually by phone to confirm their participation. Interviews occurred between September 2021 and November 2021.

Study Measures

Quantitative: Daily Step Count

The mobile phone app reads steps counted by the phone's built-in sensors and transfers the data to the cloud. Pedometer users were asked to upload their data daily or as frequently as possible. More information on the program is cited elsewhere [20,27-30].

Data extracted from the SIH web database included the available active members who had uploaded their steps before, during, and after the COVID-19–related restrictions phases. Demographic data, self-reported upon registration, were extracted from participants' SIH profiles, which included age, sex, and nationality, as well as body height and mass. BMI was calculated as the ratio between mass (kg) and height squared (m^2). Participants were then grouped according to their sex (ie, male and female), age (ie, <50 years and ≥50 years), nationality (ie, Qatari and non-Qatari), BMI classification (ie, normal weight with values <25, overweight with values between 25 and <30, and obesity with values of ≥30) and the type of device used (ie, pedometer or mobile phone app).

Qualitative: In-Depth Interviews

This study used qualitative methods to explore participants' experiences, attitudes, and perceptions regarding PA during the COVID-19 pandemic in Qatar. A total of 9 in-depth interviews were conducted remotely via online platforms or phone calls with participants of the SIH program to examine behavioral changes, barriers, and motivators related to PA during all restrictive waves of the COVID-19 pandemic. The sample included 5 male and 4 female individuals, selected to ensure diverse perspectives.

All interviews were conducted in English or Arabic, based on the participant's preference. Interviewers were fluent in both languages, and interview questions were translated into Arabic for Arabic-speaking participants. Prior to the interviews, interviewers received comprehensive training on the interview protocol to ensure consistency and clarity in question delivery.

The interviews were semistructured, focusing on key themes such as the importance of PA, changes in exercise habits prepandemic and postpandemic, the impact of COVID-19 restrictions on PA, barriers to staying active, and motivators for exercise. Participants were also asked about their experiences with home-based exercise routines, the role of the media in shaping their activity choices, and their intentions regarding future PA post restrictions.

Each interview lasted between 30 and 60 minutes. Interviews were conducted virtually or in-person, adhering to the appropriate COVID-19 safety protocols for in-person sessions. All interviews were audio-recorded for quality assurance, and transcripts were created from the recordings. Deidentified transcripts were stored for record-keeping and coding purposes.

Timeline of the 3 COVID-19 Waves in Qatar

Table 1 details the main COVID-19 phases considered during the 3 waves relating mainly to the restrictions that were imposed and lifted by the Qatar's Ministry of Public Health. A pre–COVID-19 period of over a full year (T0) is used as a reference to annual changes in daily step count. The data presented in Table 1 were compiled from the Ministry of Public Health website and other reliable news sources from the written press.

XSL•FO
RenderX

Table 1. Timeline of implementation and lifting of COVID-19 restrictions during the 3 COVID-19 waves in Qatar. T0 represents a pre–COVID-19 reference phase of more than 1 year.

Phase	Start date	End date	Description
T0	January 1, 2019	February 28, 2020	Pre–COVID-19
T1	February 29, 2020	March 11, 2020	First confirmed case in Qatar until the start of lockdown
T2	March 12, 2020	June 14, 2020	Wave 1: Qatar begins to implement lockdowns
T3	June 15, 2020	June 30, 2020	Wave 1: Phase 1 gradual lifting of restrictions
T4	July 1, 2020	July 31, 2020	Wave 1: Phase 2 gradual lifting of restrictions
T5	August 1, 2020	August 31, 2020	Wave 1: Phase 3 gradual lifting of restrictions
T6	September 1, 2020	February 3, 2021	Wave 1: Phase 4 gradual lifting of restrictions
T7	February 4, 2021	May 27, 2021	Wave 2: Qatar to reimpose COVID-19 restrictions
T8	May 28, 2021	June 17, 2021	Wave 2: Phase 1 gradual lifting of restrictions
T9	June 18, 2021	July 8, 2021	Wave 2: Phase 2 gradual lifting of restrictions
T10	July 9, 2021	August 5, 2021	Wave 2: Phase 3 gradual lifting of restrictions
T11	August 6, 2021	October 2, 2021	Wave 2: Phase 3.2 gradual lifting of restrictions
T12	October 3, 2021	February 1, 2022	Wave 2: Phase 4 gradual lifting of restrictions
T13	February 2, 2022	April 1, 2022	Wave 3: Qatar reimposes COVID-19 restrictions
T14	April 2, 2022	May 20, 2022	Wave 3: gradual lifting of restrictions: 100% employees back to work
T15	May 21, 2022	July 6, 2022	Wave 3: gradual lifting of restrictions: mask required in health care
T16	July 7, 2022	August 31, 2022	Wave 3: reimpose mask in all settings
T17	September 1, 2022	October 31, 2022	Wave 3: gradual lifting of restrictions: mask required in health care
T18	November 1, 2022	February 28, 2023	Wave 3: gradual lifting of restrictions: Ehteraz[a] no longer required

[a]Ehteraz is the mandatory mobile app used in Qatar to track the population's health status during COVID-19.

Statistical Analysis

Quantitative Analysis

Data were analyzed using the SPSS software (version 21.0; IBM Corp). A descriptive analysis presented continuous variables (eg, age and BMI) and daily step count as mean, SD, and 95% upper and lower CIs. Distribution (count and frequencies) was used to present participants' characteristics on categorical variables like nationality and sex. Time series 7-day moving average on daily step count was computed to plot the trend of PA across the study period stratified by age, sex, BMI status, and nationality. Linear mixed models were used to examine changes in daily step counts over all the 19 studied time phases (Table 1), based on independent factors such as sex, age, BMI classification, nationality, and device. All factors and their interaction with factor time were included in the linear mixed models. Pairwise comparisons were used to examine significant changes in step count between the time phases studied. By adding subject ID as a random effect with a random intercept, we accounted for individual-level variation and within-subject correlation. An unstructured covariance structure was found to be the best fit.

Qualitative Analysis

The qualitative data from the in-depth interviews were analyzed thematically to identify key patterns and insights related to participants' experiences, attitudes, and perceptions of PA during the COVID-19 pandemic. Transcripts of the interviews were reviewed and coded inductively, allowing for the emergence of recurring themes associated with changes in PA behaviors, perceived barriers, and motivators, and the broader psychological and emotional impacts of the pandemic.

The qualitative analysis was conducted by 1 researcher (SS), who has training and experience in qualitative research. A thematic analysis was carried out following Braun and Clarke's [33] 6-step framework (ie, familiarization with data, coding, generating themes, reviewing themes, defining/naming themes, and reporting). This inductive process enabled the identification of recurring patterns. A phenomenological approach was then used to interpret the lived experiences of participants, focusing on the subjective meaning they attributed to disruptions and adaptations in their PA routines during the pandemic. Although a single researcher conducted the coding, credibility was strengthened through repeated immersion in the transcripts, detailed documentation of coding decisions, and reflective discussions with the wider research team to ensure interpretations were consistent and trustworthy.

To integrate the qualitative and quantitative findings, methodological triangulation was used. Quantitative step count data and qualitative interview data were first analyzed independently, then compared to identify areas of convergence and divergence. Data triangulation was also incorporated by drawing on step count measurements from 2 different devices (pedometers and mobile apps) alongside qualitative accounts.

XSL•FO
RenderX

This approach enhanced the depth and validity of the findings by linking device-measured activity with participants' lived experiences and contextual factors.

Results

Quantitative Results

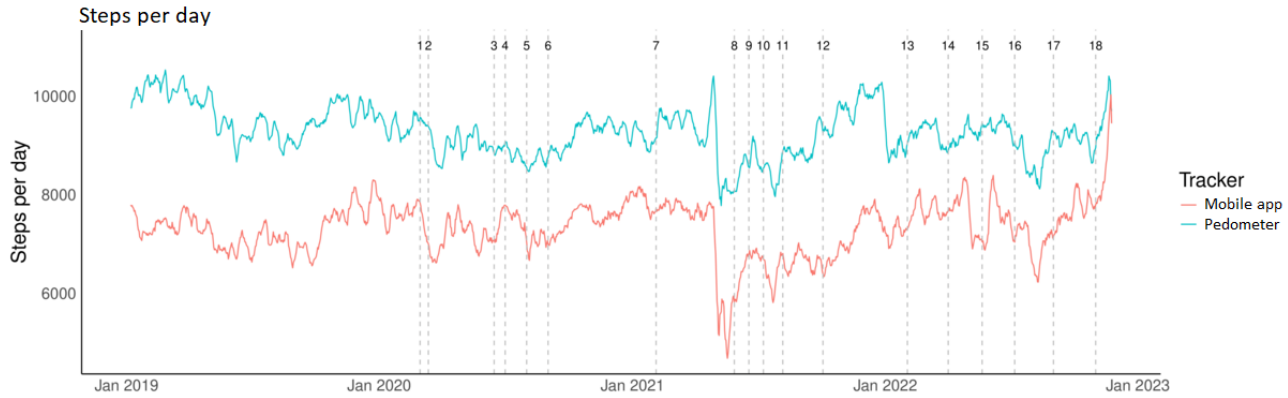
Population's Characteristics

A total of 362 participants, including 60 (16.6%) female and 302 (83.4%) male individuals, met the inclusion criteria, demonstrating consistent tracking of their daily steps throughout the entire pandemic period. On average, participants were aged 47.4 (SD 9.0) years and presented a BMI of 27.9 (SD 4.4) kg/m^2. About a third of the studied population (n=137, 37.8%) was aged 50 years or older, while 75.1% (272/362) were classified either with overweight (n=184, 50.8%) or obesity (n=88, 24.3%). Qatari nationals represented 30.7% (n=111) of the studied population, which reflects a relatively good representation compared to the overall population of Qatar. From the entire sample, 170 (47%) participants were pedometer users and 192 (53%) used the mobile app (ie, n=58, 16% Android and n=134, 37% Apple iPhone) to track their activity level.

Effect of COVID-19 Phases on Step Count

The analyses revealed a significant effect of time ($F_{18, 5027.03}$=5.901; P<.001), device ($F_{1, 1598.03}$=6.275; P=.01), and an interaction between device and time ($F_{18, 5157.67}$=1.70; P=.03) on step count (Figure 1). In comparison to the pre–COVID-19 period, significant decreases in step count ranging from 689 to 1013 steps were observed following implementation of restrictive measures for each of the 3 COVID-19 waves (Figure 1). Specifically, restrictions reimposed at the start of wave 2 brought about the largest decreases in daily step count. With the lifting of restrictions, significant increases in step count were found at the end of each wave (ie, between 609 and 1147 steps; Figure 1). On average, pedometer users presented 690 additional daily steps as compared to those using the mobile app (P=.01; Figure 1). Main patterns of change in step count with time were comparable between pedometer and mobile users, with some interactions noted. For instance, after September 2022 (ie, T17 and T18, end of wave 3) the increase in step count among mobile users was larger than that observed in pedometer users with the final lifting of restrictions (Figure 1). In the following sections, data from pedometer and mobile users are presented separately.

Figure 1. The 7-day moving average of daily step count for pedometer and mobile app users within the studied period covering 1 year before COVID-19 (T0: January 1, 2019, to February 28, 2020) and the 3 COVID-19 waves (T1-T18: February 29, 2020, to February 28, 2023). Details on the 19 time points reflecting the implementation and lifting of restrictions are found in Table 1.



Effect of COVID-19 Phases on Step Count by Sex and Nationality Groups

Although the main effect of sex ($F_{1, 360.74}$=0.125; P=.72) and its interaction with device ($F_{1, 2384.61}$=0.888; P=.35) or time ($F_{18, 5167.25}$=1.39; P=.12) were not statistically significant, post hoc comparisons revealed noticeable significant sex differences, mostly evident for mobile users (Figure 2). For mobile users (Figure 2), female individuals recorded an average of 1764 fewer daily steps as compared to male individuals at the pre–COVID-19 period (P=.03). The only other large difference between male and female groups was detected at the lifting of restrictions in wave 3, which corresponded to the return to work of 100% of the employees in Qatar (T14, 2227 less steps in female compared to male individuals; P=.01; Figure 2). However, step count for female mobile users was not significantly affected by the first or second waves of COVID-19, where even significant increases in values were recorded at the end of wave 1 (+1059 steps; P=.046) and at the last studied phase (+2219 steps; P=.01) compared to pre–COVID-19 (Figure 2). Male mobile users recorded significant decreases in steps at each of the first 2 waves as compared to the pre–COVID-19 period (–932 steps; P<.001; –954 steps; P<.001; respectively); however, they were less affected by the third wave, where a significant increase in values was observed at the last studied phase compared to pre–COVID-19 (+871 steps; P=.008; Figure 2). For pedometer users, sex differences were less evident at the pre–COVID-19 period; however, female individuals recorded 1736 more steps at the time when 100% of employees returned to work during wave 3 (T14; P=.05; Figure 2). The only significant change in step count of female pedometer users was seen in wave 2 with a decrease of –1443 steps compared to pre–COVID-19 (P=.03), after which step count went back to pre–COVID-19 values (Figure 2). Step count of male pedometer users was more significantly affected by the restrictions imposed during all COVID-19 waves, where significant declines compared to pre–COVID-19 of 993 steps (wave 1; P<.001), 1099 steps (wave 2; P<.001) and 1499 steps

XSL•FO
RenderX

(wave 3; $P<.001$) were observed (Figure 2). Values remained significantly lower throughout wave 3 and were still 708 steps lower than pre–COVID-19 at the last studied phase ($P=.06$).

Figure 2. The 7-day moving average of daily step count for male and female groups within the studied period covering 1 year before COVID-19 (T0: January 1, 2019, to February 28, 2020) and the 3 COVID-19 waves (T1-T18: February 29, 2020, to February 28, 2023). Details on the 19 time points reflecting the implementation and lifting of restrictions are found in Table 1. Data from pedometer and mobile app users are presented separately.

No significant differences in daily step count were observed between the nationality groups ($P>.05$). Please refer to Tables S2 and S3 in Multimedia Appendix 1 for the average daily step count of each of the 19 studied periods for pedometer and mobile users, respectively, among all subgroups.

Effect of COVID-19 Phases on Step Count by Age Group

A significant effect of age ($F_{1, 9233.62}=4.919$; $P=.03$) was evident on daily step count (Figure 3) in the studied period. Generally, the older group (ie, >50 years) moved less, with an average of 301 fewer steps than the younger group (ie, ≤ 50 years; Figure 3). A significant interaction effect between age and device type ($F_{1, 9031.59}=24.764$; $P<.001$) and a more detailed post hoc comparison showed a different pattern of change in step count between age groups according to the device used (Figure 3). In mobile users, the older group had initially higher step counts at the pre–COVID-19 period compared to the younger group (1344 steps) and were less affected by the first wave of COVID-19 (Figure 3). Step count of the older group, however, was significantly decreased at the second wave (–1545 steps at T11 as compared to pre–COVID-19; $P<.001$), which persisted through the entire second wave and most of the third wave. At the end of the studied period (T18), both groups regained pre–COVID-19 values, with the younger even presenting 1562 additional daily steps as compared to pre–COVID-19 (Figure 3). For pedometer users, a significant decrease in step count was recorded at the start of wave 1 for younger (–838 steps; $P=.003$) and older groups (–756 steps; $P=.005$). While the younger group's step count regained pre–COVID-19 values at the end of waves 2 and 3 with gradual increases recorded until the last phase (+699 steps compared to T0), the older group did not regain initial pre–COVID-19 values following restrictions of waves 2 and 3, with further decreases recorded at wave 3 (–1585 steps as compared to pre–COVID-19; $P<.001$) and significantly lower values at the last period (–768 steps) as compared to the pre–COVID-19 period ($P=.48$; Figure 3). Please refer to Tables S2 and S3 in Multimedia Appendix 1 for the average daily step count of each of the 19 studied periods for pedometer and mobile users, respectively, in all subgroups.

XSL•FO
RenderX

Figure 3. The 7-day moving average of daily step count for the younger (ie, ≤ 50 years) and older (ie, >50 years) groups within the studied period covering 1 year before COVID-19 (T0: January 1, 2019, to February 28, 2020) and the 3 COVID-19 waves (T1-T18: February 29, 2020, to February 28, 2023). Details on the 19 time points reflecting the implementation and lifting of restrictions are found in Table 1. Data from pedometer and mobile app users are presented separately.

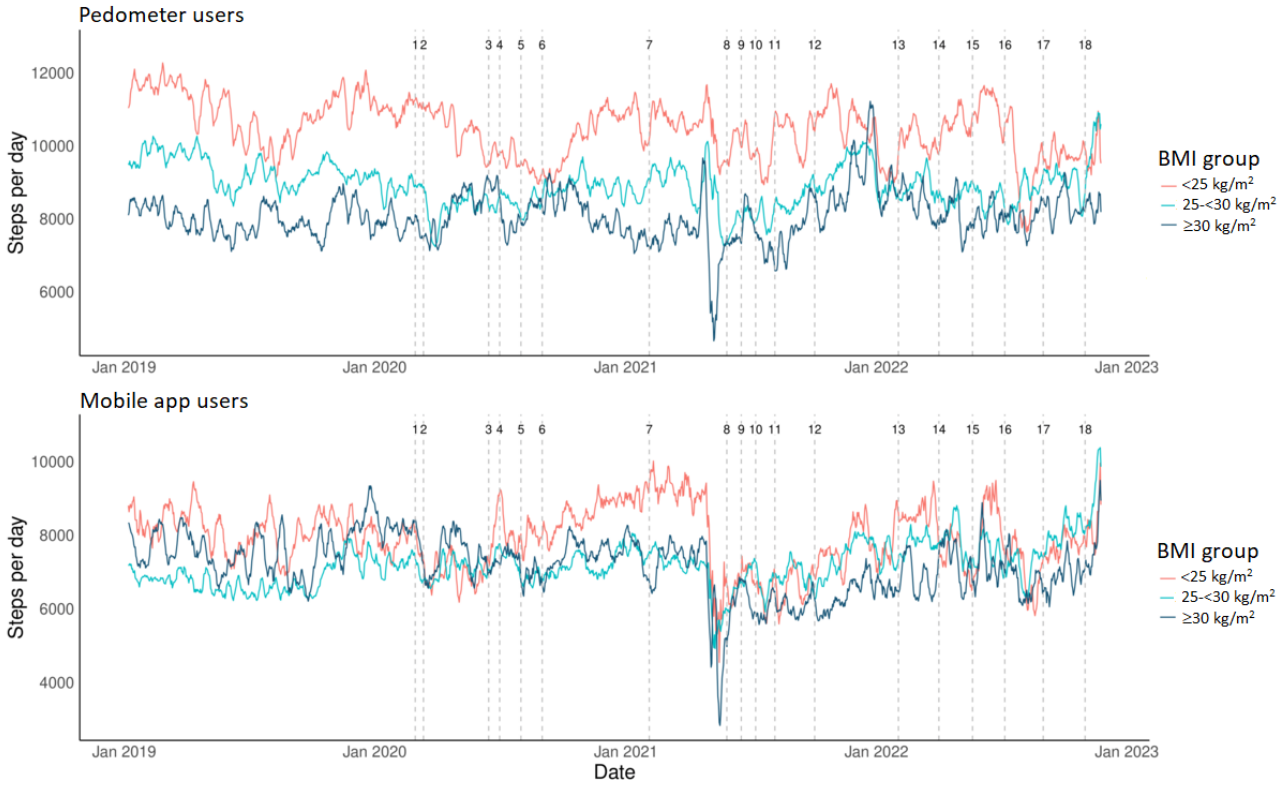
Effect of COVID-19 Phases on Step Count by BMI Groups

A significant effect of BMI ($F_{2, 362.63}$=4.210; *P*=.02) was shown on average daily step count in the studied time period (Figure 4). Findings indicated that the normal weight group had an average of 1191 and 1557 daily steps more than the overweight (*P*=.04) and obesity (*P*=.02) groups, respectively. No significant overall differences in daily step count were revealed between the overweight and obesity groups (*P*=.85). Although no significant interaction effect was found between BMI and device type ($F_{2, 507.93}$=1.459; *P*=.23), post hoc comparisons showed noteworthy differences between BMI groups using different devices (Figure 4). For instance, among mobile users, the normal weight group showed the highest fluctuations, with the different phases revealing significant decreases in step count following implementation of restrictions for wave 1 (–1276 steps compared to pre–COVID-19; *P*=.02) and a trend for wave 2 (–1070 steps compared to pre–COVID-19; *P*=.07), while no significant changes were recorded at wave 3 compared to the pre–COVID-19 period. At the end of each wave, the normal weight group recovered similar (slightly but not significantly higher) pre–COVID-19 values (Figure 4). For mobile users with overweight and obesity, less variations in step count were recorded with time. Specifically, the overweight group recorded an initial significant decrease in step count at the start of wave 1 (–691 steps compared to pre–COVID-19; *P*=.02) with no further changes until values significantly increased at the last phase of wave 3 when all restrictions were waived (+1744 step compared to pre–COVID-19; *P*<.001). The group with obesity only recorded a slight decrease of 769 steps at the start of wave 1 (compared to pre–COVID-19; *P*=.06) that persisted and became significant during wave 2 (–1259 steps compared to pre–COVID-19; *P*=.005). Slight increases in step count were seen for the group with obesity toward the end of the studied period, where daily step count regained their pre–COVID-19 values (*P*=.99; Figure 4). For pedometer users, the normal weight group recorded an initial significant decrease in values at the start of wave 1 (–1398 steps compared to pre–COVID-19; *P*=.001) and another larger significant decrease at the start of wave 2 (–1858 steps compared to pre–COVID-19; *P*<.001; Figure 4). After wave 2, values remained significantly lower until T17 (–1157 steps compared to pre–COVID-19; *P*=.02), and at the last studied period (T18), 1124 less steps were still noted compared to pre–COVID-19 (*P*=.08). The overweight group presented a similar pattern of change when compared to the pre–COVID-19 period. At the start of waves 1, 2, and 3, decreases of 995 steps (*P*=.002), 1145 steps (*P*=.002) and 1088 steps (*P*=.009) were respectively noted for the overweight group; however, an increase in values was also observed at the end of each wave with similar values noted at the end of the studied period as compared to pre–COVID-19 (Figure 4). Finally, no significant variations in step count were revealed throughout all waves for the group with obesity using the pedometer. Please

XSL•FO
RenderX

refer to Tables S2 and S3 in Multimedia Appendix 1 for the average daily step count of each of the 19 periods studied for pedometer and mobile users, respectively, in all subgroups.

Figure 4. The 7-day moving average of daily step count for groups of normal weight, overweight, and obesity within the studied period covering 1 year before COVID-19 (T0: January 1, 2019, to February 28, 2020) and the 3 COVID-19 waves (T1-T18: February 29, 2020, to February 28, 2023). Details on the 19 time points reflecting the implementation and lifting of restrictions are found in Table 1. Data from pedometer and mobile app users are presented separately.



Qualitative Results

Overview

This study explored the experiences of 9 participants regarding their PA habits during the COVID-19 pandemic. Participants provided insights into the impact of the pandemic on their PA routines, health behaviors, and motivations. The thematic analysis identified several key themes that were consistent across the transcripts. The pandemic had a significant impact on participants' PA behaviors, perceptions, and preferences.

The Importance and Awareness of PA

All participants acknowledged the importance of PA for both physical and mental health. Many participants emphasized that exercise is vital for managing stress and maintaining overall well-being, especially during stressful times like the COVID-19 pandemic.

> *It's very important to me, especially, you know, for stress relief. PA makes me feel better.* [Participant 1, Male, 30-39 years old]
>
> *It's vital for keeping your body and mind in good shape ... in these times, it's even more important for me.* [Participant 4, Female, 30-39 years old]
>
> *PA helps to reduce stress.* [Participant 5, Male, 40-49 years old]
>
> *Very important to me and to everybody. The benefits, for example, nowadays, with everybody being full of stress, walking helps relief that.* [Participant 8, Female, 40-49 years old]

While all participants acknowledged the importance of PA, there were differences in how they prioritized exercise. Several participants, particularly those who had prepandemic gym routines, viewed PA as a key part of their lifestyle. Others, however, indicated that their exercise habits were less structured during the pandemic due to limited time or motivation.

Disruptions to Routine and Shifting to Alternatives

The pandemic led to significant disruptions in participants' usual exercise routines. A total of 8 (89%) participants reported difficulty maintaining regular PA due to gyms closures, restricted access to outdoor spaces, and limited group fitness classes. Many of those participants transitioned to home-based exercises, including using fitness apps, virtual classes, and simple home exercises such as walking and rope jumping. However, limitations like space, lack of equipment, and safety concerns impacted the intensity and variety of exercises. The most notable difference was that participants who had relied heavily on gyms and group fitness classes before the pandemic, such as cycling and CrossFit, experienced a more significant disruption in their PA habits. In contrast, participants who engaged in outdoor activities, such as walking or running, were somewhat less affected.

XSL•FO
RenderX

Before COVID, I used to go to the gym almost every day. After restrictions started, it became difficult to continue my routine. [Participant 2, Female, 40-49 years old]

I couldn't go to the gym during the lockdown, and gyms were closed, so I had to stop everything. [Participant 3, Male, >50 years old]

My usual routine was walking and cycling, but with the restrictions, I had to adjust to home-based workouts. [Participant 6, Female, >50 years old]

COVID impacted my PA a lot. All gyms closed, and I was stuck at home. [Participant 9, Male, >50 years old]

I started doing home workouts with an app. It helped me stay active, even if it was hard sometimes. [Participant 1, Male, 30-39 years old]

I found some virtual classes and YouTube tutorials. I joined a few online workouts but did mostly walking. [Participant 4, Female, 30-39 years old]

During the pandemic, I worked out at home mostly, with some virtual classes. It was a bit hard to stay motivated at times. [Participant 5, Male, 40-49 years old]

I was doing indoor exercises with an app. It was quite easy since I had enough time at home. [Participant 9, Male, >50 years old]

Impact of COVID-19 on Lifestyle and Sedentary Behavior

A total of 7 (79%) participants reported an increase in their sedentary behavior during the pandemic. Working from home, limited outdoor activities, and lockdowns led to an overall reduction in physical movement, contributing to a more sedentary lifestyle for some. Participants noted spending longer hours in front of screens, either working or watching TV, which led to a more sedentary lifestyle. Participants who had previously been more active reported a noticeable decrease in PA, with some struggling to maintain their prepandemic routines. On the other hand, 2 (22%) participants who were less active before the pandemic reported that the lockdowns did not significantly alter their sedentary behaviors.

I didn't move much during COVID. Most of my time was spent on my bed, working and watching TV. [Participant 5, Male, 40-49 years old]

I found myself sitting for long periods while working from home. I wasn't as active as before. [Participant 7, Male, 30-39 years old]

During the lockdown, I stayed at home most of the time, and there were no real opportunities for exercise. [Participant 9, Male, >50 years old]

Sedentary behavior increased during the pandemic ... I was mostly indoors, working, and didn't get as much exercise as I wanted. [Participant 2, Female, 40-49 years old]

Motivation and Support System

Motivation was a key factor in maintaining regular PA during the pandemic. A total of 6 (67%) participants reported relying on external sources for motivation, such as family members, social media, or online fitness communities to continue exercising. However, 3 (33%) participants struggled with maintaining consistency due to isolation, lack of support, or insufficient motivation. Participants who had a structured workout routine before COVID-19 were often able to maintain higher levels of activity during the lockdown.

Social media really helped me stay motivated. I followed a few trainers and exercise groups that kept me going. [Participant 3, Male, >50 years old]

My family was supportive. They helped me stick to my routine. My wife did exercises with me at times. [Participant 8, Female, 40-49 years old]

I tried to stay motivated, but it was hard at times. I had to find ways like using apps and following trainers online. [Participant 4, Female, 30-39 years old]

Sometimes I felt bored, and that's when I used YouTube videos or watched football to keep me motivated. [Participant 8, Female, 40-49 years old]

Perceptions of Media and Communication During COVID-19

The media played an important role in shaping participants' perceptions of COVID-19 restrictions and the importance of PA. A total of 8 (89%) participants mentioned that the media played an important role in shaping their perceptions of COVID-19 restrictions and the importance of PA. Social media influenced many participants by providing inspiration for home workouts, although some expressed feeling overwhelmed by conflicting advice on exercise and health.

The media kept reminding us how important it is to stay active even at home. That kept me going. [Participant 3, Male, >50 years old]

Social media was very influential. I saw many people post about their home workouts and how to stay healthy. That motivated me to try new things. [Participant 9, Male, >50 years old]

The media coverage was helpful in showing us what exercises we could do at home to stay active during the pandemic. [Participant 2, Female, 40-49 years old]

Media has helped to remind us about the importance of being active during COVID. But sometimes it was overwhelming and confusing. [Participant 8, Female, 40-49 years old]

Barriers to PA During the Pandemic

Barriers such as lack of space, absence of equipment, and restrictions on outdoor activities were significant obstacles for participants trying to maintain regular exercise during the pandemic. A total of 7 (78%) participants reported facing significant barriers to maintaining PA during the pandemic.

XSL•FO
RenderX

These barriers included lack of space, absence of equipment, restrictions on outdoor activities, and psychological factors like reduced motivation or feeling overwhelmed by the stress of the pandemic. Those with access to outdoor spaces or home gyms had fewer barriers compared to those living in smaller apartments or with limited space. Additionally, participants who faced health issues or stress had more difficulty maintaining an active lifestyle during the pandemic.

> *I was missing my usual gym routine ... at home, I didn't have the right equipment, and the space was limited.* [Participant 5, Male, 40-49 years old]
>
> *The biggest barrier for me was not being able to go outside or go to the gym ... I didn't have enough space or equipment at home.* [Participant 7, Male, 30-39 years old]
>
> *It was hard to get motivated at home. And with all the stress of the pandemic, sometimes I just didn't feel like working out.* [Participant 4, Female, 30-39 years old]
>
> *The restrictions on outdoor activities were a big problem. It really affected my PA levels.* [Participant 8, Female, 40-49 years old]

Changes in PA Behavior After Lifting of Restrictions

As COVID-19 restrictions were eased, 6 (67%) participants expressed a desire to return to their previous exercise routines, such as going back to the gym or engaging in group activities. However, 3 (33%) participants were still hesitant, preferring home-based workouts or smaller, less crowded exercise options due to lingering safety concerns.

> *Now that restrictions are lifted, I prefer going back to the gym. I missed the group classes.* [Participant 6, Female, >50 years]
>
> *I am still not comfortable going back to gyms. I prefer to continue exercising at home for now.* [Participant 4, Female, 30-39 years old]
>
> *I would like to use the gym, but you know, I have a gym at home. And the gym facilities outside are too expensive.* [Participant 8, Female, 40-49 years old]
>
> *I've started going back to the gym, but I am more cautious about the crowds. I prefer small group classes now.* [Participant 9, Male, >50 years old]

Reluctance to Return to Prepandemic Routines

The long-term impact of COVID-19 on participants' PA habits was mixed. A total of 5 (56%) participants planned to continue with home-based exercises and more flexible routines post pandemic. Conversely, 4 (44%) participants were uncertain about how to reintegrate into their prepandemic activities. The pandemic also led some to reassess their health and fitness priorities. Mainly, those who had struggled to stay active during the pandemic showed uncertainty about returning to high-intensity or group activities.

> *I think after COVID, I will stick to some home exercises. It's more convenient for me.* [Participant 3, Male, >50 years old]
>
> *I don't think I will return to the same routine ... maybe I will walk more, but I don't see myself going back to the gym regularly.* [Participant 5, Male, 40-49 years old]
>
> *After all of this, I plan to keep working out at home. It fits my lifestyle better.* [Participant 6, Female, >50 years old]
>
> *I've been working out more at home during COVID, and I'll probably continue that for now.* [Participant 9, Male, >50 years old]

Triangulation of Results

The use of methodological triangulation allowed for the integration of quantitative and qualitative strands, strengthening the interpretation of results. Patterns observed in step count data (such as declines during restriction phases, subgroup differences, and device-related discrepancies) were consistent with themes identified in interviews regarding disruptions to PA routines, motivational challenges, and adaptations to home-based exercise. Data triangulation further enriched the analysis by contrasting findings across devices (pedometer vs mobile phone) with participants' perceptions. This combined approach provided a more comprehensive understanding of PA behaviors during the COVID-19 pandemic and increased the rigor of the conclusions.

Increased sedentary behavior was noted in our analyses, linked to lockdown restrictions, working from home, and limited outdoor activity. Motivation emerged as a key factor in maintaining activity, with external support from social media and family playing a significant role. These qualitative insights were reflected in the quantitative data, which indicated that engagement with online communities helped sustain exercise habits.

Barriers such as limited space, lack of equipment, and psychological factors were consistent across both data sources. Qualitative findings further highlighted how these barriers were exacerbated by stress and mental health challenges, offering a deeper context to the quantitative results. Post pandemic, participants expressed mixed intentions about returning to prepandemic routines, with some preferring home-based workouts, which was also reflected in the quantitative data.

The quantitative finding that wave 2 produced the largest declines in step count, especially among older adults, was echoed in interview data: participants described heightened fear and caution, reporting that "with all the stress of the pandemic, sometimes I just didn't feel like working out" (Female, 30-39 years old). Similarly, discrepancies between pedometer and mobile users were contextualized by participants' reports of not always carrying phones during exercise or preferring home-based workouts, which may have reduced recorded steps ("I was doing indoor exercises with an app. It was quite easy since I had enough time at home" [Male, >50 years old]). On the other hand, while female mobile users recorded fewer steps than male users in the pre–COVID-19 phase, some women reported higher motivation after returning to work, which corresponded with the observed increase in their step counts in wave 3 ("My family was supportive. They helped me stick to

XSL•FO
RenderX

my routine" [Female, 40-49 years old]). These examples show where the qualitative themes supported or nuanced the quantitative results, highlighting both convergences and divergences.

The integration of findings through triangulation enhanced interpretive validity and underscored the value of mixed methods approaches in capturing both behavioral trends and contextual complexities. This convergence of evidence supports a more holistic understanding of PA adaptation during public health crises and informs future interventions that address both structural and psychosocial determinants.

Discussion

Main Findings

This study examined the impact of COVID-19 phases of restrictions on PA levels as measured by daily step count during the entire pandemic period spanning over the 3 waves. Results offer valuable insights into behavioral adaptations to restrictive measures while considering the influence of sex, age, nationality, and BMI, as well as the impact of the device used to track step count. Generally, the findings confirm significant declines in daily step count at the implementation of restrictions within each COVID-19 wave followed by a recovery of values toward the end of each wave. The latter is consistent with previous reports focusing on the first COVID-19 wave in Qatar [20] and elsewhere using community-based longitudinal interventions [34,35]. Barriers reported by participants included lack of space, absence of equipment, restrictions on outdoor activities, and psychological factors, such as reduced motivation, stress, or feeling overwhelmed.

The most pronounced decreases in PA were observed at the second wave, coinciding with the arrival of the Omicron variant in Qatar [36]. This aligns with reports of peak severity, criticality, and fatality in locally confirmed cases between March and May 2021 [37]. Heightened public fear and/or perceived risk, potentially leading to stricter adherence to restrictions, seemed to act as important barriers to PA behavior in the given context. Indeed, in addition to several barriers reported by interviewed participants, the stress of the pandemic and relative caution came up as limiting factors ("... and with all the stress of the pandemic, sometimes I just didn't feel like working out"; "... but I am more cautious about the crowds. I prefer small group classes now"). This interpretation might indicate that severity, criticality, and fatality of cases strongly advertised on social media platforms and news outlets might have played a critical role in PA level, rather than the reported number of daily confirmed cases that peaked at a different time. Additionally, the period between April 12 and May 12, 2021, characterized with the most severe drops in daily step counts, also coincided with the month of Ramadan, where Muslims fast from dawn to sunset and naturally decrease their daily steps [38]. Large vaccination campaigns and advances in case management were among the factors that favored a 95% drop in the rates of severe, critical, and fatal cases with a turning point recorded in March 2022 [37] where increases in daily step counts were noted.

A primary factor to consider in the interpretation is the discrepancy found in the patterns of change in recorded steps between pedometer and mobile phone users. Pedometer users recorded an average of 690 daily steps more than mobile users, consistent with previous literature. A previous study showed that step count was 12% lower in iPhone users as compared to pedometer users [39]. Although underestimations seen in mobile devices are mainly attributed to noncarrying time that is influenced by the setting or context (eg, being at home or outside), other factors are worth considering. Mobile carrying habits could explain sex differences found between device groups, as women might tend to favor carrying their phones in their bags and men in their pockets [39]. Additionally, pedometers are dedicated to activity tracking unlike mobile phones, and the simple fact of carrying one can alter behaviors. Findings from a previous study on the SIH population indicated that pedometer users were significantly more committed to uploading their data and presented higher retention to the program as compared to mobile users [30]. Authors suggested that offering a free pedometer might have enhanced feelings of responsibility and motivation toward the program [40]. Finally, mobile phones fail to account for steps during activities where it is impractical to carry a phone (eg, gym training or playing team or racket ball sports) [41] which could be favored by younger groups. A recurrent observation from the qualitative analysis is that most PA routines were shifted to home-based exercises (ie, "My usual routine was walking and cycling, but with the restrictions, I had to adjust to home-based workouts"; "I started doing home workouts with an app"; "I found some virtual classes and YouTube tutorials"; "During the pandemic, I worked out at home mostly, with some virtual classes"; "I was doing indoor exercises with an app"). It is reasonable to assume that tracking activity within a limited space at home is less practical with a mobile phone as compared to a dedicated activity tracker.

In this study, differences in step count between pedometer and mobile users varied from 609 to 1264 steps according to the studied phase, with the largest differences noted toward the end of all restrictions (ie, wave 3) where 100% of employees were back to work. These observations further confirm that steps recorded by mobile phones can easily underestimate actual daily step counts in certain contexts. Moreover, pedometer users might also represent the most motivated and health-conscious portion of the population. Regardless of the device used to track steps, results should be interpreted with caution, as findings could be conflicting when trying to compare data from pedometer and mobile over an extended period of time among different subgroups, each presenting specific characteristics that might either enhance or hinder step count. These findings show that while device-driven data present an increased objectivity as compared to self-reported measures, bias in activity tracking still exists based on the presented limitations.

Overall, the qualitative analysis points toward a reluctance to go back to pre–COVID-19 PA routines and shifting to exercise within the home context (ie, "I think after COVID, I will stick to some home exercises. It's more convenient for me"; "I don't think I will return to the same routine ... maybe I will walk more, but I don't see myself going back to the gym regularly"; "After

XSL•FO
RenderX

all of this, I plan to keep working out at home. It fits my lifestyle better"; "I've been working out more at home during COVID, and I'll probably continue that for now"). While this study examines perceptions of a population that is considered health-conscious, it reveals long-term or lasting implications of the pandemic on behaviors and perceptions, where leaving the house or being active in outdoor settings are less favored. This shift in behavior, pointing to a tendency to remain within a confinement "state of mind," could have significant implications for physical and mental health, especially with the emergence of the greater role of outdoor compared to indoor PA [42]. It is important to note, however, that while interpretations of potential long-term effects are limited to November 2022, this study is not able to report and/or speculate on any longer-term impacts on behavior beyond the studied time range. Findings do not represent current or ongoing trends and future research using more recent data is needed to determine whether observed changes persist over time.

While the literature is well supported by evidence on the effects of sex, age, and BMI on step count, or more recently on the impact of COVID-19 restrictions on step count, potential interactions between both categories of factors remain limited. In addition, discrepancies between data obtained from mobiles and pedometers add a layer of complexity to the interpretation that was not previously addressed in this context. Understanding these interactions is important to develop appropriate interventions and policies targeting the most vulnerable subgroups. Our findings point toward larger significant variations/fluctuations in step counts through restriction phases among male (compared to female) individuals, pedometer users (compared to mobile users), and normal weight participants (compared to individuals with overweight and obesity). At the pre–COVID-19 phase, these subgroups presented higher step count and were more affected by (or sensitive to) restrictive measures, which is in line with previous findings on Qatar's first COVID-19 wave [20]. Interestingly, male individuals, older individuals, and normal weight groups using pedometers did not recover their initial pre–COVID-19 step count (ie, –708, –768, and –1124 steps, respectively), while female individuals, younger individuals, and those living with overweight and obesity recovered their values recorded using pedometers. In contrast, all subgroups using their mobile regained their initial step count, and most presented significantly higher values at the end of wave 3 compared to pre–COVID-19, which is a novel result emerging from the inclusion of 2 assessment devices in this study.

In sum, the use of mobile apps to track daily steps warrants a deeper interpretation of the exact use of mobile during home confinement and possible PA opportunities. The perceptions of younger vs older individuals, and female vs male, are crucial in explaining discrepancies. Different patterns of home and outdoor activities and varying preferences for gym, walking, and other sports/exercises might explain differences between device users during this phase. Notably, women were possibly more occupied at home with childcare and household responsibilities and may have been more positively affected by the "return to work" at the end of the third wave, which is reflected in a substantial increase in their step counts at this time.

The triangulation of both qualitative and quantitative results reinforces these insights, with qualitative data providing deeper context on how different factors (eg, motivation, support systems, or media influences) shaped participants' PA behaviors during the pandemic. Ultimately, qualitative findings point at an important role of media and social media in favoring active lifestyles during periods of pandemic ("The media kept reminding us how important it is to stay active even at home"; "Social media was very influential. I saw many people post about their home workouts and how to stay healthy. That motivated me to try new things"; "The media coverage was helpful in showing us what exercises we could do at home to stay active during the pandemic"). It seems that remaining active was a common concern on social media that played an important motivating role; however, it was also perceived as overwhelming and confusing ("Media has helped to remind us about the importance of being active during COVID. But sometimes it was overwhelming and confusing"). Clear instructions and content on PA and exercise are essential during such periods and should be made available from reliable sources.

Finally, findings could be interpreted from the perspective of the social cognitive theory (SCT), which focuses on the reciprocal interaction between personal, behavior, and environmental factors [43,44]. According to this theoretical framework, the reduction in daily step count seen during restrictive phases or even during Ramadan supports the impact of environmental barriers on individual and collective behavior. Conversely, the gradual recovery of PA after the lifting of restrictions demonstrates the adaptive ability of motivated individuals to revert to usual PA as environments became more supportive. This goes in line with the SCT concept that environmental facilitators favor self-regulatory behavior. The characteristics of the studied population (eg, consistent engagement in tracking and uploading step count during the COVID-19 period and knowledge on the benefits of PA for health), which is likely to possess higher levels of self-efficacy as compared to the general population of Qatar, reflect an important determinant of PA behavior. Return to normal PA behavior with efforts to exercise from home show behavioral resilience throughout the pandemic in the sample studied. The reliance on social media for motivation and guidance, highlighted in the qualitative findings, aligns with SCT's concept of observational learning, where individuals learn behaviors through exposure to peer or media models. The availability of peer activity, health promotion campaigns, and online guidance for exercise during the COVID-19 have likely reinforced participants' motivation, therefore providing behavioral models at a time when usual opportunities were limited [45]. Overall, this interpretation is consistent with previous studies showing that access to supportive environments and modeled behaviors play an important role in sustaining PA during disruptive events [46]. Future research should explicitly measure SCT constructs such as self-efficacy, outcome expectations, social support, and observational learning to identify correlates and predictors of resilience or vulnerability in PA behavior [47], and to provide

XSL•FO
RenderX

stronger public health interventions based on theoretical foundations.

Limitations

In addition to the main limitation noted earlier in the step count bias associated with the use of pedometers and mobile phones (eg, noncarrying time, carrying habits, type of activities, and settings), several other limitations should be acknowledged. While the exclusion criteria applied for this study are necessary for the analysis, the retained sample comprised health-conscious individuals who continued tracking their step count through the entire COVID-19 period and regularly uploaded their activity data. Therefore, the studied sample differs from the general population of Qatar and potentially from the initial SIH cohort. A previous SIH study on 16,711 participants reported a 50% retention rate from 2012 to 2019 and a 16% adherence rate to daily data uploading [30]. The SIH is a self-managed and voluntary program designed to increase PA among the adult population of Qatar. The program has already proved its effectiveness in increasing average daily step count at the 12th month of enrollment compared to baseline data [28], that may represent more closely PA levels of the general population of Qatar. Together, this information suggests that the requirement for continuous valid step data across the 3 COVID-19 waves likely produced a subgroup of the most active participants characterized by the highest retention and adherence to the program. These factors limit the generalizability of our findings to the wider population of Qatar and should only represent the most active portion of the population. Moreover, it is important to consider the unique environmental and cultural context of Qatar, which may further limit the generalizability of the results to other countries outside of the Gulf Peninsula region. Extreme summer temperatures, often exceeding 40 °C, can substantially restrict outdoor PA and may have influenced step count at both baseline and periods of COVID-19 restrictions. Cultural norms, including gender-specific expectations for outdoor activity and preferences for indoor vs outdoor exercise, further shape patterns of PA in the studied population of Qatar.

A further limitation is that participants were recruited for the interviews through random selection rather than purposive sampling informed by device type, activity level, or BMI. While this strategy minimized researcher bias and still yielded diversity in age and sex, it may have missed key perspectives. The interviews' small sample size (n=9) reached thematic saturation for the main themes identified, but limits the depth of subgroup-specific insights. Future studies could purposively recruit participants based on characteristics identified in the quantitative strand to strengthen the integration of mixed methods findings. Although participants' narratives sometimes referred to their use of social media for exercise motivation, no social media tracking or collection of identifiable data occurred. This reliance on self-reported perceptions avoids privacy risks but also limits the depth of analysis on media use.

Conclusions

The findings from this study emphasize the need for caution when interpreting step count data from mobile devices for research purposes, as measurement accuracy may vary by context, individual factors (eg, age, sex, and BMI), user behavior, and preference of certain types of activities. With the advancements of wearable multisensor devices, future studies should account for potential bias and limitations associated with the use of mobile phones and pedometers as PA data collection tool. The COVID-19 pandemic has taught us a lesson supported by the data of this study: developing tailored interventions should aim to encourage outdoor activities after confinement periods and provide reliable social media sources for PA and exercise across diverse populations. It is crucial to consider the contextual (eg, severity, criticality, and fatality of cases) and personal factors (eg, psychological, age, and sex) that influence PA behavior during public health crises such as the COVID-19 pandemic, where long-term consequences are observed in PA behaviors.

Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary information file.

Funding

Funding for the thematic analysis and interviews' recording was obtained through a Qatar University High Impact Grant led by Eirini Theodoropoulou and co-led by LM. The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Authors' Contributions

Conceptualization: MC, ASA-M
Methodology: MC, SS, AF, LM, ASA-M
Formal analysis: AF, LM, SS
Interview administration and curation: FD
Writing of the original draft preparation: LM, SS
Project administration: AF, ASA-M, SS, LM
All authors have contributed to the review and editing and have read and agreed to the published version of the manuscript.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Tables S2 and S3 present the mean ± SD of daily step counts for pedometer and mobile users across population subgroups before COVID-19 (T0) and during the 3 pandemic waves (T1-T18), 2019-2023.
[DOCX File , 24 KB - publichealth_v12i1e76860_app1.docx]

References

1. Coronavirus disease (COVID-19) pandemic. World Health Organization. 2019. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019 [accessed 2023-11-19]
2. Director-general's opening remarks at the media briefing on COVID-19 - 11 March. World Health Organization. 2020. URL: https://www.who.int/ [accessed 2023-11-19]
3. COVID-19 updates. Ministry of Public Health, Qatar. 2023. URL: https://www.moph.gov.qa [accessed 2023-11-19]
4. Physical activity. World Health Organization. 2023. URL: https://www.who.int/news-room/fact-sheets/detail/physical-activity [accessed 2023-11-19]
5. World Health Organization. WHO Guidelines on Physical Activity and Sedentary Behaviour. Geneva: World Health Organization; 2020.
6. Qatar National Physical Activity Guidelines. (2021). 2nd edition. Aspetar. 2021. URL: https://www.aspetar.com/AspetarFILEUPLOAD/UploadCenter/637736948034432931_QATAR%20NATIONAL%20PHYSICAL%20ACTIVITY%20GUIDELINES_ENGLISH.pdf [accessed 2025-05-02]
7. Tison GH, Avram R, Kuhar P, Abreau S, Marcus GM, Pletcher MJ, et al. Worldwide effect of COVID-19 on physical activity: a descriptive study. Ann Intern Med 2020;173(9):767-770. [doi: 10.7326/m20-2665]
8. Ammar A, Brach M, Trabelsi K, Chtourou H, Boukhris O, Masmoudi L, et al. Effects of COVID-19 home confinement on eating behaviour and physical activity: results of the ECLB-COVID19 international online survey. Nutrients 2020;12(6):1583 [FREE Full text] [doi: 10.3390/nu12061583] [Medline: 32481594]
9. Yang Y, Koenigstorfer J. Determinants of physical activity maintenance during the COVID-19 pandemic: a focus on fitness apps. Transl Behav Med 2020;10(4):835-842 [FREE Full text] [doi: 10.1093/tbm/ibaa086] [Medline: 32926160]
10. Wunsch K, Kienberger K, Niessner C. Changes in physical activity patterns due to the COVID-19 pandemic: a systematic review and meta-analysis. Int J Environ Res Public Health 2022;19(4):2250 [FREE Full text] [doi: 10.3390/ijerph19042250] [Medline: 35206434]
11. Ruíz-Roso MB, de Carvalho Padilha P, Matilla-Escalante DC, Brun P, Ulloa N, Acevedo-Correa D, et al. Changes of physical activity and ultra-processed food consumption in adolescents from different countries during COVID-19 pandemic: an observational study. Nutrients 2020;12(8):2289. [doi: 10.3390/nu12082289]
12. Castañeda-Babarro A, Arbillaga-Etxarri A, Gutiérrez-Santamaría B, Coca A. Physical activity change during COVID-19 confinement. Int J Environ Res Public Health 2020;17(18):6878. [doi: 10.3390/ijerph17186878]
13. Di Sebastiano KM, Chulak-Bozzer T, Vanderloo LM, Faulkner G. Don't walk so close to me: physical distancing and adult physical activity in Canada. Front Psychol 2020;11:1895 [FREE Full text] [doi: 10.3389/fpsyg.2020.01895] [Medline: 32849110]
14. Zheng C, Huang WY, Sheridan S, Sit CHP, Chen XK, Wong SHS. COVID-19 pandemic brings a sedentary lifestyle in young adults: a cross-sectional and longitudinal study. Int J Environ Res Public Health 2020;17(17):6035 [FREE Full text] [doi: 10.3390/ijerph17176035] [Medline: 32825092]
15. Pišot S, Milovanović I, Šimunič B, Gentile A, Bosnar K, Prot F, et al. Maintaining everyday life praxis in the time of COVID-19 pandemic measures (ELP-COVID-19 survey). Eur J Public Health 2020;30(6):1181-1186 [FREE Full text] [doi: 10.1093/eurpub/ckaa157] [Medline: 32750114]
16. Wilke J, Rahlf AL, Füzéki E, Groneberg DA, Hespanhol L, Mai P, et al. Physical activity during lockdowns associated with the COVID-19 pandemic: a systematic review and multilevel meta-analysis of 173 studies with 320,636 participants. Sports Med Open 2022;8(1):125 [FREE Full text] [doi: 10.1186/s40798-022-00515-x] [Medline: 36219269]
17. Nieman DC. Exercise is medicine for immune function: implication for COVID-19. Curr Sports Med Rep 2021;20(8):395-401. [doi: 10.1249/jsr.0000000000000867]
18. Hermassi S, Sellami M, Salman A, Al-Mohannadi AS, Bouhafs EG, Hayes LD, et al. Effects of COVID-19 lockdown on physical activity, sedentary behavior, and satisfaction with life in Qatar: a preliminary study. Int J Environ Res Public Health 2021;18(6):3093 [FREE Full text] [doi: 10.3390/ijerph18063093] [Medline: 33802891]
19. Abed MA, Abdeen S, Kehyayan V, Bougmiza I. Impact of staying at home measures during COVID-19 pandemic on the lifestyle of Qatar's population: perceived changes in diet, physical activity, and body weight. Prev Med Rep 2021;24:101545 [FREE Full text] [doi: 10.1016/j.pmedr.2021.101545] [Medline: 34493967]

XSL•FO
RenderX

20. Al-Mohannadi AS, Farooq A, Salman A, Ghram A, Al-Harahsheh ST, Majed L, et al. The effects of the first wave of COVID-19 restrictions on physical activity: a longitudinal study from "Step Into Health" program in Qatar. Front Public Health 2024;12:1333546 [FREE Full text] [doi: 10.3389/fpubh.2024.1333546] [Medline: 38510355]
21. Althumiri NA, Basyouni MH, BinDhim NF. Exploring the association between physical activity and risk of mental health disorders in Saudi Arabian adults: cross-sectional study. JMIR Public Health Surveill 2021;7(4):e25438 [FREE Full text] [doi: 10.2196/25438] [Medline: 33851932]
22. Cheikh Ismail L, Osaili T, Mohamad M, Al Marzouqi A, Jarrar A, Zampelas A, et al. Assessment of eating habits and lifestyle during the coronavirus 2019 pandemic in the Middle East and North Africa region: a cross-sectional study. Br J Nutr 2021;126(5):757-766 [FREE Full text] [doi: 10.1017/S0007114520004547] [Medline: 33198840]
23. Bifolco G, Cardinali L, Mocini E, Duradoni M, Baldari C, Ciampi M, et al. Long-term effects of COVID-19 pandemic on physical activity and eating behaviour of the Italian population: a longitudinal study. Endocrine 2024;86(3):1003-1013. [doi: 10.1007/s12020-024-03950-w] [Medline: 39177928]
24. Ripley-Gonzalez JW, Zhou N, Zeng T, You B, Zhang W, Liu J, et al. The long-term impact of the COVID-19 pandemic on physical fitness in young adults: a historical control study. Sci Rep 2023;13(1):15430 [FREE Full text] [doi: 10.1038/s41598-023-42710-0] [Medline: 37723197]
25. Katoue MG, Cerda AA, García LY, Jakovljevic M. Healthcare system development in the Middle East and North Africa region: challenges, endeavors and prospective opportunities. Front Public Health 2022;10:1045739 [FREE Full text] [doi: 10.3389/fpubh.2022.1045739] [Medline: 36620278]
26. Adli N, Bala M, Bougmiza MI, Al-Kuwari MG, Al-Khenji AH, Selim N. Post-COVID-19 syndrome, who at risk and why: an insight from Qatar 2022. Front Public Health 2024;12:1368568. [doi: 10.3389/fpubh.2024.1368568] [Medline: 38680936]
27. Sayegh S, Van Der Walt M, Al-Kuwari MG. One-year assessment of physical activity level in adult Qatari females: a pedometer-based longitudinal study. Int J Womens Health 2016;Volume 8:287-293. [doi: 10.2147/ijwh.s99943]
28. Al-Kuwari MG, Al-Mohannadi AS, Sayegh S. Effectiveness of "Step into Health" program in Qatar: a pedometer-based longitudinal study. J Sports Med Phys Fitness 2017;57(11):1513-1518. [doi: 10.23736/s0022-4707.16.06716-5]
29. Al-Mohannadi AS, Sayegh S, Ibrahim I, Salman A, Farooq A. Effect of a pedometer-based walking challenge on increasing physical activity levels amongst hospital workers. Arch Public Health 2019;77:40 [FREE Full text] [doi: 10.1186/s13690-019-0368-7] [Medline: 31572606]
30. Chrismas BCR, Majed L, Al-Mohannadi AS, Sayegh S. Adherence and retention to the self-managed community-based step into health program in Qatar (2012-2019). Front Public Health 2022;10:927386 [FREE Full text] [doi: 10.3389/fpubh.2022.927386] [Medline: 36199850]
31. Tudor-Locke C, Burkett L, Reis JP, Ainsworth BE, Macera CA, Wilson DK. How many days of pedometer monitoring predict weekly physical activity in adults? Prev Med 2005;40(3):293-298. [doi: 10.1016/j.ypmed.2004.06.003] [Medline: 15533542]
32. Yao J, Tan CS, Lim N, Tan J, Chen C, Müller-Riemenschneider F. Number of daily measurements needed to estimate habitual step count levels using wrist-worn trackers and smartphones in 212,048 adults. Sci Rep 2021;11(1):9633 [FREE Full text] [doi: 10.1038/s41598-021-89141-3] [Medline: 33953288]
33. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2008;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
34. To QG, Duncan MJ, Van Itallie A, Vandelanotte C. Impact of COVID-19 on physical activity among 10,000 steps members and engagement with the program in Australia: prospective study. J Med Internet Res 2021;23(1):e23946 [FREE Full text] [doi: 10.2196/23946] [Medline: 33449907]
35. Hino K, Asami Y. Change in walking steps and association with built environments during the COVID-19 state of emergency: a longitudinal comparison with the first half of 2019 in Yokohama, Japan. Health Place 2021;69:102544 [FREE Full text] [doi: 10.1016/j.healthplace.2021.102544] [Medline: 33714180]
36. AlNuaimi AA, Chemaitelly H, Semaan S, AlMukdad S, Al-Kanaani Z, Kaleeckal AH, et al. All-cause and COVID-19 mortality in Qatar during the COVID-19 pandemic. BMJ Glob Health 2023;8(5):e012291 [FREE Full text] [doi: 10.1136/bmjgh-2023-012291] [Medline: 37142299]
37. Chemaitelly H, Ayoub HH, Faust JS, Coyle P, Tang P, Hasan MR, et al. Turning point in COVID-19 severity and fatality during the pandemic: a national cohort study in Qatar. BMJ Public Health 2023;1(1):e000479. [doi: 10.1136/bmjph-2023-000479] [Medline: 40017867]
38. Lessan N, Saadane I, Alkaf B, Hambly C, Buckley AJ, Finer N, et al. The effects of Ramadan fasting on activity and energy expenditure. Am J Clin Nutr 2018;107(1):54-61 [FREE Full text] [doi: 10.1093/ajcn/nqx016] [Medline: 29381798]
39. Amagasa S, Kamada M, Sasai H, Fukushima N, Kikuchi H, Lee I, et al. How well iPhones measure steps in free-living conditions: cross-sectional validation study. JMIR Mhealth Uhealth 2019;7(1):e10418 [FREE Full text] [doi: 10.2196/10418] [Medline: 30626569]
40. Bravata DM, Smith-Spangler C, Sundaram V, Gienger AL, Lin N, Lewis R, et al. Using pedometers to increase physical activity and improve health: a systematic review. JAMA 2007;298(19):2296-2304. [doi: 10.1001/jama.298.19.2296] [Medline: 18029834]

XSL•FO
RenderX

41. Althoff T, Sosič R, Hicks JL, King AC, Delp SL, Leskovec J. Large-scale physical activity data reveal worldwide activity inequality. Nature 2017;547(7663):336-339 [FREE Full text] [doi: 10.1038/nature23018] [Medline: 28693034]
42. Thompson Coon J, Boddy K, Stein K, Whear R, Barton J, Depledge MH. Does participating in physical activity in outdoor natural environments have a greater effect on physical and mental wellbeing than physical activity indoors? A systematic review. Environ Sci Technol 2011;45(5):1761-1772. [doi: 10.1021/es102947t] [Medline: 21291246]
43. Bandura A. Social Foundations of Thought and Action. New Jersey: Englewood Cliffs; 1986:23-28.
44. McAlister AL, Perry CL, Parcel GS. How Individuals, Environments, and Health Behaviors Interact. USA: Health Behavior and Health Education; 2008:169-188.
45. Marchant G, Bonaiuto F, Bonaiuto M, Guillet Descas E. Exercise and physical activity eHealth in COVID-19 pandemic: a cross-sectional study of effects on motivations, behavior change mechanisms, and behavior. Front Psychol 2021;12:618362 [FREE Full text] [doi: 10.3389/fpsyg.2021.618362] [Medline: 33692722]
46. Goodyear VA, Boardley I, Chiou SY, Fenton SAM, Makopoulou K, Stathi A, et al. Social media use informing behaviours related to physical activity, diet and quality of life during COVID-19: a mixed methods study. BMC Public Health 2021;21(1):1333 [FREE Full text] [doi: 10.1186/s12889-021-11398-0] [Medline: 34229651]
47. Rhodes RE, Liu S, Lithopoulos A, Zhang CQ, Garcia-Barrera MA. Correlates of perceived physical activity transitions during the COVID-19 pandemic among Canadian adults. Appl Psychol Health Well Being 2020;12(4):1157-1182 [FREE Full text] [doi: 10.1111/aphw.12236] [Medline: 33006279]

Abbreviations

PA: physical activity
SCT: social cognitive theory
SIH: Step Into Health program

Edited by A Mavragani, T Sanchez; submitted 02.May.2025; peer-reviewed by K Di Sebastiano, ES Darmawan, K Jermsittiparsert; comments to author 18.Aug.2025; revised version received 25.Oct.2025; accepted 27.Oct.2025; published 23.Mar.2026.

Please cite as:
Majed L, Sayegh S, Dalansi F, Al-Mohannadi AS, Cardinale M, Farooq A
Longitudinal Analysis of Variations in Daily Step Counts and Long-Term Implications of COVID-19 Waves and Restriction Phases in Qatar's Step Into Health Program: Mixed Methods Study
JMIR Public Health Surveill 2026;12:e76860
URL: https://publichealth.jmir.org/2026/1/e76860
doi:10.2196/76860
PMID:

©Lina Majed, Suzan Sayegh, Feriel Dalansi, Abdulla Saeed Al-Mohannadi, Marco Cardinale, Abdulaziz Farooq. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 23.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Long-Term Survival and Life-Sustaining Device Use in Survivors of First and Second Out-of-Hospital Cardiac Arrest: Retrospective Cohort Study

Chih-Hsien Lin[1], MD; Cheng-Yi Fan[1,2,3], MD; Edward Pei-Chuan Huang[1,2], MD, MS; Chih-Wei Sung[1,2], MD, PhD

[1]Department of Emergency Medicine, National Taiwan University Hospital Hsin-Chu Branch, Hsinchu, Taiwan

[2]Department of Emergency Medicine, College of Medicine, National Taiwan University, Taipei, Taiwan

[3]Institute of Molecular Medicine, National Tsing Hua University, Hsinchu, Taiwan

Corresponding Author:
Chih-Wei Sung, MD, PhD
Department of Emergency Medicine
National Taiwan University Hospital Hsin-Chu Branch
Hsinchu
Taiwan
Phone: 886 35326151 ext 524200
Fax: 886 35348443
Email: chihweisung@ntu.edu.tw

Abstract

Background: Long-term surveillance of out-of-hospital cardiac arrest (OHCA) survivors is increasingly important, but patients who survive a second OHCA are rarely characterized because of the scarcity of such cases. The nationwide claims-based health data provide an opportunity to identify this uncommon survivor population and evaluate postdischarge outcomes at the population level. Understanding the prognosis and care needs of second-time OHCA survivors may help inform postarrest surveillance, risk stratification, and long-term care planning.

Objective: This study aimed to compare the clinical characteristics of adult patients who survived a single OHCA versus those who survived a second OHCA during a 5-year follow-up and assess and contrast their 5-year survival outcomes.

Methods: We conducted a nationwide retrospective cohort study using data from the National Health Insurance Research Database (2010-2020) in Taiwan. Survivors at discharge were classified as 1-OHCA survivors, defined as patients who survived exactly 1 OHCA episode and had no documented recurrent OHCA during follow-up, or 2-OHCA survivors, defined as patients who experienced exactly 2 documented OHCA episodes and survived to discharge after both events. Patients with more than 2 OHCA episodes were excluded. Postdischarge survival was measured from the qualifying discharge date: discharge after the first OHCA for 1-OHCA survivors and discharge after the second OHCA for 2-OHCA survivors. Variables included demographics, comorbidities, health care use, and life-sustaining device status. Five-year mortality was analyzed using the Kaplan-Meier and log-rank tests. Adjusted hazard ratios were derived from multivariable Cox models. Device use patterns were compared using Cochran-Armitage trend tests.

Results: Among 239,929 OHCA cases, 229,047 were eligible; 15,617 survived to hospital discharge. Five-year mortality was substantially higher among 2-OHCA survivors than among 1-OHCA survivors (191/201, 95% vs 6850/14,494, 47.3%; log-rank $P<.001$). In patients with 1 OHCA, mortality increased with age, male sex, and nasogastric tube use, whereas >3 outpatient visits, Foley catheter use, and tracheostomy or ventilation were associated with lower observed postdischarge mortality. In the 2-OHCA group, the rate of device-free status was lower, while the triple-device use rate was higher.

Conclusions: In this nationwide claims-based surveillance study, patients who survived discharge after a second OHCA represented a rare and clinically vulnerable postarrest population with poor subsequent 5-year survival and greater life-sustaining device dependence. Device-related variables should be interpreted as claims-based markers of postdischarge dependency, care setting, survivorship, and care intensity rather than evidence that device placement itself improves survival. As 2-OHCA survivor status is conditional on survival to a second OHCA and discharge after that event, these findings should be interpreted as descriptive and hypothesis-generating rather than causal. Population-level health data may support long-term postarrest surveillance and care planning for rare OHCA survivor populations.

XSL•FO
RenderX

(JMIR Public Health Surveill 2026;12:e90416) doi:10.2196/90416

KEYWORDS

out-of-hospital cardiac arrest; second cardiac arrest; life-sustaining device; risk stratification; National Health Insurance Research Database

Introduction

Out-of-hospital cardiac arrest (OHCA) remains a global health burden with persistently low survival rates, despite advancements in emergency care. For example, in the United States, survival to hospital discharge following emergency medical service–treated OHCA is approximately 9% to 10% [1,2], while in Asia, the survival rate ranges between 3% and 8% [3,4]. With the increasing implementation of guideline-based resuscitation, a rising number of OHCA patients survive until hospital discharge. These patients require long-term follow-up [5].

Recognizing this clinical shift, the 2020 American Heart Association Guidelines introduced "recovery" as the sixth link in the chain of survival, emphasizing the importance of long-term surveillance and care for OHCA survivors [6]. However, there has been limited attention to the long-term trajectories of these survivors, particularly those who experience second OHCA events postdischarge [7,8]. Of the OHCA survivors, 6% to 10% experience a second cardiac arrest, often within the first year of discharge [7,9]. While a second OHCA is associated with exceedingly high mortality, data regarding patients who suffer a second OHCA and survive remain scarce. No study has systematically compared the survival patterns and baseline characteristics of patients who survive 1 OHCA with those who survive 2.

Survivors of 1-OHCA face increased risks of neurological injury, cardiovascular events, and long-term mortality [10,11]. Surviving 2 OHCA episodes may lead to greater physiological compromise, including cumulative hypoxic brain injury and myocardial dysfunction, which could profoundly influence long-term survival. Understanding whether the trajectories of these 2 groups diverge in terms of both survival curves and comorbidity profiles is crucial for planning postresuscitation care, developing rehabilitation strategies, and allocating resources.

Identifying patients who survive 2 OHCA events poses a substantial methodological challenge owing to the rarity of such events. Indeed, most single-center or regional databases lack sufficient sample sizes to permit meaningful analyses. To overcome this limitation, we leveraged Taiwan's National Health Insurance Research Database (NHIRD), a comprehensive nationwide claims-based database that captures nearly all medical events among 23 million residents [12].

From a public health surveillance perspective, 2-OHCA survivors represent a rare but high-risk population that is difficult to characterize using single-center registries or conventional clinical cohorts. Nationwide administrative claims databases can function as longitudinal health surveillance infrastructure by enabling identification of uncommon survivor populations, monitoring of long-term mortality, and assessment of postdischarge care dependency at the population level. Using a nationwide claims-based surveillance database, this study aimed to compare clinical characteristics, health care use, life-sustaining device dependence, and subsequent postdischarge survival between adult 1-OHCA and 2-OHCA survivors.

Methods

Study Design and Setting

This retrospective study used data from the Taiwan NHIRD. Established in 1995 under the nation's single-payer National Health Insurance program, the NHIRD contains comprehensive claims information for more than 23 million citizens, covering >99% of Taiwan's population [13]. These data include registries of insured individuals, diagnostic codes, outpatient and inpatient visits, prescriptions, and other relevant health care variables, forming a powerful resource for population-based and epidemiological research [13,14]. Additionally, the NHIRD can be linked to disease-specific registries and other government datasets, substantially augmenting real-world evidence generation [15]. To protect patient privacy, rigorous deidentification procedures and ethical oversight protocols are in place, ensuring that individual identities remain anonymous while allowing researchers to conduct robust retrospective analyses.

Ethical Considerations

This study was approved by the institutional review board of the National Taiwan University Hospital (approval number 202502001RINC). The requirement for informed consent was waived by the institutional review board because this was a retrospective analysis using anonymized data. To protect patient privacy and confidentiality, all data used in the analysis were anonymized before access by the study investigators, and no directly identifiable personal information, such as names, medical record numbers, or other individual identifiers, was included in the analytic dataset or reported in the manuscript. The data were used solely for the purposes approved by the institutional review board. Because this study involved secondary analysis of existing anonymized data and did not involve direct contact with patients or prospective participant enrollment, no compensation was provided to participants. All study procedures adhered to relevant ethical standards and regulations, including the principles outlined in the Declaration of Helsinki [16].

Study Population and Eligibility Criteria

We identified patients who experienced OHCA between 2010 and 2020 by querying the NHIRD. To be included, individuals had to present to the emergency department (ED) with a triage level of 1 and documented *International Classification of Diseases* (*ICD*) codes reflecting cardiac arrest (*ICD-9*: 427.4, 427.5, 798, 798.1, 798.2, 798.9, and 799.9; *ICD-10*: I46, I46.2,

XSL•FO
RenderX

I46.8, I46.9, and I49.0). Both *ICD-9* and *ICD-10* were used owing to the nationwide transition from the former coding system to the latter, which occurred after 2016. We excluded cases in which OHCA was attributed to traumatic events, individuals aged <18 years, and instances in which the recorded date of death preceded the OHCA. These criteria were designed to focus on the analysis of nontraumatic OHCA and ensure proper assessment of the risk factors for subsequent OHCA among survivors.

Cohort Definition: 1-OHCA and 2-OHCA Survivors

First-event OHCA survivors (1-OHCA survivors) were defined as patients who had exactly 1 documented nontraumatic OHCA episode during the study period, survived to hospital discharge after that episode, and had no subsequent documented OHCA during follow-up. Second event OHCA survivors (2-OHCA survivors) were defined as patients who had exactly 2 distinct documented nontraumatic OHCA episodes during the study period and survived to hospital discharge after both events. In this study, survival to hospital discharge was operationalized as the absence of a death record in the NHIRD at the time of discharge. These definitions ensured that the 2 analytic groups were mutually exclusive: patients classified as 2-OHCA survivors were not counted as 1-OHCA survivors. Patients with 3 or more documented OHCA episodes were not included in the analytic cohort because such cases were exceedingly rare, typically limited to single-digit counts, and were subject to NHIRD privacy protection restrictions for rare and potentially identifiable cases. Accordingly, we used the term 2-OHCA survivors, rather than ≥2-OHCA survivors, throughout the revised manuscript to accurately reflect the analytic cohort.

Covariates and Operational Definitions

For all covariates used in the mortality models, the anchor time was defined as the discharge date after the qualifying OHCA event: discharge after the first OHCA for 1-OHCA survivors and discharge after the second OHCA for 2-OHCA survivors. All covariates were entered into the Cox models as fixed baseline variables and were not updated during follow-up.

The variables were grouped into 4 main categories—demographic characteristics, comorbidities, health care use, and daily functional status—to characterize differences between 1-OHCA and 2-OHCA survivors and to evaluate factors associated with postdischarge mortality within each survivor group. Demographic characteristics included age and sex. Comorbidities included diabetes mellitus, hypertension, hepatitis, chronic kidney disease, asthma, peptic ulcer disease, heart disease, cancer, cerebrovascular disease, hyperlipidemia, chronic obstructive pulmonary disease, and liver cirrhosis, each identified using insurance claims and prescription histories before the anchor date. Health care use variables were assessed using prespecified look-back windows before the anchor date, including intensive care unit (ICU) admission within the previous 365 days, renal dialysis for more than 3 consecutive months, more than 3 outpatient visits within the previous 30 days, and more than 2 ED or inpatient visits within the previous 30 days. Daily functional status was approximated using claims-based indicators of nasogastric tube placement, Foley catheter use, and tracheostomy with long-term mechanical ventilation before or at the qualifying discharge, which served as surrogate markers of functional dependency and ongoing care needs at the start of mortality follow-up.

Statistical Analysis

Descriptive statistics were used to summarize baseline characteristics, health care use, comorbidities, and life-sustaining device use. Continuous variables were compared using Student's *t* test, and categorical variables were compared using the chi-square test or Fisher exact test, as appropriate. We performed Kaplan-Meier analyses to describe postdischarge survival trajectories among 1-OHCA and 2-OHCA survivors, and differences in survival distributions were assessed using the log-rank test [17]. For 1-OHCA survivors, follow-up began at discharge after the first OHCA; for 2-OHCA survivors, follow-up began at discharge after the second OHCA. As the 2-OHCA survivor status is conditional on surviving the recurrent event and discharge, the survival curves were interpreted as descriptive postdischarge prognostic comparisons rather than causal estimates of the effect of recurrent OHCA.

To identify factors associated with postdischarge mortality within each survivor group, we first performed univariate Cox proportional hazards regression analyses for candidate variables. Variables with a univariate *P* value <.10 were considered for inclusion in the multivariable Cox proportional hazards models. Clinically important covariates, including age and sex, were also considered for inclusion when appropriate. The 1-OHCA and 2-OHCA groups were modeled separately, and adjusted hazard ratios (aHRs) with 95% CIs were reported. As the 2-OHCA survivor subgroup was relatively small, we used a parsimonious modeling strategy to reduce overfitting and model instability. Variables with sparse distributions, complete or near-complete separation, or unstable estimates were not forced into the final 2-OHCA model. The proportional hazards assumption was assessed for retained covariates using Schoenfeld residuals and graphical inspection of log-minus-log survival plots. No major violation of the proportional hazard assumption was identified [18].

On the basis of clinical considerations, ICU admission and more than 2 ED or ward hospitalizations were not included in the multivariable models because these variables were closely related to disease severity, ventilator use, and subsequent care intensity, and therefore could introduce collinearity or overadjustment. Device use patterns were compared between groups using Cochran-Armitage trend tests. All statistical analyses were performed using SAS software (version 9.4; SAS Institute). Statistical significance was defined as a 2-sided *P* value of <.05.

Results

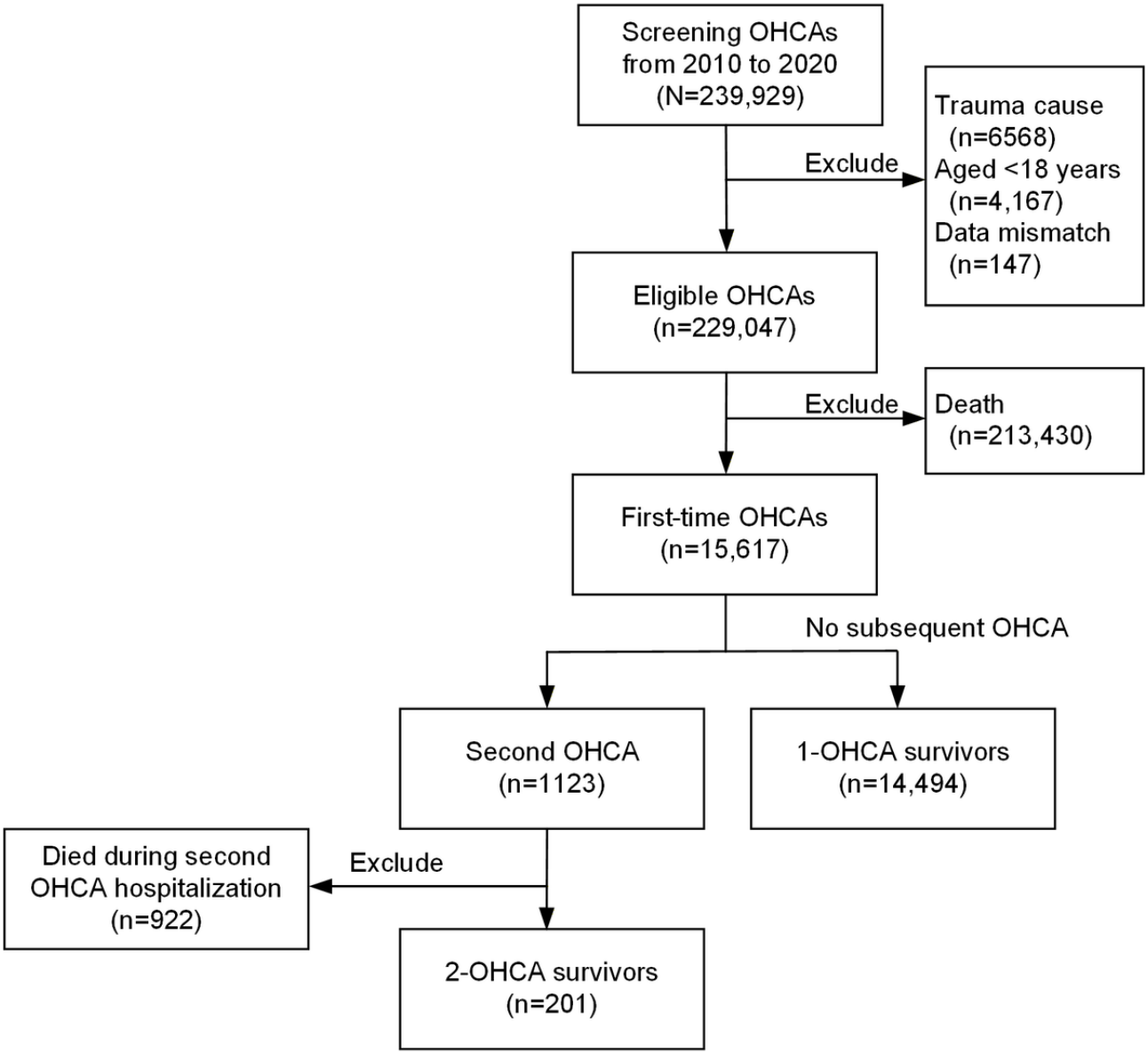
Overview

Figure 1 illustrates the patient selection process for the study cohort from 2010 to 2020. Among the 15,617 patients who survived to discharge after a first OHCA, 14,494 had no documented recurrent OHCA during follow-up and were classified as 1-OHCA survivors. A total of 1123 patients experienced a second OHCA. After excluding patients who died

XSL•FO
RenderX

during the hospitalization for the second OHCA event (n=922), 201 patients who survived to discharge after both OHCA events were classified as 2-OHCA survivors.

Figure 1. Flow diagram of cohort selection and group classification.



Baseline Characteristics of the Groups

Table 1 presents the demographics, health care use, life-sustaining interventions, and comorbidities between 1-OHCA and 2-OHCA survivors. Patients in the 2-OHCA group were older (mean 63.13, SD 13.64 years vs mean 60.95, SD 17.42 years; *P*=.02), but sex distribution was similar between groups (127/201, 63.18% vs 8563/14,494, 59.08%; *P*=.24). Use of health care facilities was consistently higher in the 2-OHCA group, including ICU admission (103/201, 51.24% vs 4897/14,494, 33.79%; *P*<.001) and ED visits and/or ward hospitalizations >2 times (117/201, 58.21% vs 7351/14,494, 50.72%; *P*=.03). The rate of life-sustaining device use was significantly higher in the 2-OHCA cohort: nasogastric tube (148/201, 73.63% vs 6752/14,494, 46.58%; *P*<.001), Foley catheter (138/201, 68.66% vs 6339/14,494, 43.74%; *P*<.001), and tracheostomy with ventilator support (92/201, 45.77% vs 3185/14,494, 21.97%; *P*<.001). However, no significant differences between the groups were found in preexisting comorbidities such as diabetes mellitus, hypertension, chronic kidney disease, and liver cirrhosis.

XSL•FO
RenderX

Table 1. Comparison of characteristics between 1 out-of-hospital cardiac arrest (OHCA) and 2-OHCA survivors[a].

Variables	1-OHCA survivors (n=14,494)	2-OHCA survivors (n=201)	*P* value
Demographics			
Age (y), mean (SD)	60.95 (17.42)	63.13 (13.64)	.02
Sex, male, n (%)	8563 (59.08)	127 (63.18)	.24
Health care use, n (%)			
ICU[b] admission[c]	4897 (33.79)	103 (51.24)	<.001
Hemodialysis	752 (5.19)	13 (6.47)	.41
Outpatient clinics >3[d]	8707 (60.07)	98 (48.76)	.001
ED[e] visits and/or ward hospitalizations >2[d]	7351 (50.72)	117 (58.21)	.03
Life-sustaining indwelling catheter, n (%)			
Nasogastric tube	6752 (46.58)	148 (73.63)	<.001
Foley catheter	6339 (43.74)	138 (68.66)	<.001
Tracheostomy and ventilator support	3185 (21.97)	92 (45.77)	<.001
Preexisting comorbidities, n (%)			
Diabetes mellitus	2931 (20.22)	46 (22.89)	.35
Hypertension	4167 (28.75)	68 (33.83)	.11
Hepatitis	1780 (12.28)	27 (13.43)	.13
Chronic kidney disease	3485 (24.04)	50 (24.88)	.17
Asthma	427 (2.95)	6 (2.99)	.97
Peptic ulcer disease	1027 (7.09)	15 (7.46)	.83
Heart disease	2737 (18.88)	42 (20.9)	.47
Cancers	506 (3.49)	5 (2.49)	.44
Cerebrovascular disease	1243 (8.58)	14 (6.97)	.41
Hyperlipidemia	1724 (11.89)	25 (12.44)	.81
Liver cirrhosis	539 (3.72)	8 (3.98)	.50

[a]The qualifying discharge date was discharge after the first OHCA for 1-OHCA survivors and discharge after the second OHCA for 2-OHCA survivors. 1-OHCA survivors: survivors with only one OHCA event. 2-OHCA survivors: survivors with exact 2 OHCA events.

[b]ICU: intensive care unit.

[c]Assessed within the 365 days before the qualifying discharge date.

[d]Assessed within the 30 days before the qualifying discharge date.

[e]ED: emergency department.

Multivariable Cox Regression Models for Postdischarge Mortality

Table 2 summarizes the multivariable Cox regression analyses evaluating factors associated with postdischarge mortality separately among 1-OHCA and 2-OHCA survivors. In the 1-OHCA group, mortality risk increased with older age (aHR 1.033 per year, 95% CI 1.031-1.035; *P*<.001), male sex (aHR 1.332, 1.268-1.399; *P*<.001), nasogastric tube placement (aHR 2.839, 95% CI 2.622-3.074; *P*<.001), diabetes mellitus (aHR 1.248, 95% CI 1.172-1.328; *P*<.001), chronic kidney disease (aHR 1.358, 95% CI 1.254-1.471; *P*<.001), and cancer (aHR 1.268, 95% CI 1.136-1.414; *P*<.001). Conversely, >3 outpatient visits within 30 days (aHR 0.236, 95% CI 0.224-0.249; *P*<.001), Foley catheter use (aHR 0.905, 95% CI 0.842-0.972; *P*=.006), and tracheostomy or ventilator support (aHR 0.786, 95% CI 0.742-0.832; *P*<.001) were associated with lower observed postdischarge mortality in the multivariable model. These associations should be interpreted as statistical associations rather than evidence of causal protection.

XSL•FO
RenderX

Table 2. Factors associated with postdischarge mortality among 1 out-of-hospital cardiac arrest (OHCA) and 2-OHCA survivors[a].

Variables	1-OHCA survivors		2-OHCA survivors	
	aHR[b] (95% CI)	*P* value	aHR (95% CI)	*P* value
Demographics				
Age (y)	1.033 (1.031-1.035)	<.001	Not included[c]	Not included
Sex, male	1.332 (1.268-1.399)	<.001	Not included	Not included
Health care use				
ICU[d] admission[e]	Not included	Not included	Not included	Not included
Hemodialysis	1.061 (0.941-1.196)	.33	Not included	Not included
Outpatient clinics >3[f]	0.236 (0.224-0.249)	<.001	0.356 (0.259-0.488)	<.001
ED[g] visits and/or ward hospitalizations >2[f]	Not included	Not included	Not included	Not included
Life-sustaining indwelling catheter				
Nasogastric tube	2.839 (2.622-3.074)	<.001	1.210 (0.699-2.097)	.49
Foley catheter	0.905 (0.842-0.972)	.006	0.469 (0.277-0.795)	.005
Tracheostomy and ventilator support	0.786 (0.742-0.832)	<.001	0.563 (0.392-0.809)	.002
Pre-existing comorbidities				
Diabetes mellitus	1.248 (1.172-1.328)	<.001	Not included	Not included
Hypertension	0.977 (0.923-1.034)	.42	Not included	Not included
Hepatitis	Not included	Not included	Not included	Not included
Chronic kidney disease	1.358 (1.254-1.471)	<.001	1.146 (0.725-1.813)	.55
Asthma	1.002 (0.881-1.141)	.97	Not included	Not included
Peptic ulcer disease	1.038 (0.954-1.131)	.38	Not included	Not included
Heart disease	0.998 (0.939-1.062)	.96	Not included	Not included
Cancers	1.268 (1.136-1.414)	<.001	Not included	Not included
Cerebrovascular disease	1.043 (0.968-1.123)	.26	Not included	Not included
Hyperlipidemia	0.938 (0.869-1.012)	.10	Not included	Not included
Liver cirrhosis	Not included	Not included	Not included	Not included

[a]The qualifying discharge date was discharge after the first OHCA for 1-OHCA survivors and discharge after the second OHCA for 2-OHCA survivors. 1-OHCA survivors: survivors with only 1 OHCA event. 2-OHCA survivors: survivors with exact 2 OHCA events. Variables were retained in each multivariable Cox model based on univariate screening, clinical relevance, and model stability. Sparse variables with unstable estimates were not included in the final 2-OHCA model.

[b]aHR: adjusted hazard ratio.

[c]Indicates variables not retained in the final multivariable Cox model.

[d]ICU: intensive care unit.

[e]Assessed within the 365 days before the qualifying discharge date.

[f]Assessed within the 30 days before the qualifying discharge date.

[g]ED: emergency department.

In the 2-OHCA group, although age and sex were clinically important covariates, they were not significantly associated with postdischarge mortality in the univariate Cox regression analysis among 2-OHCA survivors. Given the limited sample size of this subgroup, we did not force these variables into the final model to avoid overfitting and unstable estimates. Lower observed postdischarge mortality was associated with >3 outpatient visits (aHR 0.356, 95% CI 0.259-0.488; *P*<.001), Foley catheter use (aHR 0.469, 95% CI 0.277-0.795; *P*=.005), and tracheostomy or ventilator support (aHR 0.563, 95% CI 0.392-0.809; *P*=.002). No comorbidity was significantly associated with mortality in this subgroup.

The final Cox model for 1-OHCA survivors included 6850 mortality events among 14,494 patients (47.3%). The final Cox model for 2-OHCA survivors included 191 mortality events among 201 patients (95.0%). As the 2-OHCA survivor subgroup had a small overall sample size despite a high number of mortality events, we used a parsimonious modeling strategy to reduce overfitting and model instability. Variables with sparse

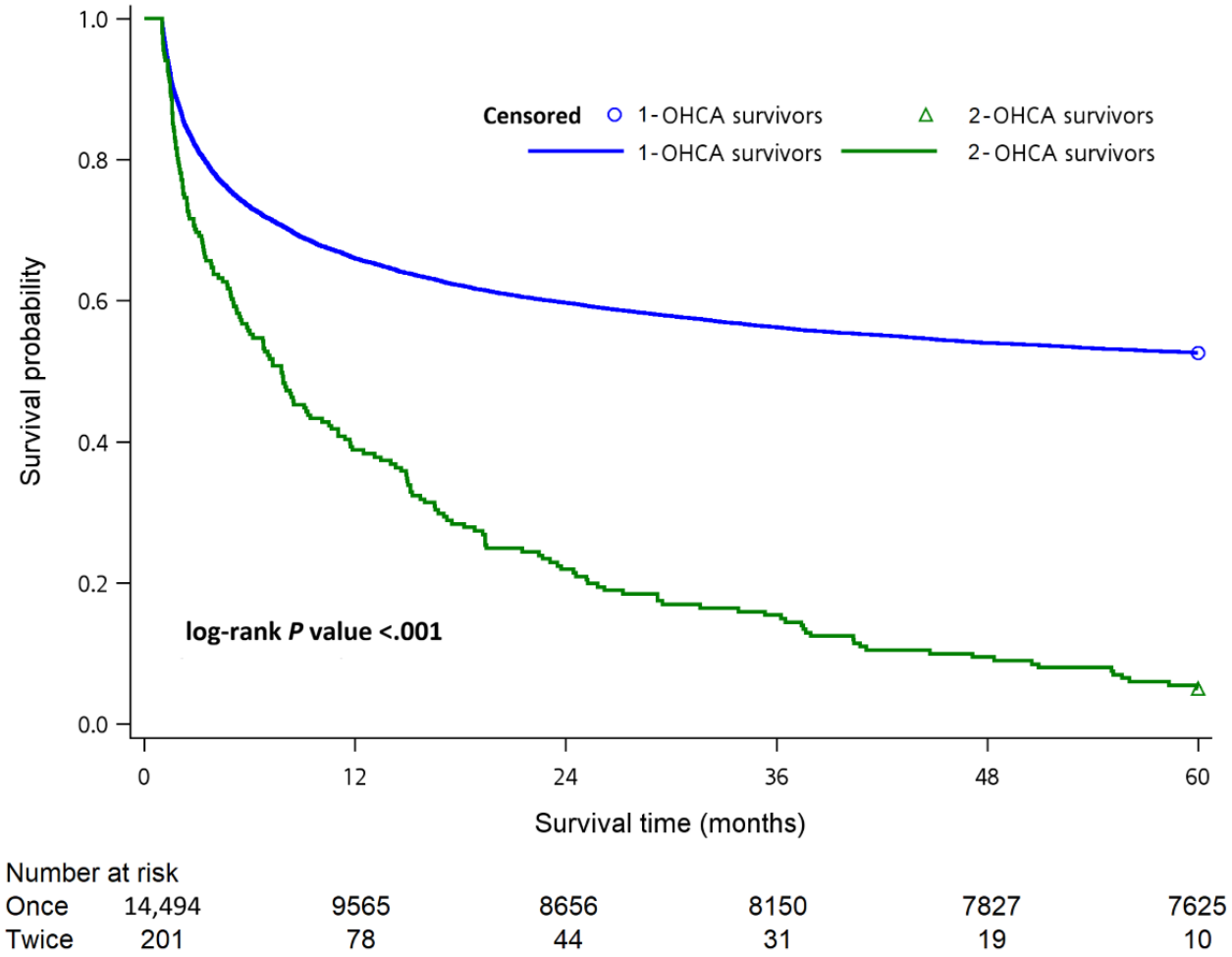
XSL•FO
RenderX

distributions, complete or near-complete separation, or unstable estimates were not forced into the final 2-OHCA model.

Survivors of 2-OHCA Had Significantly Worse Long-Term Survival

Figure 2 shows the Kaplan-Meier curves describing survival after discharge from the qualifying OHCA event. For 1-OHCA survivors, follow-up began at discharge after the first OHCA; for 2-OHCA survivors, follow-up began at discharge after the second OHCA. The 1-OHCA survivor cohort demonstrated a consistently higher postdischarge survival probability than the 2-OHCA survivor cohort. The survival curve of the 2-OHCA group declined steeply during the first year after discharge from the second OHCA. By 60 months, survival remained substantially lower in the 2-OHCA group than in the 1-OHCA group. The log-rank test showed a statistically significant difference between the survivor-defined groups (*P*<.001). As 2-OHCA group membership is conditional on surviving the second OHCA hospitalization, this comparison should be interpreted as a descriptive prognostic comparison rather than a causal estimate of the effect of recurrent OHCA.

Figure 2. Kaplan-Meier survival curves for 1 out-of-hospital cardiac arrest (OHCA) and 2-OHCA survivors.



The Distribution of Life-Sustaining Indwelling Device Patterns Among Survivors

Table 3 compares patterns of life-sustaining indwelling device use between 1-OHCA and 2-OHCA survivors. The rate of only Foley catheter use was lower in the 2-OHCA survivors compared with 1-OHCA survivors (5/201, 2.49% vs 704/14,494, 4.86%; *P*=.02). The rate of combined nasogastric tube and Foley catheter use was higher in 2-OHCA survivors (50/201, 24.88% vs 2938/14,494, 20.27%; *P*=.02). Most notably, the rate of using all 3 indwelling devices was significantly higher in 2-OHCA survivors (81/201, 40.30% vs 2642/14,494, 18.23%; *P*<.001), while the rate of patients requiring none of the catheters was markedly lower (45/201, 22.39% vs 6922/14,494, 47.76%; *P*<.001). Other patterns, including nasogastric tube-only, ventilator-only, and 2-device combinations, did not significantly differ between the groups.

XSL•FO
RenderX

Table 3. Patterns of life-sustaining indwelling catheter in 1 out-of-hospital cardiac arrest (OHCA) versus 2-OHCA survivors.

	1-OHCA survivors, n (%)	2-OHCA survivors, n (%)	*P* value
Only nasogastric tube	745 (5.14)	9 (4.48)	.55
Only Foley catheter	704 (4.86)	5 (2.49)	.02
Only ventilator	61 (0.42)	1 (0.50)	.81
Nasogastric tube combined with Foley catheter	2938 (20.27)	50 (24.88)	.02
Foley catheter combined with ventilator	55 (0.38)	2 (1.00)	.05
Nasogastric tube combined with ventilator	427 (2.95)	8 (3.98)	.22
Combination of all 3 life-sustaining indwelling catheters	2642 (18.23)	81 (40.30)	<.001
None of the above	6922 (47.76)	45 (22.39)	<.001

Discussion

Principal Findings

In this nationwide cohort, patients who survived discharge after a second OHCA had substantially poorer subsequent postdischarge survival than patients who survived a single OHCA and had no documented recurrence. As 2-OHCA survivor status was conditional on survival after the first OHCA, occurrence of a second OHCA, and survival to discharge after the second event, this comparison should be interpreted as a descriptive postdischarge prognostic comparison between survivor-defined groups rather than as evidence that the second OHCA itself caused poorer long-term survival. Moreover, the factors statistically associated with mortality in this group differed from those observed in 1-OHCA survivors and included claims-based indicators of life-sustaining device use. These device-related variables should be interpreted as markers of postdischarge dependency, care setting, and survivorship rather than as causal determinants of survival. Marked disparities in device dependence were also observed, whereas a substantial proportion of 1-OHCA survivors lived without supportive devices, and fewer than one-fourth of 2-OHCA survivors achieved device-free survival.

Clinical Registry Versus Population-Level Registry

The study by Awad et al [19] examined 1-year survival among hospital discharge OHCA survivors using linked clinical registry and administrative data in British Columbia and were able to adjust for key OHCA characteristics and prehospital variables, including arrest location, witnessed status, bystander cardiopulmonary resuscitation, initial rhythm, and in-hospital interventions. In contrast, the NHIRD provides nationwide longitudinal coverage that allowed us to identify the extremely rare 2-OHCA survivor population, but it lacks granular resuscitation variables and direct neurological outcomes. This contrast highlights the trade-off between clinical granularity and population-level scale. Additionally, these findings highlight the value of population-level claims data for postarrest surveillance, particularly for rare survivor-defined groups that are difficult to evaluate in traditional clinical datasets.

Explaining the Divergent Regression Models

The 2-OHCA survivors in our cohort constitute an even rarer, doubly selected population: those who survived the initial OHCA experienced a subsequent event and remained observable during the 5-year follow-up. This conditioning introduces a "depletion of susceptibles" effect [20] and attenuates the prognostic value of baseline comorbidities. Accordingly, traditional covariates (eg, age, male sex, diabetes, chronic kidney disease, and cancer) maintain strong associations with mortality in 1-OHCA survivors, whereas in 2-OHCA survivors, these associations are largely diminished or no longer significant. Instead, the current physiological status and context of care have become dominant determinants.

The pattern of device-related variables illustrates this shift. Among 1-OHCA survivors, nasogastric tube use is independently associated with higher mortality, likely reflecting dysphagia, frailty, and aspiration risk. In contrast, among 2-OHCA survivors, the prognostic impact of nasogastric tube use disappears, and survival is more closely associated with stable, organized support systems. Notably, the inverse associations observed for Foley catheter use and tracheostomy with ventilator support should not be interpreted as evidence that device use causally reduces mortality. More plausibly, these device-related variables reflect survivorship conditioning and postdischarge care context. Patients who were stable enough to receive long-term device support and survive to hospital discharge had already passed through the highest-risk phase after OHCA, whereas patients who died early would not have been observed as chronic device users in the postdischarge setting. Therefore, Foley catheter use and tracheostomy with ventilator support may serve as claims-based markers of survivorship rather than protective interventions.

Increased Device Dependence After a Second OHCA

We also observed a significant shift toward increased dependence on life-sustaining indwelling devices in the 2-OHCA group. Single device use has been increasingly replaced by the concurrent use of multiple devices. Among the few individuals who remained device-free after recurrence, the absence of life-sustaining indwelling devices may indicate lower postdischarge dependency. However, because neurological outcomes were not available in the NHIRD, favorable neurological recovery cannot be directly inferred from these claims-based indicators. These observations are consistent with the regression results: in 2-OHCA survivors, mortality risk was more closely associated with current support status rather than with baseline comorbidities.

XSL•FO
RenderX

The distribution of device use patterns significantly differed between 1-OHCA and 2-OHCA survivors. Device dependence is an informative claims-based marker of ongoing vulnerability among 2-OHCA survivors. However, causality cannot be inferred. The prominence of device-related variables in the regression models may reflect disability severity, survivorship, care setting, care intensity, or unmeasured confounding rather than a direct effect of the devices themselves. These findings suggest that postdischarge dependency and care needs should be incorporated into risk stratification and follow-up planning, but they should not be interpreted as evidence that device placement is beneficial. Our findings highlight the importance of incorporating the device status into post-OHCA risk stratification and follow-up planning. The underlying mechanisms linking device dependence, care setting, and long-term survival require further investigation.

Limitations

In this nationwide, population-based dataset, the scale and completeness enabled survival analyses that are rarely feasible even in large multicenter cohorts and clarified how risk structures differ between 1-OHCA and 2-OHCA cohorts. However, several methodological limitations should be emphasized. First, classification as a 2-OHCA survivor was conditional on surviving long enough after the first OHCA to experience a second OHCA and subsequently survive to discharge after the second event. Therefore, 2-OHCA status is inherently time-dependent, and comparison with 1-OHCA survivors may be affected by selection bias, survivorship conditioning, and guarantee-time or immortal-time-like bias. Accordingly, the Kaplan-Meier curves were interpreted as descriptive postdischarge survival trajectories after the qualifying OHCA event, rather than as causal estimates of the effect of a second OHCA. Second, device-related variables were derived from administrative claims and should be interpreted as surrogate indicators of functional dependency, care setting, care intensity, or unmeasured clinical status, rather than as direct measures of physiological benefit or care quality. Therefore, the inverse associations observed for Foley catheter use and tracheostomy or ventilator support should not be interpreted as causal or protective effects of device use. Third, the 2-OHCA survivor group was small, which limited statistical precision and increased the risk of sparse data bias, model instability, and overfitting. The absence of statistically significant associations for baseline comorbidities should not be interpreted as evidence that these conditions are clinically irrelevant, and the subgroup-specific Cox model should be viewed as exploratory and hypothesis-generating. Future studies using larger datasets, time-dependent Cox models, landmark analyses, or multistate models are needed to validate these findings and better estimate the dynamic effect of recurrent OHCA. Finally, confounding by indication is also possible, particularly for device-related variables. For example, tracheostomy with ventilator support may represent necessary ongoing life-sustaining care among chronically dependent patients rather than a directly protective intervention. Additionally, we did not have the granular OHCA clinical variables in the NHIRD, including initial rhythm, witnessed status, bystander cardiopulmonary resuscitation, prehospital resuscitation details, no-flow or low-flow time, and neurological outcome at discharge. These variables are among the strongest predictors of survival after cardiac arrest and could confound the observed associations between comorbidities, device-related variables, and mortality. It is a trade-off because the strength of the NHIRD is its nationwide coverage over a 10-year period, which enabled the identification of the extremely rare group of patients who survived exactly 2 OHCA episodes; however, this population-level scale comes at the cost of limited clinical granularity.

Conclusions

In this nationwide cohort, patients who survived discharge after a second OHCA represented a rare, highly selected, and clinically vulnerable survivor population with poor subsequent 5-year survival. In exploratory subgroup analyses, mortality among 2-OHCA survivors appeared to be more closely associated with life-sustaining device status than with baseline comorbidities; however, this finding should be interpreted cautiously given the small subgroup size, model-selection constraints, and potential sparse data bias.

Acknowledgments

During the preparation of this work, the authors used ChatGPT (GPT-5, OpenAI) solely to assist with grammar, spelling, and language fluency. The tool was not used to generate or analyze data, perform statistical analyses, interpret results, or draw scientific conclusions. All artificial intelligence–assisted text was critically reviewed, edited, and verified by the authors, who take full responsibility for the accuracy, integrity, and final content of the manuscript.

Funding

This study was funded by the National Science and Technology Council (grant 113-2314-B-002-182-MY3) and the National Taiwan University Hospital Hsin-Chu Branch (grant 115-GH008).

Data Availability

The individual-level National Health Insurance Research Database data used in this study are subject to governmental data access restrictions and privacy regulations. These data are not publicly available and cannot be distributed by the authors. Access requires an application to and approval by the appropriate data governance authority in Taiwan. Aggregated results or analytic code may be available from the corresponding author upon reasonable request, where permitted by regulations.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: CHL, CWS
Data curation: CYF
Formal analysis: CYF
Funding acquisition: CWS
Investigation: CHL, CYF, CWS
Methodology: CHL, CWS
Project administration: CWS
Resources: CYF
Software: CYF
Supervision: EPCH
Validation: CYF
Visualization: CHL
Writing—original draft: CHL
Writing—review and editing: EPCH, CWS

Conflicts of Interest

None declared.

References

1. Benjamin EJ, Muntner P, Alonso A, Bittencourt MS, Callaway CW, Carson AP, et al. Heart disease and stroke statistics-2019 update: a report from the American Heart Association. Circulation 2019 Mar 05;139(10):e56-528 [FREE Full text] [doi: 10.1161/CIR.0000000000000659] [Medline: 30700139]
2. Tsao CW, Aday AW, Almarzooq ZI, Alonso A, Beaton AZ, Bittencourt MS, et al. Heart disease and stroke statistics-2022 update: a report from the American Heart Association. Circulation 2022 Feb 22;145(8):e153-e639 [FREE Full text] [doi: 10.1161/CIR.0000000000001052] [Medline: 35078371]
3. Ong ME, Shin SD, De Souza NN, Tanaka H, Nishiuchi T, Song KJ, et al. Outcomes for out-of-hospital cardiac arrests across 7 countries in Asia: the Pan Asian Resuscitation Outcomes Study (PAROS). Resuscitation 2015 Nov;96:100-108. [doi: 10.1016/j.resuscitation.2015.07.026] [Medline: 26234891]
4. Kitamura T, Iwami T, Kawamura T, Nagao K, Tanaka H, Hiraide A. Nationwide public-access defibrillation in Japan. N Engl J Med 2010 Mar 18;362(11):994-1004 [FREE Full text] [doi: 10.1056/NEJMoa0906644] [Medline: 20237345]
5. Wissenberg M, Folke F, Hansen CM, Lippert FK, Kragholm K, Risgaard B, et al. Survival after out-of-hospital cardiac arrest in relation to age and early identification of patients with minimal chance of long-term survival. Circulation 2015 May 05;131(18):1536-1545. [doi: 10.1161/CIRCULATIONAHA.114.013122] [Medline: 25747933]
6. Merchant RM, Topjian AA, Panchal AR, Cheng A, Aziz K, Berg KM, et al. Part 1: executive summary: 2020 American Heart Association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care. Circulation 2020 Oct 20;142(16_suppl_2):S337-S357 [FREE Full text] [doi: 10.1161/CIR.0000000000000918] [Medline: 33081530]
7. Lam TJ, Yang J, Poh JE, Ong ME, Liu N, Yeo JW, et al. Long term risk of recurrence among survivors of sudden cardiac arrest: a systematic review and meta-analysis. Resuscitation 2022 Jul;176:30-41. [doi: 10.1016/j.resuscitation.2022.04.027] [Medline: 35526728]
8. Nehme Z, Andrew E, Nair R, Bernard S, Smith K. Recurrent out-of-hospital cardiac arrest. Resuscitation 2017 Dec;121:158-165. [doi: 10.1016/j.resuscitation.2017.08.011] [Medline: 28822726]
9. Held EP, Reinier K, Chugh H, Uy-Evanado A, Jui J, Chugh SS. Recurrent out-of-hospital sudden cardiac arrest: prevalence and clinical factors. Circ Arrhythm Electrophysiol 2022 Dec;15(12):e011018 [FREE Full text] [doi: 10.1161/CIRCEP.122.011018] [Medline: 36383377]
10. Phelps R, Dumas F, Maynard C, Silver J, Rea T. Cerebral performance category and long-term prognosis following out-of-hospital cardiac arrest. Crit Care Med 2013 May;41(5):1252-1257. [doi: 10.1097/CCM.0b013e31827ca975] [Medline: 23388519]
11. Hellsén G, Rawshani A, Skoglund K, Bergh N, Råmunddal T, Myredal A, et al. Predicting recurrent cardiac arrest in individuals surviving out-of-hospital cardiac arrest. Resuscitation 2023 Mar;184:109678 [FREE Full text] [doi: 10.1016/j.resuscitation.2022.109678] [Medline: 36581182]
12. Hsieh CY, Su CC, Shao SC, Sung SF, Lin SJ, Kao Yang YH, et al. Taiwan's National Health Insurance Research Database: past and future. Clin Epidemiol 2019 May 3;11:349-358 [FREE Full text] [doi: 10.2147/CLEP.S196293] [Medline: 31118821]
13. Lin LY, Warren-Gash C, Smeeth L, Chen PC. Data resource profile: the National Health Insurance Research Database (NHIRD). Epidemiol Health 2018;40:e2018062 [FREE Full text] [doi: 10.4178/epih.e2018062] [Medline: 30727703]
14. Hsing AW, Ioannidis JP. Nationwide population science: lessons from the Taiwan National Health Insurance Research Database. JAMA Intern Med 2015 Sep;175(9):1527-1529. [doi: 10.1001/jamainternmed.2015.3540] [Medline: 26192815]

XSL•FO
RenderX

15. Chen YC, Yeh HY, Wu JC, Haschler I, Chen TJ, Wetter T. Taiwan's National Health Insurance Research Database: administrative health care database as study object in bibliometrics. Scientometrics 2010 Sep 9;86:365-380. [doi: 10.1007/s11192-010-0289-2]
16. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
17. Goel MK, Khanna P, Kishore J. Understanding survival analysis: Kaplan-Meier estimate. Int J Ayurveda Res 2010 Oct;1(4):274-278 [FREE Full text] [doi: 10.4103/0974-7788.76794] [Medline: 21455458]
18. Bradburn MJ, Clark TG, Love SB, Altman DG. Survival analysis part II: multivariate data analysis--an introduction to concepts and methods. Br J Cancer 2003 Aug 04;89(3):431-436 [FREE Full text] [doi: 10.1038/sj.bjc.6601119] [Medline: 12888808]
19. Awad E, Fordyce CB, Grunau B, Christenson J, Helmer J, Humphries K. One-year survival after out-of-hospital cardiac arrest: sex-based survival analysis in a Canadian population. J Am Coll Emerg Physicians Open 2023 May 9;4(3):e12957 [FREE Full text] [doi: 10.1002/emp2.12957] [Medline: 37180956]
20. Ray GT, Lewis N, Klein NP, Daley MF, Lipsitch M, Fireman B. Depletion-of-susceptibles bias in analyses of intra-season waning of influenza vaccine effectiveness. Clin Infect Dis 2020 Mar 17;70(7):1484-1486 [FREE Full text] [doi: 10.1093/cid/ciz706] [Medline: 31351439]

Abbreviations

aHR: adjusted hazard ratio
ED: emergency department
ICD: International Classification of Diseases
ICU: intensive care unit
NHIRD: National Health Insurance Research Database
OHCA: out-of-hospital cardiac arrest

Edited by C Argyropoulos; submitted 27.Dec.2025; peer-reviewed by T Nojima, A Elmahrouk, E Awad; comments to author 13.Apr.2026; accepted 16.May.2026; published 02.Jun.2026.

Please cite as:
Lin CH, Fan CY, Huang EPC, Sung CW
Long-Term Survival and Life-Sustaining Device Use in Survivors of First and Second Out-of-Hospital Cardiac Arrest: Retrospective Cohort Study
JMIR Public Health Surveill 2026;12:e90416
URL: https://publichealth.jmir.org/2026/1/e90416
doi:10.2196/90416
PMID:

©Chih-Hsien Lin, Cheng-Yi Fan, Edward Pei-Chuan Huang, Chih-Wei Sung. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 02.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Real-World Mapping of Multiple Primary Carcinoma Combinations and Survival Outcomes in Shanghai, China: Retrospective Registry-Based Study

Jianwei Shi[1*], PhD; Yichen Chen[2*], MPH; Chen Chen[3*], MPH; Xiaopan Li[4,5], PhD; Hua Jin[6,7], PhD

[1]School of Public Health, Shanghai Jiaotong University School of Medicine, Shanghai, China
[2]Centers for Disease Control and Prevention, Shanghai, China
[3]Shanghai Jing'an District Jiangning Road Community Health Service Centre, Shanghai, China
[4]Department of Health Management Center, Zhongshan Hospital, Shanghai Medical College of Fudan University, Shanghai, China
[5]Shanghai Key Laboratory of Meteorology and Health, Shanghai, China
[6]Department of General Practice, Yangpu Hospital, Tongji University School of Medicine, Shanghai, China
[7]Shanghai General Practice and Community Health Development Research Center, Shanghai, China
[*]these authors contributed equally

Corresponding Author:
Hua Jin, PhD
Department of General Practice, Yangpu Hospital, Tongji University School of Medicine
450 Tengyue Road
Shanghai, 200090
China
Phone: 86 13764505535
Email: drjinhua@tongji.edu.cn

Abstract

Background: Multiple primary carcinomas (MPC) represent a clinically significant yet underexplored phenomenon, where patients develop more than one distinct primary malignancy. While prior studies have examined MPC within specific cancer types, comprehensive real-world patterns of primary malignancies and their subsequent primary malignancies remain limited. Moreover, the survival outcomes associated with these MPC patterns, particularly in relation to demographic and clinical characteristics, are not well characterized.

Objective: This study aimed to establish the patterns and combinations of MPC across a wide range of cancers and to assess whether the mortality status of patients with MPCs varies according to their demographic characteristics and disease status.

Methods: We conducted a retrospective analysis of 1560 patients with MPC in Shanghai, China, from 2002 to 2015. Data were extracted from the Shanghai Cancer Registry, with follow-up until December 2017. Cause of death was ascertained through linkage with the Shanghai Vital Registration System. The distribution of the frequency and proportion of primary carcinoma (PC) combinations were depicted, and a life table was used to calculate the 1- to 5-year survival rates. Cox regression analysis was performed to analyze the survival risk factors of the first and second PCs.

Results: Among the 1560 patients (809/1560, 51.86% male and 751/1560, 48.14% female), the most frequent first PCs were colorectal, breast, and stomach cancers, while the most frequent second PCs were lung, colorectal, and stomach cancers. The most common combinations included colorectal and lung, colorectal and stomach, and colorectal and prostate. Survival rates were lowest for first PCs of skin (5 years=46.95%) and lung (5 years=41.54%) cancers, and for second PCs of pancreatic (5 year=9.13%) and liver (5 years=14.19%) cancers. A latency period of 12 months between PC diagnoses was associated with significantly higher cancer-specific mortality for both the first primary cancer (hazard ratio [HR] 3.539, 95% CI 2.822-4.438; $P<.001$) and the second primary cancer (HR 1.369, 95% CI 1.103-1.699; $P=.004$). Older age (>65 years) and advanced tumor stage (III+IV) were also significant independent risk factors for poor survival in both first PC (age: HR 2.049, 95% CI 1.689-2.485; $P<.001$; stage: HR 1.496, 95% CI 1.315-1.703; $P<.001$) and second PC (age: HR 1.575, 95% CI 1.242-1.996; $P<.001$; stage: HR 3.933, 95% CI 3.182-4.861; $P<.001$) analyses.

Conclusions: This study provides a comprehensive, real-world map of MPC patterns and highlights important findings: high-risk cancer combinations and key factors associated with poorer survival, including a short interdiagnosis interval (12 months),

XSL•FO
RenderX

advanced age, and advanced tumor stage. Comprehensive prevention and control strategies for MPC should be developed, and clinicians should be aware of the risks of MPC in vulnerable populations during the early diagnosis stage.

(JMIR Public Health Surveill 2026;12:e82355) doi:10.2196/82355

KEYWORDS

multiple primary carcinoma; site; survival rate; risk; real-world maps

Introduction

In recent years, studies have repeatedly corroborated the fact that patients with cancer frequently have comorbid conditions, including cardiovascular disease, obesity, metabolic disease, mental health problems, and musculoskeletal disease [1,2]. Beyond these, the development of a second primary malignant neoplasm has been increasingly recognized as a distinct challenge for survivors with cancer [3,4]. A second primary carcinoma is a new primary cancer that is histologically or clonally unrelated to the first primary carcinoma, and is associated with a poorer prognosis that can substantially compromise long-term survival [5]. There has been significant progress in the prevention, diagnosis, and treatment of various metastatic malignant tumors [6]. However, the existence of primary malignancies and subsequent primary malignancies (multiple primary carcinomas [MPCs]), which differ clinically from metastatic malignancies, is a specific comorbid condition and a current hot topic in research. According to existing studies, patients with MPC are not a recent occurrence; indeed, in 1921, a report found that 4.7% of 3000 cases of malignancy showed “multiple growth” [4,7]. In some epidemiological studies, the frequency of certain MPC was reported to be in the range of 2% to 17% [8,9]. According to the study by Buikhuisen et al [10], which was based on the nationwide Netherlands Cancer Registry from 1989 to 2018, among 2933 patients with primary carcinoma (PC), 425 consecutive primary malignancies were observed in 376 patients [10]. Clinically, MPCs are considered to be of increasing relevance and importance. However, most existing studies in Asian and Western countries analyzed the occurrence of primary cancer of a specific type of cancer, such as lung, renal cell carcinoma, and squamous cell carcinoma of the esophagus [11]. Despite the reports on the frequency of MPC for certain types of cancer, the overall pattern of MPC based on the real-world data is not well established, which may result from the relatively small survey data and sometimes includes inadequate information on diagnosis and treatment [12].

While concerns about recurrence and metastatic disease are prevalent among survivors of common cancers such as lung, colorectal, gastric, and breast cancer [13], emerging research has focused on the investigation of survival outcomes and risk factors specifically associated with second primary cancers. Using the National Cancer Institute’s Surveillance, Epidemiology, and End Results (SEER) database from 1973 to 2002, Das et al. [14] found that 216,751 patients developed primary colorectal cancer (CRC), and over a follow-up period of 1,250,687 person-years, 5,595 developed a second primary CRC, with a standardized incidence ratio of 1.36 (95% CI 1.32-1.39). Similarly, based on the SEER data, Donin et al [15] revealed that among the 1,450,837 survivors of nonpulmonary cancers, 25,472 developed second primary lung cancers at a mean follow-up of 5.7 (SD 3.6) years, and increasing age and being divorced, widowed, or separated were independent risk factors for second primary lung cancer in most primary cancer types [15]. Moreover, the period of diagnosis of the first primary CRC was also found to be an independent risk factor for a subsequent primary CRC, with a relative hazard ratio (HR) of 1.18 (95% CI 1.06-1.31) [14]. However, by reviewing the existing studies, information related to the risks of mortality for the first PC and the second PC is seldom communicated, which may reveal more opportunities in the prevention or detection of subsequent primary cancers in their early stages [16]. In China, despite some efforts to study MPC, existing research is mostly limited to high-incidence tumors such as primary lung cancer and liver cancer [17]. Comprehensive, population-based analyses delineating the spectrum of MPC across all cancer sites are scarce, with few studies leveraging long-term data from cancer registries. Similarly, there is a lack of information on the comprehensive patterns and common combinations of MPC, as well as their specific survival rates and risk factors [18].

Therefore, it is important to provide comprehensive information on the prevalence of MPC patterns, their corresponding survival rates of various cancer sites, and the risk factors for the first and second site MPC. This study aimed to (1) establish the patterns and combinations of MPC across a wide range of cancers using data from the Cancer Registry in Shanghai, China; (2) analyze the survival rates across common MPC categories in the first and second PC sites; and (3) assess whether the mortality status of patients with MPC varied according to the individuals’ characteristics and disease situation.

Methods

Data Source

In our study, data on the diagnosis and survival rates of MPC were collected from the Cancer Registration System of Shanghai, which collects data from 86 public health institutions in Shanghai, China. In China, the uniform standard of diagnosis and survival information of patients with cancer has been collected since the 1970s. The Cancer Registration System was gradually established and achieved complete population coverage for Shanghai residents starting in 2002 [19].

The inclusion of patient data into the registry followed a systematic procedure. According to the registry’s operational protocol, cancer cases were reported by participating institutions. Patients or their families provided informed consent for their data to be included in the population-based cancer registry and for participation in routine follow-up surveys. The survey was conducted annually to check the survival status of patients via

XSL•FO
RenderX

telephone or home visit in accordance with standardized epidemiologic protocols [20]. Additionally, the survival information of patients was obtained from the coroner's registrar and was restricted to Shanghai residents. Cause of death was ascertained through routine linkage with the Shanghai Vital Registration System, and cancer-specific mortality was defined accordingly. The quality of data was checked and evaluated based on the Guidelines for the Chinese Cancer Registration [21] and International Agency for Research on Cancer and International Association of Cancer Registries' data quality criteria [22]. The research team had full access to the deidentified data from the Shanghai Cancer Registry for the purpose of this study.

Data Collection

In this study, the data of patients with MPC were extracted through a systematic, multistep process according to the *ICD-10* (*International Statistical Classification of Diseases and Related Health Problems 10th Revision*) from the Cancer Registration System of Shanghai. The data included patients who were diagnosed with primary cancer (*ICD-10*: C97) in whom the second primary cancer occurred during 2002-2015.

The start year of the first primary cancer was retrospectively established as 1988, considering the data quality of the information of patients with cancer. In detail, the Shanghai Cancer Registration System began collecting complete mortality data in 1988. However, the data from 1988 to 2001 do not include full population registration data. In 2002, the Shanghai Cancer Registration System collected data from all reports of hospitals with tumor diagnostic qualifications and medical records obtained from autopsies and community follow-ups, and the data were further verified by the local Center for Disease Control and Prevention. Given that the supplementary data from 1988 to 2001 were conducive to the analysis of the first primary cancer in this article, these data were incorporated. Data cleaning also confirmed that the quality for this earlier period was acceptable, with minimal missing critical information, justifying its inclusion for a more complete analysis of the first primary cancer.

The starting year was 2002 for the second PC because the system was deemed complete at that point, and the end year was 2015. Due to the time lag in information collection and data quality control, follow-up of deaths ended on December 31, 2017. The Shanghai Cancer Registry has achieved full population coverage, and death registration is a legal requirement related to the cancellation of a resident's household registration and property inheritance. Therefore, the survival status of patients with cancer can be accurately traced to a specific date. Consequently, in our current analysis of patients with cancer from 2002 to 2015, with the endpoint of survival outcome set as December 31, 2017, we effectively avoided underreporting cases and had no cases of survival or loss to follow-up.

Electronic data were used to compare the incidence, mortality, and survival of cancers. Detailed variables, including age, sex, cancer type, date of diagnosis, pathology, treatment, date and cause of death, tumor, node, and metastases stage (TNM stage), and registered residence, were collected. All cancers from the registration system were combined, and 26 cancer types were identified according to the *ICD-10*.

Variables

MPC

MPC refers to the occurrence of two or more independent primary malignant tumors in the same individual, either simultaneously or sequentially. MPCs were identified according to established Warren and Gates criteria, which require that (1) each tumor is histologically confirmed as malignant, (2) each tumor is distinct in its pathological morphology (or, if morphologies are similar, the tumors are anatomically separate), and (3) each tumor is neither a metastasis nor a recurrence of any prior tumor. According to the temporal sequence, MPCs were categorized based on the interval between diagnoses [23].

In this study, we observed two types of MPCs: patients with 2 PCs (first PC [site A] + second PC [site B]), and those with 3 PCs (first PC [site A] + second PC [site B] + third PC [site C]). Due to the low number of PCs (3), we focused on sites A and B when analyzing the pattern, combination, and survival rate using Cox regression.

Demographic Variables of Patients With MPC

We collected data on the following demographics: sex (male and female), age (65 years and 65 years, indicating the older adult group and non–older adult group, respectively), TNM (tumor, node, metastases) stage of cancer (in stage of III and IV, I, and II), diagnosis time (2002-2008, 2009-2015; 2008 was chosen because it was the median of this period), the level of the therapeutic hospital where the treatment was received (secondary hospital vs tertiary hospital in China, all patients with cancer are treated at secondary or tertiary hospitals when they were first diagnosed, since the large hospitals have sufficient expertise in diagnosing tumors), latency between the first and second PC (12 months or 12 months), and number of PCs (2 or 3).

Statistical Analysis

The distribution of the frequency and proportion of first and second PCs was depicted using Adobe Illustrator CS5. A life table method was used to calculate the 1- to 5-year survival rates. Kaplan-Meier analysis was used to estimate the patients' survival status, and Cox regression was performed to identify the risk factors associated with survival, for both first and second PCs. These conventional methods were selected as the primary analytical tools given the descriptive nature of this real-world mapping study. Competing risk regression was not applied, as our primary aim was to describe MPC combination patterns and identify associated survival risk factors in this real-world cohort, rather than to estimate the cumulative incidence function. Cases with missing data for the TNM stage variable were excluded from the Cox regression analyses.

All statistical analyses were 2-sided and performed using the Statistical Package for the Social Sciences software (version 20.0; SPSS, Inc). Statistical significance was set at $P<.05$ (2-sided).

XSL•FO
RenderX

Reporting Guidelines

Results are presented in accordance with the TRIPOD (Transparent Reporting of a multivariable Prediction model for individual Prognosis or Diagnosis) guidelines. STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) and RECORD (Reporting of studies Conducted using Observational Routinely collected health Data) guidelines for observational studies and studies using routinely collected health data were also considered. The study was conducted in accordance with relevant institutional guidelines. The completed RECORD checklist is provided in Multimedia Appendix 1.

Ethical Considerations

This retrospective registry-based study used existing, deidentified data from the Shanghai Cancer Registry. The original data collection for the registry was conducted after the approval of the Institutional Review Board of Shanghai Municipal Center for Disease Control and Prevention, as part of the legally mandated population-based cancer surveillance system authorized by Shanghai public health authorities. For the purpose of this specific research analysis, the study protocol conformed to the ethical guidelines of the 1975 Declaration of Helsinki and was approved by the Ethics Committee of Fudan University, Shanghai, China (IRB00002408 and FWA00002399). The need for additional informed consent for this retrospective analysis was waived by the Ethics Committee, as the study involved the analysis of anonymized data collected for public health purposes. All data were handled confidentially, and no compensation was provided.

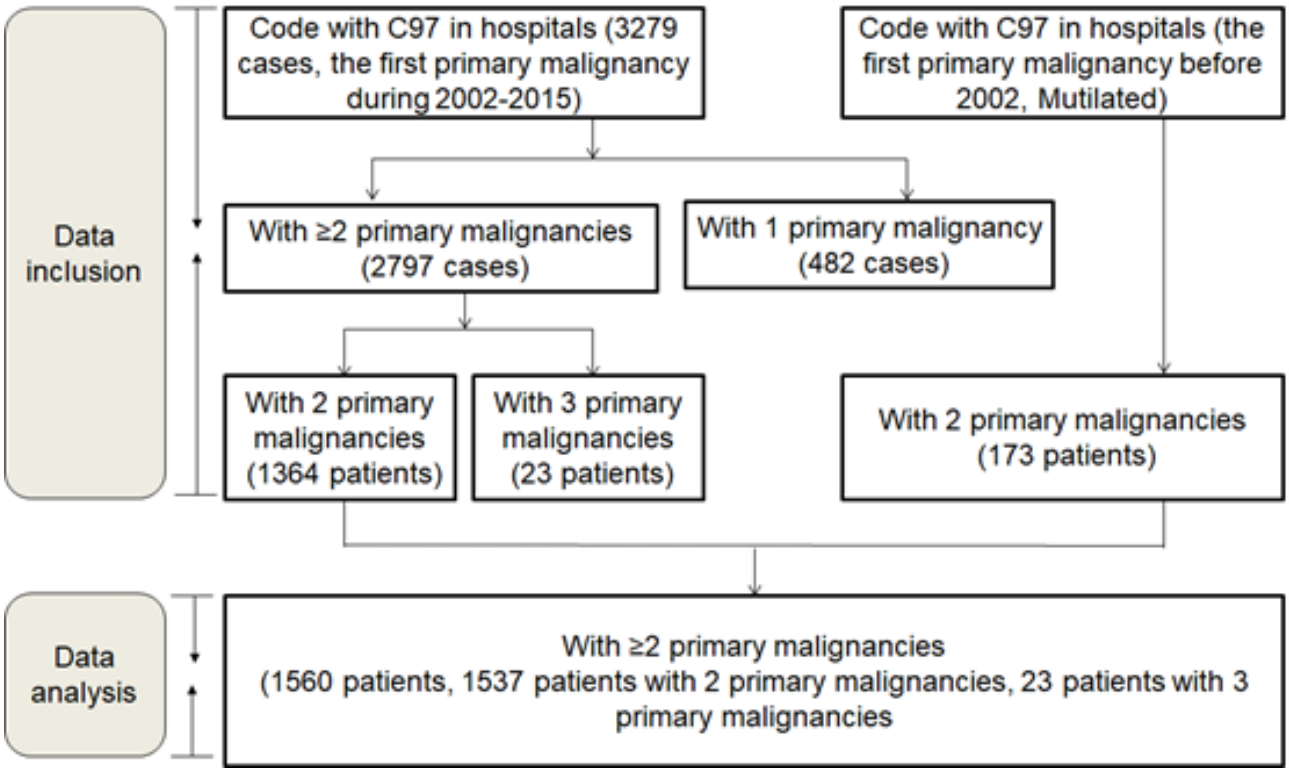
Results

Overview

In this study, data from 1387 patients (2797 cancer cases; cases describing the site of cancer) included the type of the first PC. In addition, 1364 patients (2728 cases) had 2 PCs, and 23 patients (69 cases) had 3 PCs. In total, 173 patients (all of whom had 2 PCs) had their first PC before 2002. Our dataset included a total of 1560 patients with ≥ 2 PCs ([1364 with 2 PCs from 2002 to 2015] + 23 [with 3 PCs from 2002 to 2015] + 173 [with 2 PCs, with the first occurring before 2002]) for the analysis of MPC patterns and survival rates. The median follow-up time for the first and second PC groups was 6.86 years (IQR 4.42-12.18) years and 2.23 (IQR 1.46-5.39) years, respectively (Figure 1).

As shown in Table 1, among the 1560 patients with MPC (1537 with 2 PCs and 23 with 3 PCs), the distribution of male (809/1560, 51.86%) and female (751/1560, 48.14%) patients were similar. The proportion of patients >65 years of age was greater in the 2 PCs group (≤65 years, 590/1560, 38.39%; >65 years, 947/1560, 61.61%; $P<.001$). There was a higher number of patients with I + II with 2 PCs (652/1537, 42.42%) and 3 PCs (11/23, 47.83%) compared to those with III + IV PCs (2 PCs, 249/1537, 16.20%, $P=.004$; 3 PCs, 4/23, 17.39%, $P=.88$). In both groups with 2 and 3 PCs, there were higher proportions of patients who were diagnosed in 2009-2015 (2 PCs, 944/1537, 61.42%, $P=.40$; 3 PCs, 15/23, 65.22%, $P=.02$) and in secondary hospitals (2 PCs, 1067/1537, 69.42%, $P=.44$; 3 PCs, 15/23, 65.22%, $P=.88$). In addition, our results showed that 74.50% (1145/1537) of patients had a 12-month interval between the first and second PC diagnosis in the 2 PCs group and 78.26% (18/23) of patients in the 3 PCs group.

Figure 1. Flowchart of patient selection and data sources for the retrospective registry-based study of multiple primary carcinoma in Shanghai, China.



XSL•FO
RenderX

Table 1. Baseline demographic and clinical characteristics of 1560 patients with multiple primary carcinoma in Shanghai, China (2002-2015).

Variables	With two primary carcinomas (n=1537)					With three primary carcinomas (n=23)					Total (n=1560)
	Male	Female	Total	Chi-square (*df*)	*P* value	Male	Female	Total	Chi-square (*df*)	*P* value	Total
Age (years), n (%)				64.19 (1)	<.001				0.17 (1)	.28	
≤65 years	230 (28.82)	360 (48.71)	590 (38.39)			3 (27.27)	7 (58.33)	10 (43.48)			600 (38.46)
>65 years	568 (71.18)	379 (51.29)	947 (61.61)			8 (72.73)	5 (41.67)	13 (56.52)			960 (61.54)
TNM[a] stage, n (%)				8.18 (2)	.004				0.02 (2)	.88	
III+IV	143 (17.92)	106 (14.34)	249 (16.20)			2 (18.18)	2 (16.67)	4 (17.39)			253 (16.22)
I+II	305 (38.22)	347 (46.96)	652 (42.42)			5 (45.45)	6 (50)	11 (47.83)			663 (42.50)
Unknown	350 (43.86)	286 (38.70)	636 (41.38)			4 (36.36)	4 (33.33)	8 (34.78)			644 (41.28)
Diagnosis time (years) , n (%)				0.73 (1)	.40				0.02 (1)	.88	
2009-2015	482 (60.40)	462 (62.52)	944 (61.42)			7 (63.64)	8 (66.67)	15 (65.22)			959 (61.47)
2002-2008	316 (39.60)	277 (37.48)	593 (38.58)			4 (36.36)	4 (33.33)	8 (34.78)			601 (38.53)
Therapeutic hospital level , n (%)				0.61 (1)	.44				0.02 (1)	.88	
Secondary	561 (70.30)	506 (68.47)	1067 (69.42)			7 (63.64)	8 (66.67)	15 (65.22)			1082 (69.36)
Tertiary	237 (29.70)	233 (31.53)	470 (30.58)			4 (36.36)	4 (33.33)	8 (34.78)			478 (30.64)
Latency , n (%)				2.89 (1)	.09				0.38 (1)	.54	
≤12 months	580 (72.68)	565 (76.45)	1145 (74.50)			8 (72.73)	10 (83.33)	18 (78.26)			1163 (74.55)
>12 months	218 (27.32)	174 (23.55)	392 (25.50)			3 (27.27)	2 (16.67)	5 (21.74)			397 (25.45)

[a]TNM: tumor, node, and metastases stage.

MPC Combination Patterns in the First and Second PC Sites

Figure 2 indicates the specific cancer distribution among the 1560 patients with MPC. The results show that the top three first PCs were accumulated in the colon (204/1560, 13.08%), breast (181/1560, 11.60%), and stomach (145/1560, 9.29%). With CRC as the first PC, the second PCs included the lung (49/1560, 24.02%), stomach (27/1560, 13.24%), prostate (14/1560, 6.86%), and breast (14/1560, 6.86%). For patients with breast cancer as the first PC, thyroid (34/1560, 18.78%), colorectal (31/1560, 17.13%), and lung (26/1560, 14.36%) cancers accounted for the highest proportions of second PCs. Stomach cancers as the first PC most commonly occurred with colorectal (41/1560, 28.28%), lung (27/1560, 18.62%), and skin (10/1560, 6.90%) cancer as second PCs. The patterns of all cancers are shown in Multimedia Appendix 2.

The top three secondary PCs were lung cancer (274/1560, 17.56%), colorectal cancer (254/1560, 16.28%), and stomach cancer (142/1560, 9.10%), which collectively accounted for 42.94% (670/1560) of all secondary PCs. The most frequently observed combinations with lung cancer as the secondary PC were colorectal and lung cancer (49/274, 17.88%), stomach and lung cancer (27/274, 9.85%), and breast and lung cancer (26/274, 9.49%). Overall, stomach and colorectal cancer (41/254, 16.14%) and breast and colorectal cancer (31/254, 12.20%) were the most frequently observed MPC combinations. With stomach cancers as the secondary PC, colorectal and stomach (27/142, 19.01%), lung and stomach (18/142, 12.68%), and bladder and stomach (12/142, 8.45%) cancers accounted for the highest percentage. The complete combination patterns of all cancers are shown in Multimedia Appendix 3.

As depicted in Figure 3, the colon, stomach, and lung were the most frequently observed first PC sites for both male and female patients. Cancers of the reproductive system were also frequently diagnosed, with prostate cancer ranked the third most frequently observed (86/809, 10.63%) in male patients, and breast cancer ranked first (178/751, 23.70%) in female patients. Moreover, thyroid (ranked third), brain and central nervous system (ranked sixth), corpus uteri (ranked seventh), and ovary (ranked eighth)

XSL•FO
RenderX

cancers were more common in female patients, while bladder (ranked fifth), kidney (ranked sixth), and liver cancers (ranked seventh) were more common in male patients. Lung cancer was the most frequently diagnosed secondary PC in men (166/809, 20.49%) and the second most frequent in women (108/751, 14.40%). Male patients also demonstrated a higher rate of esophageal cancer (ranked eighth) as a second PC than as a first PC. In women, liver (ranked sixth), pancreatic (ranked seventh), and gallbladder (ranked eighth) cancers were more commonly observed as second PCs than first PCs.

Figure 2. Distribution of first and second primary carcinoma sites among 1560 patients with multiple primary carcinoma in Shanghai, China (2002-2015).

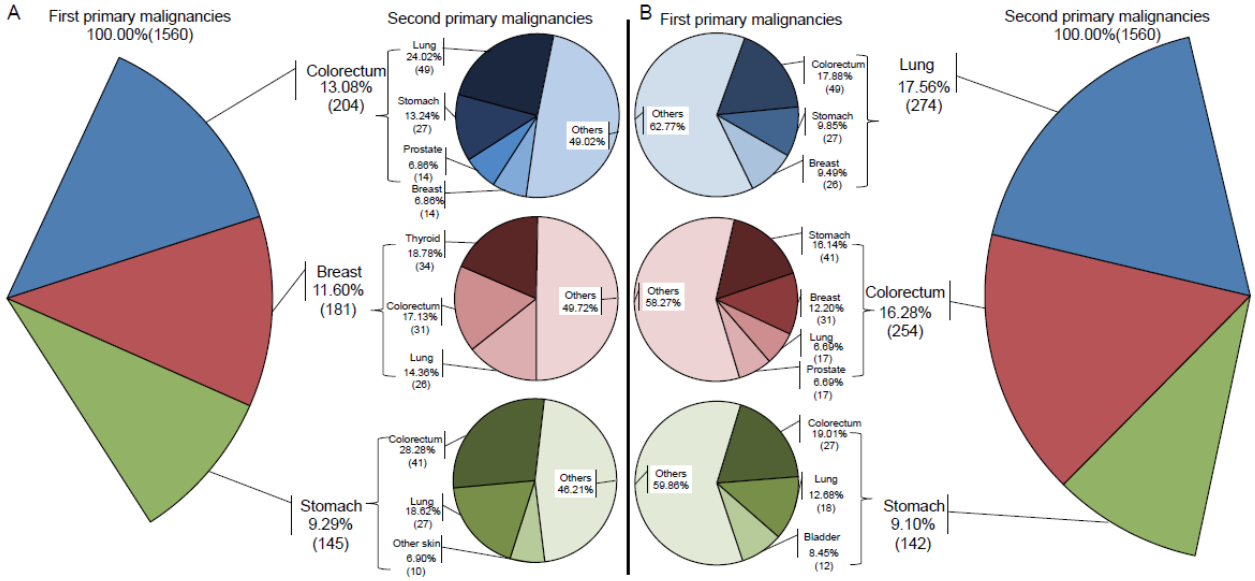
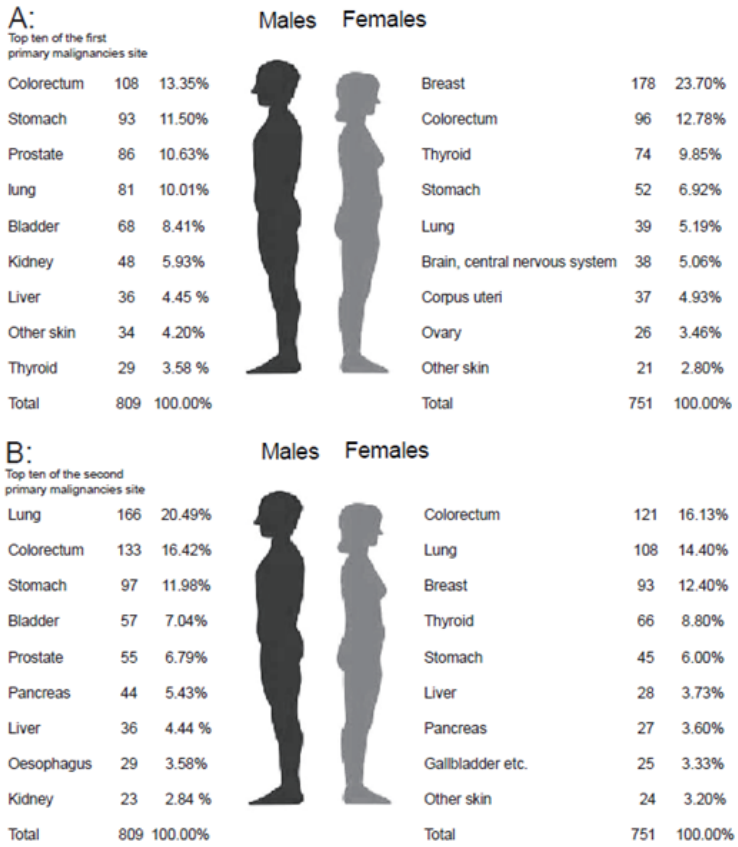


Figure 3. Sex-specific distribution of first and second primary carcinoma sites among 1560 patients with MPC in Shanghai, China (2002-2015). A indicates the first PC, and B indicates the second PC.



A:
Top ten of the first primary malignancies site

Males			Females		
Colorectum	108	13.35%	Breast	178	23.70%
Stomach	93	11.50%	Colorectum	96	12.78%
Prostate	86	10.63%	Thyroid	74	9.85%
lung	81	10.01%	Stomach	52	6.92%
Bladder	68	8.41%	Lung	39	5.19%
Kidney	48	5.93%	Brain, central nervous system	38	5.06%
Liver	36	4.45 %	Corpus uteri	37	4.93%
Other skin	34	4.20%	Ovary	26	3.46%
Thyroid	29	3.58 %	Other skin	21	2.80%
Total	809	100.00%	Total	751	100.00%

B:
Top ten of the second primary malignancies site

Males			Females		
Lung	166	20.49%	Colorectum	121	16.13%
Colorectum	133	16.42%	Lung	108	14.40%
Stomach	97	11.98%	Breast	93	12.40%
Bladder	57	7.04%	Thyroid	66	8.80%
Prostate	55	6.79%	Stomach	45	6.00%
Pancreas	44	5.43%	Liver	28	3.73%
Liver	36	4.44 %	Pancreas	27	3.60%
Oesophagus	29	3.58%	Gallbladder etc.	25	3.33%
Kidney	23	2.84 %	Other skin	24	3.20%
Total	809	100.00%	Total	751	100.00%

XSL•FO
RenderX

Survival Rates According to the Type of Cancer in the First and Second PC Sites

The survival rates for the first and second PCs are displayed in Table 2. Focusing on the first PC site, patients with breast cancer had the highest 1-year survival rate (96.69%). Patients with thyroid cancer as the first PC had the longest 2-year (93.12%), 3-year (90.01%), 4-year (86.76%), and 5-year (83.09%) survival rates. Although ranked 10th most frequently observed, patients with brain and central nervous system cancer as the first PC had relatively high 1- and 2-year survival rates. Patients with kidney cancer as the first PC had high 3-, 4-, and 5-year survival rates. It is worth noting that skin and lung cancers were associated with the lowest survival rates in patients.

Similarly, patients with thyroid cancer as the second PC had the highest survival rates (1-year, 95.04%; 2-year, 95.04%; 3-year, 90.77%; 4-year, 84.61%; 5-year, 84.61%), followed by breast cancer. Comparatively, patients with pancreatic cancer as the second PC had the lowest survival rates (1-year, 16.90%; 2-year, 14.08%; 3-year, 14.08%; 4-year, 11.74%; 5-year, 9.13%), which were similar to those of liver cancer as the second PC (1-year, 32.81%; 2-year, 26.08%; 3-year, 17.74%; 4-year, 17.74%; 5-year, 14.19%).

Table 2. One- to five-year survival rates for first and second primary carcinoma sites among patients with multiple primary carcinomas in Shanghai, China (2002-2015).

	First primary carcinoma site		Survival rates (%)					Second primary carcinoma site		Survival rates (%)				
Rank	Sites	N	1-year	2-year	3-year	4-year	5-year	Sites	N	1-year	2-year	3-year	4-year	5-year
1	Colorectum	204	88.24	85.28	81.26	73.40	66.20	Lung	274	44.14	35.77	31.05	27.78	25.99
2	Breast	181	96.69	92.81	90.01	85.99	81.72	Colorectum	254	66.39	55.04	47.21	40.36	37.37
3	Stomach	145	84.04	80.47	76.04	71.50	69.17	Stomach	142	37.67	33.75	29.12	25.76	20.61
4	Lung	120	69.70	61.91	57.36	49.65	41.54	Breast	98	85.49	84.35	78.95	75.59	68.62
5	Thyroid	103	95.11	93.12	90.01	86.76	83.09	Thyroid	81	95.04	95.04	90.77	84.61	84.61
6	Prostate	86	85.99	77.69	67.98	62.89	57.60	Bladder	72	62.43	55.86	51.64	51.64	44.52
7	Bladder	80	83.75	76.25	76.25	69.73	67.02	Pancreas	71	16.90	14.08	14.08	11.74	9.13
8	Kidney	65	92.31	89.23	87.64	82.43	80.56	Liver	64	32.81	26.08	17.74	17.74	14.19
9	Other skin	55	76.36	67.27	60.00	54.46	46.95	Prostate	55	67.13	58.74	53.27	37.49	37.49
10	Brain, CNS[a]	51	94.12	90.11	86.02	81.92	75.78	Other skin	43	74.10	63.12	42.08	37.40	32.06
__[b]	Total	1560	85.28	80.35	76.41	70.85	65.62	Total	1560	56.18	49.03	43.06	39.30	36.19
—	Patients with 2 PCs[c]	1537	85.19	80.18	76.25	70.96	65.78	Patients with 2 PCs	1537	55.84	48.64	42.73	39.38	36.67
—	Patients with 3 PCs	23	91.30	91.30	86.96	65.22	56.52	Patients with 3 PCs	23	73.91	69.57	60.59	41.20	24.72

[a]CNS: central nervous system.

[d]Not applicable.

[c]PC: primary carcinoma.

Possible Risk Factors for Survival Status in the First and Second PC Sites

The Kaplan-Meier analyses show that there were significant differences in overall and cancer-specific survival rates with regard to the first PC between sex, age, TNM, and diagnosis time and latency (P<.001; Figure S1A in Multimedia Appendix 4). The survival disparity between latency >12 months and < 12 months was the most sizeable. When assessing the second PC (Figure S1B in Multimedia Appendix 4), there was no significant survival difference with diagnosis time as the variable (P>.05), but the survival rates differed significantly based on the level of hospital where treatment was sought (P<.001).

Table 3 demonstrates the results of Cox regression analysis, showing that for the first PC site, there was a significantly higher HR of cancer-specific mortality in the >65-year-old group (HR 2.049, 95% CI 1.689-2.485; P<.001) in the patients with stage III+IV (HR 1.496, 95% CI 1.315-1.703; P<.001), with diagnosis times between 2009 and 2015 (HR 2.377, 95% CI 1.874-3.015; P<.001). Similar results were observed when assessing all-cause mortality. For the second PC, patients with an age >65 years (HR 1.575, 95% CI 1.242-1.996; P<.001) and at stage III+IV (HR 3.933, 95% CI 3.182-4.861; P<.001) also showed a higher risk of cancer-specific mortality, as well as all-cause mortality. There were also fewer male patients with higher HR cancer-specific mortality (HR 0.691, 95% CI 0.568-0.840;

XSL•FO
RenderX

P<.001). Notably, patients with an interdiagnosis interval of ≤12 months had a significantly higher risk of both cancer-specific and all-cause mortality from the first primary cancer (cancer-specific: HR 3.539, 95% CI 2.822-4.438; all-cause: HR 3.533, 95% CI 2.824-4.420) and from the second primary cancer (cancer-specific: HR 1.369, 95% CI 1.103-1.699; all-cause: HR 1.339, 95% CI 1.081-1.659), compared to those with an interval of >12 months. Regarding the number of primary cancers, there was no significant difference between the mortality of patients with primary malignancies in the first site and second site (*P*>.05).

Table 3. Cox regression analysis of risk factors for cancer-specific and all-cause mortality in first and second primary carcinoma sites among patients with multiple primary carcinoma[a].

Variables	Cancer-specific mortality		All-cause mortality	
	HR (95% CI)	*P* value	HR (95% CI)	*P* value
In the first primary carcinoma site (n=916)				
Sex (male vs female)	0.842 (0.705-1.005)	.06	0.849 (0.713-1.011)	.07
Age (>65 vs ≤65 years)	2.049 (1.689-2.485)	<.001	2.092 (1.728-2.532)	<.001
TNM[b] stage (III + IV vs I + II)	1.496 (1.315-1.703)	<.001	1.497 (1.318-1.700)	<.001
Diagnosis time (2009-2015 vs 2002-2008)	2.377 (1.874-3.015)	<.001	2.302 (1.820-2.913)	<.001
Therapeutic hospital level (secondary vs tertiary)	1.143 (0.937-1.395)	.19	1.149 (0.945-1.399)	.16
Latency (≤12 months vs >12 months)	3.539 (2.822-4.438)	<.001	3.533 (2.824-4.420)	<.001
Number of PCs[c] (with 3 PCs vs with 2 PCs)	1.253 (0.667-2.353)	.48	1.212 (0.645-2.275)	.55
In the second primary carcinoma site (n = 844)				
Sex (male vs female)	0.691 (0.568-0.840)	<.001	0.705 (0.581-0.855)	<.001
Age (>65 vs ≤65 years)	1.575 (1.242-1.996)	<.001	1.562 (1.236-1.973)	<.001
TNM stage (III + IV vs I + II)	3.933 (3.182-4.861)	<.001	3.837 (3.114-4.729)	<.001
Diagnosis time (2009-2015 vs 2002-2008)	1.188 (0.918-1.536)	.19	1.130 (0.879-1.453)	.34
Therapeutic hospital level (secondary vs tertiary)	1.033 (0.848-1.258)	.75	1.035 (0.851-1.259)	.73
Latency (≤12 months vs >12 months)	1.369 (1.103-1.699)	.004	1.339 (1.081-1.659)	.007
Number of PCs (with 3 PCs vs with 2 PCs)	1.041 (0.565-1.917)	.90	1.024 (0.556-1.885)	.94

[a]Note: In the Cox regression for the first primary carcinoma site, 644 samples were missing values for the TNM stage variable (1560-916), so only 916 samples were included. For the second primary cancer, 716 samples (1560-844) were missing data; thus, 844 samples were analyzed.

[b]TNM: tumor, node, and metastases stage.

[c]PC: primary carcinoma.

Discussion

Principal Findings

This study provides the first comprehensive population-based mapping of MPC patterns in Shanghai, China, analyzing their occurrence, combinations, and survival using real-world registry data. By extending follow-up through 2017 and retrospectively including cases dating back to 1988, we obtained long-term observation that offers valuable insight into these relatively rare events. Moreover, the observed patterns and risk factors reflect fundamental biological and demographic influences, supporting their continued clinical relevance beyond the study period.

Our analysis of real-world data revealed that colorectal, stomach, and reproductive cancers (breast cancers in women and prostate cancers in men) were the most frequently diagnosed first PC. A previous study, which compared global cancer statistics data from China to other countries, found that among patients of both sexes, China had the largest proportion of lung, colorectal, and female breast cancer (38.9% of total cases; United Kingdom: 34.8%, United States: 29%), and also had a greater proportion of infection-attributable cancers (17.8% of liver and stomach cancers) [24]. Similarly, the results of our study indicate that the prevalence of digestive tract cancers, such as colorectal and stomach cancers, were higher [25]. Moreover, Feng et al [25] study on countrywide cancer prevalence found that 36.4% of Chinese cancer-related deaths were due to digestive tract cancers (stomach, liver, and esophageal cancer), compared to ≤5% of the total cancer-related deaths in the United States or the United Kingdom [25]. We concluded that primary lung cancer ranked first as the second PC site, but was the fourth most frequently observed PC in the first site, which does not agree with the findings of Feng et al [25] or with other previous studies, in which lung cancer has always ranked first. One reason for the difference in incidence rates between this study and prior studies may be the inadequate number of data sources and research on varying MPCs; thus, further study remains necessary. Additionally, sex-specific differences in MPC distribution were also observed. When organized by sex, the order of the top 10 PCs existed as the first and second PCs between male and female

XSL•FO
RenderX

patients were distinct. Men had higher rates of infection-attributable cancers (stomach and liver), but women had higher rates of thyroid cancer. These patterns reflect the underlying differences in cancer etiology between sexes and may inform site-specific surveillance strategies for survivors with cancer.

When analyzing the PC combinations, we found that PCs including colorectal and lung, colorectal and stomach, colorectal and prostate, and breast and thyroid, breast and colorectal, stomach and colorectal, and stomach and lung were the most frequently observed. Moreover, it was common to observe comorbidities of the same system; for instance, there was a higher probability of cooccurrences of multiple digestive tract cancers (stomach and liver) [26]. Our results revealed patterns among the MPC that cooccurred; for instance, breast cancer appeared frequently with thyroid cancer. However, it is important to note that these observed patterns represent clinical cooccurrences whose frequency may be influenced by the high background incidence or survival rate of the involved cancers. Therefore, future studies incorporating standardized incidence ratios are needed to determine whether these observed combinations represent an elevated risk beyond chance. According to the study by Buikhuisen et al [10], concomitant diagnoses in the first year mainly comprised lung and renal cancer. Metachronous malignancies beyond the first year were most frequently observed for patients with breast, colorectal, prostate, and lung cancer [10]. Furthermore, a study of patients hospitalized at the Department of Oncology from 2003 to 2009 found that the incidence of MPC was 2.4%, and of the 103 patients with MPC, 97 had two, and 6 had three primary tumors. Moreover, the most frequently observed tumor combinations in men were prostate cancer-digestive system malignancies (especially colorectal cancer) and vice versa, as well as hematological malignant tumor-digestive system malignancies. In women, the most frequently observed multiple tumor occurrences were breast cancer, cancer of the contralateral breast, and hematological malignant tumors (especially non-Hodgkin lymphoma)-breast cancer. The observed associations between certain cancers are likely multifactorial in origin. Potential contributing mechanisms include shared hormonal pathways, common genetic susceptibilities, and environmental exposures. Additionally, the potential long-term effects of treatments received for the first PC should also be considered as a possible risk factor for subsequent malignancies in particular organs. Therefore, clinicians should be aware of these frequently observed combinations. Future analytical studies using standardized incidence ratios are warranted to confirm whether these pairs represent an excess risk beyond that expected by chance. Moreover, in this study, the 1-, 2-, and 3-year survival rates of patients with first or second PCs were higher, but the 4-year and 5-year survival rates were lower, which underscores the critical importance of sustained, long-term surveillance beyond the initial treatment period. Moreover, the survival rates of patients with breast and thyroid MPC were relatively high, which is consistent with the findings of previous studies [27]. Building on the patterns of metachronous cancer identified, our findings advocate for a paradigm shift in survivorship care, from a model focused solely on the primary tumor to a comprehensive, lifelong strategy. This necessitates a multidisciplinary approach that moves beyond the initial organ site to implement coordinated surveillance for the most likely subsequent cancers, based on the primary site, shared risk factors, and treatment history, thereby enabling earlier detection and improving long-term prognosis for survivors. For instance, given the high frequency of lung cancer as a second primary, future research should evaluate the potential role of enhanced surveillance strategies, such as low-dose chest computed tomography, in high-risk survivor populations.

Our results showed that, among the first PCs, skin and lung cancers had the lowest survival rates, which is consistent with the findings of other studies. For example, in Barton et al [28] systematic review, they found that the survival rates of nonmelanoma skin cancer, across both squamous cell carcinoma and basal cell carcinoma, tended to be poorer when it existed as a second PC. However, further study is required to more precisely characterize these associations and elucidate the underlying mechanisms. Patients with lung cancer as the first primary cancer also had relatively low survival rates, possibly because lung cancer is often not detected until it has spread; this finding is consistent with previous countrywide studies [29]. However, it should be noted that these survival estimates may be influenced by the retrospective design of our study. The Shanghai Cancer Registry achieved full population coverage in 2002, and patients with a first PC diagnosed before 2002 were required to survive until at least 2002 to be included. This introduced immortal time bias, which may lead to overestimation of survival for first primary cancers, especially for those diagnosed in earlier periods. Therefore, the absolute survival rates for first PCs should be interpreted with this limitation in mind. And future studies should use prospective designs or more uniformly enrolled cohorts to provide more precise survival estimates.

We observed low incidences of pancreatic and liver cancers as second primary cancers (ninth and eighth rankings, respectively), and the lowest survival rates among the most commonly occurring cancers. Pancreatic cancer often presents at an advanced stage, which contributes to poor 5-year survival rates of approximately 2%-9%, ranking firmly last among all cancer sites in terms of prognostic outcomes for patients [30]. Regarding liver cancer mortality in China, the mortality rate (422.1/1000) was lower than that of lung (610.2/1000) and stomach (498/1000) cancer, but higher than that of colorectal cancer (191/1000). Previous studies have found that liver cancer mortality is especially high in parts of Eastern China, such as Shanghai [22].

Notably, our study found that when the latency between the first and second PC diagnosis was ≤12 months, patients had a higher risk of mortality than those with an interval of >12 months. This finding indicates that the faster the new primary cancer occurs, the higher the mortality risk. Several mechanisms may underlie this association. First, a short latency period may reflect aggressive tumor biology or shared genetic susceptibility predisposing to the rapid development of a second PC. Second, treatment-related effects, especially the radiotherapy and chemotherapy administered for the first PC, may contribute to the development of second PC, with shorter latency potentially

XSL•FO
RenderX

indicating greater genotoxic susceptibility or higher cumulative exposure. Third, surveillance bias and reverse causation may contribute, as patients with more frequent follow-up or more aggressive first PC are more likely to have a second PC detected earlier or incidentally. However, this seems to differ for nonprimary cancers, as patients must receive early diagnoses followed by chemotherapy and radiotherapy, which can help delay the occurrence of secondary cancer and mortality [31]. Further clinical analysis is needed to determine if there are differences in trends between different types of cancers.

This finding was different from those of other studies that researched nonprimary cancers, indicating that more attention should be given to the distribution of cancers by sex for the prevention of subsequent primary cancers. Moreover, female patients had higher mortality rates than male patients in the second PCs, but there was no significant difference between the first PCs. This finding appears to be different from previous studies on general Chinese cancer statistics, which have indicated that male patients have higher overall mortality rates [25]. In this study, patients >65 years old had higher mortality regardless of whether they existed in the first or second PCs than the younger. A prior study concluded that increased age correlated with higher rates of cancer diagnosis. Moreover, in recent years, the incidence of cancer in people >65 years old has been 10 times higher than that of people aged <65 years in the United States [32]. Furthermore, Ershler et al [33] found that cancer mortality rates were higher for older adults, with approximately 70% of all cancer-related deaths occurring in people aged >65 years. Consistent with other studies, our results showed that patients with stage III + IV cancer in both first and second sites had higher mortality. This is mainly due to the inability of these late-stage-diagnosed patients to receive early detection and treatment [34]. Also, it was found that patients diagnosed between 2009 and 2015 had a higher mortality. Correspondingly, there were more patients diagnosed between 2009 and 2015 in the III+IV stage, and more patients diagnosed between 2002 and 2008 in the I+II stage. However, due to the limited number of samples with multiple primary tumors, only 23 patients had third primary tumors. Therefore, we calculated the time based on the first diagnosis of each primary tumor. In addition, we found no significant differences between patients with 2 PCs and 3 PCs when comparing both cancer-specific and all-cause mortality. This finding is unique compared to previous non–primary cancer studies by Sarfati et al [31], which concluded that more cancer comorbidities led to higher mortality. However, further research is still needed to explore the effects of various kinds of cancer.

There are several limitations in this study. First, our sample was based only on data from the Shanghai Cancer Registry System, with a relatively small size [35-37], and stringent data governance policies after 2016 limited access to more recent cases. Therefore, future studies should include larger, updated samples as data sharing frameworks evolve. Second, the current registry data lack detailed information on lifestyle factors and genetic predisposition, which limits comprehensive risk factor analysis. Future studies integrating such data would help determine whether the observed cancer combinations share common hereditary or environmental causes. Third, a substantial amount of data, particularly on TNM staging, was missing from the registry, primarily due to delayed pathology reporting relative to cancer diagnosis. This led to the exclusion of cases from the Cox regression analyses, potentially introducing selection bias and reducing the statistical power of these models. Additionally, due to the insufficiency of the Cancer Registration System, patients with a first cancer before 2002 had to survive until at least 2002 to be included in the dataset, which may cause an overestimation of survival for first primary cancers and an overrepresentation of less aggressive cancer types in our observed MPC patterns. Fourth, our survival analyses employed standard Kaplan-Meier and Cox proportional hazards models. These methods may overestimate cancer-specific mortality in the presence of competing risks. Future studies with more granular cause-of-death data are encouraged to apply competing risk models to obtain more precise estimates. Additionally, in line with the descriptive aim of profiling cancer combinations, our models for the second primary cancer did not adjust for the site of the first PC, which did not isolate the independent effect of the second cancer itself. Future etiological studies may consider such adjustments.

Conclusions

In this study, we provide comprehensive information synthesis from large, long-term datasets to improve the understanding of the prevalence of common MPC patterns and factors affecting survival status between various PC sites. Our results highlight specific cancer combinations with low survival rates, as well as risk factors, including a short interdiagnosis interval (≤12 months), advanced age (>65 years), and late tumor stage (III+IV). These results underscore the need for comprehensive prevention and control strategies toward the highly occurring MPC combinations in China and other countries, and also highlight the importance of clinical awareness directed toward the identified high-risk populations to enable earlier detection and intervention at an early stage of diagnosis.

Acknowledgments

The authors are grateful to the health institutions for contributing to the data collection.

Authors XL and HJ are co-corresponding authors of this paper. XL can be reached at xiaopanli0224@126.com.

Funding

This study was supported by the National Natural Science Foundation of China (grant numbers 71603182, 72104183), and the Project of the 2025 Open Fund of the Shanghai Key Laboratory of Meteorology and Health (No. QXJK202509 to XL). The funders had no involvement in the study design, data collection, analysis, interpretation of results, or the writing of the manuscript.

XSL•FO
RenderX

Data Availability

The datasets analyzed during this study are available from the corresponding author on reasonable request and with permission of the related department.

Authors' Contributions

All authors contributed to data analysis and drafting or revising the article. All authors gave their final approval of the version to be published and agree to be accountable for all aspects of the work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
RECORD checklist.
[DOCX File , 21 KB - publichealth_v12i1e82355_app1.docx]

Multimedia Appendix 2
The patterns of all cancers.
[XLSX File (Microsoft Excel File), 10 KB - publichealth_v12i1e82355_app2.xlsx]

Multimedia Appendix 3
The complete combination patterns of all cancers.
[XLSX File (Microsoft Excel File), 10 KB - publichealth_v12i1e82355_app3.xlsx]

Multimedia Appendix 4
Kaplan-Meier curves of survival status and risk factors for first and second primary carcinoma (PC) patients.
[PDF File (Adobe PDF File), 1513 KB - publichealth_v12i1e82355_app4.pdf]

References

1. Ding R, Zhu D, He P, Ma Y, Chen Z, Shi X. Comorbidity in lung cancer patients and its association with medical service cost and treatment choice in China. BMC Cancer 2020 Mar 24;20(1):250 [FREE Full text] [doi: 10.1186/s12885-020-06759-8] [Medline: 32209058]
2. Feng J, Mu X, Ma L, Wang W. Comorbidity patterns of older lung cancer patients in northeast china: An association rules analysis based on electronic medical records. Int J Environ Res Public Health 2020;17(23):9119 [FREE Full text] [doi: 10.3390/ijerph17239119] [Medline: 33291317]
3. Dagenais GR, Leong DP, Rangarajan S, Lanas F, Lopez-Jaramillo P, Gupta R, et al. Variations in common diseases, hospital admissions, and deaths in middle-aged adults in 21 countries from five continents (PURE): a prospective cohort study. Lancet 2020 Mar 07;395(10226):785-794. [doi: 10.1016/S0140-6736(19)32007-0] [Medline: 31492501]
4. Vogt A, Schmid S, Heinimann K, Frick H, Herrmann C, Cerny T, et al. Multiple primary tumours: Challenges and approaches, a review. ESMO Open 2017;2(2):e000172 [FREE Full text] [doi: 10.1136/esmoopen-2017-000172] [Medline: 28761745]
5. Choi E, Luo SJ, Aredo JV, Backhus LM, Wilkens LR, Su CC, et al. The survival impact of second primary lung cancer in patients with lung cancer. J Natl Cancer Inst 2022;114(4):618-625 [FREE Full text] [doi: 10.1093/jnci/djab224] [Medline: 34893871]
6. Jørgensen TL, Hallas J, Friis S, Herrstedt J. Comorbidity in elderly cancer patients in relation to overall and cancer-specific mortality. Br J Cancer 2012 Mar 27;106(7):1353-1360 [FREE Full text] [doi: 10.1038/bjc.2012.46] [Medline: 22353805]
7. LJ O. MULTIPLE MALIGNANT NEOPLASMS. JAMA 1921 May 14;76(20):1329-1333. [doi: 10.1001/jama.1921.02630200001001]
8. Coyte A, Morrison DS, McLoone P. Second primary cancer risk - the impact of applying different definitions of multiple primaries: results from a retrospective population-based cancer registry study. BMC Cancer 2014 May 18;14(1):272 [FREE Full text] [doi: 10.1186/1471-2407-14-272] [Medline: 24742063]
9. Weir HK, Johnson CJ, Thompson TD. The effect of multiple primary rules on population-based cancer survival. Cancer Causes Control 2013;24(6):1231-1242 [FREE Full text] [doi: 10.1007/s10552-013-0203-3] [Medline: 23558444]
10. Buikhuisen WA, Steinbusch LC, Kodach LL, Tesselaar MET, Damhuis RA. Risk of second primary malignancies among patients with carcinoid of the lung. Lung Cancer 2021 Jan;151:5-7 [FREE Full text] [doi: 10.1016/j.lungcan.2020.11.017] [Medline: 33278670]
11. Dranka-Bojarowska D, Lewiński A. Multiple primary cancers in patients treated for squamous cell carcinoma of the esophagus. Pol Przegl Chir 2019;91(5):51-57 [FREE Full text] [doi: 10.5604/01.3001.0012.8371] [Medline: 31702567]

XSL•FO
RenderX

12. Mu XM, Wang W, Wu FY, Jiang YY, Ma LL, Feng J. Comorbidity in older patients hospitalized with cancer in northeast china based on hospital discharge data. Int J Environ Res Public Health 2020;17(21):8028 [FREE Full text] [doi: 10.3390/ijerph17218028] [Medline: 33142785]
13. Shin DW, Cho JH, Yoo JE, Cho J, Yoon DW, Lee G, et al. Conditional survival of surgically treated patients with lung cancer: A comprehensive analyses of overall, recurrence-free, and relative survival. Cancer Res Treat 2021;53(4):1057-1071 [FREE Full text] [doi: 10.4143/crt.2020.1308] [Medline: 33705624]
14. Das A, Chak A, Cooper GS. Temporal trend in relative risk of second primary colorectal cancer. Am J Gastroenterol 2006;101(6):1342-1347. [doi: 10.1111/j.1572-0241.2006.00580.x] [Medline: 16771959]
15. Donin NM, Kwan L, Lenis AT, Drakaki A, Chamie K. Second primary lung cancer in United States Cancer Survivors, 1992-2008. Cancer Causes Control 2019;30(5):465-475. [doi: 10.1007/s10552-019-01161-7] [Medline: 30900065]
16. Ganz PA, Casillas JN. Incorporating the risk for subsequent primary cancers into the care of adult cancer survivors: Moving beyond 5-Year survival. JAMA 2020;324(24):2493-2495. [doi: 10.1001/jama.2020.23410] [Medline: 33351026]
17. Chinese Anti-Cancer Association Committee on Multiple Primary and Unknown Primary Tumors. China Anti-Cancer Association guideline for diagnosis and treatment of cancer of multiple and unknown primaries (2023 edition). China Oncology 2023;33(4):403-422 [FREE Full text] [doi: 10.19401/j.cnki.1007-3639.2023.04.011]
18. Li F, Zhong WZ, Niu FY, Zhao N, Yang JJ, Yan HH, et al. Multiple primary malignancies involving lung cancer. BMC Cancer 2015 Oct 14;15(3):696-156 [FREE Full text] [doi: 10.1186/s12885-015-1733-8] [Medline: 26466785]
19. Chen W, Zheng R, Zhang S, Zeng H, Xia C, Zuo T, et al. Cancer incidence and mortality in China, 2013. Cancer Lett 2017 Aug 10;401:63-71 [FREE Full text] [doi: 10.1016/j.canlet.2017.04.024] [Medline: 28476483]
20. Ji J, Luo Z, Chen Y, Xu X, Li X, Liu S, et al. Characteristics and trends of childhood cancer in Pudong, China, 2002-2015. BMC Public Health 2020 Oct 21;20(1):1430 [FREE Full text] [doi: 10.1186/s12889-020-09493-9] [Medline: 32958056]
21. Health M. Chinese cancer registry annual report. In: Beijing,China. Beijing: Military Medical Science Press; 2012.
22. Ferlay JBC, Whelan S, Parkin DM. IARC technical report No 42. Lyon: International Agency for Research on Cancer; 2005.
23. Swaroop VS, Winawer SJ, Kurtz RC, Lipkin M. Multiple primary malignant tumors. Gastroenterology 1987 Oct;93(4):779-783. [doi: 10.1016/0016-5085(87)90440-9] [Medline: 3623021]
24. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China, 2015. CA Cancer J Clin 2016;66(2):115-132 [FREE Full text] [doi: 10.3322/caac.21338] [Medline: 26808342]
25. Feng R, Zong Y, Cao S, Xu R. Current cancer situation in China: good or bad news from the 2018 Global Cancer Statistics? Cancer Commun (Lond) 2019 May 29;39(1):22 [FREE Full text] [doi: 10.1186/s40880-019-0368-6] [Medline: 31030667]
26. Romaszko AM, Doboszyńska A. Multiple primary lung cancer: A literature review. Adv Clin Exp Med 2018 May;27(5):725-730 [FREE Full text] [doi: 10.17219/acem/68631] [Medline: 29790681]
27. Hill DA, Friend S, Lomo L, Wiggins C, Barry M, Prossnitz E, et al. Breast cancer survival, survival disparities, and guideline-based treatment. Breast Cancer Res Treat 2018 Jul;170(2):405-414 [FREE Full text] [doi: 10.1007/s10549-018-4761-7] [Medline: 29569018]
28. Barton V, Armeson K, Hampras S, Ferris LK, Visvanathan K, Rollison D, et al. Nonmelanoma skin cancer and risk of all-cause and cancer-related mortality: a systematic review. Arch Dermatol Res 2017 May;309(4):243-251 [FREE Full text] [doi: 10.1007/s00403-017-1724-5] [Medline: 28285366]
29. Schabath MB, Cote ML. Cancer Progress and Priorities: Lung Cancer. Cancer Epidemiol Biomarkers Prev 2019 Oct;28(10):1563-1579 [FREE Full text] [doi: 10.1158/1055-9965.EPI-19-0221] [Medline: 31575553]
30. McGuigan A, Kelly P, Turkington RC, Jones C, Coleman HG, McCain RS. Pancreatic cancer: A review of clinical diagnosis, epidemiology, treatment and outcomes. World J Gastroenterol 2018 Dec 21;24(43):4846-4861 [FREE Full text] [doi: 10.3748/wjg.v24.i43.4846] [Medline: 30487695]
31. Sarfati D, Koczwara B, Jackson C. The impact of comorbidity on cancer and its treatment. CA Cancer J Clin 2016 Jul;66(4):337-350 [FREE Full text] [doi: 10.3322/caac.21342] [Medline: 26891458]
32. Ritchie CS, Zhao F, Patel K, Manola J, Kvale EA, Snyder CF, et al. Association between patients' perception of the comorbidity burden and symptoms in outpatients with common solid tumors. Cancer 2017 Oct 01;123(19):3835-3842 [FREE Full text] [doi: 10.1002/cncr.30801] [Medline: 28608952]
33. Ershler WB. Cancer: a disease of the elderly. J Support Oncol 2003;1(4 Suppl 2):5-10. [Medline: 15346994]
34. Noer MC, Antonsen SL, Ottesen B, Christensen IJ, Høgdall C. Type I Versus Type II Endometrial Cancer: Differential Impact of Comorbidity. International Journal of Gynecological Cancer 2018 Mar;28(3):586-593 [FREE Full text] [doi: 10.1097/igc.0000000000001184]
35. Chen L, Xia T, Rampatige R, Li H, Adair T, Joshi R, et al. Assessing the Diagnostic Accuracy of Physicians for Home Death Certification in Shanghai: Application of SmartVA. Front Public Health 2022;10:842880 [FREE Full text] [doi: 10.3389/fpubh.2022.842880] [Medline: 35784257]
36. Li X, Deng Y, Tang W, Sun Q, Chen Y, Yang C, et al. Urban-Rural Disparity in Cancer Incidence, Mortality, and Survivals in Shanghai, China, During 2002 and 2015. Front Oncol 2018;8:579 [FREE Full text] [doi: 10.3389/fonc.2018.00579] [Medline: 30560091]

XSL•FO
RenderX

37. Li X, Zhou Y, Luo Z, Gu Y, Chen Y, Yang C, et al. The impact of screening on the survival of colorectal cancer in Shanghai, China: a population based study. BMC Public Health 2019 Jul 29;19(1):1016 [FREE Full text] [doi: 10.1186/s12889-019-7318-8] [Medline: 31357981]

Abbreviations

CRC: colorectal cancer
HR: hazard ratio
***ICD-10*:** *International Statistical Classification of Diseases and Related Health Problems 10th Revision*
MPC: multiple primary carcinoma
PC: primary carcinoma
RECORD: Reporting of studies Conducted using Observational Routinely collected health Data
SEER: Surveillance, Epidemiology, and End Results
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
TNM: tumor, node, and metastases stage
TRIPOD: Transparent Reporting of a multivariable Prediction model for individual Prognosis or Diagnosis

Edited by A Mavragani, T Sanchez; submitted 13.Aug.2025; peer-reviewed by F Liu, F Calabro; comments to author 02.Jan.2026; revised version received 15.Feb.2026; accepted 24.Feb.2026; published 06.Jul.2026.

Please cite as:
Shi J, Chen Y, Chen C, Li X, Jin H
Real-World Mapping of Multiple Primary Carcinoma Combinations and Survival Outcomes in Shanghai, China: Retrospective Registry-Based Study
JMIR Public Health Surveill 2026;12:e82355
URL: https://publichealth.jmir.org/2026/1/e82355
doi:10.2196/82355
PMID:

©Jianwei Shi, Yichen Chen, Chen Chen, Xiaopan Li, Hua Jin. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 06.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Epidemiological Distribution Characteristics of Tuberculosis Among Older Adults in Chongqing (2020-2024): Spatial-Temporal Analysis

Bojie Gao[1*], MPH; Yu Xin[2*], BS; Wenping Liao[2], BS; Lin Shi[2], MPH; Chengguo Wu[2], MPH; Jun Fan[2], MPH; Qingya Wang[2], MPH; Shanrong Huang[1], MPH; Xinyuan Yi[1], MPH; Yong Li[1], MPH; Wen Zhang[2], MPH; Chuan Pu[1], MPH

[1]School of Public Health, Chongqing Medical University, No 1, Yixueyuan Road, Chongqing, China

[2]Chongqing Institute of Tuberculosis Control and Prevention, Chongqing, Chongqing, China

[*]these authors contributed equally

Corresponding Author:
Chuan Pu, MPH
School of Public Health, Chongqing Medical University, No 1, Yixueyuan Road, Chongqing, China

Abstract

Background: With global aging, the burden of tuberculosis (TB) among older adults escalates, yet spatial studies on this group are scarce. In Chongqing, where 18.87% of the population are aged 65 years and older and TB burden is high, controlling older adult TB remains a major challenge.

Objective: This study analyzed the spatiotemporal patterns of TB among adults aged 65 years and older in Chongqing, China, to inform local prevention and control strategies.

Methods: The study data were obtained from the Tuberculosis Information Management System of China. Global and local spatial autocorrelation analyses were conducted using ArcGIS (version 10.7) to identify high-risk spatial clusters and visualize their distribution. Spatiotemporal scan statistics were performed using SaTScan (version 10.3.2) to detect clusters of TB cases among the older adult population. Statistical significance was set at $P<.05$.

Results: The average annual incidence of TB among older adults in Chongqing was 69.59 per 100,000 population, with peaks occurring in spring and summer. The global Moran I ranged from 0.618 to 0.756 ($P<.001$ in all cases), indicating significant clustering. Persistent high-risk areas were identified in the northeastern and southeastern parts of Chongqing. Spatiotemporal scan statistics detected 1 most likely cluster (relative risk=3.52, 95% CI 3.37 - 3.68; log-likelihood ratio=1017.43; $P<.001$) and 3 secondary clusters.

Conclusions: Significant seasonal patterns of TB among older adults were observed in Chongqing. High-risk areas were predominantly concentrated in the northeastern and southeastern parts of the municipality. More targeted public health interventions are imperative.

(JMIR Public Health Surveill 2026;12:e89671) doi:10.2196/89671

KEYWORDS

older adults; tuberculosis; epidemiology; spatial autocorrelation analysis; spatiotemporal clusters

Introduction

Tuberculosis (TB), a chronic infectious disease primarily affecting the lungs, is caused by *Mycobacterium tuberculosis*. It ranks among the top 10 causes of death worldwide and represents a significant public health and social challenge worldwide [1]. According to the World Health Organization's *Global Tuberculosis Report 2024* [1], an estimated 10.8 million individuals developed TB worldwide in 2023, resulting in approximately 1.25 million deaths. Among the 30 high–TB burden countries, China bears the third highest burden, with 741,000 new TB cases reported in 2023, accounting for 6.8% of the global incidence. Among these cases, 25,000 fatalities were attributed to TB. Although China has witnessed a declining trend in TB incidence over the years, the situation remains challenging due to its large population [2].

With increasing life expectancy and declining fertility rates, the pace of global population aging has accelerated markedly in the new century [3]. Older adults are more susceptible to TB due to immunosenescence, a higher prevalence of comorbid chronic diseases, and an increased risk of latent infection reactivation [4-6]. Consequently, the older adult population has become a critical challenge for TB control efforts. In the Western Pacific Region of the World Health Organization, the highest number of TB cases is observed in the age group of 65 years and older, largely driven by the demographic aging in China [1]. Located in southwestern China, Chongqing is a major populous municipality where the reported TB incidence has consistently ranked among the highest nationally. A study in Chongqing

XSL•FO
RenderX

indicated that the average annual reported TB incidence from 2018 to 2022 was 65.01 per 100,000, exceeding the national average of 52.00 per 100,000 during the same period, with the peak incidence observed among individuals aged 65 years and older [1,7]. Another study reported that the spatial distribution of TB cases in Chongqing is heterogeneous, predominantly clustered in the northeastern and southeastern areas [8]. According to statistical data [9], the population aged 65 years and older in Chongqing reached 6.02 million by 2024, accounting for 18.87% of the municipality's total population. Furthermore, TB management in older adults is compounded by challenges including delayed diagnosis, poor treatment adherence, and higher rates of adverse outcomes, presenting substantial obstacles to public health control efforts [10-12].

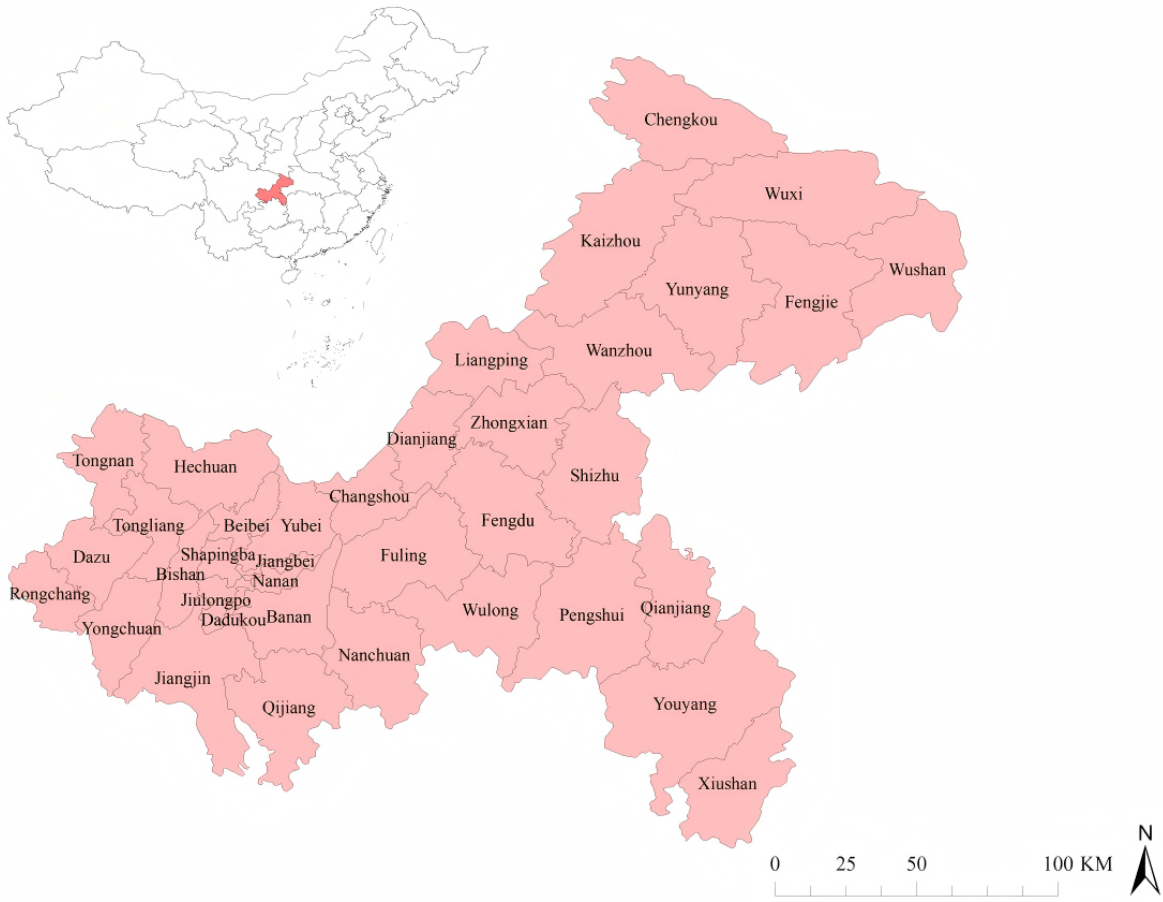
In recent years, spatiotemporal analysis has been widely used to characterize the distribution patterns and transmission dynamics of TB across space and time [13,14]. Studies conducted in China [15,16] and other countries [17-19] have demonstrated that TB exhibits highly complex dynamics with significant spatial heterogeneity at provincial, national, and international levels. The Chongqing municipality, characterized by its complex topography, significant urban-rural disparities, and frequent population mobility, demonstrates distinct spatial heterogeneity in its TB epidemic. However, few studies have investigated the spatiotemporal characteristics of TB distribution among the older adult population at the county level within Chongqing. Therefore, using reported TB data from 2020 to 2024 across Chongqing's 38 counties, this study aimed to examine the geographical distribution and spatiotemporal trends of TB among individuals aged 65 years and older. The findings are expected to provide a scientific basis for targeted TB control, resource allocation, and policy formulation.

Methods

Overview of the Study Area

Chongqing is located in the southwest of China in the upper reaches of the Yangtze River, between 28.10°N to 32.13°N and 105.11°E to 110.11°E, and is one of China's 4 direct-controlled municipalities. It spans approximately 470 km from east to west and 450 km from north to south, covering a total area of 82,400 km^2. The municipality comprises 38 counties, with a resident population of 31.90 million in 2024. Chongqing experiences a humid subtropical monsoon climate characterized by a mean annual temperature of 17.7 °C and an average relative humidity of 79.9% (Figure 1) [20].

Figure 1. Geographical location and county-level administrative divisions of the Chongqing municipality, China.



XSL•FO
RenderX

Data Collection and Management

The data for this study were obtained from the Tuberculosis Information Management System (TBIMS) of China, which collects case reports for all patients with TB during a specified period. TB incidence statistics were compiled based on patients' current residential addresses. This study statistically analyzed the incidence of cases according to the patients' current residences, collecting data on newly diagnosed older adult patients with TB aged 65 years and older reported in the system by medical institutions in the 38 counties of Chongqing from January 1, 2020, to December 31, 2024. To protect patient privacy and prevent duplicate reporting, each reported TB case was assigned a unique identifier. Each patient record contained demographic and clinical information, including age, sex, current residential address, patient type, diagnosis date, and diagnostic results. The diagnosis of TB followed the standards outlined in the WS 288-2017 guidelines issued by the National Health Commission of the People's Republic of China [21] and the *Technical Guidelines for Tuberculosis Prevention and Control in China* [22]. From 2020 to 2024, population data for individuals aged 65 years and older in each county-level area of the Chongqing municipality were obtained from the records for the corresponding years in the 2024 Chongqing Statistical Yearbook [20] and the 2024 Chongqing Statistical Bulletin on National Economic and Social Development [9]. County-level administrative boundary vector maps for China and specifically for the Chongqing municipality were obtained from the National Platform for Common Geospatial Information Services. All spatial data were obtained in the WGS 84 geographic coordinate system.

Spatial Trend Surface Analysis

Trend surface analysis uses a 3D coordinate system (x, y, z) to generate a 3D projection of the data, thereby graphically illustrating the spatial distribution trends of TB. In this coordinate system, the x- and y-axes correspond to the longitude and latitude of the geometric center of each study unit, respectively, whereas the z-axis represents the incidence rate. A higher z value indicates a greater TB incidence in the corresponding area. Thus, any point (x, y, or z) within this 3D space corresponds to the TB incidence level within a specific study region. By projecting the longitude and latitude coordinates of each county in Chongqing and their corresponding TB incidence rates onto the x-z and y-z planes and subsequently performing curve fitting using a polynomial model [23], the characteristics of incidence rate trends in relation to geographical coordinates can be further analyzed in 2D.

Spatial Autocorrelation Analysis

Spatial autocorrelation measures the correlation of a single variable across different spatial locations, quantifying the degree of clustering for an attribute value among spatial units. It is typically categorized into global and local spatial autocorrelation analyses. This study used the global Moran I index to assess global spatial autocorrelation. The value of the Moran I ranges from −1 to 1. A Moran I value greater than 0 indicates positive spatial autocorrelation, with higher values signifying more pronounced spatial clustering. Conversely, a Moran I value of less than 0 indicates negative spatial autocorrelation, suggesting a dispersed pattern; values closer to −1 denote greater spatial disparity. A Moran I of 0 implies the absence of spatial clustering, consistent with a random distribution [24]. The statistical significance was evaluated using the standardized z score.

Global spatial autocorrelation was used to determine whether TB incidence was correlated across the entire study area, but it cannot pinpoint the specific locations of clusters. In contrast, local spatial autocorrelation analysis, conducted by generating cluster maps based on local indicators of spatial association (LISA), measures the concentration of high or low values. This method detects spatial hot spots and cold spots of disease occurrence, thereby revealing spatial heterogeneity within the data. The LISA cluster patterns are categorized into 4 types: high-high, high-low, low-high, and low-low. Among these, high-high and low-low clusters represent positive spatial autocorrelation, whereas high-low and low-high clusters indicate negative spatial autocorrelation. The high-high clusters, identified as hot spots, are typically prioritized areas for infectious disease control and were the primary focus of the spatial analysis of TB in the older adult population [25]. To assess robustness, we excluded 8 counties with populations of less than 100,000 (Dadukou, Yuzhong, Jiangbei, Nan'an, Liangping, Jiulongpo, Tongnan, and Shapingba) and reran LISA for 2020 to 2024 using the remaining 30 counties.

In this study, the spatial weight matrix for spatial autocorrelation analysis was constructed using the inverse distance method, with weights calculated as the reciprocal of the Euclidean distance between the centroids of each district and county. No distance threshold was set, meaning that all districts and counties were considered neighbors, and the weight matrix was row standardized before analysis.

Spatial-Temporal Scan Statistic

Spatiotemporal scan statistics were used to identify spatial clusters of TB incidence among older adults in Chongqing and analyze their temporal evolution. The analysis was performed using the spatiotemporal scan statistic by Kulldorff [26], with a Poisson probability model applied to the number of reported TB cases, resident population data, and geographical coordinates. In this method, the scanning window is a dynamic cylinder whose base represents a circular geographic area and whose height corresponds to a time interval. As the center, radius of the base, and height of the cylinder vary, a log-likelihood ratio (LLR) is computed for each possible window to compare the risk of TB inside the window against the risk outside. The relative risk (RR) was calculated to estimate the magnitude of risk within each identified cluster. The 95% CIs for the RR were derived using the exact Poisson method based on the observed and expected case counts. The statistical significance of the LLR statistic was assessed using the Monte Carlo randomization method, with 999 iterations performed to derive the P value. A larger, statistically significant LLR value indicates a stronger clustering effect. The window with the maximum significant LLR was identified as the most likely cluster, whereas other windows with statistically significant LLRs were classified as secondary clusters. For this study, the scan period spanned January 1, 2020, to December 31, 2024,

XSL•FO
RenderX

with 1 year defined as the basic temporal unit. The maximum spatial and temporal scan window sizes were both set to the default values, with the maximum spatial cluster size set to 50% of the at-risk population and the maximum temporal cluster size set to 50% of the study period to ensure that persistent aggregation phenomena could be detected and that the reported time range reflected the actual duration of increased risk [27]. Clusters were ranked based on their LLR values. The geographic coordinates (longitude and latitude) used for spatiotemporal scanning were the centroid coordinates of each county-level administrative district in Chongqing obtained from the National Geographic Information Public Service Platform of China. These coordinates represent the spatial location of each district for aggregated case data analysis rather than individual patient addresses.

Statistical Software

Data collection, collation, and descriptive analysis for TB cases among individuals aged 65 years and older were performed using Microsoft Excel 2019. Joinpoint regression analysis (version 6.0.0.0; National Cancer Institute) was used to assess temporal trends in TB notification rates. Global and local spatial autocorrelation analyses, including the calculation of the Moran *I* index, spatial trend surface analysis, and visualization of incidence rates, were conducted using the ArcGIS software (version 10.7; Esri). Spatiotemporal scanning and the calculation of LLR values were performed using the SaTScan software (version 10.3.2). A threshold of $P<.05$ was applied to determine statistical significance for all analyses.

Ethical Considerations

The use of TB case data in this study was approved by the ethics committee of the Chongqing Tuberculosis Control and Prevention Institute (audit 2025008) on April 14, 2025. This was a retrospective approval for the use of deidentified data collected from January 1, 2020, to December 31, 2024, from the routine TBIMS. The ethics committee waived the requirement for individual informed consent because the data were anonymized prior to analysis. The authors have obtained formal permission to access and use the data for this research. All personal identifiers were removed from the dataset prior to data analysis and reporting.

Results

Descriptive Analysis of TB Cases

A total of 20,265 new TB cases in individuals aged 65 years and older were reported from 2020 to 2024, yielding an average of 4053 cases annually. Among these 20,265 patients, 15,098 (74.5%) were male, and 5167 (25.5%) were female, resulting in a male-to-female ratio of 2.92:1. The average annual incidence rate was 69.59 per 100,000 population. The specific annual rates were 74.48, 72.04, 65.19, 68.05, and 68.68 per 100,000 population in 2020, 2021, 2022, 2023, and 2024 (Figure 2). Joinpoint regression analysis identified a statistically significant change in trend occurring in 2022. From 2020 to 2022, the notification rate decreased sharply, with an annual percent change (APC) of −6.26% (95% CI −9.94% to −2.09%; $P<.001$). From 2023 to 2024, the rate increased slightly, with an APC of 2.11% (95% CI −2.31% to 6.43%; $P=.28$). The average APC over the entire 2020 to 2024 period was −2.17% (95% CI –5.05% to 0.72%; $P=.16$), indicating that the overall incidence did not change significantly over the 5 years. Regarding age distribution, 35.99% (7293/20,265) of the patients were aged 65 to 69 years, 30.22% (6125/20,265) were aged 70 to 74 years, and 33.79% (6847/20,265) were aged 75 years or older. Patients were mainly referred, accounting for 71.22% (14,432/20,265) of the total, followed by direct visits with 12.22% (2476/20,265) and tracking with 11.56% (2343/20,265; Table 1). From 2020 to 2024, the number of TB cases showed seasonal variation, with the peak of reported cases mainly concentrated in spring and summer and a small peak occurring every year in March (Figure 3).

XSL•FO
RenderX

Figure 2. Number of cases and overall notification rate of tuberculosis among older adults by sex in the Chongqing municipality, 2020 to 2024.

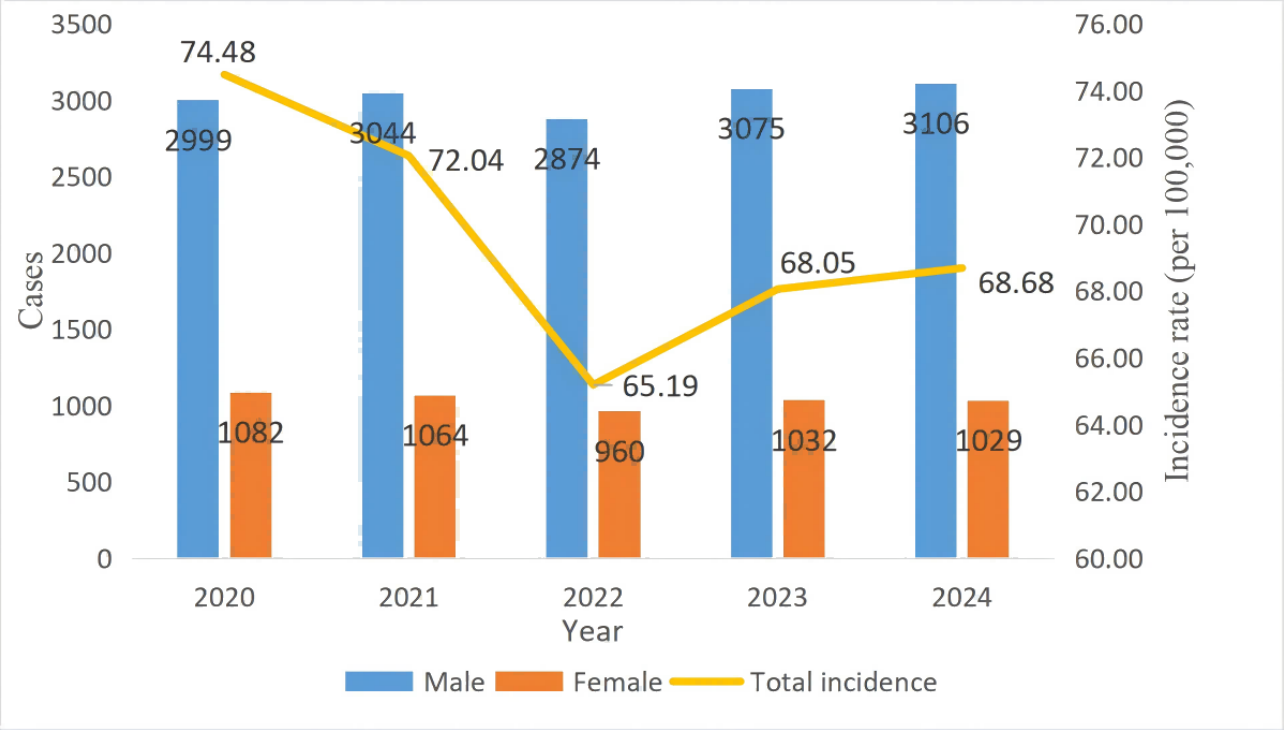
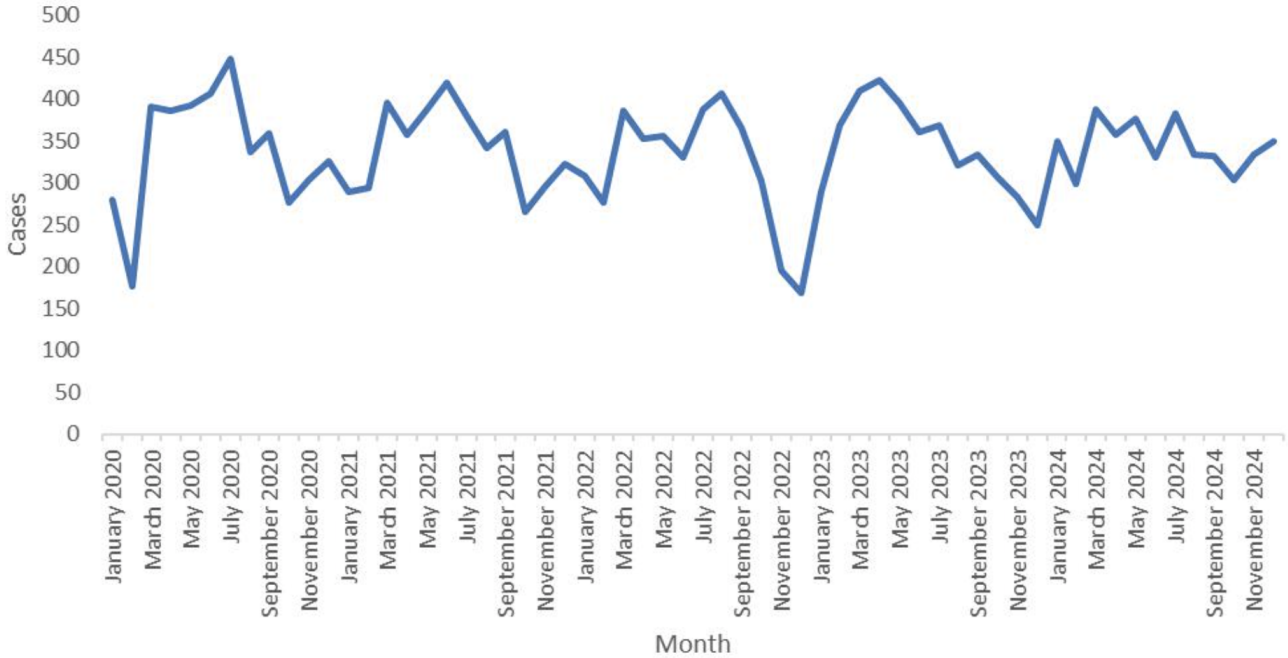


Table . Epidemiological characteristics of tuberculosis among older adults in Chongqing, 2020 to 2024 (N=20,265).

Characteristic	Patients, n (%)
Sex	
Male	15,098 (74.5)
Female	5167 (25.5)
Age group (y)	
65-69	7293 (35.99)
70-74	6125 (30.22)
≥75	6847 (33.79)
Patient sources	
Health examination	51 (0.25)
Recommended due to symptoms	526 (2.6)
Direct consultation	2476 (12.22)
Active screening	436 (2.15)
Referral	14,432 (71.22)
Tracking	2343 (11.56)
Other	1 (0)

XSL•FO
RenderX

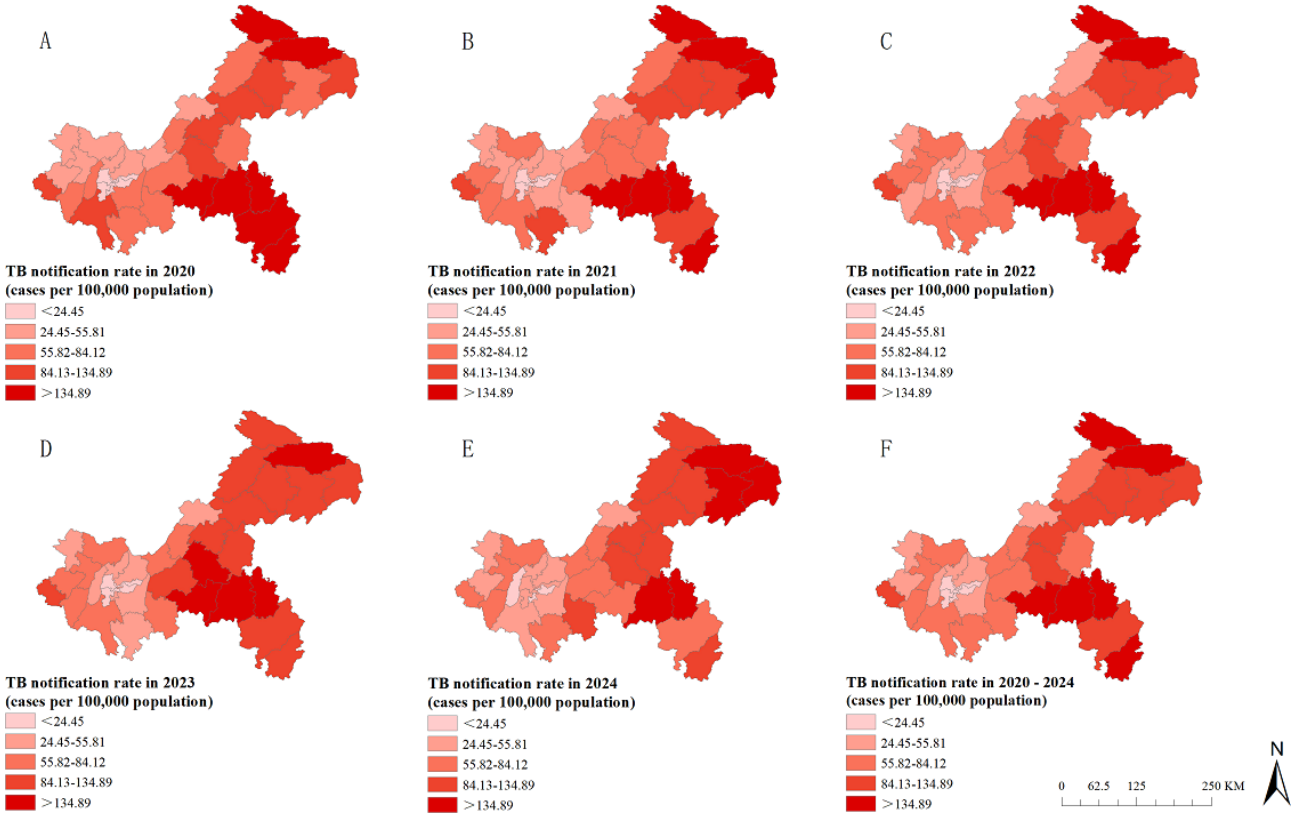
Figure 3. Monthly distribution of tuberculosis cases among older adults in the Chongqing municipality, 2020 to 2024.



Spatial Characteristics of TB Cases

Spatial distribution maps of reported TB incidence rates for the 38 counties in Chongqing (2020 - 2024) were generated using ArcGIS. From a spatial perspective, the incidence of TB was mainly distributed in the northeast and southeast of Chongqing. The top 5 regions with the highest reported TB incidence rates among individuals aged 65 years and older were Pengshui (255.38 per 100,000 population), Qianjiang (192.33 per 100,000 population), Wuxi (180.20 per 100,000 population), Wulong (175.44 per 100,000 population), and Xiushan (174.26 per 100,000 population; Figure 4).

Figure 4. Spatial distribution map of tuberculosis (TB) notification rates among the older adult population in the Chongqing municipality, 2020 to 2024.

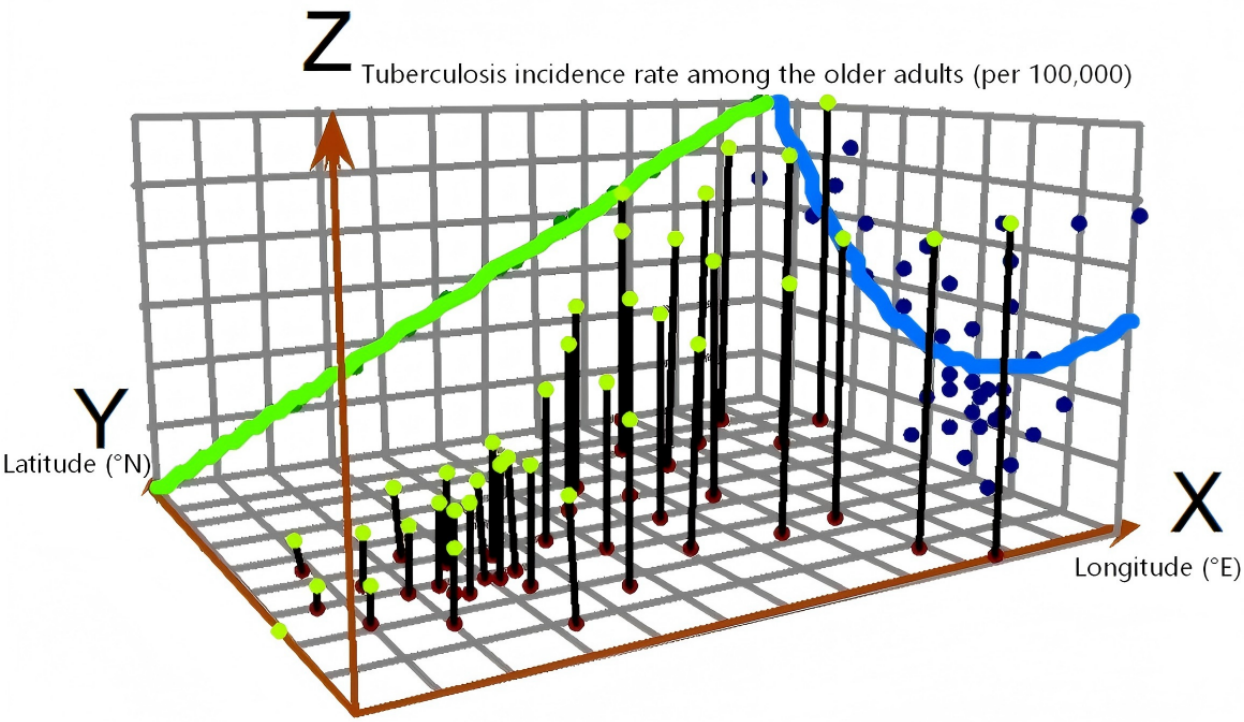


XSL•FO
RenderX

Spatial Trend Surface Analysis

Using the reported incidence rate of TB among older adults in the 38 counties of Chongqing from 2020 to 2024 as the dependent variable (z-axis) and the latitude and longitude of the geometric center of the study area as the independent variables (x- and y-axes), the analysis revealed a clear east-west gradient, with incidence rates gradually increasing from west to east. In the north-south direction, the rates exhibited a U-shaped pattern, initially decreasing from north to south before rising rapidly further south. Overall, the trend surface indicated higher incidence rates in the eastern and northern regions than in the western and southern parts of the municipality (Figure 5).

Figure 5. Spatial 3D trend surface analysis of tuberculosis among older adults in the Chongqing municipality, 2020 to 2024.



Spatial Autocorrelation Analysis

The results of the global spatial autocorrelation analysis of reported TB incidence among older adults in Chongqing from 2020 to 2024 showed that the global Moran *I* indexes were 0.636, 0.618, 0.756, 0.713, and 0.668, respectively, with all Moran *I* values greater than 0 and statistically significant ($P<.001$; Table 2). This indicates a positive global spatial autocorrelation in TB incidence from 2020 to 2024, meaning that there was spatial clustering. The spatial clustering of TB in older adults was highest in 2022.

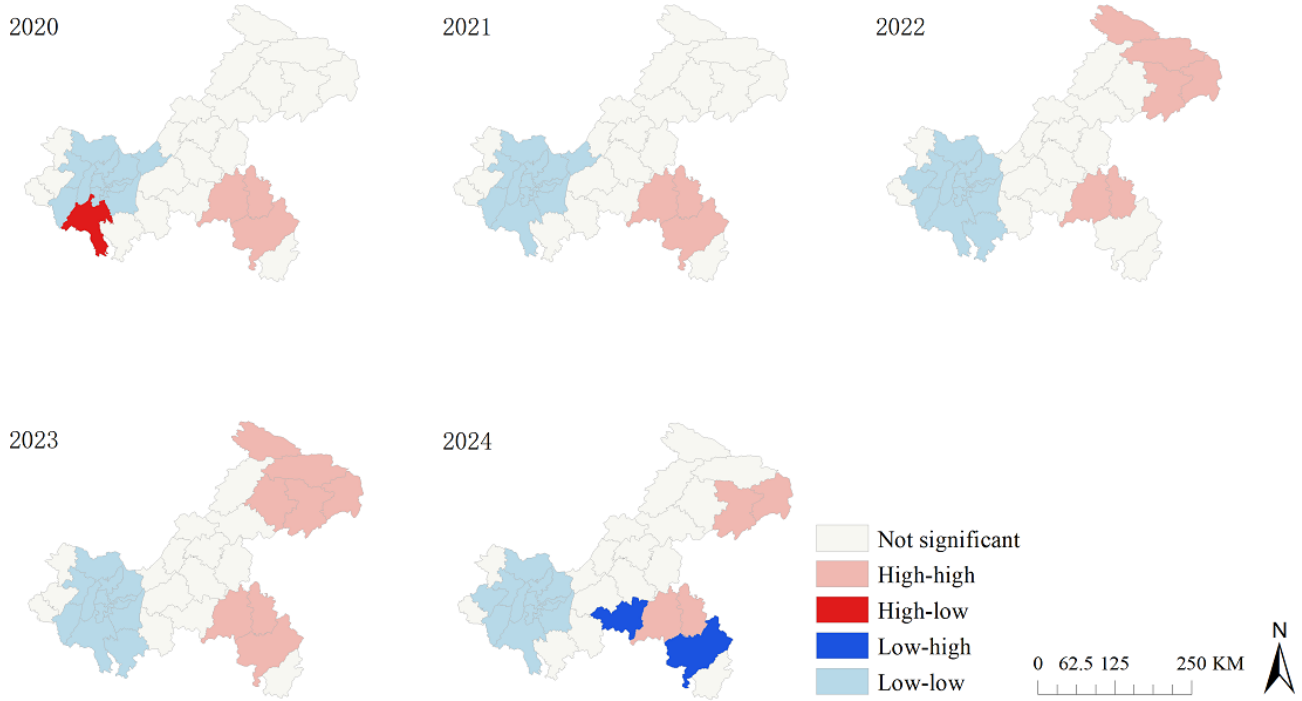
Table . Global spatial autocorrelation analysis of tuberculosis notification rates among older adults in the Chongqing municipality, 2020 to 2024.

Year	Moran *I*	*z* score	*P* value
2020	0.636	6.613	<.001
2021	0.618	6.560	<.001
2022	0.756	7.476	<.001
2023	0.713	7.086	<.001
2024	0.668	6.844	<.001

The LISA cluster maps from the local spatial autocorrelation analysis for the period from 2020 to 2024 revealed 4 types of spatial clustering patterns: high-high, high-low, low-high, and low-low clusters. Specifically, no low-high clusters were identified in 2020, neither high-low nor low-high clusters were observed from 2021 to 2023, and no high-low clusters were found in 2024. In 2020 and 2021, high-high clusters were located in the southeastern region, including Pengshui, Qianjiang, and Youyang. Starting in 2022, the high-high clusters expanded from the southeast to the northeast, encompassing Pengshui, Qianjiang, Youyang, Wushan, Fengjie, Wuxi, Chengkou, and Yunyang. Throughout the entire study period, Pengshui and Qianjiang in southeastern Chongqing consistently remained hot spots for TB among older adults. Cold spots were primarily distributed in the central urban districts and their surrounding counties (Figure 6).

XSL•FO
RenderX

Figure 6. Local spatial autocorrelation analysis of tuberculosis notification rates among older adults in the Chongqing municipality, 2020 to 2024.



The sensitivity analysis results show that the southeastern hot spots (Pengshui and Qianjiang) remained stable; northeastern hot spots began to appear in 2022, which is almost consistent with the LISA cluster patterns shown in Figure 6, in which high-high clusters were concentrated in southeastern Chongqing and expanded to the northeast after 2022. The counties that were excluded were all cold spots and had no impact on the core results (Multimedia Appendix 1).

Spatial-Temporal Clustering Analysis Using SaTScan

The spatiotemporal scan analysis identified 1 most likely cluster and 3 secondary clusters of reported TB incidence among older adults in Chongqing from 2020 to 2024. The most likely cluster covered 5 counties (Xiushan, Youyang, Qianjiang, Pengshui, and Wulong) in the southeastern region, with a clustering period from January 2020 to December 2021. This cluster contained 1970 confirmed cases (RR=3.52, 95% CI 3.37-3.68; LLR=1017.43; P<.001), indicating the highest risk of TB among older adults in these areas. Secondary cluster 1 encompassed 17 counties (Wanzhou, Yunyang, Zhongxian, Kaizhou, Liangping, Shizhu, Fengjie, Dianjiang, Fengdu, Wuxi, Chengkou, Qianjiang, Changshou, Wushan, Pengshui, Fuling, and Wulong) primarily distributed across the northeastern and southeastern regions. Its clustering period was from January 2023 to December 2024, involving 4623 cases (RR=1.70, 95% CI 1.65-1.75; LLR=454.45; P<.001). Secondary cluster 2 included 3 counties (Yongchuan, Jiangjin, and Rongchang) and was observed from January 2020 to December 2021, with 924 cases identified (RR=1.21, 95% CI 1.13-1.29; LLR=14.55; P<.001). Secondary cluster 3 covered a single county, Qijiang, during the period from January 2020 to December 2021, with 312 cases (RR=1.27, 95% CI 1.14-1.42; LLR=8.31; P=.03; Table 3). It is worth noting that the detected clusters span multiyear periods (2020 - 2021 and 2023 - 2024) rather than single years. This is because the spatiotemporal scan analysis allowed for a maximum temporal cluster size of 50% of the study period (ie, 2.5 years), a standard setting that enables the identification of both short-term and persistent aggregations. Therefore, the 2-year duration of these clusters reflects sustained elevated risk in the affected regions rather than isolated annual events.

XSL•FO
RenderX

Table . Spatiotemporal scan cluster analysis of tuberculosis cases among older adults in the Chongqing municipality, 2020 to 2024.

	Most likely cluster	Secondary cluster 1	Secondary cluster 2	Secondary cluster 3
Time frame	January 1, 2020, to December 31, 2021	January 1, 2023, to December 31, 2024	January 1, 2020, to December 31, 2021	January 1, 2020, to December 31, 2021
Cluster districts and counties	Xiushan, Youyang, Qianjiang, Pengshui, and Wulong	Wanzhou, Yunyang, Zhongxian, Kaizhou, Liangping, Shizhu, Fengjie, Dianjiang, Fengdu, Wuxi, Chengkou, Qianjiang, Changshou, Wushan, Pengshui, Fuling, and Wulong	Yongchuan, Jiangjing, and Rongchang	Qijiang
Center/radius	28.49°N, 109.01°E/160.58 km	30.71°N, 108.40°E/162.39 km	29.29°N, 105.87°E/40.42 km	28.88°N, 106.72°E/0 km
Observed cases, n	1970	4623	924	312
Expected cases	601.49	3000.75	772.53	245.81
LLR[a]	1017.43	454.45	14.55	8.31
RR[b] (95% CI)	3.52 (3.37-3.68)	1.70 (1.65 - 1.75)	1.21 (1.13 - 1.29)	1.27 (1.14 - 1.42)
P value	<.001	<.001	<.001	.03

[a]LLR: log-likelihood ratio.

[b]RR: relative risk.

Discussion

Principal Findings

This study used spatial epidemiological methods to analyze the spatiotemporal characteristics of reported TB incidence among the population aged 65 years and older across 38 counties of Chongqing from 2020 to 2024, thereby delineating the epidemiological distribution of TB within this demographic.

This study found that, from 2020 to 2024, the average annual incidence rate among older people in Chongqing was 69.59 per 100,000 population, higher than the average TB incidence rate in China of 52.00 per 100,000 population [1] and the annual incidence rate of TB among people aged 65 years and older in Xuzhou of 51.27 per 100,000 population [28]. Furthermore, joinpoint regression analysis revealed a nonlinear trend in TB incidence among older adults in Chongqing, with a significant decline from 2020 to 2022 followed by a nonsignificant increase from 2023 to 2024. The average APC over the entire study period was −2.17% (95% CI −5.05% to 0.72%; *P*=.16). Although this decline did not reach statistical significance, the point estimate was substantially lower than the reported national average annual decline of −3.55% for older adult TB in China [29]. This finding may be attributable to Chongqing's high aging population, with 18.87% of people aged 65 years and older, which is higher than the national average of 15.40% [30]. Moreover, Chongqing is mostly mountainous, and the aging and empty nest phenomena are serious, resulting in lower health awareness among the local older adults left behind and a lack of knowledge about TB prevention and control [31]. Despite the nonlinear pattern identified by the joinpoint analysis, the overall incidence among older adults in Chongqing decreased from 74.48 per 100,000 population in 2020 to 68.68 per 100,000 population in 2024, a net decline consistent with observations from other Chinese provinces [32,33] and other countries [34,35]. The overall declining trend may be attributed to factors such as national socioeconomic development, improvements in health care infrastructure, and the implementation of effective public health interventions [36,37]. Concurrently, this trend aligns with a series of TB control policies enacted by the Chongqing municipal government and health administration. For instance, the Chongqing Action Plan for Tuberculosis Control (2019-2022) issued in 2019 [38] promoted initiatives including the Hundreds, Thousands, and Tens of Thousands of Volunteers campaign for TB knowledge dissemination, the integration of TB education into school health programs, and the comprehensive incorporation of TB control for migrant populations into localized management. These measures ensured seamless processes from screening and treatment to cross-regional coordination.

The turning point detected in 2022 by the joinpoint analysis closely mirrors the trajectory of the COVID-19 pandemic. It is important to acknowledge that the COVID-19 pandemic introduced significant disruptions to TB surveillance worldwide, which may have influenced case detection and reporting during our study period. A growing body of literature has documented the substantial impact of the pandemic on TB notifications worldwide. According to the World Health Organization, TB case notifications fell considerably below prepandemic levels from 2020 to 2021, with a global reduction of 18% in 2020 [39]. Studies from Brazil reported significant decreases in TB notifications in 2020 (13%) and 2021 (9%), with partial recovery in 2022 [40]. Similar disruptions were observed across multiple countries, attributed to health care system strain, reallocation of TB diagnostic resources and personnel to COVID-19 response, lockdown measures that limited health care access, and reduced health care–seeking behavior due to fear of infection [41,42]. The overlap of respiratory symptoms between TB and COVID-19 may also have contributed to diagnostic delays and underreporting [43].

XSL•FO
RenderX

In the context of these global disruptions, it is noteworthy that large-scale preventive measures implemented in China during the pandemic, such as disinfection and isolation protocols, may have helped limit TB transmission [44]. Simultaneously, reduced social activities and increased adherence to protective measures such as mask wearing among older adults likely diminished TB exposure risks [45]. These factors may have partially offset the surveillance disruptions observed elsewhere. However, the brief upward trend from 2022 to 2024 requires further consideration. This increase may be related to the relaxation of pandemic restrictions in China after 2022, which led to increased outdoor activities among older adults [46]. Additionally, some older adults who had been infected with COVID-19 may have experienced temporary immune impairment [47], potentially increasing their susceptibility to TB. This combination of behavioral and immunological factors likely contributed to the observed temporary increase against the backdrop of ongoing postpandemic recovery in health care use.

A marked male predominance was observed in TB cases among older adults, a finding consistent with those of numerous previous studies [48,49]. This disparity is primarily attributed to gender-specific differences in lifestyle behaviors and genetic susceptibility [50,51]. High-risk behaviors more prevalent in male individuals, such as smoking, alcohol consumption, and poorer adherence to medical care, can increase susceptibility to TB infection. Concurrently, male individuals may exhibit a higher probability of specific genetic variations that elevate the risk of contracting TB. The predominant patient source being referral was likely because older adults are a key population for TB control. When primary health care institutions detect suspected patients, they are typically transferred to designated TB facilities to mitigate the risk of clustered transmission [52]. Furthermore, our analysis revealed a distinct seasonal variation in TB incidence among older adults in Chongqing, with peaks predominantly occurring in spring and summer. This pattern aligns with observations from multiple regions, including the Hong Kong special administrative region, China [53], Guangxi [54], Turkey [55], and Japan [56]. This seasonality may be associated with care seeking delays among Chinese older adults. This may be related to delayed medical visits among older people in China. Winter in China is from December to February, during which people prepare to celebrate the Chinese New Year. Some patients with mild symptoms may choose to seek medical treatment after the New Year; therefore, medical visits are avoided or postponed during this period, leading to a decrease in the reported number of cases [57].

The spatial analysis confirmed that TB among older adults in Chongqing is not randomly distributed but exhibits significant clustering. The identification of persistent hot spots in both southeastern and northeastern regions aligns with areas previously characterized by mountainous terrain, ethnic minority populations, and limited health care access [58]. These geographic and socioeconomic factors likely contribute to delayed diagnosis and ongoing transmission, particularly among older adult residents with lower health literacy [8,20].

Sensitivity analysis revealed marked regional heterogeneity in the robustness of these spatial patterns. Southeastern hot spots (eg, Pengshui and Qianjiang) demonstrated exceptional stability, persisting even after excluding low-population counties, suggesting that risk in this region is driven by entrenched, population-independent factors. In contrast, northeastern hot spots exhibited greater temporal variability; although the expansion of risk to the northeast since 2022 was reproduced in the sensitivity analysis, the detection of individual counties was more sensitive to spatial weight structures. These findings have practical implications for TB control. The southeastern region is highly stable and requires continuous, population-wide interventions, including enhanced active case detection, routine screening, and health education campaigns [59,60]. However, the northeastern region experiences considerable fluctuations in the epidemic and requires localized strategies, such as conducting large-scale screenings regularly in newly identified high-incidence counties and detecting new cluster cases through real-time monitoring. To further investigate the spatiotemporal clustering of reported TB incidence among older adults in Chongqing and delineate the specific scope and size of the clusters, this study conducted a spatiotemporal scan analysis. The results identified the most likely cluster in 5 southeastern counties: Xiushan, Youyang, Qianjiang, Pengshui, and Wulong. Secondary clusters were distributed across 17 districts and counties in both the northeastern and southeastern regions. These findings are consistent with the results of the spatial analyses. No significant clusters were detected in 2022, which may reflect pandemic-related fluctuations in case detection or a more dispersed spatial distribution during that year. Therefore, for future prevention and control efforts, it is imperative to sustain and enhance surveillance and interventions in these key areas. TB control strategies should be adjusted and refined according to the specific clustering patterns observed in Chongqing to effectively reduce the risk of TB transmission and the overall disease burden.

Limitations

Nevertheless, this study has several limitations. First, the data were derived from passively reported TB cases, which may not capture all incidents due to underreporting. In some remote areas, factors such as limited mobility among older adults and insufficient diagnostic capacity in primary health care institutions could lead to missed cases and incomplete registration. Second, the research period (2020 - 2024) largely overlapped with the COVID-19 pandemic, which may have affected TB case reporting through multiple pathways, such as health care system pressures, lockdown measures, and reduced health care–seeking behavior [3,4]. Although our data came from the mandatory national TB reporting system (TBIMS) with standardized procedures, we cannot rule out a certain degree of underreporting or diagnostic delays during the peak of the pandemic (2020-2022). At the same time, several pandemic-related factors may have introduced difficult-to-quantify confounding effects. The potential impact of these confounding effects on the spatiotemporal distribution characteristics also cannot be entirely ruled out. Despite these potential impacts, the spatiotemporal distribution patterns we identified—particularly the persistent clustering in the northeast and southeast regions of Chongqing—are consistent with prepandemic research results, indicating that the underlying epidemiological patterns remain valid. Future studies that

XSL•FO
RenderX

integrate multisource data, such as active surveillance data, postpandemic follow-up surveys, or health care use data, will help verify trends in true incidence and further clarify the confounding effects caused by the pandemic. Third, the absence of detailed geographical information at the township level for each county precluded a finer-scale spatiotemporal cluster analysis. Finally, as the older adult population is highly vulnerable to TB, its incidence is influenced by multiple factors, including socioeconomic status, health care accessibility, climatic conditions, and the intensity of policy support. The specific impacts of these factors on TB in older adults warrant further investigation.

Conclusions

This study delineated the spatiotemporal distribution patterns and clustering characteristics of TB among individuals aged 65 years and older in Chongqing from 2020 to 2024. The incidence of TB among the older adult population showed a significant downward trend from 2020 to 2022 and no significant change from 2023 to 2024, with cases predominantly observed in older men and peaks occurring during spring and summer. Spatial analysis identified persistent high-cluster areas in both the northeastern and southeastern regions of Chongqing. Spatiotemporal scan analysis further indicated a higher TB burden and increased transmission risk in the southeastern part of the municipality. Therefore, it is crucial to sustain and strengthen TB control measures in these high-risk areas by adopting more targeted outbreak containment strategies, enhancing the diagnostic and treatment capabilities of health care professionals, and reducing underreporting of TB cases among older adults.

Acknowledgments

The authors gratefully acknowledge the Chongqing Tuberculosis Control and Prevention Institute for providing the data used in this study. They also extend their appreciation to all health care workers and staff who have contributed to tuberculosis control and prevention efforts in Chongqing. This work was supported by the Chongqing Public Health Key Special Project (Discipline) Construction Fund, the Scientific Research Project of Chongqing Center for Disease Control and Prevention (2026JKXM037), and the Chongqing High-Level Young and Middle-Aged Medical Talents Project (YXGD202439).

Funding

This research did not receive any specific grant from funding agencies.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to privacy and ethical restrictions (the patient-level data come from the Chinese Tuberculosis Information Management System) but are available from the corresponding author on reasonable request. Data can also be obtained from the Chongqing Tuberculosis Control and Prevention Institute upon reasonable request by contacting the author.

Authors' Contributions

BG, YX, and CP conceptualized and designed the study. WL, LS, and CW assisted in obtaining data. BG, YX, JF, and QW conducted the statistical analysis. BG, YX, SH, XY, and YL assisted in interpreting the data and drafting the manuscript. WZ and CP carried out the literature search and manuscript review. All authors read and approved the final manuscript. CP served as the primary corresponding author responsible for editorial correspondence, whereas WZ served as the co-corresponding author.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sensitivity analysis of tuberculosis in the older adult population in Chongqing from 2020 to 2024 based on population thresholds. [PNG File, 149 KB - publichealth_v12i1e89671_app1.png]

References

1. Global Tuberculosis Report 2024. : World Health Organization; 2024 URL: https://www.who.int/teams/global-programme-on-tuberculosis-and-lung-health/tb-reports/global-tuberculosis-report-2024 [accessed 2025-10-11]
2. Sun R, Wang L, Xia H. Disease burden of tuberculosis in China from 1990 to 2021 and its prediction to 2036. Front Public Health 2025;12:1506266. [doi: 10.3389/fpubh.2024.1506266] [Medline: 39839415]
3. Ageing and health. World Health Organization. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2025-10-11]
4. Allué-Guardia A, Torrelles JB, Sigal A. Tuberculosis and COVID-19 in the elderly: factors driving a higher burden of disease. Front Immunol 2023;14:1250198. [doi: 10.3389/fimmu.2023.1250198] [Medline: 37841265]

XSL•FO
RenderX

5. Di Gennaro F, Vittozzi P, Gualano G, et al. Active pulmonary tuberculosis in elderly patients: a 2016-2019 retrospective analysis from an Italian referral hospital. Antibiotics (Basel) 2020 Aug 7;9(8):489. [doi: 10.3390/antibiotics9080489] [Medline: 32784552]
6. Li SJ, Li YF, Song WM, et al. Population aging and trends of pulmonary tuberculosis incidence in the elderly. BMC Infect Dis 2021 Mar 25;21(1):302. [doi: 10.1186/s12879-021-05994-z] [Medline: 33765943]
7. Pang Y, Fan J, Li T, Xia L, Zhang T. Spatial distribution of pulmonary tuberculosis in Sichuan and Chongqing, 2018–2022 [Article in Chinese]. Dis Surveill 2024;39(7):852-856. [doi: 10.3784/jbjc.202306210319]
8. Yu Y, Wu B, Wu C, Wang Q, Hu D, Chen W. Spatial-temporal analysis of tuberculosis in Chongqing, China 2011-2018. BMC Infect Dis 2020 Jul 22;20(1):531. [doi: 10.1186/s12879-020-05249-3] [Medline: 32698763]
9. Statistical communiqué of the People's Republic of China on the 2024 national economic and social development. National Bureau of Statistics of China. 2025. URL: https://english.www.gov.cn/archive/statistics/202503/01/content_WS67c2695cc6d0868f4e8f02ae.html [accessed 2025-10-11]
10. Taniguchi J, Aso S, Jo T, Matsui H, Fushimi K, Yasunaga H. Factors affecting in-hospital mortality in patients with miliary tuberculosis: A retrospective cohort study. Respir Investig 2024 Jul;62(4):520-525. [doi: 10.1016/j.resinv.2024.04.004] [Medline: 38636244]
11. Wang J, Zhang Y, Rao Q, et al. Factors affecting the readiness for hospital discharge of initially treated pulmonary tuberculosis patients in China: a phenomenological study. BMC Public Health 2024;24(1):2312. [doi: 10.1186/s12889-024-19793-z]
12. Yang Y, Li Z, Liu L, et al. Healthcare-seeking pathway and delay analysis of rifampicin-resistant tuberculosis patient in southwestern China. BMC Public Health 2025;25(1):2019. [doi: 10.1186/s12889-025-23288-w]
13. Arcêncio RA, Berra TZ, Terena NF, et al. Spatial clustering and temporal trend analysis of international migrants diagnosed with tuberculosis in Brazil. PLoS One 2021;16(6):e0252712. [doi: 10.1371/journal.pone.0252712] [Medline: 34106968]
14. Zhu B, Fu Y, Liu J, Mao Y. Spatial distribution of 12 class B notifiable infectious diseases in China: a retrospective study. PLoS One 2018;13(4):e0195568. [doi: 10.1371/journal.pone.0195568] [Medline: 29621351]
15. Guo J, Liu C, Liu F, et al. Tuberculosis disease burden in China: a spatio-temporal clustering and prediction study. Front Public Health 2025;12:1436515. [doi: 10.3389/fpubh.2024.1436515] [Medline: 39839385]
16. Wang L, Li Y, Chen J, Zhang D, Zhang Z, Li X. Spatiotemporal distribution and detection of spatial clusters of tuberculosis in Hubei Province, China using FleXScan (2017–2023). BMC Public Health 2025;25(1):2758. [doi: 10.1186/s12889-025-23620-4]
17. Liyew AM, Gebreyohannes EA, Python A, et al. Mapping tuberculosis prevalence in Africa using a Bayesian geospatial analysis. Commun Med (Lond) 2025 May 27;5(1):194. [doi: 10.1038/s43856-025-00831-9] [Medline: 40425819]
18. Romanyukha AA, Karkach AS, Borisov SE, Belilovsky EM, Sannikova TE, Krivorotko OI. Small-scale stable clusters of elevated tuberculosis incidence in Moscow, 2000-2015: discovery and spatiotemporal analysis. Int J Infect Dis 2020 Feb;91:156-161. [doi: 10.1016/j.ijid.2019.11.015] [Medline: 31759169]
19. Wolde HF, Clements ACA, Alene KA. Local progress towards achieving the End TB targets in Ethiopia: a geospatial analysis. Int J Epidemiol 2025 Aug 18;54(5):dyaf157. [doi: 10.1093/ije/dyaf157] [Medline: 40907466]
20. Statistical yearbook. Chongqing Municipal Bureau of Statistics. URL: https://tjj.cq.gov.cn/zwgk_233/tjnj/ [accessed 2025-10-13]
21. Diagnosis of pulmonary tuberculosis (replacing WS 288—2008). National Health Commission of the People’s Republic of China. 2017. URL: https://www.nhc.gov.cn/wjw/s9491/201712/a452586fd21d4018b0ebc00b89c06254.shtml [accessed 2025-10-13]
22. Technical guidelines for tuberculosis prevention and control in China. : Chinese Center for Disease Control and Prevention URL: https://www.chinacdc.cn/jkyj/crb2/yl/fjh/jswj_fjh/202410/P020241010432930570191.pdf [accessed 2025-10-13]
23. Yin S, Wu S, Huang J, Ren S, Xie W, Peng X. Spatial-temporal analysis of hepatitis B in Fujian Province, China in 2012-2021. Infect Med (Beijing) 2024;3(2):100110. [doi: 10.1016/j.imj.2024.100110] [Medline: 38974348]
24. Chen Y. Spatial autocorrelation equation based on Moran’s index. Sci Rep 2023;13(1):19296. [doi: 10.1038/s41598-023-45947-x]
25. Anselin L. Local indicators of spatial association—LISA. Geogr Anal 2010;27(2):93-115. [doi: 10.1111/j.1538-4632.1995.tb00338.x]
26. Kulldorff M. A spatial scan statistic. Commun Stat Theory Methods 1997;26(6):1481-1496. [doi: 10.1080/03610929708831995]
27. Yun W, Huijuan C, Long L, Xiaolong L, Aihua Z. Time trend prediction and spatial-temporal analysis of multidrug-resistant tuberculosis in Guizhou Province, China, during 2014-2020. BMC Infect Dis 2022 Jun 7;22(1):525. [doi: 10.1186/s12879-022-07499-9] [Medline: 35672746]
28. Liu X. Epidemiological characteristics and treatment outcomes of tuberculosis in the elderly in Xuzhou, 2019-2022 [Article in Chinese]. Jiangsu J Prevent Med 2025;36(2):230-231. [doi: 10.13668/j.issn.1006-9070.2025.02.025]
29. Liu YS, Fan MK, Zhang CY, et al. Analysis of pulmonary tuberculosis detection among patients aged 65 and older in China, 2015-2023 [Article in Chinese]. Zhonghua Liu Xing Bing Xue Za Zhi 2025 Apr 10;46(4):630-637. [doi: 10.3760/cma.j.cn112338-20241029-00671] [Medline: 40274569]

XSL•FO
RenderX

30. China statistical yearbook 2024. National Bureau of Statistics of China. URL: https://www.stats.gov.cn/sj/ndsj/2024/indexch.htm [accessed 2025-10-17]
31. Chen ZY, Deng XY, Zou Y, et al. A spatio-temporal Bayesian model to estimate risk and influencing factors related to tuberculosis in Chongqing, China, 2014–2020. Arch Public Health 2023 Mar 21;81(1):42. [doi: 10.1186/s13690-023-01044-z] [Medline: 36945028]
32. Duan Y, Cheng J, Liu Y, et al. Epidemiological characteristics and spatial-temporal analysis of tuberculosis at the county-level in Shandong Province, China, 2016–2020. Trop Med Infect Dis 2022;7(11):346. [doi: 10.3390/tropicalmed7110346]
33. Yang YB, Liu LL, Chen JO, et al. Predicting the incidence of rifampicin resistant tuberculosis in Yunnan, China: a seasonal time series analysis based on routine surveillance data. BMC Infect Dis 2024 Aug 16;24(1):835. [doi: 10.1186/s12879-024-09740-z] [Medline: 39152374]
34. Filardo TD, Feng PJ, Pratt RH, Price SF, Self JL. Tuberculosis - United States, 2021. MMWR Morb Mortal Wkly Rep 2022 Mar 25;71(12):441-446. [doi: 10.15585/mmwr.mm7112a1] [Medline: 35324877]
35. Glaziou P, Floyd K, Raviglione M. Trends in tuberculosis in the UK. Thorax 2018 Aug;73(8):702-703. [doi: 10.1136/thoraxjnl-2018-211537] [Medline: 29674388]
36. Chen X, Giles J, Yao Y, et al. The path to healthy ageing in China: a Peking University–Lancet Commission. Lancet 2022 Dec;400(10367):1967-2006. [doi: 10.1016/S0140-6736(22)01546-X]
37. Wang L, Liu J, Chin DP. Progress in tuberculosis control and the evolving public-health system in China. Lancet 2007 Feb;369(9562):691-696. [doi: 10.1016/S0140-6736(07)60316-X]
38. Notice on the issuance of the “implementation plan of Chongqing action plan to contain tuberculosis (2019-2022)”. Chongqing Municipal Health Commission. URL: https://wsjkw.cq.gov.cn/zwgk_242/zfxxgkml/zcwj/xzgfxwj2/202004/t20200402_6944117_wap.html [accessed 2025-10-19]
39. 2021 Global Tuberculosis Report. : World Health Organization; 2022 URL: https://www.who.int/zh/publications/b/58335 [accessed 2026-03-13]
40. Pontes TAA, Fernandez-Llimos F, Wiens A. Impact of COVID-19 on tuberculosis notifications. Rev Inst Med Trop Sao Paulo 2024;66:e37. [doi: 10.1590/S1678-9946202466037] [Medline: 38865575]
41. Semnani K, Esmaeili S. Nuances in the global impact of COVID-19 on tuberculosis control efforts: an updated review. Medicine (Baltimore) 2025 Apr 18;104(16):e42195. [doi: 10.1097/MD.0000000000042195] [Medline: 40258740]
42. Khojasteh-Kaffash S, Habibzadeh A, Moghaddam S, Afra F, Samieefar N, Fateh A. Tuberculosis trends in the post-COVID-19 era: is it going to be a global concern? Health Sci Rep 2025 May;8(5):e70792. [doi: 10.1002/hsr2.70792] [Medline: 40406648]
43. Aghajani J, Farnia P, Farnia P, et al. Effect of COVID-19 pandemic on incidence of mycobacterial diseases among suspected tuberculosis pulmonary patients in Tehran, Iran. Int J Mycobacteriol 2022;11(4):415-422. [doi: 10.4103/ijmy.ijmy_167_22] [Medline: 36510928]
44. Zhang Y, Zhang L, Gao W, et al. The impact of COVID-19 pandemic on reported tuberculosis incidence and mortality in China: an interrupted time series analysis. J Glob Health 2023 Oct 13;13:06043. [doi: 10.7189/jogh.13.06043] [Medline: 37824176]
45. Kim B, Kang YA, Lee J. Heterogeneous impact of Covid-19 response on tuberculosis burden by age group. Sci Rep 2022;12(1):13773. [doi: 10.1038/s41598-022-18135-6]
46. Choi NG, Choi BY, Marti CN. Changes in older adults’ frequency of going outside between 2020 and 2021: associations with health status and environmental factors. Clin Gerontol 2023;46(5):745-758. [doi: 10.1080/07317115.2023.2177573] [Medline: 36760067]
47. Grifoni A, Alonzi T, Alter G, et al. Impact of aging on immunity in the context of COVID-19, HIV, and tuberculosis. Front Immunol 2023;14:1146704. [doi: 10.3389/fimmu.2023.1146704] [Medline: 37292210]
48. Jiang H, Chen X, Lv J, et al. Prospective cohort study on tuberculosis incidence and risk factors in the elderly population of eastern China. Heliyon 2024;10(3):e24507. [doi: 10.1016/j.heliyon.2024.e24507]
49. Rickman HM, Phiri MD, Feasey HR, et al. Sex differences in the risk of mycobacterium tuberculosis infection: a systematic review and meta-analysis of population-based immunoreactivity surveys. Lancet Public Health 2025 Jul;10(7):e588-e598. [doi: 10.1016/S2468-2667(25)00120-3] [Medline: 40602858]
50. Horton KC, MacPherson P, Houben R, White RG, Corbett EL. Sex differences in tuberculosis burden and notifications in low- and middle-income countries: a systematic review and meta-analysis. PLoS Med 2016 Sep;13(9):e1002119. [doi: 10.1371/journal.pmed.1002119] [Medline: 27598345]
51. Khan AH, Sulaiman SA, Hassali MA, et al. Effect of smoking on treatment outcome among tuberculosis patients in Malaysia; a multicenter study. BMC Public Health 2020 Jun 4;20(1):854. [doi: 10.1186/s12889-020-08856-6] [Medline: 32498682]
52. Zhang L, Weng TP, Wang HY, et al. Patient pathway analysis of tuberculosis diagnostic delay: a multicentre retrospective cohort study in China. Clin Microbiol Infect 2021 Jul;27(7):1000-1006. [doi: 10.1016/j.cmi.2020.12.031] [Medline: 33421578]
53. Lau LH, Wong NS, Leung CC, et al. Seasonality of tuberculosis in intermediate endemicity setting dominated by reactivation diseases in Hong Kong. Sci Rep 2021 Oct 12;11:20259. [doi: 10.1038/s41598-021-99651-9]

XSL•FO
RenderX

54. Cui Z, Lin D, Chongsuvivatwong V, et al. Spatiotemporal patterns and ecological factors of tuberculosis notification: a spatial panel data analysis in Guangxi, China. PLoS One 2019;14(5):e0212051. [doi: 10.1371/journal.pone.0212051] [Medline: 31048894]
55. Taylan M, Dogru S, Sezgi C, Yilmaz S. Epidemiological trends and seasonal dynamics of tuberculosis in Southeastern Turkey. Niger J Clin Pract 2023 Jul;26(7):928-933. [doi: 10.4103/njcp.njcp_629_22] [Medline: 37635576]
56. Manabe T, Takasaki J, Kudo K. Seasonality of newly notified pulmonary tuberculosis in Japan, 2007-2015. BMC Infect Dis 2019 Jun 6;19(1):497. [doi: 10.1186/s12879-019-3957-8] [Medline: 31170932]
57. Li Y, Zhu L, Lu W, Chen C, Yang H. Seasonal variation in notified tuberculosis cases from 2014 to 2018 in eastern China. J Int Med Res 2020 Aug;48(8):300060520949031. [doi: 10.1177/0300060520949031]
58. Zhang L, Xing W, Zhou J, et al. Characteristics of tuberculosis patients in the integrated tuberculosis control model in Chongqing, China: a retrospective study. BMC Infect Dis 2020 Aug 5;20(1):576. [doi: 10.1186/s12879-020-05304-z] [Medline: 32758169]
59. Mahato RK, Htike KM, Koro AB, et al. Geographical patterns of tuberculosis notification rates and their association with socioeconomic factors in Nepal: a spatial cross-sectional study (2020-2023). BMJ Open 2025 Jul 22;15(7):e093858. [doi: 10.1136/bmjopen-2024-093858] [Medline: 40701587]
60. Wang N, Wu L, Liu Z, et al. Influence of tuberculosis knowledge on acceptance of preventive treatment and the moderating role of tuberculosis stigma among China's general population: cross-sectional analysis. BMC Public Health 2024 Aug 23;24(1):2300. [doi: 10.1186/s12889-024-19812-z] [Medline: 39180047]

Abbreviations

APC: annual percent change
LISA: local indicators of spatial association
LLR: log-likelihood ratio
RR: relative risk
TB: tuberculosis
TBIMS: Tuberculosis Information Management System

Edited by A Mavragani, T Sanchez; submitted 16.Dec.2025; peer-reviewed by FY Sitepu, K Sornlorm; revised version received 18.Mar.2026; accepted 07.Apr.2026; published 05.May.2026.

Please cite as:
Gao B, Xin Y, Liao W, Shi L, Wu C, Fan J, Wang Q, Huang S, Yi X, Li Y, Zhang W, Pu C
Epidemiological Distribution Characteristics of Tuberculosis Among Older Adults in Chongqing (2020-2024): Spatial-Temporal Analysis
JMIR Public Health Surveill 2026;12:e89671
URL: https://publichealth.jmir.org/2026/1/e89671
doi:10.2196/89671

© Bojie Gao, Yu Xin, Wenping Liao, Lin Shi, Chengguo Wu, Jun Fan, Qingya Wang, Shanrong Huang, Xinyuan Yi, Yong Li, Wen Zhang, Chuan Pu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 5.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Evolving Burden of Stroke in China's 832 Poverty-Alleviated Counties (2019-2024): Nationwide Spatiotemporal Analysis

Mengdi Chen[1*], MPH; Dong Xia[1*], PhD; Wenping Li[1*], PhD; Jie Wang[1], MPH; Zhiyu Lv[2], PhD; Jiapeng Chen[2], PhD; Lulu Zhang[1], PhD

[1]Department of Health Services, Naval Medical University, 800 Xiangyin Rd, Shanghai, China

[2]China Population and Development Research Center, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Lulu Zhang, PhD
Department of Health Services, Naval Medical University, 800 Xiangyin Rd, Shanghai, China

Abstract

Background: Stroke remains a leading cause of death and disability in China. Despite the nationwide poverty alleviation being achieved by 2020, residents of poverty-alleviated counties continue to experience poorer health outcomes. There is a critical knowledge gap regarding the evolving spatiotemporal patterns of stroke burden in this large postpoverty population.

Objective: This study aimed to quantify the longitudinal trends in stroke prevalence and incidence across all 832 poverty-alleviated counties in China from 2019 to 2024, stratified by age and sex; to characterize the spatiotemporal evolution and geographic clustering of the stroke burden; and to compare this burden in poverty-alleviated counties with that of the general Chinese population during the key transitional period from 2019 to 2021.

Methods: We conducted a nationwide retrospective analysis using data from China's National Health Poverty Alleviation Dynamic Management System (NHPADMS), covering 56.08 million residents. Temporal trends were analyzed using Joinpoint regression to calculate the annual percentage change and average annual percentage change (AAPC). Spatial clustering was assessed using Global and Local Moran *I*. Comparative analysis used age-standardized prevalence rates, referenced against the Global Burden of Disease standard population.

Results: Stroke prevalence surged markedly from 65.18 to 359.60 per 100,000 population (overall AAPC: +22.95%), with the steepest increase among adults aged 20 to 39 years (AAPC: +35.63%). Stroke incidence remained relatively stable from 2020 to 2023 but declined notably in 2024. While absolute prevalence remained lower, its growth rate in poverty-alleviated regions (+61.85%) far exceeded that in the nonpoverty-alleviated areas (+3.82%) during 2019 to 2021. Spatial analysis revealed persistent hyperendemic clusters in northern provinces (notably the Hebei-Heilongjiang-Jilin-Inner Mongolia corridor) and a significant north-south gradient.

Conclusions: The stroke burden in China's poverty-alleviated counties increased sharply in recorded prevalence, exhibited a worrying shift toward younger adults, and became geographically polarized. Economic advancement has not translated into proportionate health gains. These findings underscore the urgent need for precision public health interventions that are geographically targeted and demographically tailored to address this evolving epidemic.

(JMIR Public Health Surveill 2026;12:e91487) doi:10.2196/91487

KEYWORDS

stroke; poverty-alleviated counties; spatial analysis; incidence; prevalence; China

Introduction

Stroke ranks as the second leading cause of death globally, surpassed only by ischemic heart disease [1]. In China, it remains the leading cause of adult mortality and disability, exhibiting a persistently rising prevalence [2,3]. Notably, rural Chinese populations demonstrate significantly elevated stroke incidence and mortality rates among low-income groups compared to their higher-income counterparts [4].

Over the past decade, China has prioritized poverty eradication, achieving nationwide poverty alleviation across all 832 nationally designated impoverished counties by the end of 2020. Nevertheless, the health status of residents in these newly lifted-out-of-poverty counties continues to trail behind the national average [5]. Despite nationwide prevention initiatives and efforts to promote equitable health care access under the "Healthy China 2030" framework [6], effective burden reduction remains unachieved, particularly among vulnerable populations [7,8].

XSL•FO
RenderX

Prior epidemiological studies on stroke among China's poverty-alleviated populations have significant limitations: they often lack adequate population bases or are constrained by narrow scopes, such as restricting age ranges (eg, 35 - 64 y), exhibiting selection biases, relying on small and nonnationally representative samples, focusing solely on national or provincial scales, or being confined to periods during prealleviation poverty [9-13]. Critically, most foundational studies were conducted 10 to 30 years ago. Furthermore, China has undergone rapid health transitions and profound sociodemographic shifts during the poverty alleviation period [14]. Marked rises in the prevalence of hypertension, smoking, overweight or obesity, and diabetes represent contemporary shifts in key modifiable risk factors, potentially altering the stroke burden [15-17]. This creates a knowledge gap regarding how the burden of stroke has evolved among these populations following the elimination of absolute poverty, especially during the post-2020 transition studies [18,19].

To address this gap, our nationwide investigation conducts spatially explicit, age-stratified, and sex-stratified analyses of stroke incidence and prevalence trends across China's 832 poverty-alleviated counties from 2019 to 2024. This granular methodological approach distinctly identifies microscale geographic disparities typically obscured in aggregated data, thereby establishing a robust epidemiological foundation for designing precision interventions targeting high-risk clusters. This study aims to answer 3 core research questions: (1) How have the prevalence and incidence rates of stroke evolved longitudinally in poverty-alleviated counties, and do these trends exhibit variations by age, sex, or region? (2) What characterizes the spatiotemporal evolution of the stroke burden during the study period, including spatial clustering and regional diffusion patterns? (3) How does the stroke burden in poverty-alleviated counties during the transition period (2019 - 2021) compare with the concurrent burden across contemporary China?

Methods

Data Source

Incidence, prevalence, and population data (2019 - 2024) for the poverty-alleviated population were systematically extracted from China's National Health Poverty Alleviation Dynamic Management System (NHPADMS), maintained by the China Population and Development Research Center. The database encompasses 56.08 million officially registered poverty-alleviated residents across 832 county-level administrative divisions that completed national poverty eradication. Stratification followed NHPADMS age groups: 20 to 39, 40 to 64, and 65 years and older. The annual denominator population consists of these poverty-alleviated residents, which changes dynamically due to death, migration, or birth. The definition of "poverty-alleviated population" is provided by the National Administration for Rural Revitalization of China. The inclusion of "stroke" cases operates under the policy framework established by the National Health Commission of China, which lists stroke as a key disease for specialized treatment in poverty–alleviation health care initiatives. Consequently, all registered cases meet China's current clinical diagnostic criteria for cerebrovascular diseases. Stroke was diagnosed using National Clinical V.2.0 of the International Classification of Diseases, Tenth Revision, disease codes. Case ascertainment relied on passive registry linkage. As health-focused poverty-alleviation efforts have progressed, primary health care facilities in poverty-alleviated counties have carried out systematic screening and registration for key diseases, including stroke, which may have led to a significant increase in the detection rate of cases.

The county-level administrative boundaries used in this study were sourced from the National Geomatics Center of China (NGCC) [20]. All spatial analyses and the resulting distribution maps were generated using the unmodified standard basemaps provided by the NGCC (Map Audit number GS[2024]0650).

Comparative prevalence data for the general population of mainland China (2019 - 2021) were obtained from the Global Burden of Disease (GBD) 2021 study (Institute for Health Metrics and Evaluation) [3].

Temporal Analysis

Temporal trends (2019 - 2024) in stroke prevalence and incidence were quantified using the Joinpoint regression model (Joinpoint Regression Program v5.2.0, National Cancer Institute). The optimal number of joinpoints (turning points) was selected using the Weighted Bayesian Information Criterion (WBIC), which is the default model selection method in Joinpoint version 5.0 and higher. WBIC combines the classical Bayesian Information Criterion (BIC) and the BIC3 (a modified BIC with a harsher penalty) using a data-driven weight based on the partial R^2, thereby adapting to the magnitude of slope changes in the data. This approach does not require manual specification of the maximum number of joinpoints. In Joinpoint regression, the maximum number of joinpoints is inherently constrained by the number of time points; for a 6-year series, the algorithm considers models with up to 3 joinpoints. The model identifies significant temporal turning points (joinpoints), representing years when the trend (eg, prevalence or incidence) slope undergoes a statistically significant change [21]. These turning points reflect shifts in disease dynamics, potentially linked to factors like public health measures or environmental exposures. The annual percentage change (APC) was calculated for each trend segment identified by the selected model, and the average annual percentage change (AAPC) was computed as a summary measure for the entire study period, with 95% CI estimated through the empirical quantile method.

The statistical analysis used the χ^2 test to assess differences in rates (incidence or prevalence) across groups (age and sex strata), with pairwise comparisons adjusted using the Bonferroni correction (α=.05).

Spatial Analysis

Global and local spatial autocorrelation analysis was performed at the county level to identify the geographic clustering of incidence using ArcGIS 10.8 [22]. A spatial weights matrix was constructed based on the Inverse_Distance conceptualization, which defines the spatial relationship between counties as a function of the inverse distance between their geographic

XSL•FO
RenderX

centroids. To ensure the robustness of our findings, we conducted a sensitivity analysis by varying the distance threshold. Specifically, we compared results using 3 distance thresholds: (1) the default threshold automatically determined by ArcGIS, (2) a smaller threshold (50% of the default), and (3) a larger threshold (200% of the default). The statistical significance was determined using a permutation approach with 999 repetitions. For global analysis, we applied Global Moran *I* to assess the overall spatial pattern across the study area [23]. Values approximate −1 (dispersed) to 1 (clustered). A significantly positive Moran *I* (statistical significance was set at $P<.05$ for all analyses) indicates a clustered pattern where similar values are spatially aggregated, whereas a significantly negative value suggests a dispersed pattern. A value near zero implies a random spatial distribution. For local analysis, to pinpoint the specific locations and types of clusters, we conducted a Local Indicators of Spatial Association analysis using Anselin Local Moran *I* [24]. This method identifies statistically significant (statistical significance was set at $P<.05$ for all analyses) spatial units belonging to 4 categories: high-high clusters (a high-incidence county surrounded by other high-incidence counties, ie, a hotspot), low-low clusters (a coldspot), and spatial outliers (high-low or low-high units).

Comparative Analysis With National Data

To control for the influence of differing population age structures on comparability, this study used the direct age standardization method to calculate the stroke prevalence rate. The GBD world population age standard was used as the reference population structure [25]. The key standardized metric, along with its definition and computational formula, is as follows:

Age-standardized prevalence rate (ASPR) represents the prevalence rate adjusted to a standard age distribution. It reflects the stroke prevalence that would be observed if the populations being compared shared an identical age structure, allowing for valid comparisons across groups with different age compositions.

$$(1)\ ASPR=\left(\sum \frac{\text{Age-Stratum prevalent cases}}{\text{Total population}}\right)\times\lambda[\text{age}]\times 100{,}000$$

λ is the standardized proportion according to the GBD world population age standard.

The comparative analysis focused on the years 2019 to 2021 because this period captured the most dramatic initial surge in age-standardized stroke prevalence within the poverty-alleviated population, increasing from 65.18 to 172.79 per 100,000 population. This steep rise established a new baseline for subsequent trends. To determine whether this rapid increase reflected a phenomenon specific to this recently poverty-alleviated group or was part of a broader national pattern, we compared the magnitude and pace of prevalence change with that of the nonpoverty-alleviated regions (defined as the population of China excluding those residing in poverty-alleviated regions) during the same 3-year window. This period also coincided with the conclusive phase of China's targeted poverty alleviation campaign (ending in 2020) and the onset of the pandemic, providing crucial context for interpreting shifts in health care engagement and data reporting.

Ethical Considerations

This study involved a secondary analysis of deidentified national public health records. It was reviewed and confirmed to be exempt from full ethical review by the Naval Medical University Ethics Committee.

Results

Overall Trends in Prevalence and Incidence

From 2019 to 2024, the overall prevalence of stroke demonstrated a marked and sustained increase, rising from 65.18 to 359.60 per 100,000 population. The incidence rate remained relatively stable through 2023 before a notable decline in 2024. Joinpoint regression quantified a significant upward trend in prevalence for the overall population, with an AAPC of +22.95% ($P=.006$). This increasing trend was consistent across all subgroups but varied in magnitude: the most rapid rise was observed in the 20 to 39 years age group (AAPC=+35.63%, $P<.001$), followed by the 40 to 64 years (AAPC=+28.29%, $P<.001$), and the 65 years and older age groups (AAPC=+26.42%, $P<.001$). In the 20 to 39 years age group, prevalent stroke cases rose from 173 (1.07 per 100,000 population) in 2019 to 1728 (10.65 per 100,000 population) in 2024, with the absolute increase accounting for 0.90% of all new cases during the same period. Although the AAPC was highest in this group, the absolute burden remained modest. When stratified by sex, the AAPC was slightly higher for females (+24.31%, $P<.001$) than for males (+22.16%, $P<.001$; Table 1).

For overall stroke incidence, the WBIC selected a model with 2 joinpoints (2021 and 2023). Consequently, 3 trend segments were identified: 2020 to 2021 (APC=−7.99%, $P<.001$), 2021 to 2023 (APC=+2.92%, $P<.001$), and 2023 to 2024 (APC=−32.22%, $P<.001$). The steepest decrease occurred in the last year of the observation period. Cross-sectional analyses revealed consistent patterns within each individual year from 2019 to 2024. Specifically, in every year observed, both stroke prevalence and incidence increased with advancing age, from the 20 to 39 years to the 65 years and older age groups. Additionally, the rates for both metrics were consistently higher in males than in females within each annual cross-section. The most pronounced year-on-year increase in prevalence occurred between 2019 and 2020, whereas the largest annual decrease in incidence was observed between 2023 and 2024 (Figure 1).

To assess whether the observed 2024 incidence decline disproportionately influenced overall trend estimates, we repeated the Joinpoint regression analysis after excluding 2024 data (ie, using only 2020 - 2023 data). The overall incidence trend without 2024 remained stable and nonsignificant, confirming that the sharp decline in 2024 was not a statistical artifact that altered the long-term trend interpretation. In contrast, the prevalence trend without 2024 remained strongly positive (AAPC=+39.34%, 95% CI 16.47% - 65.40%, $P<.001$), consistent with the primary analysis. These results are presented in Multimedia Appendix 1.

XSL•FO
RenderX

Table . Stroke incidence and prevalence per 100,000 individuals (2020 - 2024).[a]

Stroke incidence and prevalence	Year						AC[b]	AAPC[c], % (95% CI)
	2019	2020	2021	2022	2023	2024		
Incidence	N/A[d]	107.80	99.02	102.43	104.87	71.21	−36.59	N/A
Incidence by sex								
Male	N/A	113.25	105.04	109.18	112.61	72.72	−40.53	N/A
Female	N/A	101.45	92.09	99.47	96.66	70.30	−31.15	N/A
χ^2 (df)	—[e]	173[f] (1)	228[f] (1)	125[f] (1)	327[f] (1)	11[f] (1)	—	—
Incidence by age								
20 - 39 y	N/A	2.55	1.20	2.43	2.78	1.35	−1.20	N/A
40 - 64 y	N/A	79.93	91	99.75	88.99	53.28	−26.65	N/A
≥65 y	N/A	384.01	367.29	380.17	426.96	322.54	−61.4	N/A
χ^2 (df)	—	33,100[f] (2)	79,573[f] (2)	73,819[f] (2)	66,903[f] (2)	70,905[f] (2)	—	—
Prevalence	65.18	173.11	172.79	206.63	313.74	359.60	+294.42	22.95[g] (4.72 - 43.73)
Prevalence by sex								
Male	67.86	181.43	180.19	215.54	321.37	368.52	+300.66	22.16[g] (4.65 - 42.05)
Female	62.08	163.44	159.57	195.74	304.24	349.36	+287.28	24.31[g] (3.01 - 49.41)
χ^2 (df)	68[f] (1)	250[f] (1)	348[f] (1)	258[f] (1)	127[f] (1)	139[f] (1)	—	—
Prevalence by age								
20 - 39 y	1.07	3.64	3.89	6.54	9.43	10.65	+9.58	35.63[g] (24.28 - 47.51)
40 - 64 y	42.85	123.33	131.01	175.90	266.41	299.01	+256.16	28.29[g] (10.41 - 48.71)
≥65 y	241.48	637.86	707.68	837.58	1342.21	1538.02	+1296.54	26.42[g] (10.91 - 41.60)
χ^2 (df)	54,201[f] (2)	137,355[f] (2)	156,901[f] (2)	170,572[f] (2)	274,136[f] (2)	131,100[f] (2)	—	—

[a]Stroke incidence and prevalence are based on the number of incident cases and the overall poverty-alleviated population.

[b]AC: absolute change.

[c]AAPC: average annual percentage change. Time periods for the AAPC analysis: prevalence (2020-2024).

[d]N/A: not applicable.

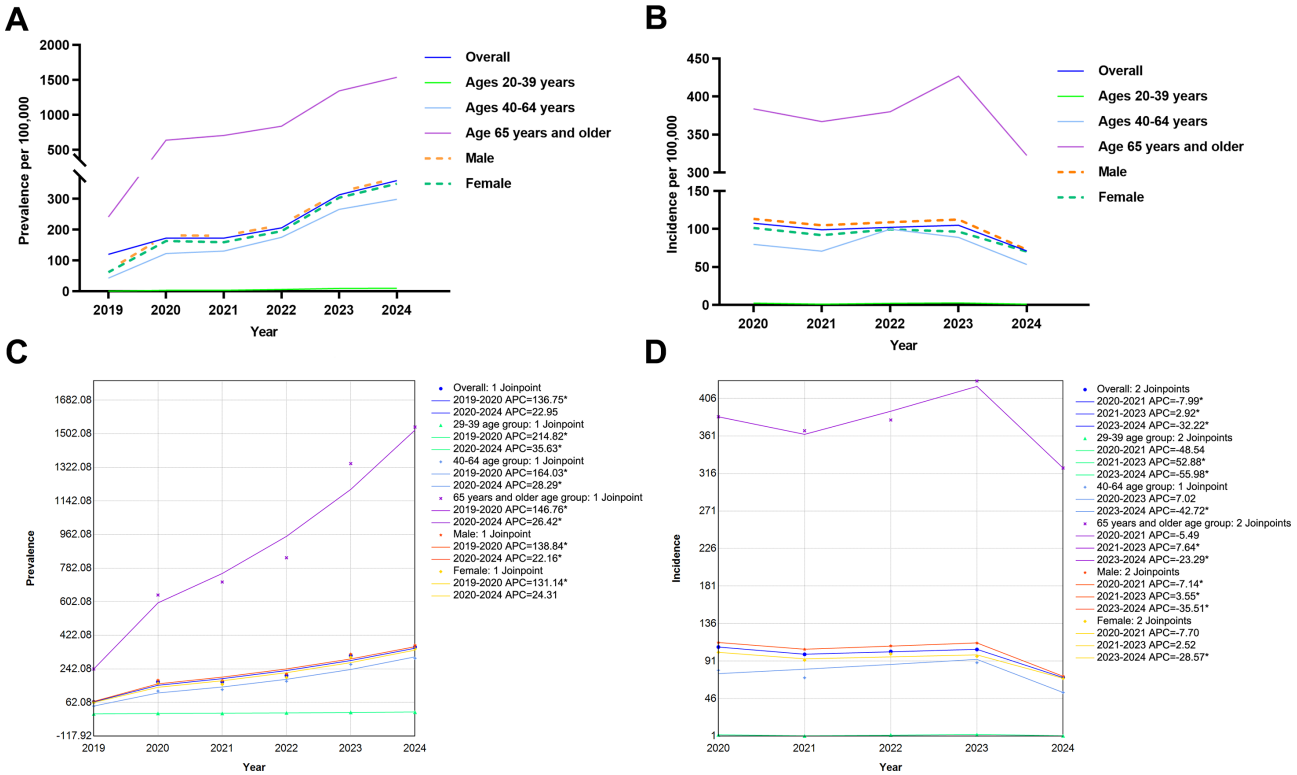
[e]Not applicable (eg, χ^2 values for AC and AAPC columns) or data not available (eg, incidence for 2019).

[f]The χ^2 test result for between-group (age and sex) comparison of rates is statistically significant at $P<.001$.

[g]The AAPC is significantly different from zero at the $\alpha=.05$ level. The AAPC in prevalence rates (2020-2024) was statistically significant across all subgroups, although the AAPC for incidence rates (2020-2024) was not.

XSL•FO
RenderX

Figure 1. Trends in prevalence and incidence rates by age and sex groups among poverty-alleviated populations. (A) Trends in prevalence rates, 2019 - 2024. (B) Trends in incidence rates, 2020 - 2024. (C and D) Joinpoint regression analysis of stroke prevalence and incidence rates by sex and age groups among poverty-alleviated populations. Annual percentage change (APC) in prevalence rates (C) and incidence rates (D). In panels A and B, solid lines represent age groups (ages 20 - 39, 40 - 64, and 65 years and older) and dashed lines represent sex groups (male, female). Data are presented as rates per 100,000 population based on the registered poverty-alleviated population (n=56.08 million) across 832 counties. In panels C and D, * indicates that the APC value is statistically significant (*P*<.05).



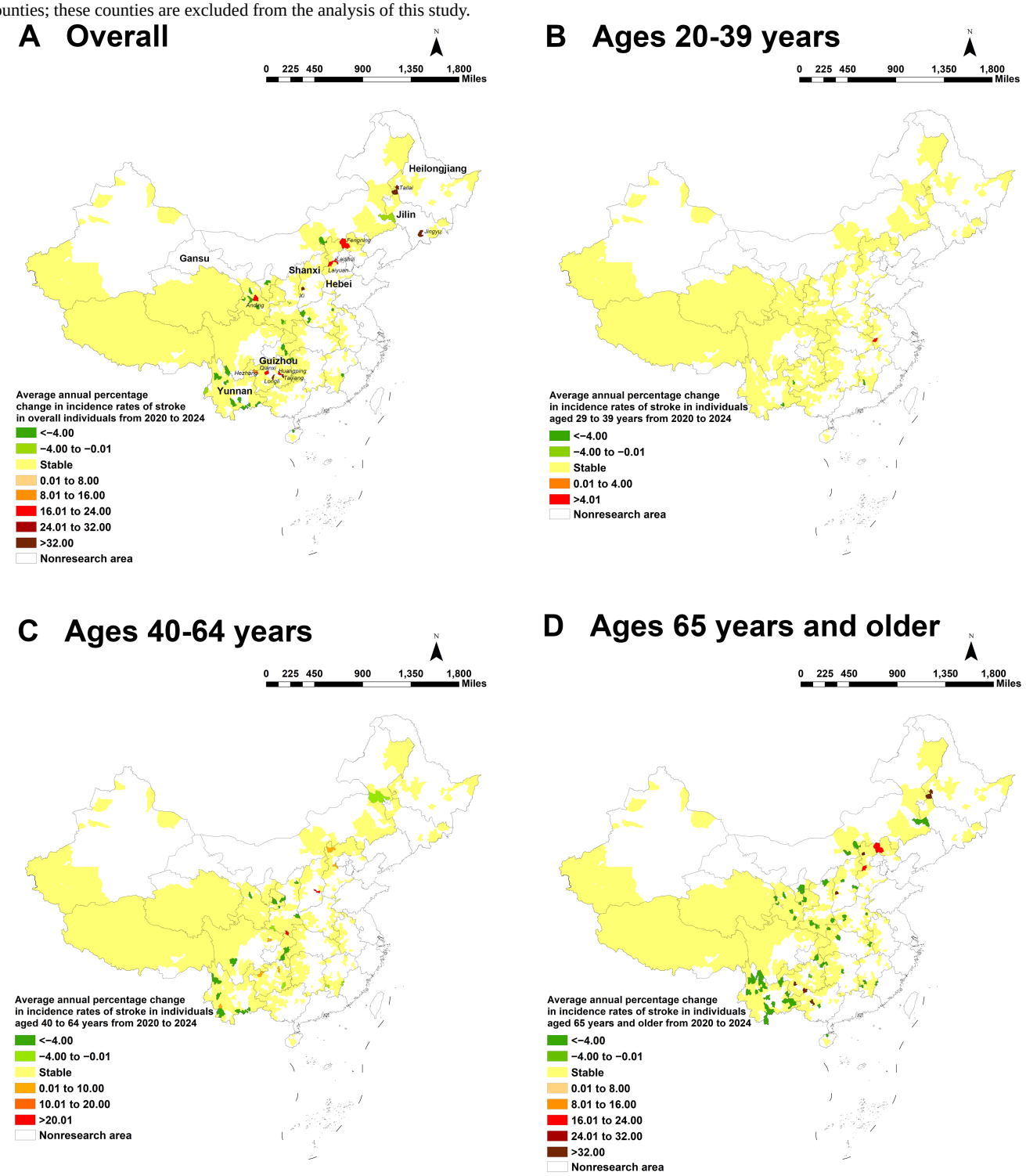
Temporal Distribution by County

Temporal trends in stroke incidence were analyzed in 832 poverty-alleviated counties between 2020 and 2024. Figure 2 presents the AAPC in incidence rates by county. Among these counties, 1.4% (n=12) showed a rising trend in stroke incidence, while 3.4% (n=28) exhibited a declining trend. No statistically significant upward or downward trends in incidence rates were observed in the remaining counties (AAPC=0, *P*>.05). The 12 counties demonstrating upward trends were distributed across 6 provinces: 5 in Guizhou (Qianxi, Hezhang, Huangping, Taijiang, and Longli), 3 in Hebei (Laishui, Laiyuan, and Fengning), and 1 each in Shanxi (Xi), Jilin (Jingyu), Heilongjiang (Tailai), and Gansu (Anding). In contrast, counties with declining trends showed a pronounced concentration in Yunnan Province in southern China, which accounted for 10 of the 28 counties.

Age-stratified analysis indicated that increasing trends were primarily observed in older age groups, with 8 and 9 counties showing rises in the 40 to 64 and 65 years and older age groups, respectively, compared to only one county in the 20 to 39 years age group. Notably, within the 65 years and older age group, Yunnan Province also had the highest number of counties (n=20), exhibiting a declining incidence trend (Figure 2).

XSL•FO
RenderX

Figure 2. The average annual percentage change (AAPC) in stroke incidence among poverty-alleviated counties in China, 2020 - 2024. (A) AAPC of stroke incidence in the overall population at the county level. The 12 counties highlighted in red exhibited a statistically significant upward trend (*P*<.05) and are annotated with their respective county and province names. (B-D) AAPC of stroke incidence stratified by age group: 20 - 39 years (B), 40 - 64 years (C), and 65 years and older (D). White areas represent counties that were never designated as impoverished and thus are not poverty-alleviated counties; these counties are excluded from the analysis of this study.

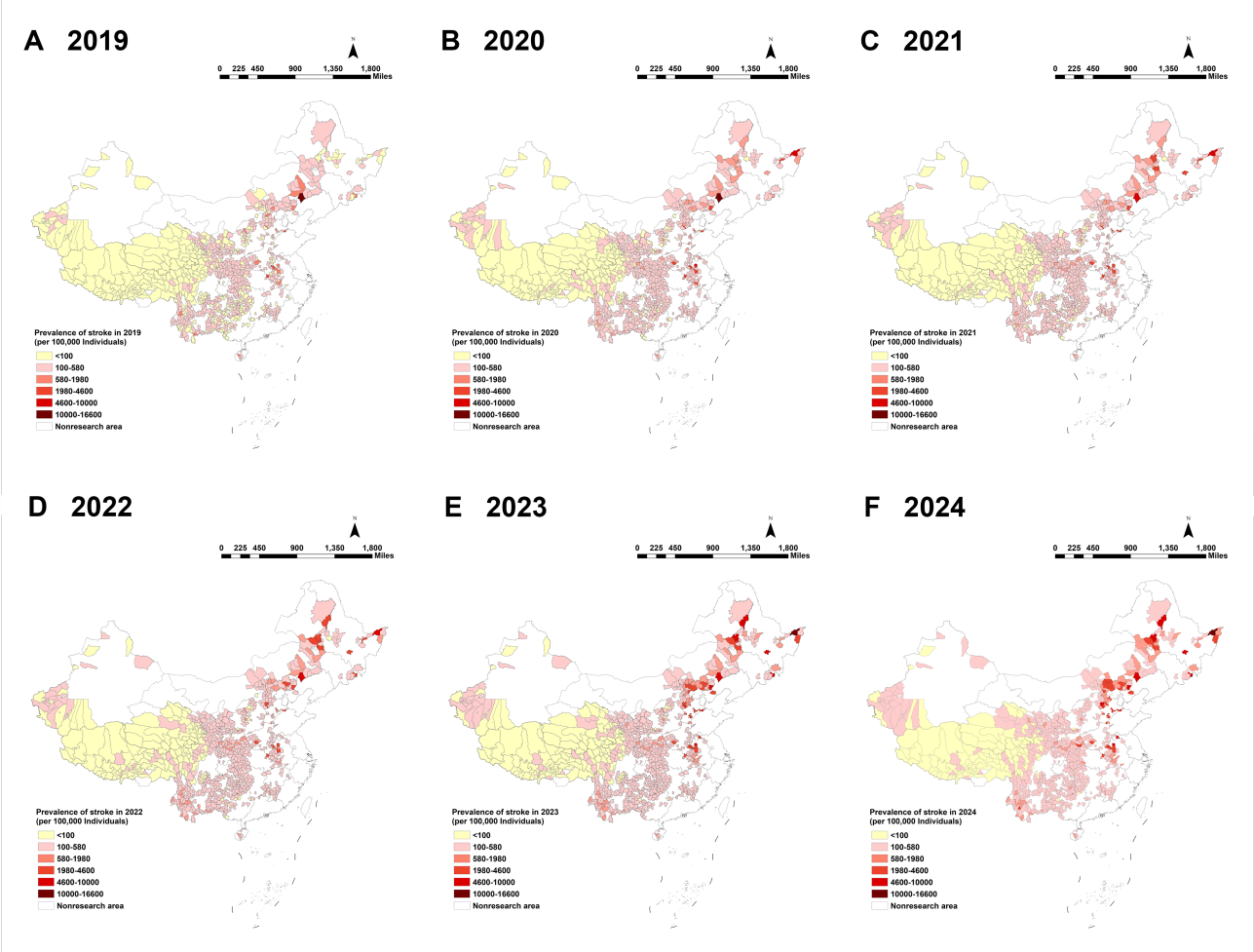


The spatiotemporal analysis from 2019 to 2024 revealed a clear and persistent geographical hierarchy in stroke burden across China's poverty-alleviated counties: high in the north, low in the south, with the most severe concentrations localized in the northeast and north China regions. Counties with the highest prevalence clustered consistently within a northern corridor encompassing Hebei, Heilongjiang, Jilin, and Inner Mongolia, followed by Anhui Province in eastern China. This spatial pattern remained stable throughout the 6-year observation period, during which stroke prevalence rates showed a marked overall increase (Figure 3). The distribution of incidence from 2020 to 2024 mirrored this pattern, with high-incidence counties concentrated in the north (eg, Hebei, Heilongjiang, Inner Mongolia, and Jilin) and low-incidence clusters in the southwest (eg, Sichuan and Qinghai). In 2024, among approximately 54 million residents of poverty-alleviated counties, 25,458 incident

XSL•FO
RenderX

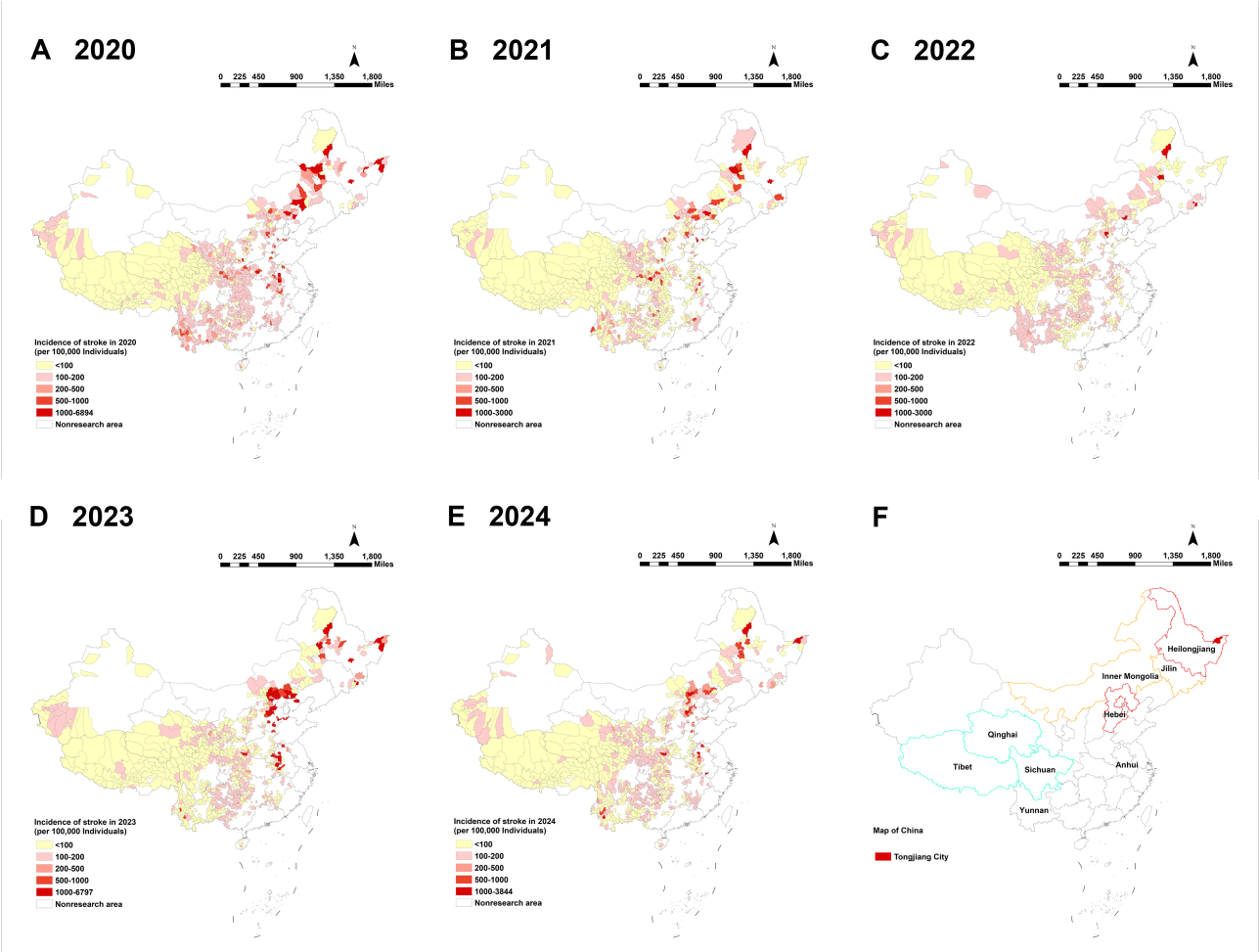
stroke cases and 194,911 prevalent cases were recorded. The highest incidence and prevalence rates nationwide were observed in Tongjiang City, Heilongjiang Province, underscoring the heightened burden in these high-burden northern regions (Figure 4).

Figure 3. Spatiotemporal distribution of stroke prevalence among 832 poverty-alleviated counties in China, 2019 - 2024. (A-F) Spatial distribution of age-standardized stroke prevalence (per 100,000 population) across 832 counties for each year from 2019 (A) to 2024 (F). White areas represent counties that were never designated as impoverished and thus are not poverty-alleviated counties; these counties are excluded from the analysis of this study.



XSL•FO
RenderX

Figure 4. Spatiotemporal distribution of stroke incidence among 832 poverty-alleviated counties in China, 2020 - 2024. (A-E) County-level stroke incidence rates (per 100,000 individuals) for each year from 2020 (A) to 2024 (E). (F) A summary map synthesizing spatial patterns across the study period. Provinces are categorized based on the persistent clustering of incidence: high-incidence provinces (eg, Hebei, Heilongjiang, Jilin, Inner Mongolia, and Anhui) and low-incidence provinces (eg, Qinghai, Sichuan, Tibet, and Yunnan). Tongjiang City (highlighted in red), Heilongjiang Province, is highlighted for recording both the highest incidence and prevalence rates in 2024. A pronounced north-to-south decreasing gradient is evident, with hyperendemic clusters concentrated in the northern Hebei-Heilongjiang-Jilin-Inner Mongolia corridor. White areas represent counties that were never designated as impoverished and thus are not poverty-alleviated counties; these counties are excluded from the analysis of this study.



Spatial Distribution by County

The results of the global spatial autocorrelation analysis are presented in Table 2. All annual Global Moran *I* indices demonstrated positive values (range: 0.09 - 0.30) with statistical significance (*P*<.001), indicating significant spatial clustering of stroke incidence in the general population. The Moran *I* exhibited progressive elevation from 2020 to 2023 (0.09→0.30), followed by a decline to 0.12 in 2024. Compared to the 20 to 39 years age group, both the 40 to 64 years and 65 years and older age cohorts exhibited higher and statistically significant Moran *I* values, revealing stronger spatial clustering of stroke incidence. An exceptional pattern emerged in the 2021 young-middle-aged cohort (20 - 39 y), demonstrating nonsignificant spatial autocorrelation (Moran *I*=0.003, *z*=0.64, *P*>.05). Both sex groups exhibited significant spatial clustering across all years (*P*<.001), though males displayed marginally higher Moran *I* values than females (eg, 2023: 0.32 vs 0.26). The sensitivity analysis using alternative distance thresholds (50% and 200% of the default) yielded Global Moran *I* values that were highly consistent with those obtained using the default threshold (Multimedia Appendix 2), confirming the robustness of our spatial clustering findings.

Statistical significance was achieved for all study years (*P*<.001) except in the 20 to 39 years subgroup. Spatial clustering intensity increased from 2020 to 2023 but declined in 2024. Positive Moran *I* values (*I*>0) indicate spatial clustering where high-incidence counties cluster with adjacent high-incidence counties and low-incidence counties cluster with adjacent low-incidence counties. Values represent Moran *I* with corresponding 95% CI (lower-upper). *P* values were computed using Monte Carlo permutation tests (999 repetitions; significance threshold α=.01). For details on spatial weight matrix construction and sensitivity analysis, see *Methods* (*Spatial Analysis* section).

We conducted a clustering and outlier analysis of stroke incidence rates across population subgroups in 832 counties that achieved poverty alleviation by 2024. The results revealed 4 distinct spatial clustering patterns in overall stroke incidence rates within these counties. High-high clusters predominantly emerged in northern Chinese cities, including Hebei, Inner Mongolia, and Heilongjiang, where stroke incidence

XSL•FO
RenderX

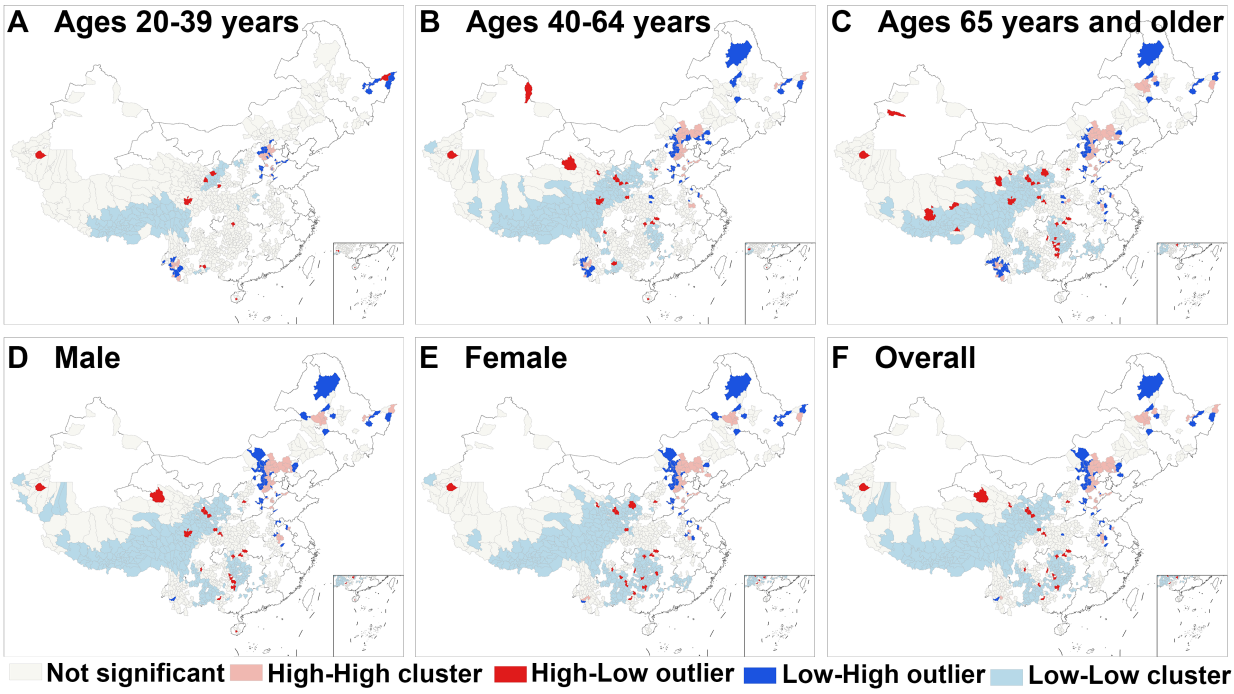
significantly exceeded that of neighboring areas and was consistent across all age-sex groups, except for those aged 20 to 39 years. This pattern may indicate high-risk disease zones. Identifying these high-cluster zones is essential for developing targeted disease prevention and control strategies; therefore, we compiled a list of county-level high-cluster zones with elevated stroke incidence across all subgroups for 2024 (see Multimedia Appendix 3 for details). Low-low clusters were concentrated in extensive contiguous areas of central and western China, such as Sichuan, Qinghai, Yunnan, and Guizhou, where stroke incidence was markedly lower than in neighboring regions. High-low clusters were scattered throughout central and western China, where counties exhibited higher stroke incidence while neighboring counties reported lower rates, potentially reflecting localized disease hotspots. Low-high clusters were sporadically distributed in northern China, where these areas demonstrated lower stroke incidence compared to neighboring regions with higher rates, possibly representing transitional zones for disease transmission (Figure 5).

Table . Global (Moran *I*) index for spatial association of stroke incidence (2020 - 2024).

Year	Overall			Ages 20 - 39 y			Ages 40 - 64 y			Ages ≥65 y			Male			Female		
	Moran *I*	*z* scores	Pattern	Moran *I*	*z* scores	Pattern	Moran *I*	*z* scores	Pattern	Moran *I*	*z* scores	Pattern	Moran *I*	*z* scores	Pattern	Moran *I*	*z* scores	Pattern
2020	0.09[a]	11.27	Clustered	0.06[a]	7.64	Clustered	0.08[a]	10.45	Clustered	0.09[a]	11.15	Clustered	0.09[a]	10.98	Clustered	0.09[a]	11.56	Clustered
2021	0.09[a]	11.31	Clustered	0.00	0.64	Random	0.08[a]	10.34	Clustered	0.07[a]	8.48	Clustered	0.10[a]	11.73	Clustered	0.08[a]	9.99	Clustered
2022	0.19[a]	24.31	Clustered	0.03[a]	6.65	Clustered	0.29[a]	34.88	Clustered	0.37[a]	43.64	Clustered	0.20[a]	24.78	Clustered	0.15[a]	19.16	Clustered
2023	0.30[a]	35.76	Clustered	0.43[a]	53.18	Clustered	0.34[a]	40.41	Clustered	0.33[a]	39.77	Clustered	0.32[a]	38.19	Clustered	0.26[a]	31.87	Clustered
2024	0.12[a]	15.16	Clustered	0.05[a]	7.93	Clustered	0.14[a]	17.57	Clustered	0.10[a]	12.06	Clustered	0.12[a]	14.70	Clustered	0.12[a]	15.46	Clustered

[a]The corresponding significance level is $P<.001$.

Figure 5. Local spatial clustering of stroke incidence among 832 poverty-alleviated counties in China, 2024. Local spatial clustering of stroke incidence in individuals aged 20 - 39 years (A), individuals aged 40 - 64 years (B), individuals aged 65 years and older (C), male (D), female (E), and overall population (F). Cluster types were classified using local Moran *I*: high-high: high-prevalence county surrounded by high-prevalence neighbors (hotspot area); low-low: low-prevalence county surrounded by low-prevalence neighbors (coldspot area); high-low: high-prevalence county surrounded by low-prevalence neighbors (spatial outlier); low-high: low-prevalence county surrounded by high-prevalence neighbors (spatial outlier). White areas represent counties that were never designated as impoverished and thus are not poverty-alleviated counties; these counties are excluded from the analysis of this study.



XSL•FO
RenderX

Comparative Analysis

A comparative analysis between poverty-alleviated regions and nonpoverty-alleviated regions of China (2019 - 2021) revealed stark differences in the temporal trend of stroke prevalence. From 2019 to 2021, the age-standardized prevalence of stroke increased markedly in poverty-alleviated regions, with an APC of +61.85% (95% CI 14.99% - 129.60%, *P*<.001). This rate of increase was considerably steeper than the concurrent trend observed in nonpoverty-alleviated regions (APC=+3.82%, 95% CI 2.31% - 5.40%, *P*<.001). This accelerating trend was most pronounced in the 20 to 39 years age group, where the APC reached +88.96% (*P*<.001) in poverty-alleviated regions, compared to a much more modest increase of +1.94% (*P*<.001) in the same age group in nonpoverty-alleviated regions (Table 3).

It is important to note that the absolute prevalence estimates between the 2 datasets are not directly comparable due to fundamental differences in data sources and population coverage (see *Limitations* section). Therefore, only the direction and magnitude of temporal change (APC) are presented for comparative purposes. The underlying absolute values are provided in Multimedia Appendix 4 for transparency.

Table . Annual percentage change (APC, 95% CI) of stroke prevalence in poverty-alleviated regions and China (2019 - 2021).[a]

Region	Overall, % (95% CI)	Ages 20 - 39 y, % (95% CI)	Ages 40 - 64 y,% (95% CI)	Ages ≥65 y, % (95% CI)
China (overall)	1.64[b] (0.32 to 3.03)	1.78[b] (1.09 to 2.47)	1.57[b] (0.67 to 2.50)	1.73[b] (0.04 to 3.62)
Poverty-alleviated regions	61.85[b] (14.99 to 129.60)	88.96[b] (26.77 to 184.81)	73.62[b] (22.90 to 147.48)	74.10[b] (26.97 to 136.55)
Nonpoverty-alleviated regions	3.82[b] (2.31 to 5.40)	1.94[b] (1.11 to 2.77)	1.57[b] (0.52 to 2.65)	0.96 (−0.72 to 2.83)

[a]The group labeled "Nonpoverty-alleviated regions" refers to data for China excluding 832 designated poverty-alleviated regions. Prevalence rates were age-standardized using the Global Burden of Disease 2021 reference population structure (Institute for Health Metrics and Evaluation) to ensure comparability.

[b]Statistically significant annual trend, with a *P*<.001 for the APC. The APC and its 95% CI were calculated using Joinpoint regression (National Cancer Institute, V.5.2.0). *P*<.001 indicates a statistically significant annual trend in prevalence. As the model failed to detect the inflection point, the average annual percentage change and APC values are identical.

Discussion

Principal Results

This study presents the first retrospective analysis of data from all 832 poverty-alleviated counties in China, identifying the spatiotemporal trends of stroke epidemiology from 2019 to 2024 and revealing significant disparities in disease burden across different age groups, sexes, and geographical regions. This represents the largest spatiotemporal analysis of postpoverty transition epidemiology, overcoming prior fragmentation.

The increased stroke burden in China's poverty-alleviated population, with the heaviest burden observed in individuals aged 65 years and older, aligns with previously documented trends in rural areas [7,26]. Stroke prevalence surged significantly in poverty-alleviated counties, with the 20 to 39 years cohort exhibiting the steepest acceleration (AAPC=+35.63%)—a trend disproportionately exceeding concurrent national growth. The rising prevalence of stroke among young adults, particularly in the age group of 20 to 39 years, has become a notable concern, especially linked to rapid urbanization in low-income regions. Two interlinked mechanisms may explain this trend: (1) behavioral risks: these regions have been experiencing a surge in traditional vascular risk factors, including hypertension, diabetes, and dyslipidemia, which are commonly associated with urbanization, such as sedentary lifestyles [27] and changes in dietary habits due to increased access to processed foods and reduced physical activity levels. For instance, one study reported an alarming rise in traditional risk factors among young adults, emphasizing how lifestyle changes linked to urbanization contribute to this trend in stroke incidence [28-30]. (2) Health care access gaps. Socioeconomic factors, including decreased access to health care and health education, leave young adults vulnerable to undetected vascular risks. It has been observed that in low-income communities, awareness and response to stroke symptoms are often inadequate, which can lead to delayed treatment and worse outcomes [31,32]. Importantly, the outmigration of working-age populations may concentrate stroke-prone individuals (eg, those with disabilities or chronic illnesses) in impoverished counties [33,34]. These factors collectively signal a failure of current preventive strategies to adapt to postpoverty health challenges.

The decline in incidence in 2024 should be interpreted cautiously. This decline may reflect a combination of factors, including registration delays (cases from late 2024 not yet entered into the system), residual postpandemic changes in health care utilization, and possibly random variation. A sensitivity analysis excluding 2024 data (Table S4 in Multimedia Appendix 4) showed that the incidence trend from 2020 to 2023 remained nonsignificant, indicating that the sharp decline in 2024 did not fundamentally alter the overall conclusion of no sustained increase in incidence. Longer-term data are required to determine whether this decline represents a genuine epidemiological shift.

Though males retained higher absolute prevalence, females experienced faster escalation (+24.31% vs +22.16%), uncovering neglected vulnerabilities. Hypertension in pregnancy, postpartum hemorrhage, and changes in vascular biomarkers [35-37] in women during middle age are women-specific risk factors that increase the risk of stroke [38]. Women in rural areas and low-income families are more susceptible to chronic diseases such as hypertension and diabetes, which further increase the

XSL•FO
RenderX

risk of stroke [39]. Women are more susceptible to stress, depression, and psychological factors that increase the risk of stroke [35,40]. Sex differences in health literacy also affect stroke prevention and management. For example, 32% of rural women recognize the risk of high blood pressure, while 48% of men do [41]. These findings demand sex-specific prevention frameworks [35].

The analysis reveals a distinct and persistent north-south divergence in stroke burden across poverty-alleviated counties. Stroke prevalence demonstrates hyperendemic clusters in northern regions (notably the Hebei-Heilongjiang-Jilin-Inner Mongolia corridor), with a clear north-to-south decreasing gradient—exemplified by Tongjiang City in Heilongjiang recording the peak prevalence intensity in 2024. Similarly, incidence forms geographically anchored high-high clusters across these northern provinces, contrasting sharply with expansive low-low belts spanning southwestern regions such as Sichuan, Yunnan, and Qinghai. Three regional drivers likely contribute to this pattern of elevated stroke burden and high incidence in northern China's poverty-alleviated counties: (1) cold climate effects: prolonged indoor confinement during the harsh northern winters disproportionately affects residents in poverty-alleviated counties. This behavior correlates with elevated blood pressure [42] and increased exposure to particulate matter (PM2.5) from biomass heating [43], a common practice in these areas. (2) Dietary sodium overload: Northeastern diets average 11.2 g of salt per day [44], far surpassing the World Health Organization recommendation of 5 g—a key modifiable risk factor. This high sodium intake is a key modifiable risk factor that directly elevates blood pressure [2,45], creating a persistent and widespread driver of stroke risk within the local populace. (3) Health care delivery bottlenecks: Northeastern regions, such as Heilongjiang Province, have a relative lack of medical resources, especially in rural and remote areas, where medical conditions are poor and the poverty-alleviated population has insufficient awareness of stroke [46], leading to inadequate preventive and treatment measures. Despite the marked rise in prevalence, an important nuance emerges when examining incidence. Overall stroke incidence across the 832 counties remained largely flat throughout the study period until 2024. However, this aggregate stability masks localized intensification: significant rises were concentrated in only 12 counties, predominantly driven by populations aged 40 years and older. This pattern suggests that while the overall entry rate of new stroke cases was stable at the macro level, specific geographic clusters—particularly those with aging demographics—are experiencing a tightening grip of the epidemic. We hypothesize that the decline in Moran *I* in 2024 may reflect temporal variation in data reporting completeness or random fluctuation, but further observation is needed. This finding shifts the focus from a universal incidence surge to the identification of high-risk geographic and demographic pockets where primary prevention efforts may be failing or where demographic aging is exerting its strongest effect.

During the pivotal postalleviation transition (2019 - 2021), a striking disparity in the temporal dynamics of stroke prevalence emerged between poverty-alleviated and nonpoverty-alleviated regions. While the recorded age-standardized stroke prevalence in poverty-alleviated regions increased at an annual rate of 61.85% over 2 years, the concurrent increase in nonpoverty-alleviated regions was only 3.82%. We caution that the absolute prevalence estimates from the 2 data sources are not directly comparable due to fundamental methodological differences (see Limitations section). Therefore, only the direction and magnitude of temporal change (APC) are comparable between the 2 populations; any comparison of absolute rates or inference about confounding factors is not valid. Nevertheless, the markedly faster pace of increase in poverty-alleviated regions suggests a distinct epidemiological trajectory during the postalleviation transition period. This disparity can be understood as a "catch-up" phenomenon driven by 2 concurrent forces. First, the conclusive phase of the national poverty alleviation campaign (ending in 2020) significantly enhanced primary health care access and systematic screening in these regions, leading to improved detection and reporting of preexisting, previously undiagnosed cases [47]. Second, despite improved access, a relative lag in population-level stroke awareness and health literacy persists [48]. This lag likely results in delayed health care–seeking for transient ischemic attacks or minor strokes, allowing many cases to progress to a more severe or recordable stage before diagnosis, thereby contributing to the rising incidence of recorded prevalent cases. Thus, the steep upward trend does not merely represent a convergence toward the level observed in nonpoverty-alleviated regions but reveals a critical phase during which previously unmet health care needs became visible. Importantly, this surge should be interpreted primarily as a "detection catch-up" of previously unregistered cases, rather than as a mere epidemiological explosion.

In conclusion, this nationwide spatiotemporally explicit analysis reveals that the stroke burden in China's poverty-alleviated counties is not merely a residual challenge of the past but a dynamically evolving epidemic characterized by a sharp increase in recorded prevalence, alarming shifts toward younger adults, and stark geographical polarization. Our findings fundamentally recalibrate the postpoverty health narrative: economic advancement has not automatically translated into commensurate health gains. Instead, it may have unmasked or even accelerated latent epidemiological risks. The concentrated burden in northern "hotspots" and among the young working-age population signals specific failures in current primary prevention strategies. Therefore, moving beyond uniform, poverty-alleviation-era health policies is imperative. Future interventions must be geographically precision-targeted to the identified high-risk clusters—hotspot identification enables precision resource allocation under China's Rural Revitalization Health Initiatives—and demographically tailored to address the unique vulnerabilities of young adults and women. This shift is crucial for mitigating the impending wave of noncommunicable diseases and achieving equitable health outcomes in the post-poverty alleviation era.

Limitations

Despite its comprehensive scope, this study has several limitations that should be considered when interpreting the findings. First, a fundamental limitation is the inability to distinguish the extent to which the observed surge in stroke

XSL•FO
RenderX

prevalence reflects a true increase in disease burden versus improved case ascertainment driven by poverty-alleviation policies. The national poverty-alleviation campaign (ending in 2020) substantially enhanced primary health care access and systematic screening in poverty-alleviated counties, likely leading to the detection and registration of previously undiagnosed prevalent cases. Consequently, the reported trends should be interpreted as a combination of genuine epidemiological change and policy-induced detection improvement. Second, reliance on retrospective surveillance data may introduce variability in diagnostic accuracy and case ascertainment across different counties and years. While we applied standardized case definitions, inconsistencies in medical record-keeping and reporting completeness, particularly in remote areas, could affect the accuracy of incidence and prevalence estimates. Third, the 832 poverty-alleviated counties are not a homogeneous group. Significant internal socioeconomic and ethnic heterogeneity exists within these counties, which our aggregate-level analysis could not dissect. Future studies should incorporate township-level data. Fourth, the comparative analysis between poverty-alleviated and nonpoverty-alleviated regions is constrained by the heterogeneity of data sources. The poverty-alleviated data were extracted from the NHPADMS (a policy-driven active registry), whereas the national comparator data were derived from the GBD 2021 study (a modeling-based estimate). These 2 sources differ substantially in case definitions, case ascertainment completeness, and population coverage. Therefore, absolute prevalence estimates are not directly comparable across the 2 datasets; the present analysis focuses only on comparing the rate of change (APC). No confounding factors (eg, age, sex, comorbidities, socioeconomic status) could be controlled for across the 2 datasets due to their heterogeneous designs. Therefore, readers should interpret any cross-dataset comparison with extreme caution, and only the temporal trend direction (APC) is considered reliable for comparative purposes. Fifth, this study did not differentiate between ischemic and hemorrhagic stroke subtypes. These 2 subtypes have distinct pathophysiological mechanisms, risk factor profiles (eg, hypertension is more strongly associated with hemorrhagic stroke, while atrial fibrillation and dyslipidemia are more linked to ischemic stroke), and treatment strategies. The NHPADMS database does not routinely record stroke subtype information, precluding subtype-specific analysis. Future studies should incorporate stroke subtype data to better tailor prevention and intervention strategies for poverty-alleviated populations.

Conclusion

This study provides the first systematic evidence that the stroke burden in China's poverty-alleviated counties is dynamically evolving, characterized by a sharp increase in recorded prevalence, a pronounced shift toward younger adults (aged 20 - 39), and distinct geographic clustering in northern regions. These findings indicate a clear disconnect between economic advancement and health gains in the postpoverty transition era. Our high-resolution analysis thus furnishes critical evidence for precision public health, underscoring the imperative to shift interventions toward geographically targeted strategies for identified high-risk clusters and tailored prevention for younger populations to achieve equitable health outcomes.

Acknowledgments

The author would like to express sincere gratitude to the China Population and Development Research Center for establishing the dataset, which provides researchers with high-quality data resources. The author declares that no generative artificial intelligence was used in the writing of this manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (72474220).

Data Availability

The data on the impoverished population used in this research are obtainable by submitting a formal data request via email to info@cpdrc.org.cn to the China Population and Development Research Center under specific licensing agreements. To download Global Burden of Disease (GBD) data used in the analyses, please visit the GBD 2021 Results Tool.

Authors' Contributions

MC and LZ developed the research aim and provided overall guidance. MC, DX, and WL analyzed the data and drafted the manuscript. JW, ZL, and JC contributed to data collection and preparation. All authors have critically reviewed and approved the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sensitivity analysis of Joinpoint regression trends for stroke incidence and prevalence after excluding 2024 data in poverty-alleviated counties, China.
[DOCX File, 23 KB - publichealth_v12i1e91487_app1.docx]

XSL•FO
RenderX

Multimedia Appendix 2
Sensitivity analysis of global spatial autocorrelation (Moran *I*) for stroke incidence across different distance thresholds in poverty-alleviated counties, China (2020-2024).
[DOCX File, 34 KB - publichealth_v12i1e91487_app2.docx]

Multimedia Appendix 3
List of counties identified as high-high spatial clusters for stroke incidence, stratified by demographic subgroups, in poverty-alleviated counties of China, 2024.
[DOCX File, 22 KB - publichealth_v12i1e91487_app3.docx]

Multimedia Appendix 4
Stroke prevalence in poverty-alleviated regions and China (2019-2021).
[DOCX File, 22 KB - publichealth_v12i1e91487_app4.docx]

References

1. GBD 2023 Causes of Death Collaborators. Global burden of 292 causes of death in 204 countries and territories and 660 subnational locations, 1990–2023: a systematic analysis for the Global Burden of Disease Study 2023. The Lancet 2025 Oct;406(10513):1811-1872. [doi: 10.1016/S0140-6736(25)01917-8] [Medline: 41092928]
2. Wu S, Wu B, Liu M, et al. Stroke in China: advances and challenges in epidemiology, prevention, and management. LANCET Neurol 2019 Apr;18(4):394-405. [doi: 10.1016/S1474-4422(18)30500-3] [Medline: 30878104]
3. Li Z, Jiang Y, Li H, Xian Y, Wang Y. China's response to the rising stroke burden. BMJ 2019 Feb 28;364:l879. [doi: 10.1136/bmj.l879] [Medline: 30819725]
4. Li B, Lou Y, Gu H, et al. Trends in incidence of stroke and transition of stroke subtypes in rural Tianjin China: a population-based study from 1992 to 2012. PLoS ONE 2015;10(10):e0139461. [doi: 10.1371/journal.pone.0139461] [Medline: 26426803]
5. Wang W, Chen K, Xiao W, Du J, Qiao H. Determinants of health poverty vulnerability in rural areas of Western China in the post-poverty relief era: an analysis based on the Anderson behavioral model. BMC Public Health 2024;24(1):459. [doi: 10.1186/s12889-024-18035-6] [Medline: 38355428]
6. Tan X, Liu X, Shao H. Healthy China 2030: a vision for health care. Value Health Reg Issues 2017 May;12:112-114. [doi: 10.1016/j.vhri.2017.04.001] [Medline: 28648308]
7. Liu J, Shi Z, Bai R, et al. Temporal, geographical and demographic trends of stroke prevalence in China: a systematic review and meta-analysis. Ann Transl Med 2020 Nov;8(21):1432-1432. [doi: 10.21037/atm-19-4342] [Medline: 33313177]
8. Ouyang F, Wang Y, Huang W, et al. Association between socioeconomic status and post-stroke functional outcome in deprived rural southern China: a population-based study. BMC Neurol 2018 Jan 25;18(1):12. [doi: 10.1186/s12883-018-1017-4] [Medline: 29370778]
9. Ning X, Sun J, Jiang R, et al. Increased stroke burdens among the low-income young and middle aged in rural China. Stroke 2017 Jan;48(1):77-83. [doi: 10.1161/STROKEAHA.116.014897] [Medline: 27924051]
10. Wang J, An Z, Li B, et al. Increasing stroke incidence and prevalence of risk factors in a low-income Chinese population. Neurology 2015 Jan 27;84(4):374-381. [doi: 10.1212/WNL.0000000000001175] [Medline: 25540314]
11. Wang J, Bai L, Shi M, et al. Trends in age of first-ever stroke following increased incidence and life expectancy in a low-income Chinese population. Stroke 2016 Apr;47(4):929-935. [doi: 10.1161/STROKEAHA.115.012466] [Medline: 26869385]
12. Yan R, Li W, Yin L, Wang Y, Bo J, PURE - China Investigators. Cardiovascular diseases and risk-factor burden in urban and rural communities in high-, middle-, and low-income regions of China: a large community-based epidemiological study. J Am Heart Assoc 2017 Feb 6;6(2):e004445. [doi: 10.1161/JAHA.116.004445] [Medline: 28167497]
13. Tu WJ, Zhao Z, Yin P, et al. Estimated burden of stroke in China in 2020. JAMA Netw Open 2023 Mar 1;6(3):e231455. [doi: 10.1001/jamanetworkopen.2023.1455] [Medline: 36862407]
14. Li W, Lv Z, Chen M, et al. Spatiotemporal inequality and financial toxicity of leukemia in post-poverty China: a national analysis of 832 counties (2019-2024). Front Public Health 2025;13:1611102. [doi: 10.3389/fpubh.2025.1611102] [Medline: 41080883]
15. GBD 2021 Adult BMI Collaborators. Global, regional, and national prevalence of adult overweight and obesity, 1990-2021, with forecasts to 2050: a forecasting study for the Global Burden of Disease Study 2021. The Lancet 2025 Mar;405(10481):813-838. [doi: 10.1016/S0140-6736(25)00355-1] [Medline: 40049186]
16. GBD 2023 Disease and Injury and Risk Factor Collaborators. Burden of 375 diseases and injuries, risk-attributable burden of 88 risk factors, and healthy life expectancy in 204 countries and territories, including 660 subnational locations, 1990-2023: a systematic analysis for the Global Burden of Disease Study 2023. Lancet 2025 Oct 18;406(10513):1873-1922. [doi: 10.1016/S0140-6736(25)01637-X] [Medline: 41092926]

XSL•FO
RenderX

17. Zhou M, Wang H, Zeng X, et al. Mortality, morbidity, and risk factors in China and its provinces, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet 2019 Sep;394(10204):1145-1158. [doi: 10.1016/S0140-6736(19)30427-1] [Medline: 31248666]
18. Ren Y, Yang J, Yin P, et al. Urban-rural disparities in mortality due to stroke subtypes in China and its provinces, 2015-2020. Chin Med J (Engl) 2025 Jun 5;138(11):1345-1354. [doi: 10.1097/CM9.0000000000003135] [Medline: 39329276]
19. Zhao Y, Hua X, Ren X, et al. Increasing burden of stroke in China: a systematic review and meta-analysis of prevalence, incidence, mortality, and case fatality. Int J Stroke 2023 Mar;18(3):259-267. [doi: 10.1177/17474930221135983] [Medline: 36274585]
20. Tianditu: national platform for common geospatial information services. National Geomatics Center of China. 2016. URL: http://bzdt.ch.mnr.gov.cn/ [accessed 2018-05-18]
21. Clegg LX, Hankey BF, Tiwari R, Feuer EJ, Edwards BK. Estimating average annual per cent change in trend analysis. Stat Med 2009 Dec 20;28(29):3670-3682. [doi: 10.1002/sim.3733] [Medline: 19856324]
22. Martins-Melo FR, Pinheiro MCC, Ramos AN Jr, Alencar CH, de Moraes Bezerra FS, Heukelbach J. Spatiotemporal patterns of schistosomiasis-related deaths, Brazil, 2000-2011. Emerg Infect Dis 2015 Oct;21(10):1820-1823. [doi: 10.3201/eid2110.141438] [Medline: 26401716]
23. Dogru AO, David RM, Ulugtekin N, Goksel C, Seker DZ, Sözen S. GIS based spatial pattern analysis: children with hepatitis A in Turkey. Environ Res 2017 Jul;156:349-357. [doi: 10.1016/j.envres.2017.04.001] [Medline: 28391174]
24. Ren Z, Yang K, Dong W. Spatial analysis and risk assessment model research of arthritis based on risk factors: China, 2011, 2013 and 2015. IEEE Access 2020;8:206406-206417. [doi: 10.1109/ACCESS.2020.3037912]
25. GBD 2021 Demographics Collaborators. Global age-sex-specific mortality, life expectancy, and population estimates in 204 countries and territories and 811 subnational locations, 1950-2021, and the impact of the COVID-19 pandemic: a comprehensive demographic analysis for the Global Burden of Disease Study 2021. Lancet 2024 May 18;403(10440):1989-2056. [doi: 10.1016/S0140-6736(24)00476-8] [Medline: 38484753]
26. Wang W, Jiang B, Sun H, et al. Prevalence, incidence, and mortality of stroke in China: results from a nationwide population-based survey of 480 687 adults. Circulation 2017 Feb 21;135(8):759-771. [doi: 10.1161/CIRCULATIONAHA.116.025250] [Medline: 28052979]
27. Cheng J, Wang W, Xu J, Yin L, Liu Y, Wu J. Trends in stroke mortality rate—China, 2004-2019. China CDC Wkly 2022 Jun 17;4(24):513-517. [doi: 10.46234/ccdcw2022.113] [Medline: 35812696]
28. Moosa A, Osama D, Alnidawi F, Algillidary S, Hussein A, Das P. Risk factors, incidence, and outcome of stroke: a retrospective cross-sectional hospital-based study comparing young adults and elderly. Cureus 2023 Jun;15(6):e40614. [doi: 10.7759/cureus.40614] [Medline: 37476123]
29. Ågesen FN, Risgaard B, Zachariasardóttir S, et al. Sudden unexpected death caused by stroke: a nationwide study among children and young adults in Denmark. Int J Stroke 2018 Apr;13(3):285-291. [doi: 10.1177/1747493017724625] [Medline: 28762897]
30. Sarfo FS, Ovbiagele B, Gebregziabher M, et al. Stroke among young west africans: evidence from the SIREN (Stroke Investigative Research and Educational Network) large multisite case-control study. Stroke 2018 May;49(5):1116-1122. [doi: 10.1161/STROKEAHA.118.020783] [Medline: 29618553]
31. Bhatt N, Malik AM, Chaturvedi S. Stroke in young adults: five new things. Neur Clin Pract 2018 Dec;8(6):501-506. [doi: 10.1212/CPJ.0000000000000522] [Medline: 30588380]
32. Mszar R, Mahajan S, Valero-Elizondo J, et al. Association between sociodemographic determinants and disparities in stroke symptom awareness among US young adults. Stroke 2020 Dec;51(12):3552-3561. [doi: 10.1161/STROKEAHA.120.031137] [Medline: 33100188]
33. Yanti ADL, Santik YDP. Risk factors associated with disability among elderly with stroke in Indonesia: a secondary data analysis of 2018 National Basic Health Research. J Public Health Trop Coast Reg 2022 Dec 29;5(3):140-154. [doi: 10.14710/jphtcr.v5i3.15694]
34. Louw Q, editor. Collaborative Capacity Development to Complement Stroke Rehabilitation in Africa: AOSIS; 2020.
35. Yoon CW, Bushnell CD. Stroke in women: a review focused on epidemiology, risk factors, and outcomes. J Stroke 2023 Jan;25(1):2-15. [doi: 10.5853/jos.2022.03468] [Medline: 36746378]
36. Wood Z, Adams S, Jones Jr J. Headache in pregnancy: preeclampsia and intracerebral hemorrhage. Cureus 2023 Jan;15(1):e34086. [doi: 10.7759/cureus.34086] [Medline: 36843722]
37. Roeder HJ, Lopez JR, Miller EC. Ischemic stroke and cerebral venous sinus thrombosis in pregnancy. Handb Clin Neurol 2020;172:3-31. [doi: 10.1016/B978-0-444-64240-0.00001-5] [Medline: 32768092]
38. Towfighi A, Saver JL, Engelhardt R, Ovbiagele B. A midlife stroke surge among women in the United States. Neurology 2007 Nov 13;69(20):1898-1904. [doi: 10.1212/01.wnl.0000268491.89956.c2] [Medline: 17581944]
39. Zhou W, Chen R, Hopkins A, et al. Association between socioeconomic status and incident stroke in China. J Epidemiol Community Health 2020 Jun;74(6):519-526. [doi: 10.1136/jech-2019-213515] [Medline: 32341052]
40. Thomas Q, Crespy V, Duloquin G, et al. Stroke in women: when gender matters. Rev Neurol (Paris) 2021 Oct;177(8):881-889. [doi: 10.1016/j.neurol.2021.01.012] [Medline: 34172293]

XSL•FO
RenderX

41. Samai AA, Martin-Schild S. Sex differences in predictors of ischemic stroke: current perspectives. Vasc Health Risk Manag 2015;11:427-436. [doi: 10.2147/VHRM.S65886] [Medline: 26251609]
42. Xia X, Yue W, Chao B, et al. Prevalence and risk factors of stroke in the elderly in northern China: data from the National Stroke Screening Survey. J Neurol 2019 Jun;266(6):1449-1458. [doi: 10.1007/s00415-019-09281-5] [Medline: 30989368]
43. Jin H, Xu Z, Li Y, et al. Seasonal variation of stroke incidence in Wujin, a city in southeast China. Health Sci Rep 2018 Apr;1(4):e29. [doi: 10.1002/hsr2.29] [Medline: 30623065]
44. Tan M, He FJ, Wang C, MacGregor GA. Twenty-four-hour urinary sodium and potassium excretion in China: a systematic review and meta-analysis. J Am Heart Assoc 2019 Jul 16;8(14):e012923. [doi: 10.1161/JAHA.119.012923] [Medline: 31295409]
45. Perera V, Allen LN, Farrand C, Kwong EJL, Liyanage I, Wickramasinghe K. Evaluating the role of salt intake in achieving WHO NCD targets in the Eurasian Economic Union: a PRIME modeling study. PLoS ONE 2023;18(7):e0289112. [doi: 10.1371/journal.pone.0289112] [Medline: 37478108]
46. Li Y, Zhang X, Sang H, et al. Urban-rural differences in risk factors for ischemic stroke in northern China. Medicine (Baltimore) 2019;98(21):e15782. [doi: 10.1097/MD.0000000000015782] [Medline: 31124971]
47. Wang X, Ye X. Impact of the targeted poverty alleviation policy on older adults' healthcare utilization: a quasi-experimental analysis from China. Soc Sci Med 2024 Sep;356:117146. [doi: 10.1016/j.socscimed.2024.117146] [Medline: 39079351]
48. Fu L, Pei T, Xu J, Han J, Yang J. Inspecting the "health poverty trap" mechanism: self-reinforcing effect and endogenous force. BMC Public Health 2024 Mar 28;24(1):917. [doi: 10.1186/s12889-024-18464-3] [Medline: 38549088]

Abbreviations

AAPC: average annual percent change
APC: annual percentage change
ASPR: age-standardized prevalence rate
GBD: Global Burden of Disease
NGCC: National Geomatics Center of China
NHPADMS: National Health Poverty Alleviation Dynamic Management System
WBIC: Weighted Bayesian Information Criterion

Edited by A Mavragani, T Sanchez; submitted 15.Jan.2026; peer-reviewed by XH Zhang, Y Li; revised version received 27.Apr.2026; accepted 30.Apr.2026; published 03.Jun.2026.

Please cite as:
Chen M, Xia D, Li W, Wang J, Lv Z, Chen J, Zhang L
The Evolving Burden of Stroke in China's 832 Poverty-Alleviated Counties (2019-2024): Nationwide Spatiotemporal Analysis
JMIR Public Health Surveill 2026;12:e91487
URL: https://publichealth.jmir.org/2026/1/e91487
doi:10.2196/91487

© Mengdi Chen, Dong Xia, Wenping Li, Jie Wang, Zhiyu Lv, Jiapeng Chen, Lulu Zhang. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 3.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Detecting Spatiotemporal Patterns of Newly Diagnosed HIV Infection in the China-Myanmar Border Region, 2010 to 2022: Longitudinal Observational Study

Jin Yang[1*], MSc; Qiyu Zhu[2*], MPH; Pei Liu[2,3], MSc; Jibao Wang[1], BA; Xing Duan[1], BA; Yikui Wang[1], BA; Shijiang Yang[1], BA; Yuecheng Yang[1], BA; Runhua Ye[1], BA; Peng Xu[2], PhD; Song Duan[1], BA; Cong Jin[2,3], PhD

[1]Dehong Prefecture Center for Disease Control and Prevention, Mangshi, China

[2]National Center for AIDS/STD Control and Prevention, Chinese Center for Disease Control and Prevention, 155 Changbai Road Changping District, Beijing, China

[3]National Key Laboratory of Intelligent Tracking and Forecasting for Infectious Diseases, National Center for AIDS/STD Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Peng Xu, PhD
National Center for AIDS/STD Control and Prevention, Chinese Center for Disease Control and Prevention, 155 Changbai Road Changping District, Beijing, China

Abstract

Background: Geospatial analysis plays an essential role in informing targeted human immunodeficiency virus (HIV) intervention. The Dai-Jingpo Autonomous Prefecture of Dehong (hereinafter referred to as Dehong), located along the China-Myanmar border in the Yunnan province, has been heavily impacted by HIV infection. Given the complex local epidemic context, particularly frequent cross-border population movement, there is an urgent need to apply spatiotemporal analytical approaches to guiding resource allocation. Existing evidence has demonstrated the substantial spatial variations of newly diagnosed HIV infection this region. However, these spatiotemporal variations have not been fully explored at a finer geographic and temporal resolution.

Objective: This study aims to comprehensively investigate the spatiotemporal variations of newly diagnosed HIV infection at a finer scale in this border region to inform targeted interventions.

Methods: Data on newly diagnosed HIV cases at the township level in Dehong were collected from 2010 to 2022. The rate of newly diagnosis HIV cases was calculated annually. GeoDetector *q* statistics were performed to assess the spatially stratified heterogeneity of the rate of newly diagnosed HIV cases. The Bayesian space-time hierarchical model was applied to detect the spatiotemporal patterns of newly diagnosed HIV infection across the region.

Results: A total of 5045 newly diagnosed HIV cases were identified in Dehong from 2010 to 2022. The rate of newly diagnosed HIV cases decreased from 57.1 cases per 100,000 population in 2010 to 13.3 cases per 100,000 population in 2022, a decrease of 76.7% over the past 13 years. The overall temporal relative risk decreased from 2.11 (95% CI 1.84 - 2.41) in 2010 to 0.48 (95% CI 0.40 - 0.56) in 2022. There was substantial spatiotemporal heterogeneity in the risk of newly diagnosed HIV infection, with townships near the China-Myanmar border having a higher spatial relative risk. Notable spatially stratified heterogeneity in the rate of newly diagnosed HIV cases when stratified by the distance of townships to the China-Myanmar border was observed (q=0.27; P=.004). Among the 51 townships in Dehong, 22 (43.1%) hotspots and 22 (43.1%) coldspots were identified. Notably, in comparison to the overall declining temporal trend, 2 hotspots and 4 coldspots exhibited a slower declining trend, suggesting that these regions may require additional intervention efforts.

Conclusions: This study comprehensively estimated the spatiotemporal risk of newly diagnosed HIV infection across Dehong, revealing high-risk areas concentrated near the China-Myanmar border. Priority should be given to implementing targeted interventions to control cross-border HIV transmission, including the establishment of cross-border HIV control mechanisms, as well as the strengthening of management measures for cross-border populations. Furthermore, this study offers methodological insights into the use of routine surveillance data and Bayesian spatiotemporal modeling to better understand HIV transmission dynamics at finer geographic scales and to support precision-oriented HIV prevention services.

(JMIR Public Health Surveill 2026;12:e81767) doi:10.2196/81767

KEYWORDS

HIV; spatial analysis; spatiotemporal heterogeneity; geographic mapping; Bayesian model; public health

XSL•FO
RenderX

Introduction

Since the acquired immune deficiency syndrome (AIDS) cases were first reported in 1981, there are now 40.8 million people living with human immunodeficiency virus (HIV) globally [1]. To prevent the transmission of HIV, the Joint United Nations Programme on HIV/AIDS (UNAIDS) has proposed the 95-95-95 targets, which aim to end HIV/AIDS as a public health threat by 2030 [2]. In China, impressive progress has been made to reduce new HIV infection and enhance the health quality of people living with HIV. The steadily increasing of antiretroviral therapy (ART) coverage in China has led to a significantly reduction in all-cause mortality, from 5.4% in 2013 to 2.7% in 2022 [3]. Achieving the goal of ending HIV/AIDS as a public health threat will increasingly depend on the implementation of precise and context-specific intervention strategies tailored to local epidemic characteristics.

The Dai-Jingpo Autonomous Prefecture of Dehong (hereinafter referred to as Dehong) in the Yunnan province, located in the China-Myanmar border, is an area severely affected by HIV infection, where the first epidemic outbreak of HIV in China was recognized [4]. Located on major illicit drug-trafficking routes from the “Golden Triangle,” a primary opium-producing area in Southeast Asia, early HIV transmission in this region was predominantly driven by injecting drug users (IDUs) [5]. After years of fighting illicit drugs and providing HIV harm reduction services to IDUs, such as syringe exchange programs and methadone maintenance treatment, sexual contact has become the dominant route of HIV transmission [6]. Consequently, the local HIV epidemic has become increasingly complex and concealed.

To achieve effective control of the epidemic, it is essential to accurately understand the current trends in HIV infection. Spatial analysis is often used to reflect the severity of the HIV epidemic at finer geographic scales and to inform the allocation of health resources and targeted interventions [7]. A previous study conducted in Dehong analyzed the spatial patterns of different years using HIV case data from 1989 to 2018 [8]. However, this previous study only performed purely spatial analyses based on annual newly diagnosed data and did not incorporate temporal terms. As local governments dynamically adjust policies and measures in response to evolving epidemic conditions, the risk and trends of HIV infection may vary over time across different regions of Dehong. Traditional spatial models, which only capture geographic variation, fail to integrate temporal dynamics of HIV transmission.

Understanding fine-scale spatiotemporal heterogeneity has become an increasingly important methodological challenge in HIV epidemiology, particularly in high-burden or highly heterogeneous settings. In this study, we employed a Bayesian space-time hierarchical model (BSTHM) to retrospectively analyze data of newly diagnosed HIV cases in Dehong from 2010 to 2022, enabling us to detect the spatiotemporal heterogeneity of HIV transmission. Compared to classical spatial analysis models, which are used to examine spatial autocorrelation and depict the spatial distribution of diseases, BSTHM is suitable for analyzing spatial data with continuous temporal variation. BSTHM effectively integrates spatiotemporal correlations and decomposes spatiotemporal data into 3 key components: common spatial patterns, common temporal patterns, and potential spatiotemporal interactions [9]. Therefore, BSTHM can be used to analyze the spatiotemporal heterogeneity and interactions of newly diagnosed HIV cases in Dehong from 2010 to 2022, providing more comprehensive and detailed data to reveal the trends of the HIV epidemic at both global and local levels over time.

This study aims to systematically describes the spatiotemporal distribution and changing characteristics of newly diagnosed HIV cases in Dehong. By applying a Bayesian spatiotemporal hierarchical modeling approach to routine HIV surveillance data, this study explores the potential of maximizing the utility of existing monitoring systems to characterize epidemic patterns at finer spatial and temporal scales. The findings will enhance the understanding of local epidemic characteristics, thereby guiding targeted interventions and resource allocation, further curbing the spread of the epidemic and effectively reducing new HIV infections in the area. In addition, this analytical approach may be applicable to understanding localized HIV transmission patterns and informing targeted intervention strategies in other regions.

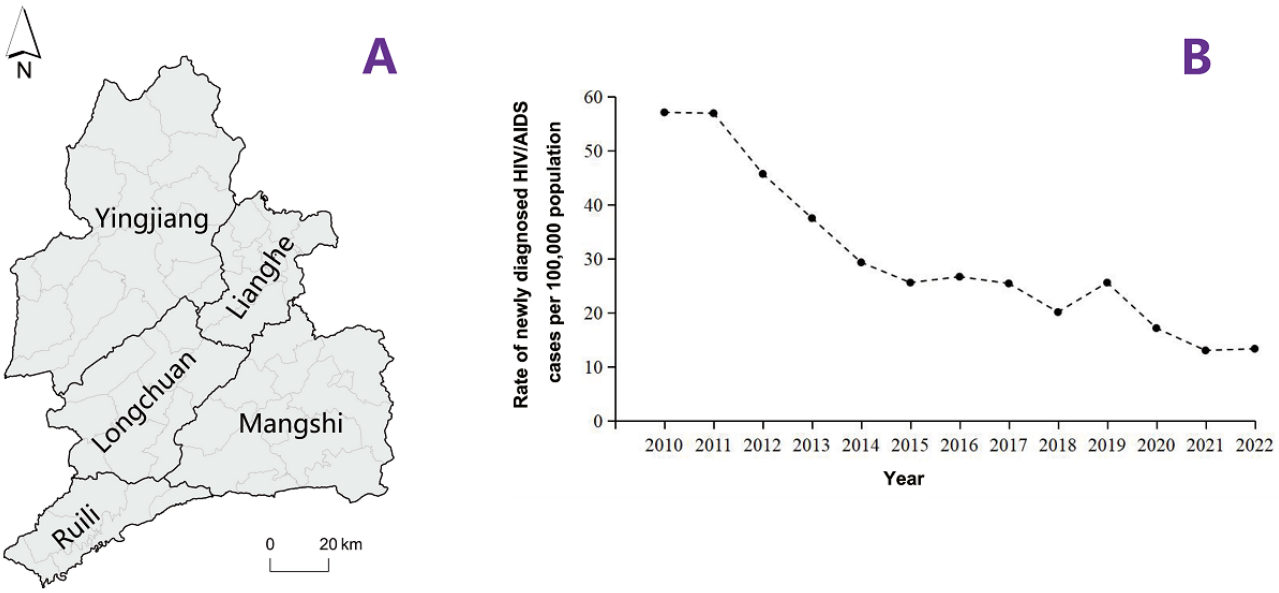
Methods

Study Area

This longitudinal observational study was conducted in Dehong (Figure 1A), which has a population of approximately 1.3 million and an area of 11,526 km^2. The study region is divided into 5 districts (Ruili, Mangshi, Lianghe, Yingjiang, and Longchuan), encompassing 51 township-level districts (6 in Ruili, 12 in Mangshi, 9 in Lianghe, 15 in Yingjiang, and 9 in Longchuan).

XSL•FO
RenderX

Figure 1. Administrative map and rate of newly diagnosed HIV/AIDS cases in the Dai-Jingpo Autonomous Prefecture of Dehong in the Yunnan province. (A) Administrative map. (B) The rate of newly diagnosed HIV/AIDS cases per 100,000 population from 2010 to 2022. AIDS: acquired immune deficiency syndrome; HIV: human immunodeficiency virus.



Data Sources

Data on newly diagnosed HIV cases at the township level in Dehong were obtained from the national HIV/AIDS Comprehensive Response Information Management System (CRIMS) by the Dehong Center for Disease Control and Prevention. The data cover the period from January 1, 2010, to December 31, 2022, and include HIV cases that met the following criteria: (1) reported in Dehong; (2) resided in the specific townships within Dehong; and (3) of Chinese nationality. Township-level population data were obtained from the 2020 China Population Census [10]. All data were collected at the township level. The distance from each township to the China-Myanmar border was defined as the shortest distance between the township administrative boundary and the international border. Based on the calculated distances, all townships were classified into 4 distance-based groups according to proximity, which served as the stratification variable in the GeoDetector analysis. The rate of newly diagnosis HIV cases was obtained by dividing the number of newly diagnosed cases by the local resident population.

GeoDetector *q* Statistics

As reported in previous studies, the HIV transmission in Dehong may be associated with cross-border population movements between China and Myanmar. We investigated the spatially stratified heterogeneity (SSH) of newly diagnosed HIV case rates, stratified by the distance of the township to the China-Myanmar border. The GeoDetector *q* statistics were calculated to measure the SSH of newly diagnosed HIV case rates across different township-level units in Dehong, and the formula is as follows [11]:

(1)q=1-1Nσ2∑h=1LNhσh2

In the formula (Eq. 1), N is the total number of townships in Dehong Prefecture, σ2 denotes the variance of newly diagnosed HIV case rates across all townships, Nh refers the number of townships in stratum h (1, 2, …, L), and σh2 represents the variance of newly diagnosed HIV case rates within the stratum h. The *q* value ranges from 0 to 1, with a value of 0 indicating complete randomness and 1 representing perfect stratified heterogeneity.

BSTHM Formulas

The Poisson and log link regression functions used to perform analyses were as follows:

(2)yit~Poisson(nirit)

(3)log (rit)=α+Si+Ui+(b0t*+vt)+b1it*+εit

In the Poisson likelihood function (Eq. 2), yit represents the number of newly diagnosed HIV cases in township i at year t, ni denotes the total population in township i, and rit is the modeled risk of newly diagnosed HIV/AIDS cases in township i at year t.

In the log link regression function (Eq. 3) that describes the risk of newly diagnosed HIV infection, α is the overall log disease risk for the entire study period, and (Si+Ui) is the disease risk in township i. Specifically, Si captures the spatially structured component, and Ui represents the spatially unstructured component. The term t* denotes the median time of the entire study period, with (b0t*+vt) and b1it* specifying the overall and local temporal trend in township i, respectively. Specifically, the term b0t* captures the linear trend, while vt accounts for the nonlinearity in the overall trend pattern. Given the observed decrease in the newly diagnosed HIV cases rate of over the study period, if b1i < 0, it indicates that the decreasing trend in township i is faster than the overall decreasing trend. Conversely, b1i > 0 suggests a slower trend relative to the overall decrease. The Gaussian noise was captured by εit.

Area Classification by BSTHM

The study area was classified into 9 categories (3 risk categories × 3 trend categories) based on the posterior probability of *P* (exp (Si) >1 |data) and *P* (b1i >0 |hi, data), following a specific

XSL•FO
RenderX

2-stage criteria [12]. In the first stage, townships were classified into 3 categories: hotspot, coldspot, and not cold/hot spot. A township is classified as a hotspot (hi=1) if *P* (exp (Si) >1 |data) is greater than 0.8, and as a coldspot (hi=2) if this probability is less than 0.2 [9]. The remaining townships were classified as not cold/hot spots (hi=3). In the second stage, temporal changes were categorized into 3 classes (faster decrease, common decrease, and slower decrease) according to *P* (b1i >0 |hi, data). A township was defined as exhibiting a slower decreasing trend compared to the overall decreasing trend if *P* (b1i >0 |hi, data) >0.8, a faster decreasing trend if *P* (b1i >0 |hi, data) <0.2, and a common decreasing trend if 0.2 <*P* (b1i >0 |hi, data) <0.8 [9]. Consequently, a total of 9 categories were established.

Hyperparameters for prior distributions were chosen following the approach of Li et al [9], which has been widely applied in BSTHM. Convergence of the Markov Chain Monte Carlo chains was evaluated using the Gelman-Rubin statistic, with values below 1.05 indicating satisfactory convergence. The variance partition coefficient was measured to assess model performance. All BSTHM analysis were conducted using WinBUGS 1.4.3.

Ethical Considerations

This study was approved by the Institutional Review Board of the National Center for AIDS/STD Prevention and Control, Chinese Center for Disease Control and Prevention (approval number: X241028828). The data used in this study were from the CRIMS. No personally identifiable information was present in the data analyzed for this study.

Results

Characteristics of Newly Diagnosed HIV Cases

Between 2010 and 2022, a total of 5045 newly diagnosed HIV cases were reported in Dehong. Of these cases, 3235 (64.1%) were male, 2570 (50.9%) were aged between 31 and 49 years at the time of diagnosis, 2120 (42.0%) were of Han ethnicity, and 2937 (58.2%) had a primary school education or less (Table 1). More than half of the cases (n=2713, 53.8%) were married or cohabiting with a partner. In terms of the transmission routes, 4211 (83.5%) were infected through heterosexual contact, followed by 562 (11.1%) through injecting drug use, and 164 (3.3%) through homosexual contact.

XSL•FO
RenderX

Table . Characteristics of newly diagnosed human immunodeficiency virus (HIV) cases in the Dai-Jingpo Autonomous Prefecture of Dehong in the Yunnan province, 2010 - 2022 (N=5045).

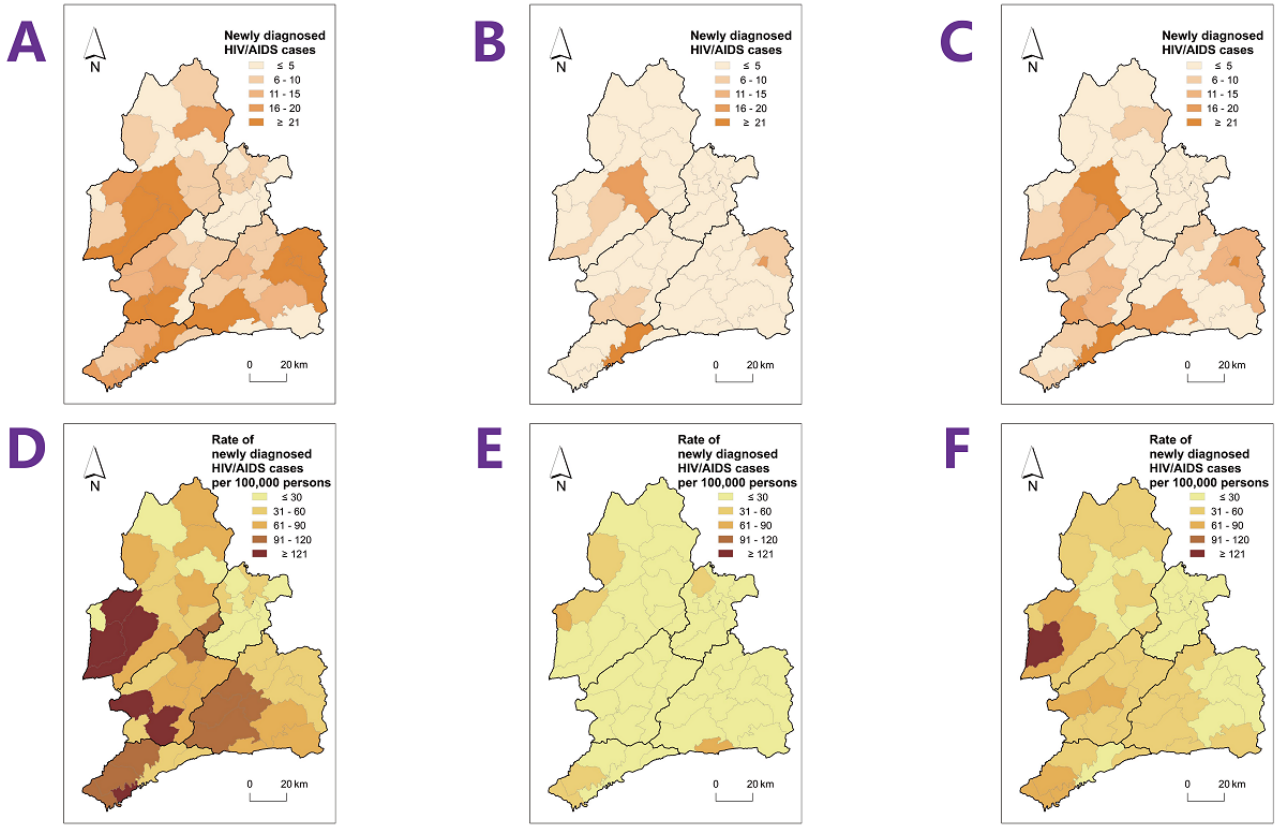
Characteristic	Cases, n (%)
Sex	
Male	3235 (64.1)
Female	1810 (35.9)
Age at HIV diagnosis (years)	
≤30	1536 (30.5)
31–49	2570 (50.9)
≥50	939 (18.6)
Ethnicity	
Han	2120 (42.0)
Dai	1509 (29.9)
Jingpo	1079 (21.4)
Others	337 (6.7)
Education	
Primary school or illiterate	2937 (58.2)
Middle school/high school/vocational school	1858 (36.9)
College or above	249 (4.9)
Unknown	1 (0.0)
Marital status	
Single	1183 (23.4)
Married or living with partner	2713 (53.8)
Divorced or widowed	1149 (22.8)
HIV transmission route	
Heterosexual contact	4211 (83.5)
Homosexual contact	164 (3.3)
Injecting drugs	562 (11.1)
Others	108 (2.1)

Dynamics of Newly Diagnosed HIV Infections

The rate of newly diagnosed HIV cases in Dehong decreased over time, from 57.1 cases per 100,000 population to 13.3 cases per 100,000 population, representing a 76.7% reduction over the past 13 years (Figure 1B). Furthermore, the proportion of townships with a rate of newly diagnosed HIV cases exceeding 60 cases per 100,000 population decreased from 52.9% (27/51) in 2010 to 3.9% (2/51) in 2022. Among the 51 townships in Dehong, the number of townships with more than 20 newly diagnosed cases decreased from 10 in 2010 to 1 in 2022 (Figure 2). GeoDetector analysis revealed notable SSH in the rate of newly diagnosed HIV case when stratified by the distance of townships to the China-Myanmar border (q=0.27; P=.004).

XSL•FO
RenderX

Figure 2. The number and rate of newly diagnosed HIV/AIDS cases in the Dai-Jingpo Autonomous Prefecture of Dehong in the Yunnan province, 2010 - 2022. (A) Number of newly diagnosed HIV/AIDS cases in 2010. (B) Number of newly diagnosed HIV/AIDS cases in 2022. (C) Annual average number of newly diagnosed HIV/AIDS cases during 2010 - 2022. The color intensity represents the number of newly diagnosed HIV/AIDS cases, with darker colors indicating higher values. (D) Rate of newly diagnosed HIV/AIDS cases per 100,000 population in 2010. (E) Rate of newly diagnosed HIV/AIDS cases per 100,000 population in 2022. (F) Annual average rate of newly diagnosed HIV/AIDS cases per 100,000 population during 2010 - 2022. The color intensity represents the rate of newly diagnosed HIV/AIDS cases per 100,000 population, with darker colors indicating higher rates. AIDS: acquired immune deficiency syndrome; HIV: human immunodeficiency virus.



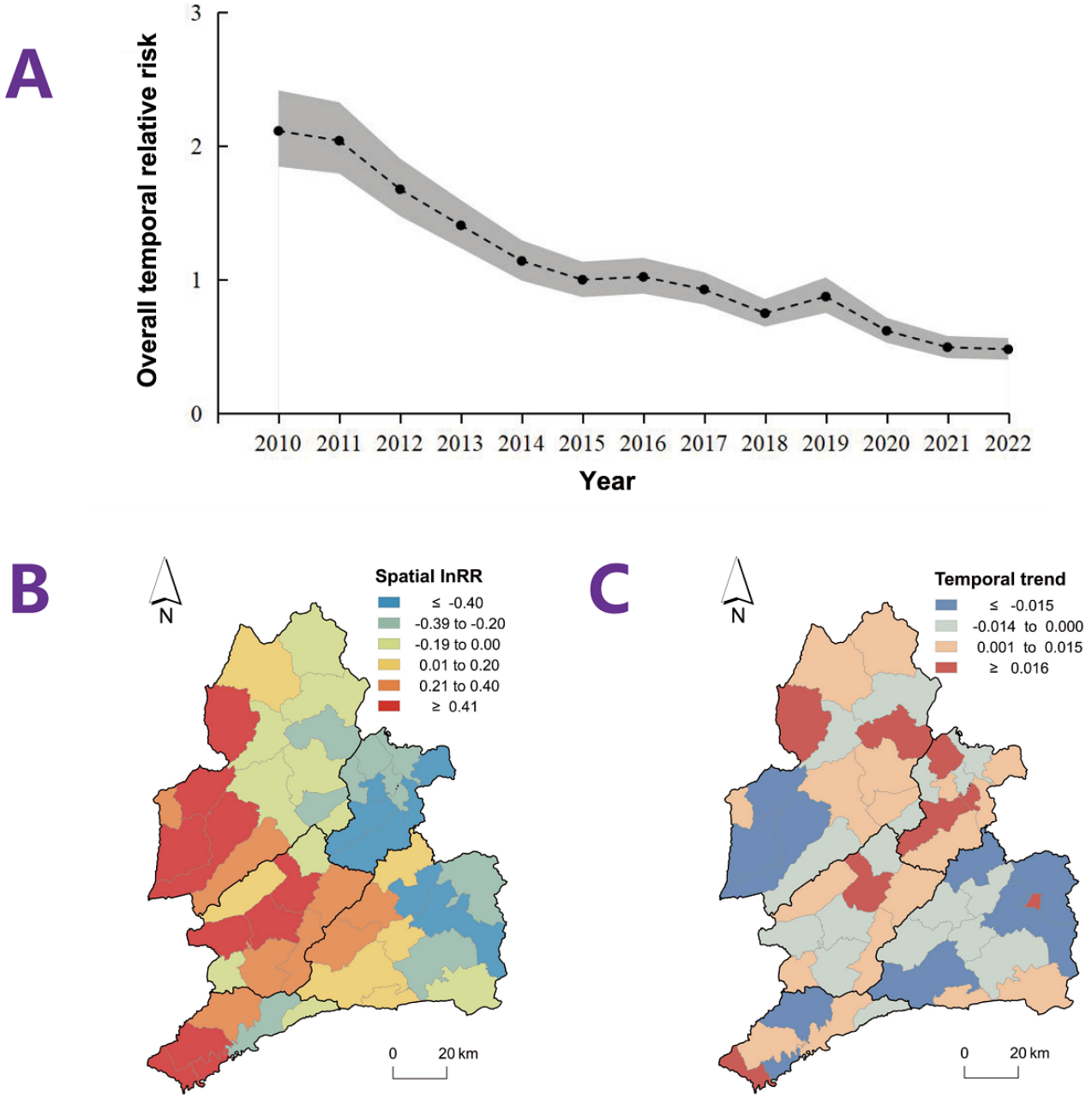
Bayesian Analysis of Newly Diagnosed HIV Infections

Overall Temporal Trend

The variance partition coefficient for BSTHM was 91.3% (95% CI 85.8% - 95.5%). The BSTHM results indicate a declining trend in newly diagnosed HIV infection over the past 13 years, with the overall temporal relative risk decreasing from 2.11 (95% CI 1.84 - 2.41) in 2010 to 0.48 (95% CI 0.40 - 0.56) in 2022 (Figure 3A).

XSL•FO
RenderX

Figure 3. The characteristics of spatiotemporal relative risk of newly diagnosed HIV/AIDS cases in the Dai-Jingpo Autonomous Prefecture of Dehong in the Yunnan province, 2010 - 2022. (A) Overall temporal relative risk. The dots represent the estimated overall temporal relative risk [exp(b0t*+vt)], with the corresponding 95% credible intervals shown as the shaded area. (B) Spatial relative risk. The color intensity represents the estimated spatial relative risk (Si+Ui) without exponential transformation, where higher values (warmer colors) indicate a relatively higher risk in township i compared to the overall average risk in the study region. (C) Local temporal trend departure from the overall trend. The color intensity represents the estimated local departure trend (b1i) without exponential transformation, where the estimated lower value (cooler colors) indicates a faster declining trend compared to the overall declining trend. AIDS: acquired immune deficiency syndrome; HIV: human immunodeficiency virus; InRR: natural logarithm of the spatial relative risk.



Geographical Distribution

Substantial geographical variations in the spatial relative risk of newly diagnosed HIV infection were observed across the study area (Figure 3B). Generally, the western region of Dehong experienced relatively higher risks. The estimated local slopes, which measure the departure of local trends from the overall declining trend, revealed that townships depicted in warm color (red and orange) exhibited slower declining trends compared to the overall trend (Figure 3C). Specifically, townships with slower local declining trends than the prefecture-wide level were predominantly located in Yingjiang, Longchuan, and Lianghe.

Area Classification

Among the 51 townships, 22 (43.1%) were identified as hotspots, 22 (43.1%) as cold spots, and 7 (13.7%) as neither hot nor cold spots (Table 2, Multimedia Appendices 1 and 2). As shown in Figure 4, the 22 hotspots were mainly distributed in the southwestern part of Dehong Prefecture, near the China-Myanmar border. Among these hotspots, 1 township in Ruili and 1 in Longchuan exhibited a slower declining trend

XSL•FO
RenderX

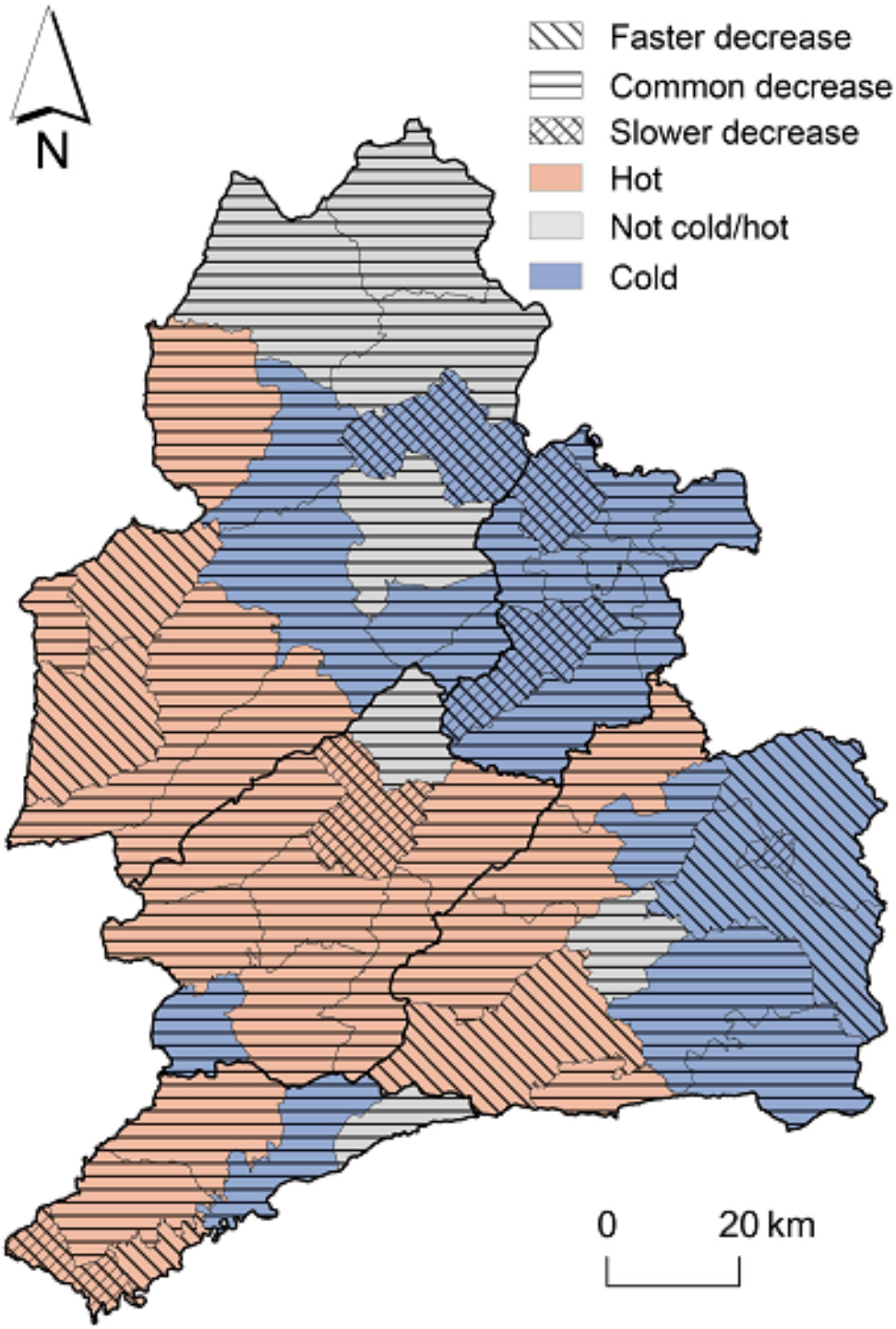
compared to the overall trend. Conversely, 2 hotspots in Yingjiang, 1 in Mangshi, and 1 in Ruili showed a faster declining trend than the overall trend. Among the 22 cold spots, 2 townships in Lianghe, 1 in Mangshi, and 1 in Yingjiang showed a slower declining trend than the overall trend. Meanwhile, 2 cold spots in Mangshi exhibited a faster declining trend than the overall trend. There were 7 townships that were neither hot nor cold spots.

Table . Area classification based on the Bayesian space-time hierarchy model in the Dai-Jingpo Autonomous Prefecture of Dehong in the Yunnan province, 2010 - 2022.

Area classification	Faster decrease	Slower decrease	Common decrease	Total, n (%)
Hotspots	4	2	16	22 (43.1)
Coldspots	2	4	16	22 (43.1)
Not hotspots/coldspots	0	0	7	7 (13.7)

XSL•FO
RenderX

Figure 4. Spatial distribution of hotspots and coldspots for newly diagnosed HIV/AIDS cases in the Dai-Jingpo Autonomous Prefecture of Dehong in the Yunnan province, 2010 - 2022. AIDS: acquired immune deficiency syndrome; HIV: human immunodeficiency virus.



Discussion

Principal Findings

In this study, using township-level data on newly diagnosed HIV infection in Dehong from 2010 to 2022, we quantitatively characterized fine-scale spatiotemporal patterns of HIV risk using BSTHM. The number of newly diagnosed HIV cases has declined significantly over the past 13 years, accompanied by notable SSH. Substantial spatiotemporal heterogeneity was observed in the risk of newly diagnosed HIV infection. By jointly modeling spatial structure and temporal dynamics, several townships were identified as areas that are likely to become or remain hotspots in the future.

In our analysis of newly diagnosed HIV infection from 2015 to 2022, IDUs accounted for 11.1% of cases, while heterosexual contact primarily dominated HIV transmission, responsible for

XSL•FO
RenderX

83.5% of cases. Since HIV was first identified among IDUs in Dehong in 1989 [4], the Chinese government has allocated massive resources to curb HIV transmission. At the provincial level, the Yunnan government has initiated multiple campaigns against HIV and illicit drugs, with remarkable achievements. A study incorporating HIV incidence assays to estimate the HIV incidence among IDUs in Dehong estimated that the HIV incidence among Chinese IDUs in Dehong decreased from 3.11% (95% CI 1.59% - 4.64%) in the 2009 - 2010 period to 0.58% (95% CI –0.06% to 1.04%) in the 2015 - 2017 period [13]. Furthermore, the proportion of IDUs among newly diagnosed male HIV cases in Dehong also showed a downward trend, decreasing from 46.7% in 2012 to 32.0% in 2016 [6]. These shifts reflect a structural transition in transmission dynamics, with heterosexual contact becoming the leading cause of local HIV infections, accounting for 84.7% of newly diagnosed HIV cases during the 2015 - 2017 period [14]. In 2017, the Yunnan government issued an implementation plan aimed at achieving the 90-90-90 HIV targets by the end of 2020. These unprecedented efforts to strengthen HIV prevention, testing, and treatment have resulted in a significant reduction of 76.7% in the rate of newly diagnosed HIV cases over the past 13 years.

Notable SSH was observed in newly diagnosed HIV cases rate across Dehong (q=0.27; P=.004). The spatial relative risk was apparently higher in townships closer to the China-Myanmar border and declined with increasing distance (Figure 3B). Consistent with the area classification analysis, most hotspots were concentrated in the cross-border region. Cross-border mobility and high HIV prevalence of Burmese individuals may play a significant role in HIV transmission. Due to lower income levels in Myanmar, many Burmese individuals migrate to China in pursuit of better economic opportunities, resulting in frequent interactions between the 2 populations. Previous studies have noticed the high HIV prevalence among Chinese-Burmese mixed couples [15,16]. Chen et al [17] recruited 617 Yunnan-Myanmar IDUs in 3 counties of Yunnan and conducted molecular transmission analysis on 117 (19.0%) people living with HIV. Their findings indicate that Yunnan-Myanmar IDUs play a pivotal role in the bidirectional cross-border HIV transmission [17]. Hu et al [18] further demonstrated the strong association between Burmese immigrants and HIV transmission in Dehong. According to a survey of 393 Burmese immigrants in Dehong in 2020, they had low levels of HIV knowledge, condom usage, and willingness to undergo HIV testing [19]. Consequently, the convergence of high background prevalence, frequent population mobility, and constrained access to HIV services likely contributes to the persistent elevated risk observed in border townships. Furthermore, in districts such as Ruili and Longchuan, where ports are in close proximity to main residential and commercial areas, the frequency of cross-border mobility and interactions is increased.

To control the cross-border transmission of HIV, several measures have been implemented in Dehong. These include strengthening comprehensive management of cross-border populations, providing free ART to Chinese-Burmese mixed couples or long-term resident immigrants living with HIV, and expanding HIV testing among immigrants engaged in long-term employment in Dehong. Nevertheless, the persistence of high-risk hotspots identified through BSTHM suggests that additional, more precisely targeted interventions are needed. The identification of these townships provides important epidemiological clues for prioritizing resource allocation. Further prevention and control measures are required for hotspot areas of cross-border HIV transmission. First, cross-border HIV prevention and control mechanisms should be established by strengthening bilateral collaborations between China and Myanmar. Second, programs should be implemented to improve health education, promote condom use, and increase HIV testing rates among Burmese immigrants. Third, appropriate HIV treatment programs should be explored for short-term Burmese immigrants, as they are currently ineligible for China's free ART program. Lastly, efforts should be made to explore ways to improve treatment adherence among HIV-infected immigrants to prevent treatment failure and the development of drug resistance.

In this study, we further estimated the local temporal trend of the spatial pattern using BSTHM, which can predict the spatial pattern shift in the future. The local trends were estimated by comparing their trend patterns with the overall trend pattern. In this study, we found that 2 hotspots exhibited a slower declining trend than the overall trend, while 4 coldspots with a slower declining trend showed a tendency to become "hot" (Table 2). These townships exhibited higher inclination of hotspots in the future and should be prioritized for targeted interventions. Further in-depth epidemiological investigations are warranted to explore the local drivers of transmission in these areas and to inform more precise and context-specific intervention strategies. To our knowledge, few studies in China have investigated the combined spatial patterns and temporal trends in the rate of newly diagnosed HIV infections. Previous spatial analysis studies primarily employed spatial autocorrelation analysis methods to examine the spatial distribution characteristics separately across different time periods [20-24] and used spatial scan statistics to detect cluster areas [21,25]. None of these studies incorporated spatiotemporal correlations. This study modeled spatial relative risk, overall temporal trends, and local temporal trends at the township level, thereby constructing a multifaceted picture of the evolving risk of newly diagnosed HIV cases across Dehong, representing a methodological advantage over traditional spatial analyses. This small area analysis has important implications for the allocation of resources in controlling HIV transmission in similar settings, as interventions and measures can be directly implemented in hotspots or potential hotspots to reduce new infections.

This study is subject to several limitations. First, since HIV diagnoses are not made immediately after HIV infection, the incidence rate could not be calculated in this study. Nevertheless, in the context of generally increasing HIV testing coverage in Dehong, the observed decline in newly diagnosed cases likely reflects, at least to some extent, a reduction in new HIV infections, indicating the effectiveness of local HIV control measures. Second, the analysis in this study was conducted at the township level; the macro-level factors potentially associated with the rate of newly diagnosed HIV infections were not investigated due to the data availability. Third, all analyses in

XSL•FO
RenderX

this study were restricted to newly diagnosed HIV cases of Chinese nationality, as data on foreign nationals were not available. Consequently, the findings primarily reflect HIV infection patterns among the local resident population and may underestimate the overall contribution of cross-border populations to the regional HIV epidemic.

Conclusions

In conclusion, the rate of newly diagnosed HIV cases in Dehong has significantly decreased from 2010 to 2022. By applying a BSTHM framework to routine surveillance panel data, this study reveals substantial variations in the spatial distribution and temporal trends of the risk of newly diagnosed HIV infections within Dehong, with high-risk areas concentrated in townships near the China-Myanmar border. Future efforts should focus on exploring effective measures to reduce new infections in these high-risk areas. This analytical approach provides a new perspective for assessing local needs, guiding resource allocation, and developing targeted interventions to control HIV infection in other high-burden or highly heterogeneous settings.

Acknowledgments

We thank all participants and investigators who contributed to the data collection.

Funding

This study was supported by the National Natural Science Foundation of China (No. 72374186) and the Project of China National Key Laboratory of Intelligent Tracking and Forecasting for Infectious Diseases. The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

The data analyzed during this study are not publicly available but are available from the corresponding author on reasonable request.

Authors' Contributions

CJ, SD, PX proposed the research question and designed the study. JY and QZ wrote the manuscript. QZ and JY performed the analyses. JW, XD, YW, SY, and YY contributed to laboratory experiments and epidemiology data collection. PL and RY provided expert knowledge. All authors have read and contributed to the final version of the manuscript.

JY and QZ contributed equally to this work and should be considered co-first authors. CJ (jinc@chinaaids.cn), SD (dhduansong@shina.com.cn), and PX (xupeng@chinaaids.cn) are co-corresponding authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Posterior probability of spatial relative risk.
[PNG File, 519 KB - publichealth_v12i1e81767_app1.png]

Multimedia Appendix 2
Posterior probability of temporal trend.
[PNG File, 221 KB - publichealth_v12i1e81767_app2.png]

References

1. AIDS, crisis and the power to transform: UNAIDS Global AIDS Update 2025. Joint United Nations Programme on HIV/AIDS. 2025. URL: https://www.unaids.org/sites/default/files/2025-07/2025-global-aids-update-JC3153_en.pdf [accessed 2026-01-26]
2. Understanding fast-track: accelerating action to end the AIDS epidemic by 2030. Joint United Nations Programme on HIV/AIDS. 2014. URL: https://www.unaids.org/sites/default/files/media_asset/JC2686_WAD2014report_en.pdf [accessed 2024-10-12]
3. Zhao Y, Wei L, Dou Z, et al. Changing mortality and patterns of death causes in HIV infected patients - China, 2013-2022. China CDC Wkly 2023 Dec 1;5(48):1073-1078. [doi: 10.46234/ccdcw2023.201] [Medline: 38058989]
4. Ma Y, Li Z, Zhang K, et al. HIV was first discovered among IDUs in China [Article in Chinese]. Chin J Epidemiol 1990;11(3):184-185 [FREE Full text]
5. Zhang J, Cheng H, Jia M, Zhang Y. Ten year experience on AIDS control in Yunnan, 1989–1998 [Article in Chinese]. Chin J Epidemiol 1999;20(6):377-380 [FREE Full text]

XSL•FO
RenderX

6. Wang J, Chen X, Duan X, et al. Epidemiological characteristics of newly reported HIV infections in Chinese and Burmese residents, during 2012-2016 in Dehong Dai and Jingpo Autonomous Prefecture, Yunnan province [Article in Chinese]. Chin J Epidemiol 2017;38(10):1372-1375. [doi: 10.3760/cma.j.issn.0254-6450.2017.10.015]
7. Ma KF, Zhang XT, Li DM. Progress on application of spatial epidemiology in HIV/AIDS control and prevention [Article in Chinese]. Chin J Epidemiol 2022 Jan 10;43(1):128-133. [doi: 10.3760/cma.j.cn112338-20210816-00645] [Medline: 35130664]
8. Yang S, Yang G, Yang Y, et al. Spatiotemporal characteristics of HIV/AIDS epidemic in Dehong Dai and Jingpo Autonomous Prefecture [Article in Chinese]. Chin J Dis Control Prev 2020;24(4):444-450. [doi: 10.16462/j.cnki.zhjbkz.2020.04.015]
9. Li G, Haining R, Richardson S, Best N. Space–time variability in burglary risk: a Bayesian spatio-temporal modelling approach. Spat Stat 2014 Aug;9:180-191. [doi: 10.1016/j.spasta.2014.03.006]
10. Tabulation on 2020 China Population Census by Township [Book in Chinese]: Office of the Leading Group of the State Council for the Seventh National Population Census; 2022.
11. Wang JF, Zhang TL, Fu BJ. A measure of spatial stratified heterogeneity. Ecol Indic 2016 Aug;67:250-256. [doi: 10.1016/j.ecolind.2016.02.052]
12. Richardson S, Thomson A, Best N, Elliott P. Interpreting posterior relative risk estimates in disease-mapping studies. Environ Health Perspect 2004 Jun;112(9):1016-1025. [doi: 10.1289/ehp.6740] [Medline: 15198922]
13. Yang YC, Shi RZ, Tang RH, et al. Estimating HIV incidence among female sex workers and injection drug users in Dehong Prefecture, 2009-2017 [Article in Chinese]. Chin J Prev Med 2018;52(12):1243-1247. [doi: 10.3760/cma.j.issn.0253-9624.2018.12.010]
14. Zhu Q, Wang Y, Liu J, et al. Identifying major drivers of incident HIV infection using recent infection testing algorithms (RITAs) to precisely inform targeted prevention. Int J Infect Dis 2020 Dec;101:131-137. [doi: 10.1016/j.ijid.2020.09.1421] [Medline: 32987184]
15. Duan S, Ding Y, Yang Y, et al. Prevalence and correlates of HIV discordance and concordance among Chinese-Burmese mixed couples in the Dehong prefecture of Yunnan province, China. Sex Health 2012 Nov;9(5):481-487. [doi: 10.1071/SH12065] [Medline: 23380199]
16. Yang YC, Li ZH, Li L, et al. Prevalence and correlates of HIV infection among cross-border families in Dehong Dai and Jingpo Autonomous Prefecture [Article in Chinese]. Chin J Epidemiol 2021;42(4):683-689. [doi: 10.3760/cma.j.cn112338-20200903-01124]
17. Chen X, Zhou YH, Ye M, et al. Burmese injecting drug users in Yunnan play a pivotal role in the cross-border transmission of HIV-1 in the China-Myanmar border region. Virulence 2018;9(1):1195-1204. [doi: 10.1080/21505594.2018.1496777] [Medline: 30001176]
18. Hu Z, Liu Y, Wang J, et al. Frequently transmission and close relationship among immigrants in the China-Myanmar border region indicated by molecular transmission analysis from a cross-sectional data. Front Med (Lausanne) 2021;8:693915. [doi: 10.3389/fmed.2021.693915] [Medline: 35572804]
19. Zhang W, Diao F, Liu G, et al. Study on the influencing factors of HIV testing awareness among Myanmar people in Dehong, Yunnan province [Article in Chinese]. Chinese J Exp Clin Virol 2021;35(2):130-134. [doi: 10.3760/cma.j.cn112866-20201224-00310]
20. Xie Z, Chen B, Duan Z. Spatiotemporal analysis of HIV/AIDS incidence in China from 2009 to 2019 and its association with socioeconomic factors: geospatial study. JMIR Public Health Surveill 2024 Jun 7;10:e56229. [doi: 10.2196/56229] [Medline: 38848123]
21. Hou YS, Jin YC, Cai C, Tang HL, Qin QQ, Lyu F. Characteristics of the HIV/AIDS epidemic among people aged ≥ 50 years in China during 2018-2021. Biomed Environ Sci 2024 Apr 20;37(4):399-405. [doi: 10.3967/bes2024.044] [Medline: 38727162]
22. Huang G, Cheng W, Xu Y, et al. Spatiotemporal pattern and its determinants for newly reported HIV/AIDS among older adults in eastern China from 2004 to 2021: retrospective analysis study. JMIR Public Health Surveill 2024 Feb 13;10:e51172. [doi: 10.2196/51172] [Medline: 38349727]
23. Wang S, Wang J, Xu P, Lyu W, Han M. Spatial analysis of HIV/AIDS cases among out-of-school youth aged 15-24 years - China, 2010-2020. China CDC Wkly 2021 Nov 26;3(48):1015-1017. [doi: 10.46234/ccdcw2021.247] [Medline: 34888117]
24. Ren N, Li Y, Wang R, et al. The distribution of HIV and AIDS cases in Luzhou, China, from 2011 to 2020: Bayesian spatiotemporal analysis. JMIR Public Health Surveill 2022 Jun 14;8(6):e37491. [doi: 10.2196/37491] [Medline: 35700022]
25. Chen M, Ma Y, Chen H, et al. Demographic characteristics and spatial clusters of recent HIV-1 infections among newly diagnosed HIV-1 cases in Yunnan, China, 2015. BMC Public Health 2019 Nov 11;19(1):1507. [doi: 10.1186/s12889-019-7557-8] [Medline: 31711447]

Abbreviations

AIDS: acquired immune deficiency syndrome
ART: antiretroviral therapy
BSTHM: Bayesian space-time hierarchical model

XSL•FO
RenderX

CRIMS: HIV/AIDS Comprehensive Response Information Management System
HIV: human immunodeficiency virus
IDU: injecting drug user
SSH: spatially stratified heterogeneity
UNAIDS: Joint United Nations Programme on HIV/AIDS

Edited by A Mavragani, T Sanchez; submitted 03.Aug.2025; peer-reviewed by M Chen, Q Duan; revised version received 26.Jan.2026; accepted 10.Mar.2026; published 20.May.2026.

Please cite as:
Yang J, Zhu Q, Liu P, Wang J, Duan X, Wang Y, Yang S, Yang Y, Ye R, Xu P, Duan S, Jin C
Detecting Spatiotemporal Patterns of Newly Diagnosed HIV Infection in the China-Myanmar Border Region, 2010 to 2022: Longitudinal Observational Study
JMIR Public Health Surveill 2026;12:e81767
URL: https://publichealth.jmir.org/2026/1/e81767
doi:10.2196/81767

© Jin Yang, Qiyu Zhu, Pei Liu, Jibao Wang, Xing Duan, Yikui Wang, Shijiang Yang, Yuecheng Yang, Runhua Ye, Peng Xu, Song Duan, Cong Jin. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 20.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Disparities in Antiemetic Prophylaxis Care Processes Predicted by Patient Neighborhood: Retrospective Cohort and Geospatial Analysis

Jiuying Han[1], PhD; Neng Wan[1], PhD; Cameron K Jacobson[2], MS; Nathan L Pace[2], MD, MStat; Cade K Kartchner[2], BS; Alexander Hohl[1], PhD; Robert B Schonberger[3], MD; Douglas A Colquhoun[4], MD; Richard P Dutton[5], MD; Michael Andreae[2], MD, PhD; John F Pearson[2,6,7], MD

[1]School of Environment, Society and Sustainability, University of Utah, 260 South Central Campus Drive, Room 4625, Salt Lake City, UT, United States

[2]Department of Anesthesiology, Perioperative and Pain Medicine, Spencer Fox Eccles School of Medicine, University of Utah, Salt Lake City, UT, United States

[3]Department of Anesthesiology, Yale School of Medicine, New Haven, CT, United States

[4]Department of Anesthesiology, University of Michigan Medical School, Ann Arbor, MI, United States

[5]US Anesthesia Partners, Dallas, TX, United States

[6]Global Change and Sustainability Center, University of Utah, Salt Lake City, UT, United States

[7]Department of Anesthesiology, Perioperative and Pain Medicine, School of Medicine, Stanford University, 300 Pasteur Drive, Stanford, CA, United States

Corresponding Author:
John F Pearson, MD
Department of Anesthesiology, Perioperative and Pain Medicine, Spencer Fox Eccles School of Medicine, University of Utah, Salt Lake City, UT, United States

Abstract

Background: Social determinants of health continue to drive persistent disparities in perioperative care. Our team has previously demonstrated racial and socioeconomic disparities in perioperative processes, notably in the administration of antiemetic prophylaxis, in several large perioperative registries. Given how neighborhoods are socially segregated in the United States, we examined geospatial clustering of perioperative antiemetic disparities.

Objective: The study aimed to determine whether disparities in perioperative antiemetic prophylaxis exhibit geographic clustering based on neighborhood-level disadvantage and whether patients from disadvantaged communities are more likely to be undertreated after adjusting for individual postoperative nausea and vomiting risk.

Methods: We conducted a retrospective cohort study of anesthetic records from the University of Utah Hospital involving 19,477 patients who met the inclusion criteria. We geocoded patient home addresses and combined them with the census block group–level neighborhood disadvantage, a composite index from the National Neighborhood Data Archive. We stratified our patients by antiemetic risk score and calculated the number of antiemetic interventions. We used Poisson spatial scan statistics, implemented in SaTScan (Information Management Services, Inc), to detect geographic clusters of undertreatment.

Results: We identified 1 significant cluster ($P<.001$) of undertreated perioperative antiemetic prophylaxis cases. The relative risk of the whole cluster was 1.44, implying that patients within the cluster were 1.44 times more likely to receive fewer antiemetics after controlling for antiemetic risk. Patients from more disadvantaged neighborhoods were more likely to receive below-median antiemetic prophylaxis after controlling for risk.

Conclusions: To our knowledge, this is the first geospatial cluster analysis of perioperative process disparities; we leveraged innovative geostatistical methods and identified a spatially defined, geographic cluster of patients whose home address census-tract level neighborhood deprivation index predicted disparities in risk-adjusted antiemetic prophylaxis.

(JMIR Public Health Surveill 2026;12:e69133) doi:10.2196/69133

KEYWORDS

anesthesia; postoperative; geospatial analysis; database; informatics; GIS; geographic information system

XSL•FO
RenderX

Introduction

Background

Social determinants of health (SDOH) continue to drive disconcerting health care disparities [1-4]. SDOH are defined as "the societal circumstances in which we are born and grow up, learn and mature, and work and age," and also includes race, ethnicity, education, wealth, insurance coverage, and health literacy; they impact equitable perioperative processes and outcomes as fundamental causes of disease [1,3-5]. Disparities can concern access to care, care processes, and health care outcomes. We focus on care processes, as the means to intervene [2,3,6]. To precisely target the underlying causes and to create, test, and implement specific countermeasures, we need to define the exact granular mechanisms leading to disparities in processes [4,5,7-10].

Health Care Disparities in Anesthesiology and Perioperative Medicine

Disparities in care processes lead to subpar health outcomes in minoritized, marginalized, and migrant populations; differences in care processes based on race and ethnicity or socioeconomic status rather than patient risk factors and comorbidities have been a longstanding and troubling aspect of care delivery in the United States [1,3]. In outpatient settings, racial and ethnic disparities impact care for diabetes, asthma, and heart disease [11,12], as well as veterans' care. For inpatients, race-based process-of-care surgical disparities [13] have been described, especially within neonatal and obstetric care [7,13-15]. Macario et al [16] and our team previously demonstrated process variability in anesthesiology, possibly tied to unconscious bias and negative stereotypes [2,5], as well as disparities in access to chronic pain treatment [10]. We previously demonstrated that individual clinicians provide fewer antiemetics to people who identify as Black, to those with lower health insurance, and to patients living in zip codes with lower median income [2,5]. One hypothesis is that this arises from the clinicians' prejudice towards their patients [17]: clinicians (sub)consciously detect the social status of their patients in their interaction, and therefore become less likely to diligently elicit for, document, and treat risk factors of postoperative nausea and vomiting (PONV) [2,5]. We discussed possible concrete drivers leading to PONV disparities elsewhere [4]. Given how neighborhoods are socially segregated in the United States, the patient's home address can provide a proxy for socioeconomic status and race [5,9]. This motivates our more granular geospatial analysis of perioperative process disparities in anesthesia care, [9]: Our primary question is whether we can identify clusters of neighborhoods with elevated social deprivation scores whose inhabitants receive less antiemetic prophylaxis after controlling for PONV risk [2,18].

Geospatial Analysis to Investigate Health Care Disparities

The use of geospatial analysis to understand social and environmental influences on perioperative outcomes is a promising emerging science, with studies suggesting zip code may predict health outcomes better than a patient's DNA [2,10]. Figure 1 illustrates the various data that can be leveraged for geospatial analysis (eg, spatial accessibility, air quality, and neighborhood deprivation). We discussed the promise and pitfalls of geospatial analysis for health systems and disparity research and the pertinent concepts previously [9]. While more mainstream in public health, very few studies have applied these innovative techniques to perioperative disparities research, especially in conjunction with large-scale perioperative registry data, such as the Multicenter Perioperative Outcomes Group (MPOG), a leading US perioperative electronic health record registry [18,19]. The integration with census-tract and block group level social determinants of health derived from publicly available databases (eg, the National Neighborhood Data Archive) further augments the power of granular electronic health registries [9]. Not only is geographically informed health care disparity research ideally suited to studying the impact of patients' SDOH on their perioperative trajectory, but the information can also be leveraged to provide individual clinicians, teams, and institutions with actionable insights on how to improve care processes [4,7,9]. While northern Utah and the Salt Lake region have a reputation for homogeneity, the National Equity Atlas ranks Salt Lake County at 208 out of 430 US counties ranked, with a diversity index of 0.95 (diversity index is a measure of racial or ethnic diversity, with a maximum score of 1.95 if all ethnic groups are equally represented), while the United States as a whole ranges from 0.22 to 1.58, making our study population comparable to a broad swath of midsize US counties [20].

Our innovation was to apply a novel geostatistical method (Poisson spatial scan statistic) [21] to identify neighborhood-level clusters of systematic under-treatment after adjusting for PONV risk. As this methodology represents an alternative statistical approach to classical statistics, we provided Table S1 in Multimedia Appendix 1 to aid readers in interpretation of our results. We chose to use geospatial clustering, widely used in disease and crime surveillance [22,23], to explore its use of this technique in the service of health care equity. We hypothesized that patients living in neighborhoods with higher levels of social disadvantage are more likely to receive fewer perioperative antiemetics after adjusting for antiemetic risk factors, and that such neighborhoods would be found to cluster geographically [2,18,24].

XSL•FO
RenderX

Figure 1. Geographic information systems are able to integrate patient home addresses, census-tract demographics, neighborhood deprivation index, and percent Hispanic to create a comprehensive view of social determinants of health.

XSL•FO
RenderX

Methods

Overview

We performed a retrospective cohort analysis of the University of Utah local MPOG electronic health registry (supplemented with University of Utah Health Epic database), enriched with geocoded SDOH for the year 2021 [25]. We adhered to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement and principles (Checklist 1) [26]. All authors approved the statistical analysis plan before analyses began. We describe patient demographic characteristics, PONV risk, and perioperative antiemetic prophylaxis in Table 1.

Table 1 details characteristics of our population before and after applying exclusion criteria and filtering of data to the Wasatch Front for the purpose of our geographic analysis. Characteristics of the final study set with home addresses in the target area are listed in the final cohort data. It should be noted that the "unknown" category is unable to be disaggregated due to limitations in the underlying dataset.

Most patients in the final cohort had a median age of 48 (IQR 63-34) years and received a general anesthetic of more than 1-hour duration, including the administration of opioid medication. Their risk classification according to the American Society of Anesthesiologists (ASA) was mostly 1-3. Only 10% reported a history of PONV, and most were nonsmokers. For a median of 4 PONV risk factors, they received on average about 2 prophylactic antiemetic interventions.

The focus on the Wasatch Front area shrunk the total number of patients from 51,809 to 19,477 in the final cohort analyzed and increased the relative proportion of non-White patients (typically clustered in urban areas in Utah); the prevalence of risk factors and the number of interventions also changed. Percentages have been rounded to the nearest 10th place with some factors being described by the mean, SD, median, and IQR of the data represented.

XSL•FO
RenderX

Table . Demographic and clinical characteristics of the initial and final study cohorts.

	Original cohort	Final cohort
Race, n (%)		
Non-White	4817 (9.3)	4254 (21.8)
White	41,189 (79.5)	15,223 (78.2)
Unknown	5803 (11.2)	0 (0)
Age		
Mean (SD)	50.73 (19.68)	48.695 (17.55)
Median (IQR)	52 (34-67)	48 (34-63)
Sex, n (%)		
Male	22,923 (44.2)	8563 (44)
Female	28,871 (55.7)	10,908 (56.0)
Smoking status, n (%)		
Nonsmoker	51,112 (98.7)	19,208 (98.6)
Anesthesia duration (>1 h), n (%)	16,270 (31.4)	7838 (40.2)
History of PONV[a] or motion sickness, n (%)	4608 (8.9)	2159 (11.1)
Opioids used for postoperative pain, n (%)	45,556 (87.9)	19,176 (98.5)
ASA PS[b] classification, n (%)		
ASA[c] class 1	4670 (9.0)	2411 (12.4)
ASA class 2	20,984 (40.5)	9157 (47.0)
ASA class 3	21,991 (42.4)	7222 (37.1)
ASA class 4	3669 (7.1)	685 (3.5)
ASA class 5	440 (0.8)	0 (0)
ASA class 6	21 (0)	0 (0)
Type of anesthesia, n (%)		
Deep sedation	21 (0)	0 (0)
General	36,614 (70.7)	18,491 (94.9)
MAC[d]	9814 (18.9)	0 (0)
Regional	4499 (8.7)	986 (5.1)
Number of risk factors		
Mean (SD)	3.46 (1.16)	3.79 (1.05)
Median (IQR)	3 (3-4)	4 (3-4)
Number of interventions		
Mean (SD)	1.88 (1.13)	2.49 (0.91)
Median (IQR)	2 (1-3)	2 (2-3)

[a]PONV: postoperative nausea and vomiting.

[b]ASA PS: American Society of Anesthesiologists Physical Status Classification.

[c]ASA: American Society of Anesthesiologists.

[d]MAC: Monitored Anesthesia Care

XSL•FO
RenderX

Ethical Considerations

This study was considered exempt on February 9, 2023, by the institutional review board at the University of Utah (reference: 00142167). All patients at the University of Utah agreed to secondary data use as part of their consent to care, and therefore, no additional informed consent was required for this retrospective study. All data were stored on secure servers at the University of Utah consistent with university policy, and no data were transmitted outside the University firewall for geocoding, as this was done using locally installed geocoding servers. No patients were compensated for this study and no identifiable data were included in this paper.

Geocoding SDOH

Patient home addresses were obtained from Epic and then geocoded to latitude and longitude with ArcGIS (ESRI, Inc) [9,27] and spatially matched to census block group (CBG), the boundaries of which were obtained from US Census TIGER/Line (Topologically Integrated Geographic Encoding and Referencing System) 2020 CBG boundaries layer. Any incomplete address information was excluded. We extracted neighborhood disadvantage (ND), a composite index of neighborhood deprivation, from the National Neighborhood Data Archive, a publicly available database, and used this as our independent variable [28]. ND is an average of 5 US Census indicators, including proportion of the following: non-Hispanic Black, female-headed families, households with public assistance or on food stamps, income below the federal poverty level, and unemployment, all derived from the US Census American Community Survey 5-year estimates from 2016 to 2020. ND is an average of these proportions; therefore, as this value increases, the relative disadvantage or social stress increases.

Inclusion and Exclusion Criteria

Following the MPOG Anesthesiology Performance Improvement and Reporting Exchange (ASPIRE) quality metric PONV05 Inclusion Criteria [29], and using information obtained from the Utah local Epic and MPOG instances, we excluded cases with ASA Physical Status Classification of 5 or 6, aged <18 years, ICU transfers, and cases performed without general anesthesia (including obstetrics, ECT electroconvulsive therapy, and bronchoscopy procedures). We limited our study area further to the Wasatch Front, a densely populated metropolitan area served by the University of Utah hospital in Salt Lake City, Utah, as this better enables geospatial clustering analysis and is more generalizable to urban areas in the United States.

Risk-Adjusted Antiemetic Prophylaxis

The novel geostatistical method (Poisson spatial scan statistic) required a characterization of patients as receiving relatively more (+1), similar (0), or fewer (−1) antiemetic interventions compared to their peers with similar PONV risk factors, leading us to define a 3-level ordinal variable of risk-adjusted antiemetic prophylaxis (RAAP), similar to the MPOG PONV-05 quality metric; RAAP is discussed in the later sections.

Antiemetic Interventions

The response variable was the administration of appropriate numbers of antiemetic prophylactic interventions. We counted up to 6 antiemetic interventions (0 - 1, 2, 3, 4, 5, and 6+) [4,24]. These interventions included the number of and class of antiemetic prophylaxis (1 point per each), and the use of total intravenous anesthesia, as the presence of a propofol infusion, as defined in PONV05, a performance metric developed by MPOG [25].

PONV Risk

Risk factors for PONV likely mediate RAAP; as they could be unevenly distributed among different racially, ethnically, and geographically defined groups; PONV risk factors could confound the analysis of equitable antiemetic administration [2,4,30]. All PONV risk factors were extracted from our local MPOG instance. We categorized the PONV risk score into 6 ordinal levels (0 - 1, 2, 3, 4, 5, and 6+), following the PONV05 standard of MPOG [29]. These PONV risks are defined by PONV05 and follow widely accepted guidelines including female sex (assigned at birth), history of PONV or motion sickness, nonsmoker, opioid use (either intraoperative or postoperative), duration of inhalational anesthesia greater than 1 hour, aged <50 years, and certain select procedures known to cause PONV.

Characterizing RAAP

First, we stratified our patients by antiemetic risk [4,24]. Within each stratum of risk, we calculated the median number of antiemetic interventions. We defined RAAP to describe above median, median, and below median RAAP. Finally, we assigned the 3 levels (below median, median, and above median) RAAP contingent on the number of prophylactic antiemetic interventions received: a high RAAP for an above median number, a median RAAP for cases with an intermediate number, and a low RAAP for cases with a below median number of antiemetic prophylactic interventions compared to other cases in the same PONV risk stratum. With RAAP, we sought to *contrast* different levels of antiemetic prophylaxis *despite* similar PONV risk, ie, process disparity driven by neighborhood-level social determinants of health [4]. Hence, RAAP serves to characterize the relative intensity of antiemetic prophylaxis after controlling for risk (not as a measure of guideline adherence) [4]. This was due to guideline variability and the use of adherence to guidelines could obscure underlying disparity within our practice, which includes physicians who trained outside of the United States, which runs counter to the purpose of the investigation.

Cluster Analysis

We used the Poisson spatial scan statistic [21,31] implemented in SaTScan (Information Management Services, Inc) [32] to detect the clusters of undertreated perioperative antiemetic prophylaxis cases in the Wasatch Front. In this study, we focused specifically on detecting clusters of low RAAP incidence, as identifying areas with undertreatment is of particular relevance for targeting systemic interventions to improve perioperative care. A cluster contains a set of neighboring regions which collectively have higher incidence

XSL•FO
RenderX

(of low RAAP) rates than expected if the low RAAP cases were evenly distributed across the study area. The Poisson spatial scan statistic uses a likelihood ratio test combined with Monte Carlo simulations to determine whether any detected clusters represent a significantly higher or lower incidence than would be expected by chance. Thus, we identify clusters only when there is statistically significant evidence of an unusually high or low rate of under-treatment relative to the expected distribution under the null model.

We used univariate and multivariate logistic regressions to investigate the characteristics of the CBG in the identified low RAAP cluster. The outcome was a binary variable indicating whether the CBG belongs to the high-rate low RAAP or not. We tested previously described or suspected social determinants of health as predictors of disparities: [2,5] percentage of non-White population, percentage of population with education less than high school, percentage of population below poverty rate, percentage of population married, percentage of older people (aged >65 y), and percentage of renters. Finally, the presence of spatial autocorrelation was tested by calculating the Moran I statistic for the residuals of the multivariate model.

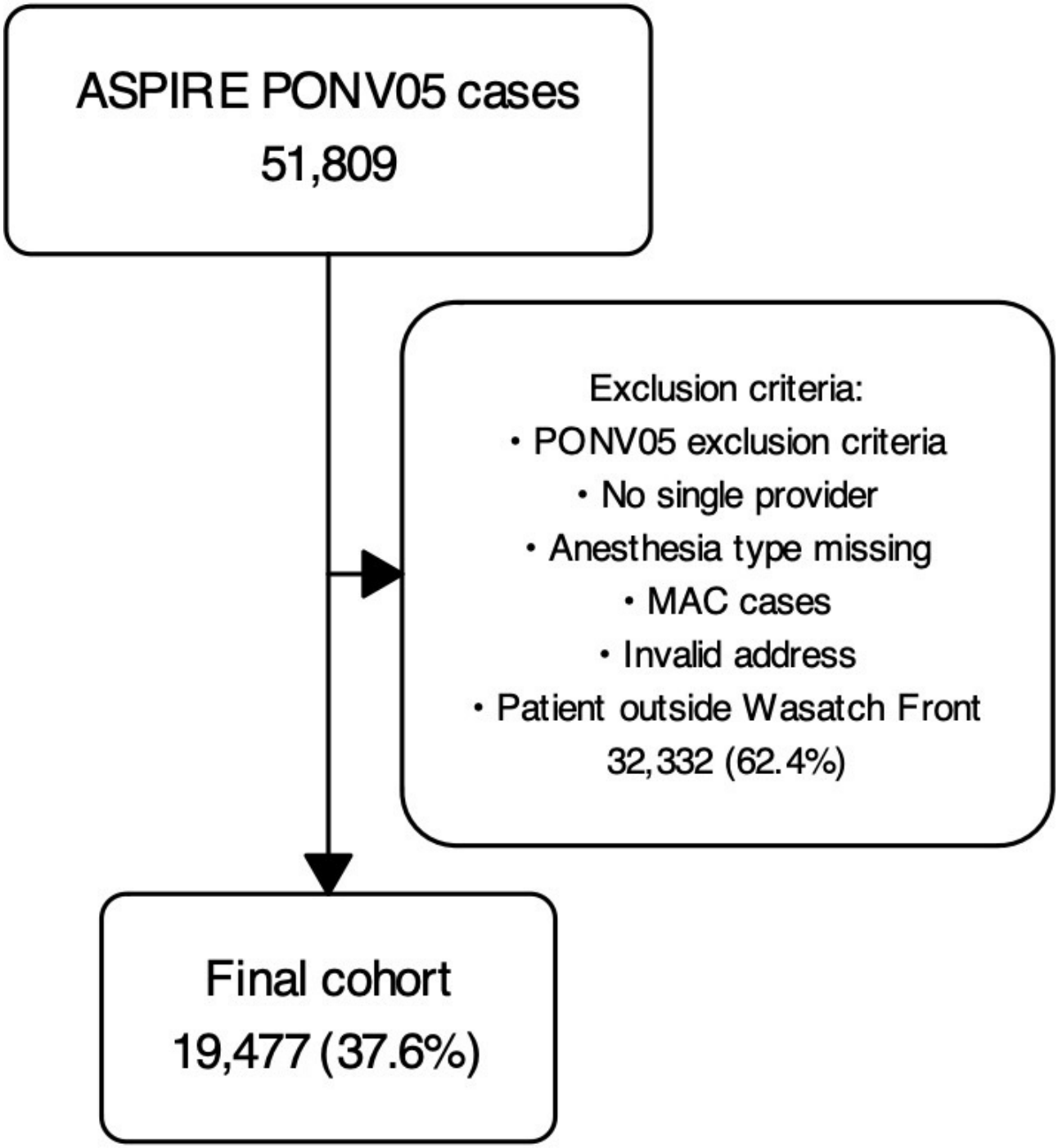
Results

Description of the Population and Study Flow Diagram

A total of 51,809 anesthetic case records of patients undergoing surgery at the University of Utah were extracted from the Utah MPOG 2021 dataset. After excluding cases with missing data, a final dataset of 19,477 cases was analyzed in the cluster analysis, as detailed in Figure 2 with missing data noted therein. Table 1 lists detailed characteristics of the original cohort as well as the final analysis cohort after applying our exclusions criteria.

XSL•FO
RenderX

Figure 2. Cohort selection flow diagram of 51,809 cases with exclusions and final analytic cohort of 19,477 cases. ASPIRE: Anesthesiology Performance Improvement and Reporting Exchange; MAC: monitored anesthesia care (excluded); MPOG: Multicenter Perioperative Outcomes Group; PONV: postoperative nausea and vomiting.



The median age in the final cohort was 48 (IQR 34‐63) years; 94.9% (18,491/19,477) of patients underwent a general anesthetic, often lasting longer than 1 hour (7838/19,477, 40.2%); the administration of opioid medication was almost always part of the anesthetic plan (19,176/19,477, 98.5%). ASA class was mostly 1‐3 (2411/19,477, 12.4%; 9157/19,477, 47.0%; and 7222/19,477, 37.1%, respectively). Approximately one-tenth of patients reported a history of PONV. Most patients were nonsmokers (19,208/19,477, 98.6%). The median number of PONV risk factors was 4 (IQR 3‐4), and patients received a median of 2 (IQR 2‐3) prophylactic antiemetic interventions.

The final cohort had approximately the same percentage of White patients (41,189/51,809, 79.5% original vs 14,223/19,477, 78.2% final; standardized mean difference [SMD] 0.03), similar sex distribution (28,871/51,809, 55.7% female original vs 10,908/19,477, female 56.0% final; SMD 0.01) and similar smoking status (51,112/51,809, 98.7% nonsmoking original vs 19,208/19,477, 98.6% nonsmoking final; SMD 0.00) while the final cohort had slightly more risk factors (3.46 vs 3.79; SMD 0.29), interventions (1.88 vs 2.44, SMD 0.55), and higher prevalence of a lower ASA (ASA1-2: 25,654/51,809, 50% original vs 11,568/19,477, 59% final; SMD 0.27).

XSL•FO
RenderX

Uni- and Bivariate Analysis

In total, we found 2260 low RAAP cases, 7358 median RAAP, and 9589 high RAAP cases in our final dataset. The association between socioeconomic neighborhood-level and demographic patient factors and RAAP is presented in a bivariate tabulation (Table 2).

Table 2 demonstrates how home address and race predicted RAAP in a descriptive, bivariate comparison. Levels of RAAP are tabulated in 3 columns: below median, median, and above median, which reflect disparities. The ND scores in the rows were separated into quartiles, with the 1st quartile Q1 being the lowest score (meaning the least disadvantaged) and the 4th quartile being the highest disadvantaged score (meaning the most disadvantaged). Populations living in more disadvantaged neighborhoods (eg, Q4=ND 4th quartile) on average received fewer antiemetic interventions after adjusting for risk, with 17.1% (835/4869) of Q4 receiving low RAAP versus 14.2% (692/4870) low RAAP for Q1 (most affluent) and similarly 41.1% (2001/8850) of Q4 receiving high RAAP versus 49.1% (2390/8850) of Q1 receiving high RAAP; hence, they were relatively undertreated compared to patients living in more affluent neighborhoods (Q1). In addition, people who self-identified as White mostly received many antiemetic interventions after controlling for risk (only 933/8850, 10.5% low RAAP) compared with unknown patients who more often received fewer interventions after controlling for risk (470/2986, 15.7% low RAAP) and non-White patients (897/7641, 11.7% low RAAP).

Table . Association of neighborhood disadvantage and race with risk-adjusted antiemetic prophylaxis (RAAP).

Category	RAAP			Total, n (%)
	High RAAP, n (%)	Median RAAP, n (%)	Low RAAP, n (%)	
Quartile				
Q1	2390 (49.07)	1788 (36.71)	692 (14.21)	4870 (25)
Q2	2331 (47.87)	1858 (38.16)	680 (13.96)	4869 (25)
Q3	2128 (43.70)	1962 (40.29)	779 (16.41)	4869 (25)
Q4	2001 (41.09)	2033 (41.75)	835 (17.14)	4869 (25)
Total	8850 (45.43)	7641 (39.23)	2986 (15.33)	19,477 (100)
Race				
Non-White	5990 (78.39)	754 (9.86)	897 (11.74)[a]	7641 (39.23)
White	7097 (80.19)	820 (9.26)	933 (10.54)[b]	8850 (45.43)
Unknown	2136 (71.53)	380 (12.73)	470 (15.74)[a]	2986 (15.33)
Total	15,223 (78.15)	1954 (10.03)	2300 (11.81)	19,477 (100)

[a]$P<.001$

[b]$P=.003$

Finally, an asymptotic generalized Pearson chi-squared test was performed to compare the distribution of RAAP values between all pairwise race categories (White, non-White, and unknown). The associated P values are as follows: White versus non-White ($P=.003$), White versus unknown ($P<.001$), and non-White versus unknown ($P<.001$).

Patient self-identified race predicted RAAP. Living in more disadvantaged neighborhoods (eg, Q4=ND 4th quartile) was associated with receiving fewer antiemetic interventions after adjusting for risk (835/4869, 17.1% low RAAP); in other words, compared to patients living in more affluent neighborhoods (Q1=1st ND quartile with 2390/8850, 49.0% high RAAP), clinicians administered fewer antiemetic interventions, even after adjusting for risk, in cases of patients living in CBGs with higher ND scores, which represent neighborhoods with higher proportions of non-Hispanic Black, female-headed families, households with public assistance or on food stamps, income below federal poverty level, and unemployment. In addition, after controlling for PONV risk, patients who self-identified as White received more antiemetic interventions (7097/8850, 80.1% high RAAP; $P=0.003$), and rarely fewer antiemetic interventions (only 933/8850 10.5% low RAAP; $P<.001$) while non-White patients more often received fewer interventions after controlling for risk (470/2986, 15.7% low RAAP; $P<.001$), compared to White patients in similar PONV risk strata. These associations are congruent with the results of our more complex statistical models.

The bivariate analyses showed that all the variables of interest, except age, were significantly related to being inside the cluster, with the percentage of non-White, percentage of male, percentage of education less than high school, percentage of population below poverty level, and the percentage of renter being positively related to belonging to a high-rate cluster, while the percentage of married population being negatively related to belonging to a cluster.

Geospatial Analysis

We identified 1 significant cluster ($P<.001$) of low RAAP cases in Wasatch Front (Multimedia Appendix 2), comprising 101 CBGs. The cluster was located around the geographic center of Salt Lake City, and included West Valley City, South Salt Lake, and Taylorsville. The cluster had a total of 1934 included

XSL•FO
RenderX

participants residing within it, or approximately 9.93% of the study population. The relative risk (RR) of the whole cluster was 1.44 (P<.001), implying that the risk of receiving less (low RAAP) was 1.44 times higher for patients living *within* the cluster than for those living in other CBGs along the Wasatch Front. The CBG level RR within the cluster ranged between 0 and 3.73 (with RR>1 being a high relative risk), meaning the highest risk of low RAAP was 3.73 times higher living in CBGs within versus outside the cluster.

After adjusting for all listed covariates, we found that percentage of age >65 years in a CBG cluster became highly significantly associated with a low RAAP cluster, which would represent guidance-congruent care as age protects against PONV. However, we similarly demonstrated that the percentage of non-White people in a geographic area was also associated with low RAAP, which represents an unexplained process disparity. The larger the percentage of age >65 years and the percentage of non-White, which in West Valley City is over 42% Hispanic [33], the more likely the CBG belonged to the low RAAP cluster. In addition, the percentage of males was positively related to belonging to a low RAAP cluster, but with borderline statistical significance (again indicating guidance-congruent care, as female sex is positively associated with PONV risk). The Moran I statistic (Moran I=0.003; P=.79) indicated no evidence of spatial autocorrelation in the model residuals.

Discussion

Principal Findings

In our retrospective cohort of patients at the University of Utah local MPOG electronic health registry [25], we leveraged a novel geostatistical method to identify a spatial cluster of patients receiving less RAAP than peers in similar PONV risk strata [32]. We identified one significant cluster (P<.001) of low RAAP cases in the Wasatch Front (Multimedia Appendix 2), comprising 101 CBGs.

The cluster identified is an area of low RAAP cases with a high ND score (reflecting structural neighborhood features such as female-headed families, households with public assistance or on food stamps, income below the federal poverty level, and unemployment), and a relatively higher percentage of Hispanic and non-White patients. Low RAAP, defined in the methods, reflects process disparities in RAAP. Our findings are consistent with a direct association between ND and RAAP. Our results remained statistically significant in multivariate analysis, where we demonstrated that non-White individuals received lower RAAP than self-identified White participants. Uni- and bivariate tabulation of our data (Table 2) corroborated our principal geospatial analysis, while the association of male sex and older age in CBGs with low RAAP was consistent with guidance-congruent care, as both patient characteristics are associated with lower PONV risk.

Interpretation of Principal Findings

Only individual patient-level PONV risk factors should be the guiding principle of RAAP [2,4,5,24]; neither patient race and ethnicity, nor ND should impede equitable RAAP [2,18,24]. By identifying geographic clusters of low RAAP, representing populations experiencing disparities in antiemetic prophylaxis after adjusting for risk, we visualized how geographically defined SDOH drive health care process disparities within anesthesiology [9]. Our finding that those within a disadvantaged geographic area, in this case the Hispanic-dominated West Valley region of Salt Lake County (where West Valley City has over 42% identify as Hispanic) [33], are over 40% less likely to receive RAAP is troubling, and it reinforces the use of geospatial analysis to identify populations at risk of process disparities. It should be further noted that the use of the spatial scan statistic examines individual CBGs in the context of neighboring CBGs: the ability to analyze beyond administrative boundaries can capture broader area characteristics, an advantage when examining subdivisions within a broader area of disparity.

Significance of Geospatial Process Disparities in the Literature

Why would a patient's home address determine how many antiemetics they receive, as we demonstrated in this geospatial analysis? Geographic analyses in perioperative medicine have been sparse [34-36], and prior work tended to focus on patient *outcomes* rather than on *process* disparities. Geographic clusters of disparities in care have been well-documented in various areas of medicine, including diabetes, asthma, heart disease, veterans care, surgical care, as well as neonatal care [7,11-15]. Our findings also build on process outcomes in pediatric anesthesiology and our own work in obstetric care pain medicine, where cancellation rates were found to be higher in communities with lower socioeconomic status [10,19,35,37,38]. Our findings also add to our prior research on disparities in RAAP driven by patient- and neighborhood-level social determinants of health (patient race, insurance, and median income in zip code) [2,5]. Our results are consistent with these prior findings but examine perioperative process disparities through a geospatial lens with granular geostatistical methods, considering proximity statistically [9,21]. In addition, compared to previous work, the cluster analysis used here offers 2 key advantages. First, zip codes, designed primarily for mail delivery, are often misaligned with meaningful health behaviors or risk boundaries and can introduce bias when used as analytic units [39,40]. Second, by applying spatial scan statistics to pinpoint geographic clusters of suboptimal perioperative care, we move beyond coarse aggregation to directly identify and quantify localized disparities. The geospatial granularity empowers public health officials to deploy interventions (eg, quality improvement initiatives, tailored community outreach, and provider education) focused precisely on those areas where care gaps are statistically significant, thus enhancing the effectiveness of efforts to improve perioperative outcomes.

Promise of Geospatial Analysis of Disparities

The innovation of our study is to perform a geospatial analysis of perioperative process disparities, leveraging equitable RAAP as a case study [2,5]. This represents an important step forward in terms of ability to ameliorate disparities, as geospatial drivers can be reevaluated over time, for example, after mitigation measures are put in place to address disparities [9]. The benefits of exploring process differences by neighborhood were also noted in studies in Scotland and California [41]; identifying

XSL•FO
RenderX

specific neighborhood clusters provided for targeted outreach to improve diabetes care.

Our analysis of local catchment areas with disparities in care *processes*, for which we are individually accountable as clinicians [5], can make for a different conversation within a department and among clinicians about potential modifications to practice patterns; these may ultimately drive many outcome disparities which some may have traditionally considered beyond our direct influence and scope of practice [7]. Outcome differences may be influenced by environmental factors (eg, pollution) while differences in risk-adjusted processes, under the primary domain of the individual anesthesia provider in the operating room, leave fewer excuses [4], though our findings are based on aggregate patterns and should not be interpreted as direct evidence of bias at the individual provider level. Despite this limitation, a focus on process provides greater insight into clinician behaviors and relationships with patients, with neighborhood structural factors potentially leading to care disparities potentially shaping the provider-patient relationship in the high-stakes and brief time of the preoperative anesthesia assessment and contributing to cumulative disadvantage. It also identifies, in our instance, an area where language barriers (due to Hispanic, Spanish-speaking population) may impede recognition of PONV risk factors and thus opens the door to potential novel avenues to improve care [4,10]. In our institution, identifying that a Hispanic, Spanish-speaking predominant neighborhood is potentially experiencing disparities is a unique way that geospatial clustering uncovers actionable, neighborhood-targeted hypotheses for intervention.

Studies from Taiwan and the United Kingdom National Health Service system show a complex interplay between care disparities and socioeconomic status of the neighborhood, demonstrating that even in systems with near-universal health care coverage, socioeconomic status can drive disparities in the delivery of care [42,43]. In contrast, other studies have suggested that neighborhood characteristics have only a modest impact on some care processes and outcomes [44], while the influence of race may overshadow neighborhood characteristics.

Limitations and Potential Bias in Our Methods

Our study is not without its limitations. First, limiting generalizability, we confined the study population to the catchment area of the University of Utah Hospital, limited to 1 health system and 1 group of anesthesiologists. However, our prior studies around race have examined multiple hospitals in the large perioperative electronic health registries with analogous disparities identified [2,5]. Furthermore, the racial and ethnic diversity of populations varies across the United States, somewhat limiting the external validity of our study to states with lower minority representation. Low RAAP areas had low socioeconomic status and higher rates of Hispanic populations. This could result in autocorrelation and bias, as it is possible that neighborhood socioeconomic status and geographic factors were less salient than race but could not be fully separated in our model due to our small patient cohort. Expanding our approach to include areas of economic diversity among the Hispanic or other minoritized populations (eg, Texas, South Florida, and California) would further elucidate this relationship. The MPOG registry and our local dataset do not provide data on the race, sex, or ethnicity of the *clinicians* assigned to a particular anesthetic case. We were therefore unable to investigate if congruence of identity characteristics between patients and clinicians for a given case had an impact on clinician adherence with the best antiemetic practices. Furthermore, the assignment of staff to specific clinical locations could introduce bias in our results, especially if a single subgroup regularly practices at a given location; however, as the vast majority of cases take place at the main University of Utah clinical campus, the potential for this bias is minimal.

Another limitation, inherent in all geographic epidemiological retrospective studies, is the ecological fallacy [9,45]; we did not have individual self-identified race for all patients, as our local instance of MPOG and Epic provides inconsistent information on race and ethnicity, thus this is a data quality gap that constrains our interpretation. Our most reliable information is aggregate data at the neighborhood CBG level. Despite joining some individual data with group-level data, the risk of ecological bias, while mitigated, is still present. We acknowledge potential *missing data bias*, since we were only able to identify 93.1% of patients' addresses to the level of the street point address, (which is, however, consistent with other geocoding results) [9,46]. We also acknowledge that limiting the study area to the Wasatch Front, which was done to ensure contiguous CBGs for geospatial clustering analysis, may have resulted in the exclusion of rural disparities. We also did not examine the race of clinicians or perform any analysis by clinician characteristics [4]. A few outliers could have skewed the cohort, though this is unlikely given the inclusion of over 100 anesthesiologists. More detailed strengths and weaknesses of our geospatial analysis approach are discussed in Multimedia Appendix 1.

Conclusions

Demonstrating the use of a novel geostatistical method in a retrospective cohort of anesthesia case records from the local MPOG electronic health registry at the University of Utah, we identified a spatial cluster of patients in the West Valley area of Salt Lake City, Utah, receiving less RAAP (Multimedia Appendix 2); they also represented a population from CBGs with relatively high neighborhood deprivation. Our results are consistent with prior work that demonstrated disparities in antiemetic prophylaxis based on *patient*-level factors (eg, race) [2], but adding a *neighborhood*-level, spatial dimension to perioperative health care disparities and health systems research, with many potential applications to elucidate geographic, social, and environmental drivers of the health care process and outcome (Figure 1) [9].

Acknowledgments

The authors would like to acknowledge Beca Chacin of the Department of Anesthesiology, University of Utah for her contribution to creating Figure 1. During the preparation of this work, the authors used ChatGPT 4, from OpenAI, in order to edit the manuscript

XSL•FO
RenderX

for clarity of language and to ensure adherence to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist. After using this tool or service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Funding

This study was funded from an intramural University of Utah Vice President's for Research Seed Grant (JFP and NW) and was partially supported by the National Institute of Health National Center for Advancing Translational Science, the National Heart Lung and Blood Institute, and the National Institute on Aging (National Institutes of Health grant UL1TR004409 to MA, K08HL159327 to DAC, and R01AG059607 to RBS). The funding sources were not involved in the study design, collection, analysis, and interpretation of the data, in writing the report, or in the decision to submit the article for publication. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the University of Utah Vice President's Office. No funder had any involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

A fully anonymized version of the dataset can be made available upon reasonable request to the authors and with appropriate permission from relevant oversight bodies at the University of Utah.

Authors' Contributions

Conceptualization: JH (lead), MA (equal), JFP (equal), NW (supporting), NLP (supporting), CKK (supporting), AH (supporting), RBS (supporting), DAC (supporting), RPD (supporting)
Data curation: JH (lead), CKJ (supporting), CKK (supporting)
Formal analysis: JH (lead), NLP (supporting)
Funding acquisition: MA (lead), JFP (equal), NW (supporting)
Investigation: JH (lead), CKJ (supporting), CKK (supporting)
Methodology: MA (lead), JFP (equal), NW (supporting), CKJ (supporting), NLP (supporting), AH (supporting)
Project administration: MA (lead), JFP (equal), NW (supporting)
Resources: MA (lead), JFP (equal), NW (supporting), NLP (supporting)
Software: JH (lead), CKJ (supporting), CKK (supporting)
Supervision: MA (lead), JFP (equal)
Validation: AH (lead), RBS (supporting), DAC (supporting), RPD (supporting)
Visualization: JH (lead)
Writing – original draft: JH (lead), MA (equal), JFP (equal), CKJ (supporting)
Writing – review & editing: JH (lead), MA (equal), JFP (equal), NW (supporting), CKJ (supporting), NLP (supporting), CKK (supporting), AH (supporting), RBS (supporting), DAC (supporting), RPD (supporting)

Conflicts of Interest

RBS reports that he owns stock in Johnson & Johnson unrelated to this work, and that his institution has received research support from Merck, Inc, for a study in which he participated unrelated to this work. DAC reports research support paid to his institution unrelated to this work from Merck, Inc, and Chiesi USA. RPD is an equity option holder in US Anesthesia Partners and a consultant for GE, Cerus, Eagle Pharmaceuticals, and Edwards Lifesciences.

Multimedia Appendix 1
Strengths and weaknesses of different spatial analysis approaches.
[DOCX File, 22 KB - publichealth_v12i1e69133_app1.docx]

Multimedia Appendix 2
Map of the Wasatch Front highlighting cluster of under-treatment in risk-adjusted antiemetic prophylaxis identified by SaTScan (Information Management Services, Inc). This corresponds to an area of Salt Lake County with a high prevalence of Hispanic population.
[PNG File, 454 KB - publichealth_v12i1e69133_app2.png]

Checklist 1
STROBE checklist.
[DOCX File, 47 KB - publichealth_v12i1e69133_app3.docx]

References

XSL•FO
RenderX

1. Link BG, Phelan J. Social conditions as fundamental causes of disease. J Health Soc Behav 1995;Spec No:80-94. [doi: 10.2307/2626958] [Medline: 7560851]
2. White RS, Andreae MH, Lui B, et al. Antiemetic administration and its association with ace: a retrospective cohort study. Anesthesiology 2023 Jun 1;138(6):587-601. [doi: 10.1097/ALN.0000000000004549] [Medline: 37158649]
3. Bailey ZD, Feldman JM, Bassett MT. How structural racism works—racist policies as a root cause of U.S. racial health inequities. N Engl J Med 2021 Feb 25;384(8):768-773. [doi: 10.1056/NEJMms2025396] [Medline: 33326717]
4. Andreae MH, Maman SR, Behnam AJ. An electronic medical record–derived individualized performance metric to measure risk-adjusted adherence with perioperative prophylactic bundles for health care disparity research and implementation science. Appl Clin Inform 2020 May;11(3):497-514. [doi: 10.1055/s-0040-1714692] [Medline: 32726836]
5. Andreae MH, Gabry JS, Goodrich B, White RS, Hall C. Antiemetic prophylaxis as a marker of health care disparities in the National Anesthesia Clinical Outcomes Registry. Anesth Analg 2018 Feb;126(2):588-599. [doi: 10.1213/ANE.0000000000002582] [Medline: 29116968]
6. Bailey ZD, Krieger N, Agénor M, Graves J, Linos N, Bassett MT. Structural racism and health inequities in the USA: evidence and interventions. Lancet 2017 Apr 8;389(10077):1453-1463. [doi: 10.1016/S0140-6736(17)30569-X] [Medline: 28402827]
7. Toyobo O, Guglielminotti J, Andreae DA, Andreae MH. Health equity research in obstetric anesthesia. Curr Opin Anaesthesiol 2022 Jun 1;35(3):306-316. [doi: 10.1097/ACO.0000000000001135] [Medline: 35671017]
8. Asare EA, Andreae MH, Stewart JH 4th. Bridging the disparity gap in surgical oncology access: Does telehealth hold a key? Ann Surg Oncol 2022 Nov;29(12):7235-7236. [doi: 10.1245/s10434-022-12262-0] [Medline: 35854032]
9. Pearson J, Jacobson C, Ugochukwu N, et al. Geospatial analysis of patients' social determinants of health for health systems science and disparity research. Int Anesthesiol Clin 2023 Jan 1;61(1):49-62. [doi: 10.1097/AIA.0000000000000389] [Medline: 36480649]
10. Brosius DJ, Chaturvedi R, Andreae MH, et al. Social determinants of health: modeling and targeting patient propensity to attend pain clinic appointments. Pain Manag 2023 Mar;13(3):151-159. [doi: 10.2217/pmt-2022-0059] [Medline: 36718774]
11. Bulger JB, Shubrook JH, Snow R. Racial disparities in African Americans with diabetes: process and outcome mismatch. Am J Manag Care 2012 Aug;18(8):407-413. [Medline: 22928755]
12. Cabana MD, Lara M, Shannon J. Racial and ethnic disparities in the quality of asthma care. Chest 2007 Nov;132(5 Suppl):810S-817S. [doi: 10.1378/chest.07-1910] [Medline: 17998345]
13. Kristensen PK, Falstie-Jensen AM, Madsen M, Johnsen SP. Patient-related healthcare disparities in the quality of acute hip fracture care: a 10-year nationwide population-based cohort study. BMJ Open 2021 Dec 30;11(12):e051424. [doi: 10.1136/bmjopen-2021-051424] [Medline: 35446782]
14. Ravi D, Iacob A, Profit J. Unequal care: Racial/ethnic disparities in neonatal intensive care delivery. Semin Perinatol 2021 Jun;45(4):151411. [doi: 10.1016/j.semperi.2021.151411] [Medline: 33902931]
15. Dagher RK, Linares DE. A critical review on the complex interplay between social determinants of health and maternal and infant mortality. Children (Basel) 2022 Mar 10;9(3):394. [doi: 10.3390/children9030394] [Medline: 35327766]
16. Macario A, Chung A, Weinger MB. Variation in practice patterns of anesthesiologists in California for prophylaxis of postoperative nausea and vomiting. J Clin Anesth 2001 Aug;13(5):353-360. [doi: 10.1016/s0952-8180(01)00283-5] [Medline: 11498316]
17. Hoberman JM. Medical racism and the rhetoric of exculpation: How do physicians think about race? New Lit Hist 2007 Jun;38(3):505-525. [doi: 10.1353/nlh.2007.0041]
18. White R, Andreae MH, Ma X, Turnbull ZA, Nachamie AS, Rosenbloom JM. Antiemetic prophylaxis as an anesthesia quality marker and its association with race in the multicenter perioperative outcomes group: a retrospective cohort study, 2004-2018. Presented at: Anesthesiology Annual Meeting; Oct 19-23, 2019.
19. Shaparin N, White R, Andreae M, Hall C, Kaufman A. A longitudinal linear model of patient characteristics to predict failure to attend an inner-city chronic pain clinic. J Pain 2014 Jul;15(7):704-711. [doi: 10.1016/j.jpain.2014.03.004] [Medline: 24747766]
20. Atlas NE. Diversity index: racial and cultural diversity create thriving and prosperous communities. National Equity Atlas. 2024. URL: https://nationalequityatlas.org/indicators/Diversity_index?geo=03000000000041620 [accessed 2024-01-10]
21. Kulldorff M. Spatial Scan Statistics: Models, Calculations, and Applications: Springer; 1999.
22. Lawson A. Disease mapping and its uses. In: Biggeri A, Boehning D, Lesaffre E, Viel J, editors. Disease Mapping and Risk Assessment for Public Health: Wiley; 1999:3-14.
23. Ceccato V. Homicide in São Paulo, Brazil: Assessing spatial-temporal and weather variations. J Environ Psychol 2005 Sep;25(3):307-321. [doi: 10.1016/j.jenvp.2005.07.002]
24. Gan TJ, Belani KG, Bergese S, et al. Fourth consensus guidelines for the management of postoperative nausea and vomiting. Anesth Analg 2020 Aug;131(2):411-448. [doi: 10.1213/ANE.0000000000004833] [Medline: 32467512]
25. Colquhoun DA, Shanks AM, Kapeles SR, et al. Considerations for integration of perioperative electronic health records across institutions for research and quality Improvement: the approach taken by the Multicenter Perioperative Outcomes Group. Anesth Analg 2020;130(5):1133-1146. [doi: 10.1213/ANE.0000000000004489]

XSL•FO
RenderX

26. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. Int J Surg 2014 Dec;12(12):1495-1499. [doi: 10.1016/j.ijsu.2014.07.013] [Medline: 25046131]
27. Nassel A, Wilson-Barthes MG, Howe CJ, et al. Characterizing the neighborhood risk environment in multisite clinic-based cohort studies: a practical geocoding and data linkages protocol for protected health information. PLoS ONE 2022;17(12):e0278672. [doi: 10.1371/journal.pone.0278672] [Medline: 36580446]
28. Melendez R, Clarke P, Khan A, Gomez-Lopez I, Li M, Chenoweth M. National Neighborhood Data Archive (NaNDA): Socioeconomic Status and Demographic Characteristics of ZIP Code Tabulation Areas, United States, 2008–2017: University of Michigan; 2020. [doi: 10.3886/E120462V1]
29. PONV-05: PONV prophylaxis: adults. Multicenter Perioperative Outcomes Group. 2022. URL: https://spec.mpog.org/Spec/Public/53 [accessed 2025-10-25]
30. White RS, Aaronson JA. Obstetric and perinatal racial and ethnic disparities. Curr Opin Anaesthesiol 2022 Jun 1;35(3):260-266. [doi: 10.1097/ACO.0000000000001133] [Medline: 35671010]
31. Bambhroliya AB, Burau KD, Sexton K. Spatial analysis of county-level breast cancer mortality in Texas. J Environ Public Health 2012;2012(959343):959343. [doi: 10.1155/2012/959343] [Medline: 22518193]
32. SaTScan. 2023. URL: http://www.satscan.org [accessed 2025-10-25]
33. QuickFacts: West Valley City, Utah; Salt Lake County, Utah. US Census Bureau. 2025 Jun 17. URL: https://www.census.gov/quickfacts/fact/table/westvalleycitycityutah,saltlakecountyutah/RHI725223 [accessed 2025-10-25]
34. Roman Kovtun LH, Mendoza D, Pearson J. Influence of air pollution on perioperative outcomes & potential for big data drive discoveries. In: Anesthesia and Analgesia: Lippincott Williams & Wilkins; 2021.
35. Alicia Adiwidjaja EK, Hsu G, Quiroz PA, Tan JM. Spatial Analysis of Telehealth Utilization in a Pediatric Pain Clinic: International Anesthesia Research Society; 2022.
36. Adiwidjaja A, Hsu G, Tan JM. Ambient air quality exposure among pediatric anesthesia patients in a children's hospital: an application of spatial analytics and population data linkage. Presented at: Society for Technology in Anesthesia; Jan 13-15, 2022 URL: stahq.org/userfiles/files/14_STA_22AM_Abstract_43_Adiwidjaja%28723%29.pdf [accessed 2025-12-22]
37. Andreae MH, Nair S, Gabry JS, Goodrich B, Hall C, Shaparin N. A pragmatic trial to improve adherence with scheduled appointments in an inner-city pain clinic by human phone calls in the patient's preferred language. J Clin Anesth 2017 Nov;42:77-83. [doi: 10.1016/j.jclinane.2017.08.014] [Medline: 28841451]
38. Andreae MH, White RS, Chen KY, Nair S, Hall C, Shaparin N. The effect of initiatives to overcome language barriers and improve attendance: a cross-sectional analysis of adherence in an inner city chronic pain clinic. Pain Med 2017 Feb 1;18(2):265-274. [doi: 10.1093/pm/pnw161] [Medline: 28204760]
39. Grubesic TH, Matisziw TC. On the use of ZIP codes and ZIP code tabulation areas (ZCTAs) for the spatial analysis of epidemiological data. Int J Health Geogr 2006 Dec 13;5:1-15. [doi: 10.1186/1476-072X-5-58] [Medline: 17166283]
40. Grubesic TH. Zip codes and spatial analysis: Problems and prospects. Socioecon Plann Sci 2008 Jun;42(2):129-149. [doi: 10.1016/j.seps.2006.09.001]
41. Bonnet JB, Sultan A. Social deprivation, healthcare access and diabetic foot ulcer: a narrative review. J Clin Med 2022 Sep 15;11(18):5431. [doi: 10.3390/jcm11185431] [Medline: 36143078]
42. Hsu CC, Lee CH, Wahlqvist ML, et al. Poverty increases type 2 diabetes incidence and inequality of care despite universal health coverage. Diabetes Care 2012 Nov;35(11):2286-2292. [doi: 10.2337/dc11-2052] [Medline: 22912425]
43. Nishino Y, Gilmour S, Shibuya K. Inequality in diabetes-related hospital admissions in England by socioeconomic deprivation and ethnicity: facility-based cross-sectional analysis. PLoS One 2015;10(2):e0116689. [doi: 10.1371/journal.pone.0116689] [Medline: 25705895]
44. Piccolo RS, Duncan DT, Pearce N, McKinlay JB. The role of neighborhood characteristics in racial/ethnic disparities in type 2 diabetes: results from the Boston Area Community Health (BACH) Survey. Soc Sci Med 2015 Apr;130(79-90):79-90. [doi: 10.1016/j.socscimed.2015.01.041] [Medline: 25687243]
45. Schwartz S. The fallacy of the ecological fallacy: the potential misuse of a concept and the consequences. Am J Public Health 1994 May;84(5):819-824. [doi: 10.2105/ajph.84.5.819] [Medline: 8179055]
46. Zandbergen PA. Geocoding quality and implications for spatial analysis. Geogr Compass 2009 Mar;3(2):647-680. [doi: 10.1111/j.1749-8198.2008.00205.x]

Abbreviations

ASA: American Society of Anesthesiologists
ASPIRE: Anesthesiology Performance Improvement and Reporting Exchange
CBG: census block group
MPOG: Multicenter Perioperative Outcomes Group
ND: neighborhood disadvantage
PONV: postoperative nausea and vomiting
RAAP: risk-adjusted antiemetic prophylaxis

XSL•FO
RenderX

RR: relative risk
SDOH: social determinants of health
SMD: standardized mean difference
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
TIGER/Line: Topologically Integrated Geographic Encoding and Referencing System

Edited by A Mavragani; submitted 25.Nov.2024; peer-reviewed by C Zhao, SY Jiang; revised version received 22.Jul.2025; accepted 23.Jul.2025; published 24.Feb.2026.

Please cite as:
Han J, Wan N, Jacobson CK, Pace NL, Kartchner CK, Hohl A, Schonberger RB, Colquhoun DA, Dutton RP, Andreae M, Pearson JF
Disparities in Antiemetic Prophylaxis Care Processes Predicted by Patient Neighborhood: Retrospective Cohort and Geospatial Analysis
JMIR Public Health Surveill 2026;12:e69133
URL: https://publichealth.jmir.org/2026/1/e69133
doi:10.2196/69133

© Jiuying Han, Neng Wan, Cameron K Jacobson, Nathan L Pace, Cade K Kartchner, Alexander Hohl, Robert B Schonberger, Douglas A Colquhoun, Richard P Dutton, Michael Andreae, John F Pearson. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 24.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Risk Factors Associated With Tuberculosis Diagnostic Delay in the Jiangsu Province, China (2011-2021): Spatiotemporal Database Analysis Study

Yifan Tang[1,2], MSc; Cheng Chen[3], PhD; Mingming Chen[1], MSc; Kai Wang[1], MSc; Sifan Wang[1], MSc; Yi Lin[1], BSc; Qiao Liu[3], PhD; Chengxiu Ling[1], PhD; Tenglong Li[1*], PhD; Limei Zhu[3*], MSc

[1]Department of Biostatistics, Academy of Pharmacy, Xi'an Jiaotong-Liverpool University, Suzhou, China

[2]Department of Mathematical Sciences, University of Liverpool, Liverpool, United Kingdom

[3]Department of Chronic Communicable Disease, Center for Disease Control and Prevention of Jiangsu Province, No.172 Jiangsu Road, Gulou District, Nanjing, China

[*]these authors contributed equally

Corresponding Author:
Limei Zhu, MSc
Department of Chronic Communicable Disease, Center for Disease Control and Prevention of Jiangsu Province, No.172 Jiangsu Road, Gulou District, Nanjing, China

Abstract

Background: Tuberculosis (TB) remains a major public health concern. Despite improved diagnostic tools, delays in TB diagnosis persist and hinder control efforts.

Objective: This study aims to investigate the spatiotemporal patterns of TB diagnostic delay and identify individual and spatial risk factors in Jiangsu Province, China, from 2011 to 2021.

Methods: This study included 332,091 patients with TB who reported in Jiangsu Province from 2011 to 2021, using data obtained from the Jiangsu TB Information Management System, and diagnostic delay was defined as an interval of more than 28 days between symptom onset and diagnosis. Logistic regression was used to evaluate individual-level factors associated with delayed status, while a Bayesian spatiotemporal Beta model was used to analyze county-level TB diagnostic delay rates and assess spatial correlation using the global Moran *I*. The panel Granger causality analysis explored the temporal dynamics of delay rate transitions.

Results: Male patients, educators, and those diagnosed at the local Centers for Disease Control and Prevention had lower odds of diagnostic delay, whereas the older adults, agricultural workers, migrants, clinically diagnosed cases, and those diagnosed at community health centers had higher odds of delay. Spatial clustering in TB diagnostic delay rates was significant from 2015 onward (Moran I=0.110-0.193; all P<.05), excluding 2018 when Moran I was 0.054. The Bayesian spatiotemporal Beta model, which accounted for 31.8% of the total variation due to spatial structure, indicated that for each 1-unit increase in the proportion of local patients and for each 100,000-person increase in resident population, the TB diagnostic delay rate decreased by 33.9% (95% CI 0.128-0.498) and 2% (95% CI 0.005-0.033), respectively. The panel Granger causality analysis indicated that TB incidence and health care technicians significantly influenced temporal changes in delay rates.

Conclusions: TB diagnostic delays in Jiangsu were influenced by both individual and spatial factors, with the proportion of local patients and resident population size contributing significantly to spatiotemporal variation. Tailored interventions targeting high-risk groups and health care settings are needed.

(JMIR Public Health Surveill 2026;12:e80052) doi:10.2196/80052

KEYWORDS

tuberculosis; diagnostic delay; Bayesian spatiotemporal model; Moran I index; integrated nested Laplace approximation

Introduction

Tuberculosis (TB) is a chronic infectious disease caused by *Mycobacterium tuberculosis* and is primarily transmitted via airborne particles [1]. In 2022, TB was the second leading cause of death from a single infectious agent, surpassed only by COVID-19 [2]. Globally, an estimated 10.8 million new TB cases occurred in 2023, with an incidence rate of 134 per 100,000 population, posing a grave threat to public health [3]. As the country with the third highest TB burden [4], China has made significant efforts in TB control over recent decades, increasing the case detection rate from 30% in the 1990s to 80% by 2005 and thereby effectively reducing transmission and

XSL•FO
RenderX

incidence [5,6]. In Jiangsu Province, China, TB remains a critical public health challenge, consistently ranking second among reported class A and B infectious diseases [7].

The World Health Organization (WHO) End TB Strategy emphasized the importance of early diagnosis and timely treatment for TB control and prevention, as well as reducing treatment costs [8]. However, most national TB control programs primarily rely on passive case finding, a practice often resulting in treatment delays exceeding 1 month in approximately 42% of patients [9]. The extent of these delays varies globally due to socioeconomic and health care disparities, with particularly severe delays observed in less developed regions. For instance, reported median total delays range significantly from 68 days in France to 104 days in Ghana and up to 366 days in Afghanistan [10]. Therefore, identifying the key factors contributing to delays in specific regions is essential for developing targeted TB interventions and control measures.

Extensive research has examined risk factors contributing to TB diagnostic delays, which arise from complex interactions of individual behaviors, social determinants, and health care system challenges [11]. The declining clinical awareness of TB among health care workers, especially in low-incidence settings, combined with the nonspecific nature of typical TB symptoms, such as persistent cough and sputum production, often leads to early misdiagnosis as common respiratory infections [12]. Meanwhile, misdiagnosis and missed cases are exacerbated by urban-rural disparities in medical resources and surges in diagnostic pressure during peak health care demand periods, such as holidays or influenza seasons [4]. Socioeconomically vulnerable populations, including migrant workers and older adults, often experience delays in seeking medical care due to limited access to health care insurance, language barriers, and varying levels of education [13,14].

Although previous analyses have identified key risk factors, they have generally failed to sufficiently account for the spatial and temporal dependence inherent in TB diagnostic delays [15,16]. Past research using descriptive statistics has highlighted spatial heterogeneity, for instance, by revealing median delays of 30 days in eastern or central China versus 41 days in the west [6] and identifying regional disparities in Portugal [17]; these studies often overlook spatial autocorrelation. To address this limitation, the Bayesian spatiotemporal model provides a rigorous framework that incorporates explanatory variables to capture large-scale trends while also accounting for residual dependencies to reveal robust spatial patterns and potential risk factors [18]. Furthermore, by integrating prior knowledge to quantify uncertainty, this approach enhances both the accuracy and interpretability of findings [19,20]. The integrated nested Laplace approximation (INLA) algorithm offers an efficient approach for implementing Bayesian inference in such complex models [21]. Notably, the spatiotemporal patterns of TB diagnostic delay have been rarely investigated using this Bayesian approach, representing a critical research gap that this study aims to address.

This study used Bayesian spatiotemporal analysis to investigate TB diagnosis delays, capturing spatiotemporal dependencies and examining patterns of temporal and spatial variation. We have the following three research goals: (1) assess the existence of spatial and temporal autocorrelation in TB diagnostic delays within Jiangsu Province, China; (2) identify individual-level risk factors associated with delayed diagnosis; and (3) determine county-level determinants of diagnostic delay rates, while accounting for potential temporal and spatial random effects.

Methods

Data and Variables

We obtained TB surveillance data from the Jiangsu Tuberculosis Information Management System (TBIMS), spanning January 1, 2011, to December 31, 2021. The original dataset contained 354,274 infection cases reported in Jiangsu Province during this period, including patient information such as names, ages, sex, occupations, sources of patients, types of diagnosis, tracking status, types of hospital, dates of birth, onset dates, diagnostic dates, and so on. A total of 332,091 patients were analyzed in this study following exclusion criteria: (1) patients diagnosed and reported between January 1, 2011, and December 31, 2021, at health care institutions outside Jiangsu Province (n=8010); (2) patients with missing critical information (n=12,889); and (3) patients whose standardized *z* value for the total diagnostic delay that exceeded 3 (ie, more than 3 SD from the mean; n=1266) [22].

Total diagnostic delay, defined as the interval from the onset of TB symptoms to formal diagnosis, comprises both patient delay and health system delay [13,23]. Patient delay refers to the time between symptom onset and the first medical consultation, while health system delay spans from the time of the first health care visit to diagnosis. This study focused exclusively on total diagnostic delay, as the dataset lacked information regarding the date of the first medical visit. At the individual level, the outcome was defined as the diagnostic delay status, which indicated whether a patient's total diagnostic delay exceeded 28 days, a commonly adopted threshold for total TB diagnostic delay based on previous studies [13]. At the county level, the outcome was the TB diagnostic delay rate, which was calculated by dividing the number of delayed patients (ie, those with more than 28 d of total diagnostic delay) by the total number of patients in each county each year.

In addition to the individual-level TB surveillance data from Jiangsu TBIMS, a total of 9 annual county-level explanatory variables were included, categorized into three distinct domains: (1) demographic factors, comprising the annual proportions of older adult patients (≥60 y), male patients, local patients, and agricultural-worker patients among reported cases; (2) socioeconomic and health care indicators, including gross domestic product (GDP) per capita (adjusted to the 2021 Consumer Price Index), resident population size, TB incidence rate (per 1000 population), and the health care technicians (professionals per 1000 population); and (3) pandemic period, a binary variable (1=2020 - 2021; 0=2011 - 2019) introduced to adjust for the potential impact of social isolation policies and health care resource diversion during the COVID-19 pandemic. Data regarding health care technicians, GDP, and resident population were sourced from annual county statistical

XSL•FO
RenderX

yearbooks, while other variables were aggregated directly from the TB surveillance data.

The Bayesian Spatiotemporal Model

To explore the spatial correlation of the TB diagnostic delay rate across 89 districts and counties in Jiangsu Province from 2011 to 2021, we calculated the global Moran *I* to measure the spatial correlation [24,25]. When the result for Moran *I* is statistically significant, a positive value for Moran *I* suggests spatial clustering, while a negative value suggests spatial dispersion. The closer the value of *I* is to 1 or −1, the stronger the spatial association, while a value near 0 indicates a random spatial distribution of TB diagnostic delay rates.

Prior to the spatiotemporal modeling, multivariable binary logistic regression was used to identify risk factors associated with individual diagnostic delay status (binary outcome: 1 if total delay >28 d, 0 otherwise). Subsequently, a logit-link Bayesian Beta regression model was then applied, incorporating fixed effects for the proportion of older adult patients, proportion of male patients, proportion of local patients, proportion of agricultural-worker patients, GDP, TB incidence rate, number of health care technicians, and resident population, along with spatial and temporal random effects.

Specifically, let y(s,t) denote the TB diagnostic delay rate in year *t* over district or county *s*, with values ranging from 0 to 1. Here, *t*=1, 2, ..., 11 represents the years 2011 to 2021, and *s*=1, 2, ..., 95 represents the 95 counties or districts in Jiangsu Province, China. We further assumed that y(s,t) follows a β distribution with mean μ(s,t)∈(0,1) varying over time across counties or districts, and a constant precision parameter φ>0. Namely,

$$\text{logit}\ \mu(s,t)=xT(s,t)\beta+\delta t+us+\ (s,t),$$

where the vector x(s,t) represents the regional-level variables in Table 3. The (s,t) term is an unstructured random effect in the model. We employed the first-order Gaussian random walks (RW1) model and the Besag-York-Mollié 2 model to capture the overall temporal random effect δt and spatial random effect us [26].

The Besag-York-Mollié 2 model was used to capture spatial random effects by combining structured and unstructured spatial components through a mixing parameter [27]. The spatial effect for area i is expressed as:

$$bi=1\tau b(1-\varphi vi*+\varphi ui*),$$

where τb is the overall precision, ui∗ is a standardized intrinsic conditional auto-regressive component capturing structured spatial dependence, and vi∗ is a standardized Gaussian noise term representing unstructured spatial variability. The mixing parameter φ∈[0,1], is a spatial smoothing parameter, measuring the proportion of the marginal variance explained by the structured random effect.

As a latent effect implemented in the R-INLA package, the first-order Gaussian random walk was used to model temporal dependence [26]. For a latent Gaussian field u=(u1,…,un)T, it is a random walk of order 1 if the increments Δui=ui−ui−1 are independent and identically distributed Gaussian random variables with zero mean and precision τ>0 (inverse variance).

Based on the delay rates estimated from the Bayesian spatiotemporal Beta model, we dichotomized at the median and then fitted a Bayesian spatiotemporal binomial model to identify factors associated with a higher likelihood of diagnostic delay. Furthermore, to explore the temporal dynamics and potential drivers of delay rate, we conducted a panel Granger causality analysis [28].

All statistical analyses were conducted using R software (version 4.4.1; R Foundation for Statistical Computing). The *INLA*, *stats*, and *lmtest* packages were used to conduct the Bayesian spatiotemporal modeling, logistic regression, and the Granger causality analysis, respectively, with default prior distributions specified for the INLA hyperparameters.

Ethical Considerations

Anonymized data were obtained from the Jiangsu TBIMS, with all personal identifiers (eg, name and ID number) removed prior to analysis. The study protocol was reviewed by the ethical review board of the Jiangsu Provincial Center for Disease Control and Prevention (Jiangsu CDC) and granted an official exemption (acceptance: SL2025-B030-01), as the study was deemed retrospective and the data were deidentified. Data access and usage were strictly governed by a formal Data Usage Agreement between the Jiangsu CDC and the study authors. Informed consent was not required for this retrospective study using anonymized data.

Results

Descriptive Statistics

Table 1 summarizes the demographic and clinical features of patients with TB with delayed (≥28 d) and nondelayed (<28 d) diagnoses, respectively, in Jiangsu Province (2011-2021). Among all patients, the majority were male participants (239,692/332,091, 72.18%), working in agriculture (204,983/332,091, 61.72%), and reported by general hospitals (148,960/332,091, 44.86%), with a high proportion of local residents (239,440/332,091, 72.10%).

XSL•FO
RenderX

Table . Demographic and clinical characteristics of patients with tuberculosis reported in Jiangsu Province, 2011-2021.

Variables	Total delay days (<28), n (%)	Total delay days (≥28), n (%)
Age (y)		
<60	132,502 (39.90)	47,358 (14.26)
≥60	79,287 (23.87)	72,944 (21.97)
Sex		
Female	57,212 (17.23)	35,187 (10.60)
Male	154,577 (46.55)	85,115 (25.63)
Occupation		
Agriculture	128,957 (38.83)	76,026 (22.89)
Education	9219 (2.78)	4523 (1.36)
Health care	913 (0.27)	554 (0.17)
Housekeeping	36,705 (11.05)	20,111 (6.06)
Officials	2264 (0.68)	1415 (0.43)
Service	3660 (1.10)	1778 (0.54)
Worker	20,304 (6.11)	9386 (2.83)
Other	9767 (2.94)	6509 (1.96)
Source of patient		
Local	158,987 (47.87)	80,453 (24.23)
Different county	45,710 (13.76)	35,262 (10.62)
Different city	4471 (1.35)	3482 (1.05)
Different province	2621 (0.79)	1105 (0.33)
Types of diagnosis		
Confirmed	78,554 (23.65)	45,019 (13.56)
Clinically diagnosed	129,551 (39.01)	74,750 (22.51)
Suspected	3684 (1.11)	533 (0.16)
Tracking status		
Recorded	74,344 (22.39)	44,062 (13.27)
Referred	92,716 (27.92)	45,878 (13.81)
Tracked	32,184 (9.69)	23,132 (6.97)
Other	12,545 (3.78)	7230 (2.18)
Types of hospital		
Designated hospital	62,425 (18.80)	37,307 (11.23)
CHC[a]	28,793 (8.67)	15,460 (4.66)
CDC[b]	26,392 (7.95)	11,298 (3.40)
General hospital	93,216 (28.07)	55,744 (16.79)
Other	963 (0.29)	493 (0.15)
COVID-19		
Pre-epidemic	182,934 (55.09)	28,855 (8.69)
Epidemic	103,886 (31.28)	16,416 (4.94)

[a]CHC: community health center.

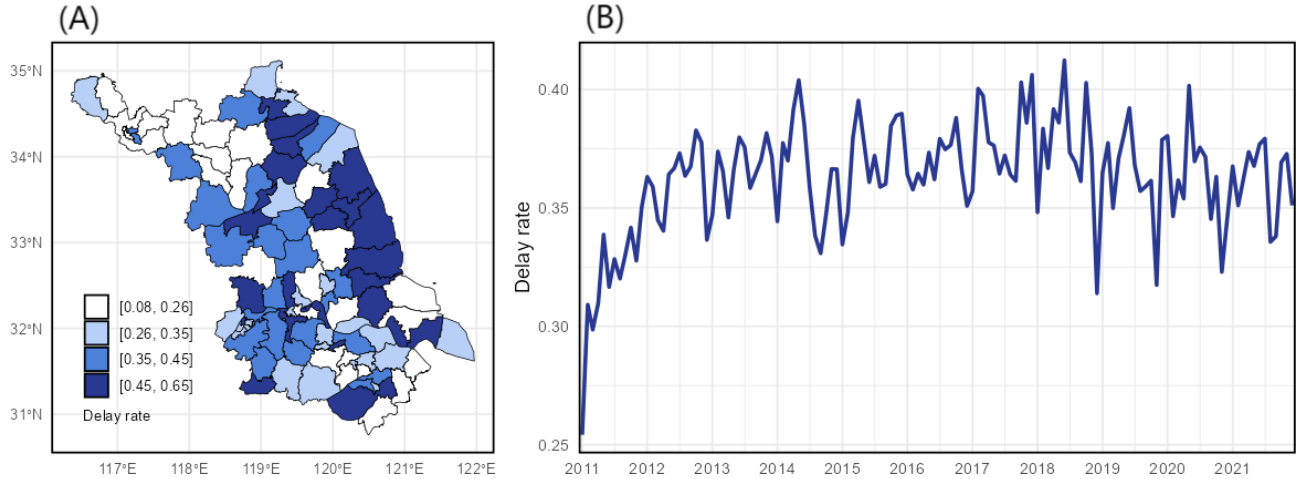
[b]CDC: Centers for Disease Control and Prevention.

Figure 1A shows the spatial distribution of average TB diagnostic delay rates in Jiangsu Province over the period from 2011 to 2021. The cities of Yancheng and Huaian generally exhibited the highest delay rates, and the Binhai district in

XSL•FO
RenderX

Yancheng recorded a peak delay rate of 93.35% in 2016. Figure 1B illustrates the temporal trend across all 95 counties. An increasing trend was evident from 2011 to 2014, followed by a plateau. Notable distinct declines in the average delay rate were observed in specific months, including December 2018, November 2019, and November 2020.

Figure 1. Annual average spatial distribution across 95 counties (A) and monthly provincial average temporal trends (B) of tuberculosis delayed diagnoses in Jiangsu Province, 2011-2021.



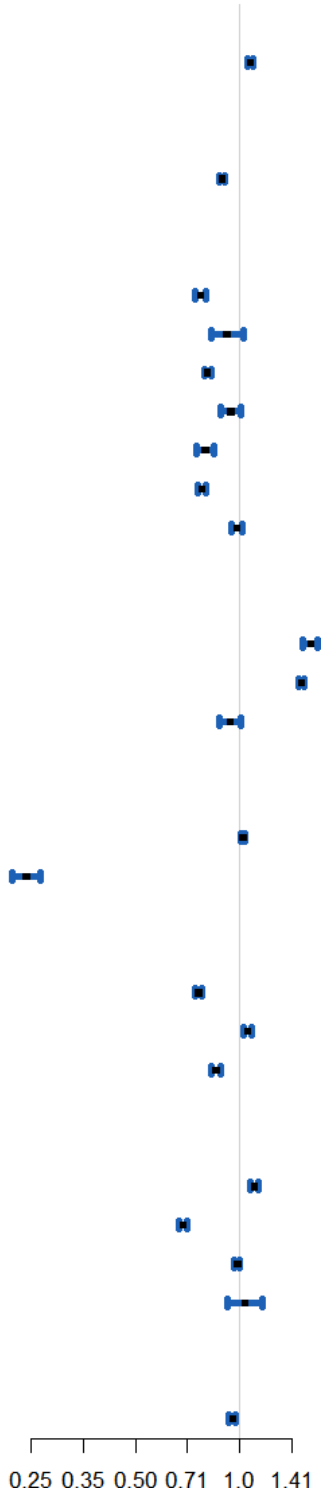
Analysis at the Individual Level

Figure 2 presents the results of a logistic regression analysis used to investigate potential risk factors associated with individual TB diagnosis delay status. Specifically, males had significantly lower odds of experiencing delayed TB diagnosis compared to females (odds ratio [OR] 0.891, 95% CI 0.874-0.906), while patients older than 60 years had slightly higher odds of delayed TB diagnosis compared to those who were younger than 60 years (OR 1.070, 95% CI 1.053-1.088). We also found that patients who worked in agriculture had higher odds of delayed TB diagnosis than patients in other occupations. Particularly, the differences between patients working in agriculture and patients of occupations in education (OR 0.769, 95% CI 0.740-0.799), housekeeping (OR 0.807, 95% CI 0.791-0.824), service (OR 0.794, 95% CI 0.749-0.843), and worker (OR 0.778, 95% CI 0.757-0.800) were statistically significant. Local patients (from the same county/district) had significantly lower odds of delayed TB diagnosis than patients from different counties (OR 1.598, 95% CI 1.525-1.675) or cities (OR 1.508, 95% CI 1.480-1.536) in Jiangsu. Regarding types of diagnosis, the suspected cases had significantly lower odds of delayed TB diagnosis compared to the confirmed cases (OR 0.242, 95% CI 0.221-0.266), while the clinically diagnosed cases had significantly higher odds of delayed TB diagnosis (OR 1.021, 95% CI 1.003-1.037) compared to the confirmed cases. Patients diagnosed at community health centers (CHCs) had higher odds of delayed diagnosis compared to those diagnosed at designated TB hospitals (OR 1.103, 95% CI 1.075-1.133). However, patients diagnosed at local CDCs had considerably lower odds of delayed diagnosis than those diagnosed at designated TB hospitals (OR 0.687, 95% CI 0.668-0.707). Finally, the COVID-19 period was associated with significantly lower odds of diagnostic delay (OR 0.954, 95% CI 0.933-0.975).

XSL•FO
RenderX

Figure 2. Odds ratios (ORs) from logistic regression model for factors affecting tuberculosis diagnostic delay status at the individual level in Jiangsu Province, 2011-2021. CDC: Centers for Disease Control and Prevention; CHC: community health center.

Variable	Category	OR (95% CI)	P value
Age			
	< 60	Reference	
	≥ 60	1.070 (1.053-1.088)	<.001
Sex			
	Female	Reference	
	Male	0.891 (0.874-0.906)	<.001
Occupation			
	Agriculture	Reference	
	Education	0.769 (0.740-0.799)	<.001
	Health care	0.918 (0.824-1.022)	.12
	Housekeeping	0.807 (0.791-0.824)	<.001
	Officials	0.941 (0.878-1.007)	.08
	Service	0.794 (0.749-0.843)	<.001
	Worker	0.778 (0.757-0.800)	<.001
	Other	0.981 (0.946-1.016)	.28
Source of patient			
	Local	Reference	
	Different county	1.598 (1.525-1.675)	<.001
	Different city	1.508 (1.480-1.536)	<.001
	Different province	0.940 (0.875-1.011)	.09
Type of diagnosis			
	Confirmed	Reference	
	Clinically diagnosed	1.021 (1.003-1.037)	.006
	Suspected	0.242 (0.221-0.266)	<.001
Tracking status			
	Recorded	Reference	
	Referred	0.760 (0.745-0.776)	<.001
	Tracked	1.052 (1.028-1.077)	<.001
	Other	0.853 (0.825-0.882)	<.001
Type of hospital			
	Designated hospital	Reference	
	CHC	1.103 (1.075-1.133)	<.001
	CDC	0.687 (0.668-0.707)	<.001
	General hospital	0.984 (0.967-1.002)	.08
	Other	1.033 (0.922-1.157)	.58
COVID 19			
	Pre-epidemic	Reference	
	Epidemic	0.954 (0.933-0.975)	<.001



Analysis at the County Level

Table 2 presents the global Moran *I* indices for TB diagnostic delay rates from 2011 to 2021. From 2011 to 2014, the Moran *I* values were close to zero (all *P*>.05), indicating no significant spatial autocorrelation. However, a marked increase in Moran *I* was observed from 2015 onwards, suggesting the emergence of substantial spatial clustering. Consequently, the Bayesian spatiotemporal Beta model was used to fully account for the spatial autocorrelation found in TB diagnostic delay rates across the years.

XSL•FO
RenderX

Table . Global Moran *I* index for tuberculosis diagnostic delay rate in Jiangsu Province, 2011-2021.

Year	Moran *I*	*P* value
2011	0.018	.33
2012	0.079	.08
2013	0.047	.19
2014	0.040	.22
2015	0.193	<.001
2016	0.167	.003
2017	0.161	.004
2018	0.054	.16
2019	0.113	.03
2020	0.110	.03
2021	0.085	.07

Table 3 presents the fixed effect estimates for the proportion of older adult patients, proportion of male patients, proportion of local patients, proportion of agricultural-worker patients, GDP, TB incidence rate, number of health care technicians, and resident population from the Bayesian spatiotemporal Beta model. For each 1-unit increase in the proportion of local patients, the TB diagnostic delay rate decreases by 33.9% (1−exp[−0.415], 95% CI 0.128-0.498). Additionally, each 100,000-person increase in resident population is associated with a 2% (95% CI 0.005-0.033) decrease in TB diagnostic delay. For random effects, the precision parameter τb was estimated as 3.411, indicating a moderate degree of spatial variation in the diagnostic delay rates across counties. In addition, the mixing parameter φ was estimated as 0.318, meaning that the structured spatial component accounts for 31.8% of the total spatial variation in TB diagnostic delay rate.

Table . The result of the Bayesian Beta spatiotemporal model of TB[a] diagnostic delay rate at the county level in Jiangsu Province, 2011-2021.

Variables	Mean (SD)	0.025 quantile	0.975 quantile
Fixed effect			
Proportion of older adult patients (%)	0.234 (0.261)	−0.281	0.747
Proportion of male patients (%)	−0.002 (0.014)	−0.030	0.026
Proportion of local patients (%)	−0.415 (0.140)	−0.691	−0.138
Proportion of agricultural-worker patients (%)	0.187 (0.184)	−0.176	0.548
GDP[b] (thousand yuan per person)	0.061 (0.037)	−0.014	0.135
TB incidence rate (%)	0.055 (0.062)	−0.068	0.180
Health care technicians (per 1000 residents)	−0.008 (0.008)	−0.024	0.007
Resident population (in 100,000)	−0.020 (0.007)	−0.034	−0.005
COVID-19 (0 - 1)	−0.049 (0.050)	−0.148	0.053
Random effect			
Precision parameter for BYM2[c]	3.411	2.389	4.828
Mixing parameter for BYM2 (φ)	0.318	0.085	0.652

[a]TB: tuberculosis.

[b]GDP: gross domestic product.

[c]BYM2: Besag-York-Mollié 2.

Figure S1 in Multimedia Appendix 1 illustrates the spatiotemporal distribution of the estimated TB diagnostic delay rate obtained from the Bayesian spatiotemporal Beta model. Higher estimated delay rates were concentrated in the northern coastal cities such as Huaian, Yancheng, and Lianyungang, while lower rates were observed in Xuzhou and Suzhou in the southern region. Among the 95 districts and counties over the 11-year period, Qingjiangpu of Huaian ranked among the top

XSL•FO
RenderX

3 with the highest estimated delay rates, with values of 0.648 in 2021, 0.645 in 2014, and 0.644 in 2018. Meanwhile, from 2011 to 2013, the delay rate exhibited an increasing temporal trend, which was particularly evident in Suqian. In contrast, Figure S2 in Multimedia Appendix 1 presents the observed delay rates, which show much larger fluctuations, reflecting random noise and potential instability due to small-area sample variation. The differences between Figures S1 and S2 in Multimedia Appendix 1 arise because the Bayesian spatiotemporal Beta model smooths random noise and incorporates both spatial and temporal dependencies. By integrating relevant covariates and modeling residual correlation through random effects, it effectively adjusts for unobserved dependence and confounding, producing spatially coherent and statistically reliable estimates.

The median of the estimated TB diagnostic delay rate was 0.352. Based on the obtained median, the results from the Bayesian spatiotemporal binomial model (Table S1 in Multimedia Appendix 1) show that counties with a higher proportion of agricultural workers and higher GDP were more likely to experience high diagnostic delays. Conversely, counties with larger shares of local patients, larger resident populations, and the COVID-19 period were associated with a lower risk of high delays. Furthermore, Table S2 in Multimedia Appendix 1 indicated that GDP, TB incidence, and health care technicians had significant Granger causal effects on the temporal changes in the delay rate ($P<.05$).

Sensitivity Analysis

The sensitivity analysis confirmed the consistency of our main results. We applied penalized complexity priors to spatial and temporal model parameters to constrain model complexity and prevent overfitting (Tables S3-S6 in Multimedia Appendix 1) [29]. Across these analyses, the direction and statistical significance of the main variables remained consistent, with a slight change in the magnitude of some coefficients. To examine the impact of risk factors associated with TB diagnostic delays over shorter periods, we divided the study period into 2 subperiods (2011 - 2015 vs 2016 - 2021), representing China's 12th and 13th Five-Year Plans for Tuberculosis Prevention and Control (Table S7 in Multimedia Appendix 1) [30]. We found that the significant associations between diagnostic delay and residency or occupation observed in the first subperiod disappeared in the second subperiod, likely due to expanded health care access and continuous public health developments in Jiangsu.

Discussion

This research analyzed 332,091 patients with TB in Jiangsu Province from 2011 to 2021, combining individual-level analysis with county-level spatiotemporal modeling to enhance understanding of diagnostic delay risk factors and their spatial and temporal patterns. At the individual level, we found that all 7 risk factors (ie, age, sex, occupation, patient source, type of diagnosis, tracking status, and type of hospital) were significant. At the county level, significant spatial clustering was observed from 2015. Drawing on such spatial dependence, we found that the proportion of local patients and the resident population were significantly and negatively associated with the TB diagnostic delay rates. Counties with higher proportions of older adults and agricultural-worker patients were more likely to experience high diagnostic delays. Moreover, GDP, TB incidence, and health care technicians exhibited significant effects on temporal changes in the delay rate.

Various individual characteristics were found to be significantly associated with the status of TB diagnostic delay. The odds of experiencing a TB diagnostic delay were significantly lower for males, consistent with the findings of previous studies [31,32]. A study in Portugal suggested that the higher overall TB burden in males (male-to-female ratio 2:1) could increase clinical suspicion and expedite diagnosis when men seek care [33]. The odds of experiencing TB diagnostic delay for education industry workers were also significantly lower, potentially attributed to strict TB screening programs for students and higher health management standards in Jiangsu and elsewhere [34]. We found older adult patients had higher odds of experiencing TB diagnostic delay, mainly due to factors such as lower education levels, poorer health awareness, lack of knowledge on TB prevention and treatment, economic difficulties, and insufficient social support [35,36]. Regarding the type of hospital, patients diagnosed by CHCs had higher odds of experiencing TB diagnostic delay, likely because these CHCs have limited resources and clinical experience [37]. For example, an artificial intelligence–assisted diagnostic platform was launched in Jiangsu Province in 2023, but this system has not been implemented at CHCs [38]. In contrast, patients diagnosed by CDCs had much lower odds of experiencing delay, underscoring the specialized knowledge needed for early TB diagnosis and thus the critical role of CDCs in TB detection. Finally, the significantly lower odds of diagnostic delay during the COVID-19 pandemic likely reflect how rigorous respiratory screening, targeting basically the same symptoms (eg, fever and cough), prompted earlier identification of TB cases that might be otherwise overlooked [2].

The spatial distribution of TB diagnostic delay exhibited clear characteristics of spatial clustering in Jiangsu Province. The global Moran I index showed no significant spatial autocorrelation in TB diagnostic delay rates from 2011 to 2014. Starting in 2015, however, significant spatial clustering emerged (Moran I=0.110-0.193; $P<.05$ for 2015-2017 and 2019-2020), indicating that the delay rates formed a stable spatial dependence pattern across counties or districts [25,39]. The mixing parameter of the Bayesian spatiotemporal Beta model was estimated at 0.318, indicating that a substantial portion of the spatial variation was attributable to structured spatial effects, also reflecting interdependency in the delay-risk patterns of neighboring areas. A possible explanation for this is the promotion of rapid drug-resistant TB molecular biological testing equipment in Jiangsu Province, which has improved TB diagnostic efficiency but may be disproportionately allocated across counties/districts due to a limited supply [40].

Our research revealed significant associations between the key factors and the TB diagnostic delay rate. For each unit increase in the proportion of local patients, the TB diagnostic delay rate decreased by 33.9%. This suggested that patients who lived permanently within the county might have had better access to

XSL•FO
RenderX

local health services, greater familiarity with the health care system, or improved continuity of care, all of which could have facilitated earlier diagnosis [41]. Additionally, each increase of 100,000 residents was associated with a 2% decrease in the TB diagnostic delay rate. Larger populations were typically found in more urbanized or economically developed counties, which tended to have better health care infrastructure, higher diagnostic capacity, and more accessible TB services [42]. When counties were classified using the median estimated delay rate, those with larger older adult populations were consistently identified as high-delay areas, which was consistent with the individual-level associations. A higher proportion of agricultural-worker patients was likewise linked to greater diagnostic delays, reflecting structural barriers such as limited health care access, seasonal labor patterns, and insufficient disease awareness [43]. Panel Granger causality tests further indicated that GDP and health care technician density were key drivers of temporal fluctuations in diagnostic delay. Higher GDP likely supported more advanced diagnostic infrastructure and resource allocation, while greater technician density enhanced diagnostic capacity and accelerated case detection [44,45].

The significance of this study lies in three main aspects. First, there had been a lack of large-scale, population-based investigations into diagnostic delay in TB in Jiangsu Province. Based on a comprehensive surveillance system covering 332,091 patients with TB, the study identified several key risk factors associated with TB diagnostic delay in the general population of Jiangsu, providing an essential empirical basis for targeted interventions. Second, the analysis confirmed the existence of spatial correlation and random effects in TB diagnostic delay rates, thereby underscoring the necessity of using a Bayesian spatiotemporal modeling approach to capture the underlying spatial and temporal dependence appropriately. Third, we conducted statistical analysis at both the individual and county levels, which accounts for geographic disparities while estimating the risk factors associated with TB diagnostic delay rates. This approach leads to more targeted and synergistic public health strategies for reducing regional TB diagnostic delay rates, as well as individual chances of experiencing TB diagnostic delays.

Our study also has several limitations. First, while diagnostic delay may affect the temporal alignment between symptom onset and reporting, the calculation of TB incidence rate was not adjusted for such delays and thereby may be biased [46]. Second, due to data limitations, this study focused on the total diagnostic delay, making it impossible to distinguish between patient delay and health system delay. Future studies with more detailed health care-seeking information are needed to explore these 2 components separately. Third, continuous socioeconomic development over the decade may introduce temporal heterogeneity in the impact of risk factors, limiting the findings to this specific developmental stage.

In conclusion, a multifaceted and targeted approach is essential for effectively reducing diagnostic delays in TB. First, efforts should be made to promote proactive health-seeking behaviors among individuals, particularly among high-risk populations such as agricultural workers, older adults, and female individuals. These groups are more vulnerable to pulmonary diseases and often underrepresented in passive case detection strategies [47]. Second, from a health system perspective, it is necessary to enhance TB screening and testing protocols for migrants by improving access to care and removing systemic barriers that can delay their diagnoses, should they have TB [48]. Third, more resources should be allocated to CHCs to enhance their diagnostic capacity and knowledge of TB, as they are often the first point of contact for patients with TB, especially in underdeveloped areas [49].

Acknowledgments

The authors thank the editor and reviewers for their valuable comments and suggestions, which greatly improved the manuscript. While preparing this work, the authors used ChatGPT (OpenAI) to assist with language editing and improve readability. The authors critically reviewed and revised the artificial intelligence–generated content to ensure accuracy and adherence to scientific standards. The authors take full responsibility for the final manuscript.

Funding

This study was supported by the following funding sources: National Natural Science Foundation of China (82574173 and 82003516), Jiangsu Provincial Natural Science Foundation (BK20251958), and Jiangsu Provincial Medical Key Discipline (ZDXK202250).

Data Availability

The data analyzed during this study are not publicly available due to restrictions imposed by the Jiangsu Provincial Center for Disease Control and Prevention in eastern China. However, access to the data can be requested from the corresponding author, LZ, upon reasonable request and with appropriate approval.

Authors' Contributions

TL, LZ, and YT developed the methodology and conceptualization. YT prepared the initial draft of the paper and programmed the code. YT and TL analyzed the main results. CC, QL, and LZ conducted the survey and data collection. YL, YT, MC, KW,

XSL•FO
RenderX

and SW processed the original data. CC, QL, LZ, MC, KW, CL, and TL commented on and revised manuscript drafts. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Estimated and observed tuberculosis diagnostic delay rates, sensitivity analyses, Bayesian spatiotemporal binomial model results, and the panel Granger causality analysis results.
[DOCX File, 384 KB - publichealth_v12i1e80052_app1.docx]

References

1. Yang H, Ruan X, Li W, Xiong J, Zheng Y. Global, regional, and national burden of tuberculosis and attributable risk factors for 204 countries and territories, 1990-2021: a systematic analysis for the Global Burden of Diseases 2021 study. BMC Public Health 2024 Nov 11;24(1):3111. [doi: 10.1186/s12889-024-20664-w] [Medline: 39529028]
2. Ntoumi F, Nachega JB, Aklillu E, et al. World Tuberculosis Day 2022: aligning COVID-19 and tuberculosis innovations to save lives and to end tuberculosis. Lancet Infect Dis 2022 Apr;22(4):442-444. [doi: 10.1016/S1473-3099(22)00142-6] [Medline: 35248166]
3. Chen Z, Wang T, Du J, et al. Decoding the WHO global tuberculosis report 2024: a critical analysis of global and Chinese key data. Zoonoses 2025;5(1):999. [doi: 10.15212/ZOONOSES-2024-0061]
4. Gilmour B, Xu Z, Bai L, Alene KA, Clements ACA. The impact of ethnic minority status on tuberculosis diagnosis and treatment delays in Hunan province, China. BMC Infect Dis 2022 Jan 26;22(1):90. [doi: 10.1186/s12879-022-07072-4] [Medline: 35081919]
5. Wang L, Liu J, Chin DP. Progress in tuberculosis control and the evolving public-health system in China. The Lancet 2007 Feb;369(9562):691-696. [doi: 10.1016/S0140-6736(07)60316-X]
6. Jiang H, Liu M, Zhang Y, et al. Changes in incidence and epidemiological characteristics of pulmonary tuberculosis in Mainland China, 2005-2016. JAMA Netw Open 2021 Apr 1;4(4):e215302. [doi: 10.1001/jamanetworkopen.2021.5302] [Medline: 33835173]
7. Notice on issuing the Jiangsu Province's "14th five-year plan" for tuberculosis prevention and control. : Jiangsu Provincial Health Commission; 2021 Nov 16 URL: https://wjw.jiangsu.gov.cn/art/2021/11/16/art_49495_10114796.html [accessed 2025-12-31]
8. Floyd K, Glaziou P, Zumla A, Raviglione M. The global tuberculosis epidemic and progress in care, prevention, and research: an overview in year 3 of the End TB era. Lancet Respir Med 2018 Apr;6(4):299-314. [doi: 10.1016/S2213-2600(18)30057-2] [Medline: 29595511]
9. Getnet F, Demissie M, Assefa N, Mengistie B, Worku A. Delay in diagnosis of pulmonary tuberculosis in low-and middle-income settings: systematic review and meta-analysis. BMC Pulm Med 2017 Dec 13;17(1):202. [doi: 10.1186/s12890-017-0551-y] [Medline: 29237451]
10. Yanogo PK, Balima C, Meda N. Total, patient and system diagnostic delays for pulmonary bacilliferous tuberculosis in the six diagnostic and treatment centers in the five health districts of the Central Region, Burkina Faso, 2018. J Epidemiol Glob Health 2022 Mar;12(1):124-132. [doi: 10.1007/s44197-021-00027-z] [Medline: 34978709]
11. Lönnroth K, Migliori GB, Abubakar I, et al. Towards tuberculosis elimination: an action framework for low-incidence countries. Eur Respir J 2015 Apr;45(4):928-952. [doi: 10.1183/09031936.00214014] [Medline: 25792630]
12. Miller AC, Arakkal AT, Koeneman S, et al. Incidence, duration and risk factors associated with delayed and missed diagnostic opportunities related to tuberculosis: a population-based longitudinal study. BMJ Open 2021 Feb 18;11(2):e045605. [doi: 10.1136/bmjopen-2020-045605] [Medline: 33602715]
13. Xiao W, Chen B, Huang D, et al. Comparison of delay in tuberculosis diagnosis between migrants and local residents in an eastern county of China: an analysis of the electronic data between 2015 and 2019. Front Public Health 2021;9:758335. [doi: 10.3389/fpubh.2021.758335] [Medline: 34869174]
14. Vigneswaran N, Parnis R, Lowbridge C, Townsend D, Ralph AP. Factors leading to diagnostic delay in tuberculosis in the tropical north of Australia. Intern Med J 2024 Apr;54(4):582-587. [doi: 10.1111/imj.16223] [Medline: 37688576]
15. Hong CY, Wang FL, Zhang YT, et al. Time-trend analysis of tuberculosis diagnosis in Shenzhen, China between 2011 and 2020. Front Public Health 2023;11:1059433. [doi: 10.3389/fpubh.2023.1059433] [Medline: 36891348]
16. Zhou Y, Wang F, Huang L, et al. Factors associated with tuberculosis care-seeking and diagnostic delays among childhood pulmonary tuberculosis in Zhejiang Province, China: a 10-year retrospective study. Sci Rep 2024;14(1):17086. [doi: 10.1038/s41598-024-68173-5] [Medline: 39048697]
17. Soares P, Aguiar A, Leite A, Duarte R, Nunes C. Ecological factors associated with areas of high tuberculosis diagnosis delay. Public Health (Fairfax) 2022 Jul;208:32-39. [doi: 10.1016/j.puhe.2022.04.010] [Medline: 35687953]

XSL•FO
RenderX

18. Wang K, Ling C, Chen Y, Zhang Z. Spatio-temporal joint modelling on moderate and extreme air pollution in Spain. Environ Ecol Stat 2023 Dec;30(4):601-624. [doi: 10.1007/s10651-023-00575-6]
19. Debusho LK, Gemechu LL. Joint spatiotemporal modelling of tuberculosis and human immunodeficiency virus in Ethiopia using a Bayesian hierarchical approach. BMC Public Health 2025 Jan 30;25(1):377. [doi: 10.1186/s12889-024-20996-7] [Medline: 39885478]
20. Chen ZY, Deng XY, Zou Y, et al. A spatio-temporal Bayesian model to estimate risk and influencing factors related to tuberculosis in Chongqing, China, 2014-2020. Arch Public Health 2023 Mar 21;81(1):42. [doi: 10.1186/s13690-023-01044-z] [Medline: 36945028]
21. Tang Y, Chen Y, Zheng J, et al. Spatio-temporal pattern and risk factors of HIV/AIDS prevalence in Zhejiang, China, from 2005 to 2022 using R-INLA. One Health 2025 Jun;20:101038. [doi: 10.1016/j.onehlt.2025.101038] [Medline: 40321630]
22. Leys C, Ley C, Klein O, Bernard P, Licata L. Detecting outliers: do not use standard deviation around the mean, use absolute deviation around the median. J Exp Soc Psychol 2013 Jul;49(4):764-766. [doi: 10.1016/j.jesp.2013.03.013]
23. Yang J, Kwon Y, Kim J, et al. Delays in the diagnosis and treatment of tuberculosis during the COVID-19 outbreak in the Republic of Korea in 2020. Osong Public Health Res Perspect 2021 Oct;12(5):293-303. [doi: 10.24171/j.phrp.2021.0063] [Medline: 34719221]
24. MORAN PAP. Notes on continuous stochastic phenomena. Biometrika 1950 Jun;37(1-2):17-23. [Medline: 15420245]
25. Xue M, Zhong J, Gao M, et al. Analysis of spatial-temporal dynamic distribution and related factors of tuberculosis in China from 2008 to 2018. Sci Rep 2023 Mar 27;13(1):4974. [doi: 10.1038/s41598-023-31430-0] [Medline: 36973322]
26. Gómez-Rubio V. Bayesian Inference With INLA: Chapman and Hall/CRC; 2020:330. [doi: 10.1201/9781315175584]
27. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. J R Stat Soc Ser B Stat Methodol 2009 Apr 1;71(2):319-392. [doi: 10.1111/j.1467-9868.2008.00700.x]
28. Shojaie A, Fox EB. Granger causality: a review and recent advances. Annu Rev Stat Appl 2022 Mar;9(1):289-319. [doi: 10.1146/annurev-statistics-040120-010930] [Medline: 37840549]
29. Simpson D, Rue H, Riebler A, Martins TG, Sørbye SH. Penalising model component complexity: a principled, practical approach to constructing priors. Statist Sci 2017;32(1):1-28. [doi: 10.1214/16-STS576]
30. Ni S, Chen G, Wang J, et al. Assessment of public literacy in TB prevention and control in the National 13th Five-Year plan for Tuberculosis Prevention and Control (2016-2020) in China. BMC Health Serv Res 2025 Jan 9;25(1):50. [doi: 10.1186/s12913-024-12155-w] [Medline: 39789635]
31. Karim F, Islam MA, Chowdhury AMR, Johansson E, Diwan VK. Gender differences in delays in diagnosis and treatment of tuberculosis. Health Policy Plan 2007 Sep;22(5):329-334. [doi: 10.1093/heapol/czm026] [Medline: 17698889]
32. Alavi SM, Bakhtiyariniya P, Albagi A. Factors associated with delay in diagnosis and treatment of pulmonary tuberculosis. Jundishapur J Microbiol 2015 Mar;8(3):e19238. [doi: 10.5812/jjm.19238] [Medline: 25861434]
33. Santos JA, Leite A, Soares P, Duarte R, Nunes C. Delayed diagnosis of active pulmonary tuberculosis—potential risk factors for patient and healthcare delays in Portugal. BMC Public Health 2021 Nov 27;21(1):2178. [doi: 10.1186/s12889-021-12245-y] [Medline: 34837969]
34. Bhardwaj B, Naik E, Casanas B, Breglia MD, Lauzardo M. Tuberculosis screening and treatment of latent tuberculosis infection among international college students. Fla Public Health Rev 2010;7(1):26-31 [FREE Full text]
35. Raghu S. Challenges in treating tuberculosis in the elderly population in tertiary institute. Indian J Tuberc 2022;69 Suppl 2:S225-S231. [doi: 10.1016/j.ijtb.2022.10.008] [Medline: 36400514]
36. Hassani S, Mohammadi Shahboulagi F, Foroughan M, Nadji SA, Tabarsi P, Ghaedamini Harouni G. Factors associated with medication adherence in elderly individuals with tuberculosis: a qualitative study. Can J Infect Dis Med Microbiol 2023;2023(1):4056548. [doi: 10.1155/2023/4056548] [Medline: 36937803]
37. Li M, Wang Z, Zhang B, Wei T, Hu D, Liu X. Job performance of medical graduates with compulsory services in underserved rural areas in China: a cohort study. Int J Health Policy Manag 2022 Dec 6;11(11):2600-2609. [doi: 10.34172/ijhpm.2022.6335] [Medline: 35184509]
38. Feng Y, Liang Y, Li P, et al. Artificial intelligence assisted detection of superficial esophageal squamous cell carcinoma in white-light endoscopic images by using a generalized system. Discov Oncol 2023 May 19;14(1):73. [doi: 10.1007/s12672-023-00694-3] [Medline: 37208546]
39. Li Y, Luo D, Zheng Y, et al. Spatiotemporal distribution and risk factors for patient and diagnostic delays among groups with tuberculous pleurisy: an analysis of 5-year surveillance data in eastern China. Front Public Health 2024;12:1461854. [doi: 10.3389/fpubh.2024.1461854]
40. Ding XY, Mao WH, Lu W, et al. Impact of multiple policy interventions on the screening and diagnosis of drug-resistant tuberculosis patients: a cascade analysis on six prefectures in China. Infect Dis Poverty 2021 Jan 19;10(1):8. [doi: 10.1186/s40249-021-00793-9] [Medline: 33468247]
41. Dreiher J, Comaneshter DS, Rosenbluth Y, Battat E, Bitterman H, Cohen AD. The association between continuity of care in the community and health outcomes: a population-based study. Isr J Health Policy Res 2012 May 23;1(1):21. [doi: 10.1186/2045-4015-1-21] [Medline: 22913949]

XSL•FO
RenderX

42. Chen K, Cheng L, Yu H, et al. Spatial-temporal distribution characteristics of pulmonary tuberculosis in eastern China from 2011 to 2021. Epidemiol Infect 2024 May 15;152:e84. [doi: 10.1017/S0950268824000785] [Medline: 38745412]
43. Liu Y, Fan M, Li Y, et al. Patient delay in the diagnosis of pulmonary tuberculosis in the elderly—China, 2015–2023. China CDC Weekly 2024;6(42):1075-1079. [doi: 10.46234/ccdcw2024.221]
44. Teo AKJ, Singh SR, Prem K, Hsu LY, Yi S. Duration and determinants of delayed tuberculosis diagnosis and treatment in high-burden countries: a mixed-methods systematic review and meta-analysis. Respir Res 2021 Sep 23;22(1):251. [doi: 10.1186/s12931-021-01841-6] [Medline: 34556113]
45. Munyangaju I, José B, Bassat Q, et al. Assessment of radiological capacity and disparities in TB diagnosis: a comparative study of Mozambique, South Africa and Spain. BMJ Public Health 2024 Dec;2(2):e001392. [doi: 10.1136/bmjph-2024-001392] [Medline: 40018597]
46. Li T, White LF. Bayesian back-calculation and nowcasting for line list data during the COVID-19 pandemic. PLoS Comput Biol 2021 Jul;17(7):e1009210. [doi: 10.1371/journal.pcbi.1009210] [Medline: 34252078]
47. Bhatia V, Rijal S, Sharma M, et al. Ending TB in South-East Asia: flagship priority and response transformation. Lancet Reg Health Southeast Asia 2023 Nov;18:100301. [doi: 10.1016/j.lansea.2023.100301] [Medline: 38028166]
48. Cudahy PGT, Andrews JR, Bilinski A, et al. Spatially targeted screening to reduce tuberculosis transmission in high-incidence settings. Lancet Infect Dis 2019 Mar;19(3):e89-e95. [doi: 10.1016/S1473-3099(18)30443-2] [Medline: 30554997]
49. Arini M, Sugiyo D, Permana I. Challenges, opportunities, and potential roles of the private primary care providers in tuberculosis and diabetes mellitus collaborative care and control: a qualitative study. BMC Health Serv Res 2022 Feb 17;22(1):215. [doi: 10.1186/s12913-022-07612-3] [Medline: 35177037]

Abbreviations

CDC: Centers for Disease Control and Prevention
CHC: community health center
GDP: gross domestic product
INLA: integrated nested Laplace approximation
OR: odds ratio
TB: tuberculosis
TBIMS: Tuberculosis Information Management System
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 07.Jul.2025; peer-reviewed by M Hu, Z He; revised version received 13.Dec.2025; accepted 17.Dec.2025; published 26.Jan.2026.

Please cite as:
Tang Y, Chen C, Chen M, Wang K, Wang S, Lin Y, Liu Q, Ling C, Li T, Zhu L
Risk Factors Associated With Tuberculosis Diagnostic Delay in the Jiangsu Province, China (2011-2021): Spatiotemporal Database Analysis Study
JMIR Public Health Surveill 2026;12:e80052
URL: https://publichealth.jmir.org/2026/1/e80052
doi:10.2196/80052

© Yifan Tang, Cheng Chen, Mingming Chen, Kai Wang, Sifan Wang, Yi Lin, Qiao Liu, Chengxiu Ling, Tenglong Li, Limei Zhu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 26.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Potential Impact of Federal Funding Cuts on Access to Pre-Exposure Prophylaxis in Atlanta, Georgia: Geographic Modeling Study

Noah Mancuso[1,2], MSPH; Patrick S Sullivan[1], DVM, PhD

[1]Department of Epidemiology, Emory University, Atlanta, GA, United States
[2]Queer Health Collaborative, Atlanta, GA, United States

Corresponding Author:
Noah Mancuso, MSPH
Department of Epidemiology
Emory University
1518 Clifton Rd NE
Atlanta, GA, 30322
United States
Phone: 1 4047278720
Email: noah.mancuso@emory.edu

Abstract

Background: Despite major biomedical advances in HIV testing, prevention, and treatment, annual HIV transmissions in the United States remain above 30,000. Geographic access to pre-exposure prophylaxis (PrEP) is critical to HIV prevention efforts, particularly in regions with high HIV burdens, such as metro-Atlanta. Community-based organizations (CBOs) play a central role in delivering culturally competent prevention services, yet many rely on federal funding that is increasingly unstable. Understanding the potential impact of CBO closures on geographic access to PrEP is essential for anticipating inequities and informing policy.

Objective: The aim of this study was to estimate how hypothetical closures of federally funded CBOs providing PrEP affect geographic access to PrEP clinics by car and public transit across metro-Atlanta and to assess whether impacts differ by community racial/ethnic composition.

Methods: We identified 71 PrEP-providing clinics in metro-Atlanta (August 2025), including 12 CBOs. Using 3 simulated closure scenarios in which 25% of CBOs were randomly closed, we calculated one-way travel times from 2466 census block group (CBG) centroids to the nearest PrEP-providing clinic. Travel times were estimated for car and public transit across 3 weekdays and timepoints and then averaged per CBG. Two-sided paired *t* tests were used to compare the change in travel time compared to baseline. Logistic regression assessed associations between racial/ethnic plurality and increased travel times.

Results: Under baseline conditions, 100% of CBGs had car access to a PrEP clinic within 30 minutes compared to only 41.6% (1027/2466) via public transit. Across closure scenarios, 732 CBGs (29.6%; representing over 1 million residents) experienced increased transit times (mean increase 1.2 minutes; range 0.0-11.6; $P<.001$), and 7 CBGs lost transit access entirely. For car travel, 1184 CBGs (48%; representing approximately 1.7 million residents) experienced increased drive times (mean increase 0.5 minutes; range 0.0-6.4; $P=.03$). Black-plurality CBGs had higher odds of increased drive times compared to White-plurality CBGs (odds ratio 1.37, 95% CI 1.15-1.63).

Conclusions: Even limited closure of CBO PrEP providers meaningfully reduces geographic access to HIV prevention services, disproportionately affecting communities already experiencing transportation and HIV-related vulnerabilities. Sustained federal investment in CBOs is essential to preserve equitable PrEP access and prevent avoidable HIV infections.

(JMIR Public Health Surveill 2026;12:e89473) doi:10.2196/89473

KEYWORDS

HIV prevention; pre-exposure prophylaxis; community health centers; health services; geographic access; federal

XSL•FO
RenderX

Introduction

We have arrived at a point in the HIV epidemic in the United States where we arguably have all of the biomedical tools we need to end the epidemic [1]. There are multiple options for HIV screening, including self-testing [2]; multiple formulations and regimens for pre-exposure prophylaxis (PrEP) to prevent HIV infection; and effective treatments that result in full healthy lives for people living with HIV and, for people living with HIV who are virally suppressed, eliminating the risk of onward transmission through sex [3]. Despite these biomedical advances, annual HIV transmissions in the United States remain above 30,000 [4].

The realties of bringing these foundational tools to scale in a coordinated response has proven to be a substantial challenge of implementation and has resulted in increased attention to and funding for implementation science approaches [5-7]. Geographic access to HIV prevention services, including PrEP, is foundational to the "Prevent" pillar of the National HIV/AIDS Strategy [8]. We have previously explored geographic access to HIV prevention and care issues during a time of more certainty and consistency about the funding mechanisms for HIV prevention and care [9-11]. We have found that inequities in geographic access to HIV prevention and care services, and often, commute times disproportionately impact communities of color for whom car ownership may be less [12,13] and for whom public transportation access and route frequency may be more limited [14,15].

As we consider the structural and socioeconomic factors that impact geographic access to HIV prevention services, we are now in an era where federal funding for HIV services is declining or under threats of future reduction or elimination [16-19]. HIV prevention services in the United States are supported through a patchwork of federal funding mechanisms, including direct service delivery grants from the Centers for Disease Control and Prevention, funds administered through the Health Resources and Services Administration like the Ryan White HIV/AIDS Program, the Ending the HIV Epidemic initiative, Title X, Medicaid reimbursement for clinical PrEP services, and time-limited federal demonstration and implementation grants. These funding streams vary in stability, allowable services, and eligibility criteria, creating differential vulnerability to funding reductions across delivery settings. Although most PrEP prescriptions in the United States are provided through traditional health care clinics, community-based organizations (CBOs) contribute disproportionately to PrEP access among marginalized populations through direct provision, partnerships with prescribing clinicians, and intensive navigation services [20]. Given that many CBOs that provide HIV prevention services are supported, at least in part, by federal funding, it is important to consider how reductions in federal funding that lead to the closure of existing prevention providers might impact geographic access to HIV prevention services. Further, given the disparities in new HIV diagnoses related to acquisitions risks (eg, men who have sex with men), racial/ethnic minority groups (eg, Black and Hispanic or Latinx people) and geographic region (eg, the US South), it is important to assess whether gaps in geographic access to prevention services are differential by these population characteristics. To address these gaps, we used public data about the locations of HIV PrEP providers and previously described methods for estimating commute times to HIV prevention care services [10] to model the potential impact of federal funding cuts on geographic access to PrEP in Atlanta, Georgia.

Methods

We examined how the closing of CBOs that provide HIV prevention services may impact access to PrEP by both car and public transit across the four counties that encompass metro-Atlanta: DeKalb, Cobb, Fulton, and Gwinnett (all of which are "Ending the HIV Epidemic"–prioritized counties for HIV prevention).

Ethical Considerations

This study used publicly available, deidentified census and clinic location data. No human participants were involved, and therefore, no reviews or approvals by an ethics committee or institutional review board were required.

Clinic Data and Closure Scenarios

We identified all PrEP-providing clinics in the metro-Atlanta area as of August 2025 by using the Centers for Disease Control and Prevention's National Prevention Information Network PrEP Locator directory [21]. Clinics in the database are screened for eligibility if they have at least one health care professional (eg, physician, nurse practitioner, physician assistant) who is qualified to prescribe PrEP and if they have confirmed that they actively prescribe it. In addition to geographic data, each clinic is categorized by its funding type (federally qualified health center, public health department, hospital, CBO, etc). Of the 71 clinics in metro-Atlanta, 12 were designated as a CBO. To simulate potential impacts of federal funding reductions that disproportionately affect CBOs, we created 3 separate closure scenarios. In each scenario, 3 CBOs (approximately 25% of all Atlanta-area CBO PrEP clinics) were randomly selected for closure, and analyses were rerun to estimate the resulting changes in geographic access.

Origin and Destination Specification

Census block groups (CBGs) served as the unit of analysis for travel time estimation. Origin coordinates were defined as the population-weighted centroid of each CBG, representing where residents are most likely to live. CBG-level sociodemographic data were obtained from the American Community Survey 5-year estimates [22]. CBGs were categorized by the racial or ethnic plurality of residents. Destination coordinates were the locations of PrEP-providing clinics identified above. This approach captures area-level variation in accessibility rather than individual travel patterns.

Travel Time Estimation

We estimated one-way travel times between each CBG centroid and the 10 nearest PrEP clinics by using the Google Maps Distance Matrix application programming interface accessed through R [9]. The application programming interface was used to calculate the shortest available route under two transportation

XSL•FO
RenderX

modes: public transit and private vehicle. For both modes, travel times were calculated across 3 weekdays (Tuesday, Friday, and Saturday) and 3 time points (8 AM, noon, and 3:30 PM).

The shortest one-way travel time for each day-time combination was selected and then averaged to provide one estimate per mode for each CBG for each closure scenario. CBGs with no available public transit route to any PrEP clinic, defined as a walking distance of more than 30 minutes to the nearest transit stop, were coded as having no transit access. Transit time estimates included walking time to any transit stops, waiting time, boarding/transfer time, and time in transit.

Statistical Analysis

We used descriptive statistics to describe mean one-way travel times to the nearest PrEP clinic before and after simulated closures, separately for car and public transit. The mean change in travel time (minutes) was averaged across the 3 closure scenarios and compared using a 2-sided paired *t* test. CBGs were dichotomized into those that saw increased travel times compared to no change for both car and public transit. We then used simple logistic regression to model associations between race/ethnicity and increased travel times. All analyses and visualizations were conducted in R software (version 4.3.2; R Foundation for Statistical Computing).

Results

Descriptive Results

A total of 2466 CBGs across metro-Atlanta were included in the analysis. More than half of the CBGs (n=1361, 55.2%) had a White plurality. An additional 943 (38.2%) CBGs had a Black plurality, 155 (6.3%) had a Hispanic/Latinx plurality, and only 2 (<1%) CBGs had a plurality of another race or ethnicity. Under baseline conditions, all CBGs had access to at least one PrEP-providing clinic by car within 30 minutes, while only 1027 (41.6%) CBGs had access via public transit within 30 minutes. A total of 567 (22.9%) CBGs did not have access to PrEP via public transit.

Public Transit Changes in Access

Across the 3 simulated CBO closure scenarios (N=2466), 732 (29.6%) CBGs experienced longer average public transit times to the nearest PrEP clinic compared with current access. These CBGs represented approximately 1,024,900 residents or about 27.8% of Atlanta's total population. The average change in one-way transit time was 1.2 (range 0.0-11.6) minutes, which was significantly longer than the baseline ($P<.001$). Seven CBGs lost access to a PrEP clinic or CBO via public transit in the modeled closure scenarios. Compared to CBGs with a White plurality, the odds of experiencing increased transit times under the CBO closure scenarios were significantly lower for CBGs with a Hispanic plurality (odds ratio [OR] 0.61, 95% CI 0.40-0.95) but no different for CBGs with a Black plurality (OR 0.99, 95% CI 0.81-1.20).

Drive Time Changes in Access

Across the 3 simulated CBO closure scenarios (N=2466), 1184 (48%) CBGs experienced longer average drive times to the nearest PrEP clinic compared with current access. These CBGs represented approximately 1,698,000 residents or about 46.4% of Atlanta's total population. The average change in one-way drive time was 0.5 (range 0.0-6.4) minutes, which was significantly longer than baseline ($P=.03$). Compared to CBGs with a White plurality, the odds of experiencing increased drive times under the CBO closure scenarios were significantly higher for CBGs with a Black plurality (OR 1.37, 95% CI 1.15-1.63) but no different for CBGs with a Hispanic plurality (OR 0.90, 95% CI 0.64-1.28).

Discussion

Our research explores the potential impact of federal funding cuts on geographic access to PrEP in Atlanta, Georgia. Whereas previous research in this area was proposed to identify geographic areas within cities that would be important locational targets for new services to decrease travel times to care [23-25], we argue that this expansion-oriented framing must be complemented by models of the closing of existing facilities due to cuts in HIV prevention funding to better understand PrEP service availability and the risk of new HIV infections. Our data suggest that increases in travel times related to the closure of community providers were greater for people who used public transit, underscoring the importance of evaluating modal dimensions of transportation time to health care facilities. This is especially relevant given that communities most vulnerable to HIV are also most likely to rely on public transportation [12,26].

Although the differences in travel times might seem modest in some settings, such differences represent meaningful barriers to seeking care among marginalized groups already facing transportation disadvantages. A given level of change in commute time may result in different willingness to travel for people who use different modes of transportation or for people who do not have employment benefits that include paid time for seeking health care [26,27]. It is also important to acknowledge that HIV prevention service providers have both located their facilities in areas with high needs for prevention services, and clinics that provide services in areas with fewer community and economic resources are often set up to meet multiple care and social service needs [28]. Colocation of PrEP services with other health and social services may facilitate engagement in prevention by reducing logistical barriers, normalizing HIV prevention within broader care settings, and leveraging established trust between providers and communities. Such organizations spend decades building trusting relationships with communities, fostering trusting relationships [29,30]. Depending on which facilities close, it is foreseeable that populations with multiple medical and social service needs might lose access to a variety of services that would increase their vulnerability to HIV and other health threats.

Other aspects of our analysis reinforce the stakes in terms of health and equity considerations that would attend closures of existing PrEP-prescribing organizations. Although any licensed medical provider can provide PrEP, CBOs are disproportionately located in high-need areas and serve populations often excluded from traditional health care systems [31]. Thus, the loss of such facilities would create a double impact on PrEP accessibility:

XSL•FO
RenderX

it might create a longer commute time to the nearest PrEP provider, and it might mean that the nearest PrEP provider would be less likely to have culturally competent services [32-34]. Further, it is unclear whether existing clinics that survive the closure would have sufficient clinical capacity to provide PrEP to more patients even if prospective PrEP users are able to commute longer distances to continue to access PrEP [35,36]. Lastly, while our analysis focused on neighborhood level demographics, we know that PrEP-seeking patterns differ across multiple identity groups. For example, White men who have sex with men are more likely to access PrEP though traditional health care systems [37], while Black cisgender women and other groups with high HIV transmission tend to rely more often on community-based and safety-net providers [38]. Therefore, the loss of CBO-based PrEP services may disproportionately impact more marginalized groups.

The eventual impact of closures on the coverage of PrEP and the impacts of limitations on PrEP services are not easy to predict with confidence. However, it is clear that there are limitations to how far people will drive to obtain services for nonemergent health problems (eg, preventive services); several publications suggest that 30 minutes is a common threshold of willingness [39,40]. In any case, distance to receipt of care has an important and converse relationships with receipt of nonurgent health care, including preventive care [41]. A recent analysis suggests that even small decreases in PrEP coverage among people with indications could have substantial impacts on new HIV infections and health care costs. For example, just a 3% reduction in PrEP coverage is forecast to result in over 8000 preventable infections, with lifetime medical costs for HIV care exceeding US $3.6 billion for those infections [42].

Our findings, while empirically derived and reflective of real-world challenges to transportation to PrEP care services, have several limitations. First, we present one-way estimates of transit times, so the actual additional transportation burden to seeking PrEP care would be double the one-way estimates presented in the analysis. Simulated closures may not reflect actual likelihoods of closure; in the absence of empiric, publicly accessible data about organization-specific funding portfolios, we were unable to weight closures by likelihood and instead selected organizations randomly in the simulation. Future work incorporating such data could identify especially high-risk providers and communities. Conversely, we did not consider the likelihood of other types of organizations that provide PrEP closing; so, our estimates of impact on PrEP service might not reflect the true additional travel burden associated with the closure of facilities. Further, our analysis did not collect new data on, or consider the service capacity of, clinics; so, even if potential PrEP users were willing to travel longer distances, it is possible that existing clinics might not be able to handle an increase in the patient load.

According to our data, the closure of even a small number of current PrEP providers would have important impacts on physical access to HIV prevention services and would be predicted to increase commute times, decrease engagement in PrEP care, and result in avoidable new HIV infections and associated care costs. Therefore, our findings support sustained investment in CBOs as critical access points for HIV prevention services.

Funding

This work was supported by the Center for AIDS Research at Emory University (P30 AI050409) and the National Institutes of Health (R01AI181799, 5R01MH122341).

Authors' Contributions

Conceptualization: PSS, NM
Methodology: NM
Formal analysis: NM
Supervision: PSS
Funding acquisition: PSS
Writing - review & editing: NM, PSS
Writing - original draft: NM

Conflicts of Interest

PSS reports a relationship with Gilead Sciences Inc that includes consulting or advisory, funding grants, and speaking and lecture fees; with Merck & Co Inc that includes consulting or advisory and funding grants; with National Institutes of Health that includes consulting or advisory and funding grants; and with Centers for Disease Control and Prevention that includes funding grants and speaking and lecture fees.

References

1. Eisinger RW, Fauci AS. Ending the HIV/AIDS pandemic. Emerg Infect Dis 2018 Mar;24(3):413-416 [FREE Full text] [doi: 10.3201/eid2403.171797] [Medline: 29460740]
2. Hecht J, Sanchez T, Sullivan PS, DiNenno EA, Cramer N, Delaney KP. Increasing access to HIV testing through direct-to-consumer HIV self-test distribution - United States, March 31, 2020-March 30, 2021. MMWR Morb Mortal Wkly Rep 2021 Sep 24;70(38):1322-1325 [FREE Full text] [doi: 10.15585/mmwr.mm7038a2] [Medline: 34555001]

XSL•FO
RenderX

3. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, Kumarasamy N, HPTN 052 Study Team. Antiretroviral therapy for the prevention of HIV-1 transmission. N Engl J Med 2016 Sep 01;375(9):830-839 [FREE Full text] [doi: 10.1056/NEJMoa1600693] [Medline: 27424812]
4. HIV surveillance supplemental report: estimated HIV incidence and prevalence in the United States, 2018-2022. Centers for Disease Control and Prevention. 2024 May 21. URL: https://stacks.cdc.gov/view/cdc/156513 [accessed 2026-02-09]
5. Parker RG, Easton D, Klein CH. Structural barriers and facilitators in HIV prevention: a review of international research. AIDS 2000 Jun;14 Suppl 1:S22-S32. [doi: 10.1097/00002030-200006001-00004] [Medline: 10981471]
6. Beyrer C, Adimora AA, Hodder SL, Hopkins E, Millett G, Mon SHH, et al. Call to action: how can the US Ending the HIV Epidemic initiative succeed? The Lancet 2021 Mar 20;397(10279):1151-1156. [doi: 10.1016/s0140-6736(21)00390-1] [Medline: 33617770]
7. Proctor EK, Bunger AC, Lengnick-Hall R, Gerke DR, Martin JK, Phillips RJ, et al. Ten years of implementation outcomes research: a scoping review. Implement Sci 2023 Jul 25;18(1):31 [FREE Full text] [doi: 10.1186/s13012-023-01286-z] [Medline: 37491242]
8. Updated to 2020 - indicator supplement (2016). National HIV/AIDS Strategy (NHAS). URL: https://files.hiv.gov/s3fs-public/nhas-indicators-supplement-dec-2016.pdf [accessed 2026-02-09]
9. Mancuso N, Sullivan PS. Methods for estimating public transit travel times to healthcare services as a measure of equitable healthcare access. Ann Epidemiol 2025 Nov;111:24-29 [FREE Full text] [doi: 10.1016/j.annepidem.2025.09.010] [Medline: 40967351]
10. Mancuso N, Weiss K, Siegler AJ, Sullivan PS. Changes in geographic drive time access to pre-exposure prophylaxis (PrEP) in the United States, 2017 to 2025. Prev Med Rep 2026 Jan;61:103367 [FREE Full text] [doi: 10.1016/j.pmedr.2025.103367] [Medline: 41550475]
11. Luan H, Ransome Y, Taggart T, Chow J, Mancuso N, Grant M, et al. A comparison of HIV pre-exposure prophylaxis accessibility by public transit- and drive-time in Dallas-Fort Worth, Texas, 2024. Health Place 2025 Nov;96:103550. [doi: 10.1016/j.healthplace.2025.103550] [Medline: 40986968]
12. Dasgupta S, Kramer MR, Rosenberg ES, Sanchez TH, Reed L, Sullivan PS. The effect of commuting patterns on HIV care attendance among men who have sex with men (MSM) in Atlanta, Georgia. JMIR Public Health Surveill 2015;1(2):e10 [FREE Full text] [doi: 10.2196/publichealth.4525] [Medline: 27227128]
13. Goswami ND, Schmitz MM, Sanchez T, Dasgupta S, Sullivan P, Cooper H, et al. Understanding local spatial variation along the care continuum: the potential impact of transportation vulnerability on hiv linkage to care and viral suppression in high-poverty areas, Atlanta, Georgia. J Acquir Immune Defic Syndr 2016 May 01;72(1):65-72 [FREE Full text] [doi: 10.1097/QAI.0000000000000914] [Medline: 26630673]
14. Anderson KF, Galaskiewicz J. Racial/ethnic residential segregation, socioeconomic inequality, and job accessibility by public transportation networks in the United States. Spat Demogr 2021 Aug 06;9(3):341-373. [doi: 10.1007/s40980-021-00093-8]
15. Borowski E, Ermagun A, Levinson D. Disparity of access: variations in transit service by race, ethnicity, income, and auto availability. 2019 Jan 13 Presented at: Transportation Research Board 98th Annual Meeting; 2019-01-13; Washington DC, United States URL: https://trid.trb.org/View/1572981
16. Chan PA, Nunn AS. Grant terminations in the US threaten HIV research and future progress. JAMA 2025 Sep 16;334(11):949-950. [doi: 10.1001/jama.2025.12688] [Medline: 40839363]
17. Elendu C, Amaechi DC, Elendu TC, Amaechi EC, Elendu ID, Akpa KN, et al. Shaping sustainable paths for HIV/AIDS funding: a review and reminder. Ann Med Surg (Lond) 2025 Mar;87(3):1415-1445. [doi: 10.1097/MS9.0000000000002976] [Medline: 40213211]
18. McClure M, Gandhi M. Threat of HIV and tuberculosis drug resistance after US funding cuts. The Lancet Infectious Diseases 2025 May;25(5):e256-e257. [doi: 10.1016/s1473-3099(25)00209-9]
19. Moore C, Berl T. Sustaining the fight: maintaining HIV service funding in Delaware. Dela J Public Health 2025 Apr;11(1):98-99. [doi: 10.32481/djph.2025.04.15] [Medline: 40331172]
20. Smith DK, Maier E, Betts J, Gray S, Kolodziejski B, Hoover KW. What community-based HIV prevention organizations say about their role in biomedical HIV prevention. AIDS Educ Prev 2016 Oct;28(5):426-439. [doi: 10.1521/aeap.2016.28.5.426] [Medline: 27710082]
21. Siegler AJ, Wirtz S, Weber S, Sullivan PS. Developing a web-based geolocated directory of HIV pre-exposure prophylaxis-providing clinics: The PrEP locator protocol and operating procedures. JMIR Public Health Surveill 2017 Sep 06;3(3):e58 [FREE Full text] [doi: 10.2196/publichealth.7902] [Medline: 28877865]
22. American community survey 2016-2020 5-year data release. US Census Bureau. URL: https://www.census.gov/newsroom/press-kits/2021/acs-5-year.html [accessed 2026-10-04]
23. Harrington KRV, Hamilton C, Alohan DI, Hudson A, Young HN, Crawford ND. The PrEP pharmacy reach study: protocol for the creation of maps to visualize the impact of expanding access to HIV prevention services through pharmacies. JMIR Public Health Surveill 2025 Oct 27;11:e75077 [FREE Full text] [doi: 10.2196/75077] [Medline: 41144689]

XSL•FO
RenderX

24. Aral SO, Torrone E, Bernstein K. Geographical targeting to improve progression through the sexually transmitted infection/HIV treatment continua in different populations. Curr Opin HIV AIDS 2015 Nov;10(6):477-482 [FREE Full text] [doi: 10.1097/COH.0000000000000195] [Medline: 26352392]
25. Kedziora DJ, Stuart RM, Pearson J, Latypov A, Dierst-Davies R, Duda M, et al. Optimal allocation of HIV resources among geographical regions. BMC Public Health 2019 Nov 12;19(1):1509 [FREE Full text] [doi: 10.1186/s12889-019-7681-5] [Medline: 31718603]
26. Kabayundo J, Ahuja M, Jadhav S, Smith LC, Bithi FA, Shahi S, et al. Transportation barriers to healthcare access: a scoping review of measurement approaches and associated health outcomes. Transportation Research Interdisciplinary Perspectives 2025 Nov;34:101697. [doi: 10.1016/j.trip.2025.101697]
27. Harrison SE, Hung P, Green K, Miller SJ, Paton M, Ahuja D, et al. Does travel time matter?: predictors of transportation vulnerability and access to HIV care among people living with HIV in South Carolina. BMC Public Health 2025 Mar 08;25(1):926 [FREE Full text] [doi: 10.1186/s12889-025-22090-y] [Medline: 40057737]
28. Robillard AG, Julious CH, Smallwood SW, Douglas M, Gaddist BW, Singleton T. Structural inequities, HIV community-based organizations, and the end of the HIV epidemic. Am J Public Health 2022 Mar;112(3):417-425. [doi: 10.2105/ajph.2021.306688]
29. Elkins JS, Dale I, Okumu E, Zinck M, Golin C, McCrimmon J, et al. Survival, trust, and compassion: persistent HIV prevention concerns faced by Black sexual and gender minority communities in the US South - a qualitative study. Front Public Health 2025;13:1695474 [FREE Full text] [doi: 10.3389/fpubh.2025.1695474] [Medline: 41323625]
30. Corbie-Smith G, Isler MR, Miles MS, Banks B. Community-based HIV clinical trials: an integrated approach in underserved, rural, minority communities. Prog Community Health Partnersh 2012;6(2):121-129 [FREE Full text] [doi: 10.1353/cpr.2012.0023] [Medline: 22820222]
31. Mantell JE, Bauman LJ, Bonett S, Buchbinder S, Hoffman S, Storholm ED, et al. Innovation in Providing Equitable Pre-exposure Prophylaxis Services in the United States: Expanding Access in Nontraditional Settings. J Acquir Immune Defic Syndr 2025 Apr 15;98(5S):e156-e169. [doi: 10.1097/QAI.0000000000003610] [Medline: 40163068]
32. Gómez W, Gomez AM, Solis S, Dimonte C, Organista KC. Provider perspectives on multi-level barriers and facilitators to PrEP access among Latinx sexual and gender minorities. J Racial Ethn Health Disparities 2025 Apr;12(2):1139-1149. [doi: 10.1007/s40615-024-01948-8] [Medline: 38381327]
33. Ross J, Betancourt GS, Andrade EA, Klein A, Marrero L, Morales GA, et al. Collaborative PrEP implementation strategies for Latino men who have sex with men: a health center-community consensus process. J Community Health 2023 Dec;48(6):994-1003 [FREE Full text] [doi: 10.1007/s10900-023-01266-w] [Medline: 37548892]
34. Self KJ, Johnson A, Craker L, Silvey R, Fallon S, Cunningham SR, et al. Strengthening PrEP services at community-based organizations for Latinx men who have sex with men: an implementation science approach. Arch Public Health 2025 Mar 20;83(1):72 [FREE Full text] [doi: 10.1186/s13690-025-01560-0] [Medline: 40114286]
35. Agovi AM, Anikpo I, Cvitanovich MJ, Craten KJ, Asuelime EO, Ojha RP. Knowledge needs for implementing HIV pre-exposure prophylaxis among primary care providers in a safety-net health system. Prev Med Rep 2020 Dec;20:101266 [FREE Full text] [doi: 10.1016/j.pmedr.2020.101266] [Medline: 33364148]
36. Ard KL, Uzoeghelu U, Bruno J, Lambert C, Mayer KH, Davis JA, et al. Readiness of US federally qualified health centers to provide HIV pre-exposure prophylaxis. Open Forum Infect Dis 2021 Oct;8(10):ofab447 [FREE Full text] [doi: 10.1093/ofid/ofab447] [Medline: 34738023]
37. Kanny D, Jeffries W, Chapin-Bardales J, Denning P, Cha S, Finlayson T, et al. Racial/Ethnic Disparities in HIV Preexposure Prophylaxis Among Men Who Have Sex with Men — 23 Urban Areas. MMWR Morb Mortal Wkly Rep. Sep 20, 2019 2017;68(37):801-806 [FREE Full text]
38. Boudreaux J, Valdebenito CM, Pichon LC. Identifying access barriers to PrEP among cisgender black/African American women in the United States: a systematic review of the literature. Healthcare (Basel) 2025 Jan 06;13(1):86 [FREE Full text] [doi: 10.3390/healthcare13010086] [Medline: 39791693]
39. Yen W. How long and how far do adults travel and will adults travel for primary care? Washington State Office of Financial Management. 2013 Apr. URL: https://ofm.wa.gov/wp-content/uploads/sites/default/files/public/legacy/researchbriefs/2013/brief070.pdf [accessed 2025-11-04]
40. Baldomero AK, Kunisaki KM, Wendt CH, Bangerter A, Diem SJ, Ensrud KE, et al. Drive time and receipt of guideline-recommended screening, diagnosis, and treatment. JAMA Netw Open 2022 Nov 01;5(11):e2240290 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.40290] [Medline: 36331503]
41. Koob C, Babatunde A, Baxter SLK, Parisi M, Griffin SF. Examining primary care utilization and distance traveled for care among Medicaid-insured children in South Carolina. J Community Health 2025 Dec;50(6):1029-1036. [doi: 10.1007/s10900-025-01492-4] [Medline: 40493317]
42. Sullivan PS, Wall KM, Juhasz M, DuBose S, Crowley JS, Breyer C, et al. Excess HIV infections and costs associated with reductions in HIV prevention services in the US. JAMA Netw Open 2025 Sep 02;8(9):e2531341 [FREE Full text] [doi: 10.1001/jamanetworkopen.2025.31341] [Medline: 40932715]

XSL•FO
RenderX

Abbreviations

CBG: census block group
CBO: community-based organization
OR: odds ratio
PrEP: pre-exposure prophylaxis

Edited by A Mavragani; submitted 12.Dec.2025; peer-reviewed by KM Anderson, H Luan; comments to author 06.Jan.2026; revised version received 25.Jan.2026; accepted 28.Jan.2026; published 13.Feb.2026.

Please cite as:
Mancuso N, Sullivan PS
The Potential Impact of Federal Funding Cuts on Access to Pre-Exposure Prophylaxis in Atlanta, Georgia: Geographic Modeling Study
JMIR Public Health Surveill 2026;12:e89473
URL: https://publichealth.jmir.org/2026/1/e89473
doi:10.2196/89473
PMID:

©Noah Mancuso, Patrick S Sullivan. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 13.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Gestational Hypertension as a Mediator of Prenatal Ozone Exposure and Term Low Birth Weight: Birth Cohort Study

Yuxin Zhang[1,2], BSc; Quanjun Liu[1,2], BSc; Mengwei Song[1,2], BSc; Qiulin Huang[1,2], BSc; Qingbo Zhao[2], BSc; Qingqing Tao[3], BSc; Chenchen Zhang[3], MMed; Qin Li[3], PhD; Qing Wang[1,2,4], PhD

[1]Department of Biostatistics, School of Public Health, Cheeloo College of Medicine, Shandong University, No.44 Wenhuaxi Road, Jinan, China

[2]National Institute of Health Data Science of China, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan, China

[3]Department of Maternal and Child Health, School of Public Health, Peking University, Beijing, China

[4]NHC Key Lab of Health Economics and Policy Research, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan, China

Corresponding Author:
Qing Wang, PhD
Department of Biostatistics, School of Public Health, Cheeloo College of Medicine, Shandong University, No.44 Wenhuaxi Road, Jinan, China

Abstract

Background: Ambient ozone (O_3) exposure has been found to be associated with gestational hypertension, which, in turn, increases the risk of term low birth weight (LBW). As such, gestational hypertension acts as a potential mechanism mediating restricted fetal growth; however, few epidemiological studies have quantified this specific mediation pathway.

Objective: This study aims to examine whether gestational hypertension serves as a mediator of the association between prenatal O_3 exposure and term LBW.

Methods: We conducted a population-based cohort study using the Cheeloo Lifespan Electronic Health Research Data-library, including 3,394,739 singleton term live births in Shandong Province, China, from January 1, 2016, to December 31, 2022. We used high-resolution spatiotemporal models based on residential addresses for exposure assessment. In addition to term LBW, we examined term small for gestational age (SGA) to capture fetal growth restriction while accounting for gestational age at birth. Given the low prevalence of these outcomes, we used logistic regression models where odds ratios approximated relative risks. A 4-step mediation analysis using logistic regression was conducted, followed by a counterfactual-based causal mediation analysis, to test the mediating role of gestational hypertension.

Results: The mean (SD) O_3 concentration was 113.90 (13.03) μg m^{-3}. Each IQR increase in O_3 was positively associated with the risks of term LBW (relative risk 1.055, 95% CI 1.034-1.077) and term SGA (relative risk 1.037, 95% CI 1.026-1.048). Using the traditional approach, gestational hypertension mediated 19.94% of the risk for term LBW and 13.41% for term SGA. Under the counterfactual framework, the contribution rates were 38.82% (term LBW) and 19.96% (term SGA) when excluding exposure-mediator interaction, and 35.15% (term LBW) and 18.82% (term SGA) when accounting for such interaction.

Conclusions: Our findings showed that gestational hypertension was a significant mediator of the association between O_3 exposure and risks of term LBW. Consequently, a multitiered strategy—encompassing stricter air quality standards, integrating O_3 risk education into routine prenatal care, and taking proactive measures to minimize personal exposure—is essential to prevent potential adverse impacts on developing fetuses and mothers.

(JMIR Public Health Surveill 2026;12:e81412) doi:10.2196/81412

KEYWORDS

ozone exposure; low birth weight; small for gestational age; gestational hypertension; causal mediation

Introduction

Low birth weight (LBW) is one of the major markers of fetal growth and stands as a predominant factor contributing to neonatal mortality and morbidity [1-3]. LBW infants are about 20 times more likely to die than heavier infants. Additionally, LBW is associated with adverse consequences in later life, such as neurobehavioral issues and respiratory, cardiovascular, cerebrovascular, and metabolic diseases [4-7]. LBW, which brings about a considerable disease burden and high medical costs, has presented significant challenges to health care systems. Consequently, it has been incorporated into the Sustainable Development Goals [8]. Nevertheless, despite years of strenuous efforts, merely a slight reduction in the prevalence of LBW has been noted [9], and factors such as multifetal pregnancy and maternal malnutrition, while recognized

XSL•FO
RenderX

contributors, cannot fully account for the persistent prevalence of fetal growth restriction leading to LBW.

In recent years, regions with a high prevalence of LBW often face severe air pollution. A series of investigations noted that exposure to air pollution, such as particulate matter (PM), could plausibly be involved in LBW via higher oxidative stress levels [10,11]. Compared to PM, ozone (O_3) is a pollutant endowed with more potent oxidative stress-inducing properties and is more challenging to mitigate through individual behavioral measures [12]. Thus, it is sensible to postulate that O_3 exposure might potentiate susceptibility to LBW. However, little is known about the association between O_3 exposure during pregnancy and the risk of LBW. Five studies have reported a positive and significant association between O_3 exposure and LBW [13-17], whereas findings from another 5 studies were inconsistent or mixed [18-22]. Exposure misclassification might, to some extent, account for the variability in the results, which stems from regionally specific evaluations of pollution exposure [13-15,20,22]. Several recent studies further used satellite retrieval data to attenuate exposure misclassification to a certain degree [16,18,19,21]. However, most of these studies failed to account for several fundamental covariates such as gestational temperature exposure, maternal BMI, and the smoking status of the husband, which may lead to biased results [13,14,16,18-21]. Thus, more information may be needed in future studies.

Another pressing question is through which mechanisms O_3 exposure during pregnancy affects fetal development in utero. Unlike acute and abruptly occurring conditions, the influence of O_3 on fetal development unfolds gradually throughout pregnancy. In this chronic process, O_3 can slow the development of the early placental vasculature and ultimately may promote maternal cardiovascular abnormalities later in gestation or intrauterine growth deficiency [23]. A wide range of studies has demonstrated that air pollution exposure could be involved in the development of hypertension through oxidative stress [24,25]. Three empirical studies focusing on the relationship between O_3 exposure and the risk of gestational hypertension achieved similar results [26-28]. Gestational hypertension is evidenced by a reduction in placental perfusion [29,30], and the occurrence of intravascular coagulation [31,32], which is considered to precipitate placental dysfunction and subsequently lead to LBW [33]. Given the pairwise associations, we propose that gestational hypertension may mediate the associations between O_3 and the risks of LBW and small for gestational age (SGA), which has yet to be studied.

To address the above gaps, our study estimated the independent effects of maternal O_3 exposure on term LBW and SGA based on a broad and highly representative sample, while controlling for a range of covariates. Additionally, we examined, for the first time, the mediating effects of gestational hypertension to provide evidence regarding the underlying mechanisms.

Methods

Study Population

This is a retrospective cohort study based on the extended version of the Cheeloo Lifespan Electronic Health Research Data-library in Shandong Province, China (Multimedia Appendix 1). Shandong is a coastal province in eastern China (34°22.9′–38°24.01′N, 114°47.5′–122°42.3′E) with approximately 100 million inhabitants and substantial geographic and socioeconomic heterogeneity across coastal and inland areas. The foundation of our data infrastructure is the Shandong Province Birth Certificate System, a national government–administered registry that integrates birth information reported by medical institutions at multiple levels across the province. This system captures standardized demographic and administrative information for nearly all newborns and their parents, including place of residence (eg, municipality or township of residence). Additional data sources include linked electronic medical records (EMRs) and electronic health records, which provide complementary information on maternal health conditions, clinical encounters, and other relevant characteristics. A unique deidentified personal identification code is systematically applied across all databases; therefore, linkage of these records enables the construction of a large population-based birth cohort. During the study period, birth records were continuously available from 16 cities and 1840 township-level administrative units across the province.

We initially identified 7,141,641 live births recorded in Shandong Province between January 1, 2016, and December 31, 2022. Among these, 6,099,731 were singleton live births. Neonatal information was routinely verified by designated quality-control personnel at delivery institutions, and data underwent consistency and range checks, with obvious logic inconsistencies corrected when possible. To improve comparability and focus on a relatively homogeneous population, we restricted the cohort to singleton firstborn live births among women aged 20 to 45 years and further limited analyses to term births (gestational age 37 - 45 wk) to reduce confounding related to prematurity, resulting in a total of 5,720,445 eligible births. We excluded records with incomplete hospitalization records or residential address information (n=2,130,147), unmarried women (n=125,465), missing O_3 exposure (n=70,066), and implausible birthweight values (n=28). The final analytical sample included 3,394,739 births (Multimedia Appendix 2).

Term LBW and SGA

Our key interests were term LBW and SGA. Term LBW was defined as birth weight less than 2500 g and gestational age greater than 37 weeks. In addition, term SGA was used in the analysis because it takes gestational length into consideration. Term SGA was determined as birth weight below the 10th percentile and gestational age greater than 37 weeks according to the INTERGROWTH-21st (International Fetal and Newborn Growth Consortium for the 21st Century) fetal growth standards [34].

XSL•FO
RenderX

O_3 Exposure

Daily concentrations of O_3, with a spatial resolution of 1 km × 1 km, were sourced from the CHAP (China High Air Pollutants) dataset. The ground-level 8-hour maximum daily average O_3 concentration was predicted through an ensemble learning approach using the space-time extremely randomized trees model, which incorporated solar radiation intensity, surface temperature, remote sensing products, and ground-based observations. The 10-fold cross-validation results for O_3 demonstrated a coefficient of determination (R^2) of 0.87 and a root mean square error of 17.10. Detailed information about the prediction is available elsewhere [35]. Address information for each participant was gathered and geocoded into longitude and latitude coordinates. We calculated the O_3 exposure levels throughout the entire pregnancy for each subject based on their home address, the date of their last menstrual period, and the date of delivery. We also calculated pregnancy exposure to PM≤2.5 µm in aerodynamic diameter ($PM_{2.5}$), PM≤10 µm in aerodynamic diameter (PM_{10}), nitrogen dioxide (NO_2), sulfur dioxide (SO_2), and carbon monoxide (CO) concentrations using similar methods.

Gestational Hypertension

Gestational hypertension was identified based on the *International Classification of Diseases, Tenth Revision* (ICD-10) [36] codes documented in integrated outpatient and inpatient medical records throughout pregnancy, which were derived from EMR. Specifically, the diagnosis was determined using ICD-10 diagnostic codes O13 (gestational [pregnancy-induced] hypertension), O14 (pre-eclampsia), and O15 (eclampsia), as well as the keyword "hypertensive disorders of pregnancy." In addition, as a nationally implemented program, China's National Essential Public Health Services includes standardized maternal follow-up management, during which pregnant women are routinely assessed for pregnancy complications such as gestational hypertension. This framework may support more complete case ascertainment across health care settings, including resource-limited areas, and help mitigate potential under-ascertainment and information loss.

Covariates

Using a directed acyclic graph (Multimedia Appendix 3) informed by the literature [37], we selected confounders that are associated with both O_3 exposure and birth outcomes and that, when adjusted for, remove all backdoor paths between the exposure and the outcome while minimizing overadjustment [38]. The models were adjusted for maternal age, maternal occupation (such as farmer, clerk, housewife, worker, and other categories), infant sex, gestational diabetes, and the smoking status of the husband (the prevalence of active smoking among pregnant women in China is exceptionally low) [13,18]. In addition, given the close association between temperature and O_3 concentration, temperature exposure during pregnancy was included. Monthly ambient temperature was obtained from the Global Surface Summary of the Month data files provided by the National Oceanic and Atmospheric Administration (NOAA), with a spatial resolution of 10 km × 10 km [39].

Ethical Considerations

This study was conducted in accordance with the Declaration of Helsinki and was approved by the ethics committee of the Shandong University School of Public Health (LL20240334). The requirement for informed consent was waived because all data were anonymized.

Statistical Analysis

We used logistic regression models to estimate the associations of an IQR increase in O_3 concentration with term LBW and SGA. Given the low prevalence of these outcomes in our study population (typically <10%), the odds ratios (ORs) derived from logistic regression were considered reasonable approximations of relative risks (RRs) [40]. Consequently, we reported the results as RRs and 95% CIs. To ensure the robustness of our findings, 3 sensitivity analyses were conducted. First, Cox proportional hazards models were fitted using gestational age (wk) as the time scale to incorporate time-to-event data. Second, 2-pollutant models were established by adjusting for $PM_{2.5}$, PM_{10}, NO_2, SO_2, and CO to isolate the independent effect of O_3. Third, maternal BMI was added to control for nutritional status; missing BMI data were imputed using the "*missForest*" algorithm prior to analysis (Multimedia Appendix 4). Fourth, to address potential biases arising from (1) incomplete residential addresses or failed hospitalization record linkages and (2) nonrandom missingness of O_3 exposure data, and to improve the representativeness of the analytic sample, we used a weighted logistic regression model supplemented by robust variance estimation. We first derived calibration weights for records with complete residential address and hospitalization information. These weights were constructed by calibrating the distribution of key demographic variables (eg, maternal age, infant sex, and gestational age) in our complete-case sample to the eligible population, thereby adjusting for potential bias due to missingness in these essential identifiers [41]. Next, to account for nonrandom missingness in O_3 exposure assessment (eg, dependent on geographic location), we estimated inverse probability of selection weights (IPSWs) [42]. The IPSWs were derived from a logistic regression model predicting inclusion in the exposure-assessed analytic sample based on maternal age, infant sex, gestational diabetes, the smoking status of the husband, and maternal occupation. The final analytic weight for each participant was obtained by multiplying the calibration weight and the IPSW. To ensure numerical stability and facilitate variance estimation, the composite weights were normalized to sum to the actual sample size (ie, mean weight=1). Additional details are provided in Multimedia Appendix 5.

Stratified analyses were conducted based on (1) maternal age (<30 y vs ≥30 y, with the latter marking the initial decline in fecundity and ovarian function [43]; (2) infant sex (female vs male) [44]; and (3) coastal residence (coastal vs inland) to account for geographic environmental differences [45,46].

The hypothesis that the association of O_3 exposure with term LBW and SGA would be mediated by gestational hypertension was tested using a 4-step analysis. The method involved testing the total effect (TE) of O_3 exposure with term LBW and SGA and then estimating how much the association was reduced by

XSL•FO
RenderX

the inclusion of gestational hypertension. Using a logistic regression model, the TE, represented as the RR of term LBW or SGA linked to an IQR rise in O_3 exposure, was divided into direct and indirect effects in line with the methodology described by Buis [47]. The mediating role of gestational hypertension was quantified and evaluated within the framework of a Cox proportional hazards model.

To gain a deeper understanding of the underlying causal processes, a mediation approach based on the counterfactual framework for causal inference was adopted. In contrast to traditional methods, this framework defines mediation effects using counterfactual scenarios. This enables a more precise exploration of how the exposure and the potential mediator jointly affect the outcome [48,49]. In this counterfactual framework, the TE was defined as the RR of term LBW or SGA related to an IQR increase in O_3 exposure, comparing the 75th percentile to the 25th percentile. This was calculated while adjusting for covariates and allowing gestational hypertension to occur naturally. Following Pearl's [50] approach for the risk difference scale, the TE was expressed as the product of the ORs for the natural direct effect and the natural indirect effect (NIE) [49]. The natural direct effect was estimated by comparing O_3 exposure at the 75th percentile versus the 25th percentile, while keeping gestational hypertension at the level it would have been under the 25th percentile of O_3 exposure. The NIE was defined as the change in odds resulting from gestational hypertension when O_3 exposure increased by an IQR. It represents the part of the causal pathway from exposure to outcome that was mediated by gestational hypertension. The proportion mediated was computed as log(NIE)/log(TE), following the method of Imai et al [51]. The details of the analysis are presented in Multimedia Appendix 6.

All statistical analyses were conducted using R (version 4.4.0; R Foundation for Statistical Computing) and Stata (version 17.0, StataCorp LLC). Bias-corrected standard errors and 95% CIs were estimated using bootstrapping with 1000 replications. All statistical tests were 2-sided, with a significance level set at α=.05.

Results

Descriptive Results

Table 1 presents the descriptive characteristics of the study population. Among the 3,394,739 term births included in this study, 1,615,993 (47.6%) were female. The overall prevalence of term LBW and SGA infants was 0.89% and 3.34%, respectively. Notably, a higher incidence of both term LBW and SGA was observed among mothers diagnosed with gestational hypertension. Regarding the air pollution levels during the study period, the summary statistics are shown in Table 2. The mean (SD) O_3 concentration during pregnancy was 113.90 (13.03) μg m^{-3}. Furthermore, O_3 exposure was correlated with other pollutants, with the Pearson correlation coefficients ranging from –0.47 to 0.75.

XSL•FO
RenderX

Table . Characteristics of 3,394,739 singleton term live births included in the retrospective birth cohort study in Shandong Province, China, 2016 - 2022 (N=3,394,739).

Characteristics	Total (N=3,394,739)	Term LBW[a] (n=30,145)	Term SGA[b,c] (n=113,382)
Infant sex, n (%)			
Male	1,778,746 (52.4)	12,141 (40.3)	56,316 (49.7)
Female	1,615,993 (47.6)	18,004 (59.7)	57,066 (50.3)
Gestational age (wk), mean (SD)	39.0 (1.1)	38.0 (1.0)	39.1 (1.1)
Maternal age (y), mean (SD)	30.0 (4.8)	29.9 (5.3)	29.0 (5.0)
Maternal occupation, n (%)			
Clerk	563,149 (16.6)	4670 (15.5)	18,194 (16)
Farmer	1,121,893 (33)	10,371 (34.4)	39,293 (34.7)
Housewife	372,462 (11)	3682 (12.2)	12,951 (11.4)
Worker	220,077 (6.5)	1846 (6.1)	6713 (5.9)
Others	1,117,158 (32.9)	9576 (31.8)	36,231 (32)
Status of husband smoking, n (%)			
Yes	116,956 (3.4)	1123 (3.7)	3921 (3.5)
No	3,277,783 (96.6)	29,022 (96.3)	109,461 (96.5)
Maternal BMI (kg/m^2), mean (SD)[d]	22.4 (4.0)	22.0 (4.1)	21.6 (3.8)
Gestational diabetes, n (%)			
Yes	349,910 (10.3)	2489 (8.3)	7850 (6.9)
No	3,044,829 (89.7)	27,656 (91.7)	105,532 (93.1)
Gestational hypertension, n (%)			
Yes	103,443 (3.1)	4915 (16.3)	8315 (7.3)
No	3,291,296 (96.9)	25,230 (83.7)	105,067 (92.7)

[a]LBW: low birth weight.

[b]SGA: small for gestational age.

[c]The sample size of term SGA was 3,394,049.

[d]The sample size of maternal BMI was 1,556,480.

XSL•FO
RenderX

Table . Distribution and correlation of air pollutant concentrations and temperature during pregnancy among the study population in Shandong Province, China (2016 - 2022).

Exposure	Mean (SD)	Min-Max	Median (IQR),	O_3[a]	$PM_{2.5}$[a,b]	PM_{10}[a,c]	NO_2[a,d]	CO[a,e]	SO_2[a,f]	Temperature[a]
O_3 (μg/m^3)	113.90 (13.03)	62.35-157.58	112.70 (103.60-23.53)	1.00	–0.47	–0.36	–0.42	–0.42	–0.39	0.75
$PM_{2.5}$ (μg/m^3)	53.01 (12.37)	17.76-111.27	52.80 (44.22-61.11)	—[g]	1.00	0.96	0.80	0.81	0.73	–0.35
PM_{10} (μg/m^3)	97.25 (19.58)	35.03-191.31	97.37 (83.29-111.10)		—	1.00	0.81	0.77	0.72	–0.26
NO_2 (μg/m^3)	34.36 (6.37)	8.81-68.55	34.61 (30.04-38.93)			—	1.00	0.75	0.62	–0.43
CO (mg/m^3)	0.98 (0.23)	0.44-2.55	0.94 (0.82-1.09)				—	1.00	0.81	–0.28
SO_2 (μg/m^3)	18.98 (8.94)	5.60-92.77	15.97 (12.36-23.45)					—	1.00	–0.14
Temperature (°C)	14.58 (3.06)	2.04-22.12	14.48 (12.09-17.22)						—	1.00

[a]These are *r* values with *P*<.001.

[b]$PM_{2.5}$: particulate matter with a diameter of 2.5 μm or smaller.

[c]PM_{10}: particulate matter with diameter of 10 μm or smaller.

[d]NO_2: nitrogen dioxide.

[e]CO: carbon monoxide.

[f]SO_2: sulfur dioxide.

[g]Not applicable.

Association Between O_3 Exposure and Term LBW or SGA

Table 3 illustrates the associations of O_3 exposure with term LBW and SGA across various adjusted models. Based on logistic regression models, each IQR increase in O_3 concentration was significantly associated with elevated risks of term LBW (RR 1.055, 95% CI 1.034-1.077) and SGA (RR 1.037, 95% CI: 1.026-1.048). The median (IQR) O_3 concentration was 112.70 (19.93) μg/m^3. These results were corroborated by Cox proportional hazards models, yielding hazard ratios of 1.071 (95% CI 1.049-1.093) and 1.081 (95% CI 1.069-1.092) for term LBW and SGA, respectively. Sensitivity analyses adjusting for maternal BMI (original and imputed) and co-pollutants confirmed the robustness of our findings. The weighted model produced RRs consistent with the main model (term LBW: RR 1.055, 95% CI 1.032 - 1.077; term SGA: RR 1.033, 95% CI 1.022 - 1.044). Additionally, stratified analyses showed stable positive associations across subgroups of maternal age, infant sex, and coastal residence (Multimedia Appendix 7).

XSL•FO
RenderX

Table . Associations between maternal ozone exposure and the risks of term low birth weight (LBW) and term small for gestational age (SGA) among singleton term births in Shandong Province, China, 2016 - 2022[a].

Model	RR[b] (95% CI)	
	Term LBW (n=3,394,739)	Term SGA (n=3,394,049)
Main model[c]	1.055 (1.034-1.077)	1.037 (1.026-1.048)
Cox model[d]	1.071 (1.049-1.093)	1.081 (1.069-1.092)
Main model adjusted for maternal BMI[e]	1.051 (1.021-1.082)	1.033 (1.018-1.049)
Main model adjusted for imputed maternal BMI[f]	1.056 (1.034-1.078)	1.040 (1.029-1.051)
Main model adjusted for $PM_{2.5}$[g]	1.058 (1.034-1.082)	1.072 (1.060-1.085)
Main model adjusted for PM_{10}[h]	1.051 (1.028-1.074)	1.063 (1.051-1.075)
Main model adjusted for NO_2[i]	1.056 (1.033-1.079)	1.048 (1.036-1.060)
Main model adjusted for SO_2[j]	1.040 (1.017-1.063)	1.047 (1.035-1.059)
Main model adjusted for CO[k]	1.061 (1.038-1.085)	1.057 (1.045-1.069)
Main model adjusted for gestational hypertension	1.034 (1.013-1.055)	1.029 (1.019-1.040)
Weighted model [l]	1.055 (1.032-1.076)	1.033 (1.022-1.044)

[a]Effect estimates were expressed per IQR increase in O_3. The median (IQR) O_3 concentration was 112.70 (19.93) μg/m³.

[b]RR: relative risk.

[c]Main model: single-pollutant model for O_3 adjusted for maternal age, infant sex, temperature, maternal occupation, gestational diabetes, and the smoking status of the husband.

[d]Cox proportional hazards model (time scale: gestational age): adjusted for maternal age, infant sex, temperature, maternal occupation, gestational diabetes, and smoking status of the husband. Results are presented as hazard ratios (HRs) with 95% CIs.

[e]The sample size of maternal BMI was 1,556,480.

[f]Missing values were imputed using the missForest algorithm.

[g]$PM_{2.5}$: fine particulate matter with a diameter of 2.5 μm or smaller.

[h]PM_{10}: particulate matter with a diameter of 10 μm or smaller.

[i]NO_2: nitrogen dioxide.

[j]SO_2: sulfur dioxide.

[k]CO: carbon monoxide.

[l]Adjusted for maternal age, infant sex, gestational diabetes, maternal occupation, smoking status of the husband, and temperature category. Final weights were calculated as ranking weights multiplied by inverse probability of selection weight (IPSW) truncated at the 99th percentile.

Notably, while O_3 exposure remained significantly associated with adverse outcomes after adjusting for gestational hypertension, the effect estimates were slightly attenuated (term LBW: RR 1.034, 95% CI 1.013-1.055; term SGA: RR 1.029, 95% CI 1.019-1.040), suggesting a potential mediating pathway, as shown in Table 4.

XSL•FO
RenderX

Table . Mediation of gestational hypertension in the association between ozone exposure and term low birth weight (LBW) and term small for gestational age (SGA) in the Shandong birth cohort (2016 - 2022).

Mediation of gestational hypertension	Term LBW	Term SGA
Four-step analysis with the Sobel approach[a]		
Total effect, RR[b] (95% CI)	1.042 (1.021-1.064)	1.034 (1.023-1.045)
Direct effect, RR (95% CI)	1.034 (1.013-1.055)	1.029 (1.019-1.040)
Indirect effect, RR (95% CI)	1.008 (1.008-1.009)	1.005 (1.004-1.005)
Mediation effect (%)	19.94	13.41
Causal mediation[a] (excluding exposure-mediator interaction)		
Marginal total effect, RR (95% CI)	1.055 (1.034-1.077)	1.037 (1.026-1.048)
Natural direct effect, RR (95% CI)	1.034 (1.013-1.055)	1.029 (1.019-1.040)
Natural indirect effect, RR (95% CI)	1.021 (1.020-1.021)	1.007 (1.007-1.008)
Mediation effect (%)	38.82	19.96
Causal mediation[a] (including exposure-mediator interaction)		
Marginal total effect, RR (95% CI)	1.055 (1.033-1.076)	1.036 (1.025-1.047)
Natural direct effect, RR (95% CI)	1.035 (1.014-1.056)	1.029 (1.019-1.040)
Natural indirect effect, RR (95% CI)	1.019 (1.018-1.021)	1.007 (1.006-1.007)
Mediation effect (%)	35.15	18.82

[a]The use of the bootstrap was proposed to estimate standard errors.

[b]RR: relative risk.

The Mediation of Gestational Hypertension

Exposure to O_3 was correlated with an elevated risk of developing gestational hypertension, with an OR of 1.162 (95% CI 1.149-1.175). Additionally, individuals who experienced gestational hypertension were more likely to deliver babies with term LBW (RR 6.499, 95% CI 6.300-6.706) and term SGA (RR 2.767, 95% CI 2.703,2.832; Multimedia Appendix 8). Table 4 presents the indirect contribution of gestational hypertension in the association between O_3 exposure and term LBW or SGA. We found that gestational hypertension partly mediated these associations. Gestational hypertension accounted for 19.94% and 13.41% of the association between O_3 exposure and term LBW or SGA, respectively. Under the framework of Cox proportional hazards models, consistent results were observed, with mediation proportions of 31.28% for term LBW and 11.35% for term SGA (Multimedia Appendix 9).

Under the counterfactual framework, the contribution of gestational hypertension to the association between prenatal O_3 exposure and term LBW was 38.82% when excluding exposure-mediator interaction, and 35.15% when accounting for such interaction. For term SGA, the corresponding contribution rates were 19.96% (without interaction) and 18.82% (with interaction), respectively.

Discussion

Principal Findings

To our knowledge, this is the first study to investigate the mediation effects of gestational hypertension in the associations of O_3 exposure during pregnancy with term LBW and SGA. Based on a birth cohort of 3,394,739 samples in China, this study provides novel findings by demonstrating the independent effect of O_3 exposure on term LBW and SGA. Additionally, mediation analysis revealed that the proportion of the effect of O_3 exposure on term LBW and SGA mediated through gestational hypertension ranges from 13.41% to 19.94%. This study holds significant implications for the surveillance and management of maternal blood pressure during prenatal care and contributes to our comprehension of the underlying mechanisms regarding the associations of O_3 exposure with term LBW and SGA.

Associations Between Prenatal O_3 Exposure and Term LBW

Previous studies have revealed that PM exposure might impact fetal growth through elevated oxidative stress levels [10,11]. As another major ambient air pollutant, O_3, with more powerful oxidative stress-eliciting characteristics [12], may act indirectly on the fetus by inducing oxidative stress, inflammatory responses, and hormonal imbalances in the mother [52]. These maternal states, over time, may disrupt normal placental function and fetal development, ultimately resulting in LBW [23]. Our study indicated a similarly notable association between O_3 exposure during pregnancy and term LBW as well as term SGA. This is consistent with nationwide studies conducted in the United States with 2,179,040 participants and Iran with 4,030,383 participants [13,16]. However, some recent studies have reported conflicting findings [18-22], which may be

XSL•FO
RenderX

attributed to differences in study design and the statistical models used.

It is well-known that variability in exposure assessment approaches is an important source of variability in effect estimates between different studies. Unlike previous studies that estimated exposure based on data from fixed monitoring stations, we used O_3 concentrations retrieved from satellite remote sensing to calculate individual O_3 exposure levels throughout pregnancy. This approach enables a more reliable estimation of the impact of O_3 exposure on adverse birth outcomes. Additionally, by leveraging enriched indicators from EMRs and electronic health records, this study incorporated maternal and paternal information into the analytical models based on an extensive sample covering the period from 2016 to 2022. These factors have been demonstrated to potentially confound the associations under investigation [53]. A series of robust analyses further validates the reliability and significance of our results. Our findings provide justification for addressing air pollution-related adverse birth risk to achieve Sustainable Development Goals.

Gestational Hypertension Mediated Partly the Associations Between Prenatal O_3 Exposure and Term LBW

Fetal growth and development in utero are a continuous process, indicating that the impact of air pollution during pregnancy is likely to be long-term and chronic. Consequently, pregnancy complications and the associated physiological changes in the mother may play a crucial role in mediating these effects. Several studies have used effect modification to explore the mechanisms by which air pollution exposure affects birth outcomes, although the process through which air pollution exposure causes LBW remains elusive. Recent studies, including ours, have suggested that O_3 or $PM_{2.5}$ exposure in the first trimester is associated with gestational hypertension [26-28,54]. There is also evidence that gestational hypertension increases the risk of LBW [55,56]. Exposure to O_3 may not only contribute to gestational hypertension but may also generate intergenerational effects, resulting in term LBW and SGA. However, none of these studies have explored whether gestational hypertension could link O_3 exposure to adverse birth outcomes.

This present study found that gestational hypertension significantly mediated 13.41% and 19.94% of the association between O_3 exposure and term LBW or SGA, respectively. These findings provide strong evidence that gestational hypertension occurring in later pregnancy, along with associated placental perfusion abnormalities, significantly mediates the effect of O_3 exposure on fetal growth restriction. After entering the body, O_3 primarily affects the respiratory and circulatory systems, where it induces oxidative stress and disrupts vascular tone regulation [57]. Mechanistically, O_3-related oxidative stress and inflammation may impair endothelial function and activate vasoactive pathways, including the renin-angiotensin system (eg, angiotensin-converting enzyme/angiotensin II) and endothelin-1 signaling, thereby contributing to elevated blood pressure [58]. These changes may provoke enhanced sympathetic activity, thereby augmenting the risk of gestational hypertension [59]. Supporting evidence from pregnant animal models shows that mid-gestational O_3 exposure can induce a mild but significant increase in maternal arterial blood pressure [57]. Gestational hypertension may adversely affect placental blood supply, leading to suboptimal perfusion [31,32,60]. This impairment restricts the delivery of oxygen and nutrients to the fetus [30]. Ultimately, such deprivation significantly increases the risk of low birth weight. In addition, experimental studies indicate that prenatal O_3 exposure can increase circulating soluble fms-like tyrosine kinase-1 and elevate uterine artery resistance, accompanied by placental transcriptional changes consistent with preeclampsia-related pathways, providing biological plausibility for placental vascular maladaptation [61]. Beyond circulatory effects, O_3 exposure may contribute to fetal growth disturbances through maternal chronic inflammation. Some inflammatory mediators, including interleukin-17A, have been shown to cross the placental barrier and potentially influence fetal developmental regulation [62].

To further quantify this mediation pathway, causal mediation analysis was conducted. When allowing for exposure-mediator interaction, the mediation effects were 35.15% for term LBW and 18.82% for term SGA. These findings suggest that, beyond acting through independent primary pathways, O_3 exposure and gestational hypertension may also exert synergistic effects that exacerbate intrauterine fetal development impairments. Moreover, the mechanisms underlying these associations appear to be more complex than previously understood. Taken together, interventions or policies addressing O_3 exposure should coordinate the health risks for mothers and neonates. Efforts to reduce the risks of LBW should be tailored to focus on mothers with gestational hypertension, while recognizing early warning signs of O_3 exposure.

Collectively, we propose a multitiered strategy to mitigate the adverse effects of O_3 exposure. At the policy level, stricter enforcement of air quality standards and the development of pregnancy-specific health warning systems are essential. Within the health care system, O_3 risk education should be integrated into routine prenatal care, particularly for high-risk populations. At the individual level, we recommend that pregnant women monitor air quality indices, minimize outdoor activities during peak pollution, use indoor air purifiers, and maintain an antioxidant-rich diet. Synergistic efforts across these sectors are required to effectively protect maternal and fetal health.

Contributions and Limitations

This study enriches the growing body of literature in 2 aspects. First, it evaluates the associations between O_3 exposure and the health risks faced by mothers and neonates, thereby enhancing understanding of the intergenerational impacts. Second, it reveals the mediating role of gestational hypertension within the aforementioned association, aiming to comprehensively illustrate the mechanisms underlying these associations.

However, our study has several limitations. First, exposure misclassification was a potential issue in our research. Specifically, using a single address to estimate O_3 exposure

XSL•FO
RenderX

may not accurately reflect personal exposure, as individuals may commute to workplaces during the day. Due to data limitations, indoor air pollution was not accounted for. However, given the large sample size, the impact of exposure misclassification primarily manifests as a dilution of the study's effect, making it unlikely to fundamentally reverse the association we observed. Second, gestational hypertension was ascertained using ICD-10 codes from pregnancy-related outpatient and inpatient EMRs, which may have under-captured milder or undiagnosed cases, particularly in rural or medically underserved settings. Such under-ascertainment could underestimate the prevalence of gestational hypertension and bias the estimated associations toward the null. Nevertheless, China's National Essential Public Health Services include standardized maternal follow-up management, during which pregnant women are routinely assessed for conditions such as gestational hypertension, which may support more complete case ascertainment across health care settings. Future studies could link more granular clinical data (eg, serial blood pressure measurements and proteinuria indicators) to construct a more sensitive phenotyping definition, thereby validating and potentially refining the effect estimates reported here. Third, although covariates were selected a priori using a directed acyclic graph, residual confounding from unmeasured factors (eg, diet or other environmental exposures not available in our database) cannot be completely ruled out. Fourth, we recognize that the process of constructing the analytic sample may have introduced selection bias, potentially affecting the generalizability of our findings. Approximately 40% of eligible births were not included in the analytic sample, largely due to unsuccessful linkages to hospitalization records, incomplete residential addresses, and non-random missingness in O_3 exposure data. Compared with the eligible population, the average maternal age in our analytic sample was slightly lower (29.85 y, SD 4.82 vs 30.16 y, SD 4.91), with a small standardized mean difference of 0.062. Infant sex was nearly identical (female 47.6% in the analytic sample vs 47.7% in the eligible population; standardized mean difference=0.003). The limited imbalance in measured age and sex distributions suggests our sample is representative. Additionally, evidence from prior hospital-based birth cohorts in China [63] suggests that excluding participants with missing address data or failed geocoding does not correlate with a higher risk of adverse birth outcomes. In summary, while selection bias cannot be entirely ruled out, its impact is likely minimal. To reduce the potential bias, a weighting analysis has been applied. Sensitivity analyses using weighting methods yielded consistent estimates, thereby reinforcing the validity and robustness of our primary conclusions. In future studies, we will refine the data collection process by strengthening multisource record linkages and improving the precision of residential history tracking to further enhance data completeness and minimize potential selection bias. Fifth, our study was restricted to Shandong Province, China; therefore, caution is warranted when extrapolating the findings to regions with different climatic conditions, demographic structures, health care access, or pollutant concentrations, which may yield heterogeneous associations. Shandong spans substantial east-west variation and shares several natural and socioeconomic characteristics with China overall, providing a reasonable basis for examining the health implications of O_3 exposure. Subgroup analyses also support the robustness of this study. However, the exposure distribution in our cohort was dominated by relatively high ambient O_3 levels. Accordingly, applying these estimates to areas with lower O_3 concentrations should be done cautiously, as the underlying exposure distribution may differ from that of our study population.

Conclusions

This study provides, for the first time, robust evidence quantifying gestational hypertension as a significant mediator between prenatal O_3 exposure and adverse birth outcomes—19.94% for term LBW and 13.41% for term SGA. Using a counterfactual mediation framework, we clarify a key biological pathway linking air pollution to impaired fetal growth. These findings highlight the need to integrate O_3 pollution control into prenatal health guidance and strengthen blood pressure monitoring during pregnancy. Future research should validate these associations across diverse populations and investigate underlying biological mechanisms, such as oxidative stress.

Acknowledgments

During the preparation of this work, the authors did not use any artificial intelligence–assisted tools or services for content generation, data analysis, or manuscript drafting.

Funding

This study was funded by National Natural Science Foundation of China (82373583 and 82003474), Natural Science Foundation of Shandong Province (grant ZR202211080088), and Research Cultivation and Talent Development Promotion on Ecological Protection and High-quality Development of the Yellow River Basin (project hhzx202507).

Data Availability

The data used in this study were obtained from the official health data platform of Shandong Province. Due to institutional and data governance policies, the authors do not have permission to publicly share the dataset.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: YZ, Q Li, QW
Data curation: YZ, Q Liu, Q Li, QW
Formal analysis: YZ
Funding acquisition: Q Li, QW
Investigation: YZ
Methodology: YZ, Q Liu, MS, QH, QZ, QT, CZ, Q Li, QW
Project administration: YZ, Q Li, QW
Resources: Q Li, QW
Software: YZ, Q Liu, QZ
Supervision: YZ, Q Li, QW
Validation: YZ
Visualization: YZ
Writing–original draft: YZ
Writing–review & editing: MS, QH, QZ, QT, CZ, Q Li, QW
YZ and Q Li are co–first authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Map of the study site in Shandong Province, China.
[PNG File, 122 KB - publichealth_v12i1e81412_app1.png]

Multimedia Appendix 2
Study flowchart for participant inclusion and exclusion.
[PNG File, 72 KB - publichealth_v12i1e81412_app2.png]

Multimedia Appendix 3
Directed acyclic graph of the associations between ozone exposure and child weight.
[PNG File, 284 KB - publichealth_v12i1e81412_app3.png]

Multimedia Appendix 4
Imputation procedure based on random forest.
[DOCX File, 15 KB - publichealth_v12i1e81412_app4.docx]

Multimedia Appendix 5
Weighting approach.
[DOCX File, 31 KB - publichealth_v12i1e81412_app5.docx]

Multimedia Appendix 6
A 4-step analysis to test the mediator of gestational hypertension.
[DOCX File, 21 KB - publichealth_v12i1e81412_app6.docx]

Multimedia Appendix 7
Associations of ozone exposure with term low birth weight and term small for gestational age, stratified by maternal age, infant sex, and coastal residence.
[DOCX File, 14 KB - publichealth_v12i1e81412_app7.docx]

Multimedia Appendix 8
Association of ozone exposure with gestational hypertension, term low birth weight, and term small for gestational age.
[DOCX File, 12 KB - publichealth_v12i1e81412_app8.docx]

Multimedia Appendix 9
Mediation analysis of gestational hypertension on the association between ozone exposure and term low birth weight and term small for gestational age based on Cox proportional hazards models.

XSL•FO
RenderX

[DOCX File, 17 KB - publichealth_v12i1e81412_app9.docx]

References

1. Ashorn P, Ashorn U, Muthiani Y, et al. Small vulnerable newborns-big potential for impact. Lancet 2023 May 20;401(10389):1692-1706. [doi: 10.1016/S0140-6736(23)00354-9] [Medline: 37167991]
2. Million Death Study Collaborators, Bassani DG, Kumar R, et al. Causes of neonatal and child mortality in India: a nationally representative mortality survey. Lancet 2010 Nov 27;376(9755):1853-1860. [doi: 10.1016/S0140-6736(10)61461-4] [Medline: 21075444]
3. Katz J, Lee AC, Kozuki N, et al. Mortality risk in preterm and small-for-gestational-age infants in low-income and middle-income countries: a pooled country analysis. Lancet 2013 Aug 3;382(9890):417-425. [doi: 10.1016/S0140-6736(13)60993-9] [Medline: 23746775]
4. Barker DJP. The developmental origins of chronic adult disease. Acta Paediatr 2004 Nov;93(s446):26-33. [doi: 10.1111/j.1651-2227.2004.tb00236.x]
5. Bianchi ME, Restrepo JM. Low birthweight as a risk factor for non-communicable diseases in adults. Front Med (Lausanne) 2021;8:793990. [doi: 10.3389/fmed.2021.793990] [Medline: 35071274]
6. Ha SU, Thompson LA, Kearney G, Roth J, Xu X. Population attributable risks of neurobehavioral disorders due to low birth weight in US children. Adv Pediatr Res 2014;1:2. [doi: 10.12715/apr.2014.1.2] [Medline: 29057332]
7. Caudri D, Wijga A, Gehring U, et al. Respiratory symptoms in the first 7 years of life and birth weight at term: the PIAMA Birth Cohort. Am J Respir Crit Care Med 2007 May 15;175(10):1078-1085. [doi: 10.1164/rccm.200610-1441OC] [Medline: 17290040]
8. McGuire S. World Health Organization. Comprehensive implementation plan on maternal, infant, and young child nutrition. Geneva, Switzerland, 2014. Adv Nutr 2015 Jan;6(1):134-135. [doi: 10.3945/an.114.007781] [Medline: 25593153]
9. Okwaraji YB, Krasevec J, Bradley E, et al. National, regional, and global estimates of low birthweight in 2020, with trends from 2000: a systematic analysis. The Lancet 2024 Mar;403(10431):1071-1080. [doi: 10.1016/S0140-6736(23)01198-4]
10. Bekkar B, Pacheco S, Basu R, DeNicola N. Association of air pollution and heat exposure with preterm birth, low birth weight, and stillbirth in the US: a systematic review. JAMA Netw Open 2020 Jun 1;3(6):e208243. [doi: 10.1001/jamanetworkopen.2020.8243] [Medline: 32556259]
11. Song S, Gao Z, Zhang X, et al. Ambient fine particulate matter and pregnancy outcomes: an umbrella review. Environ Res 2023 Oct 15;235:116652. [doi: 10.1016/j.envres.2023.116652] [Medline: 37451569]
12. Lu X, Zhang Y, Jiang R, et al. Interpregnancy interval, air pollution, and the risk of low birth weight: a retrospective study in China. BMC Public Health 2024 Sep 17;24(1):2529. [doi: 10.1186/s12889-024-19711-3] [Medline: 39289643]
13. Sun S, Wang J, Cao W, et al. A nationwide study of maternal exposure to ambient ozone and term birth weight in the United States. Environ Int 2022 Dec;170:107554. [doi: 10.1016/j.envint.2022.107554] [Medline: 36202016]
14. Wang Q, Miao H, Warren JL, et al. Association of maternal ozone exposure with term low birth weight and susceptible window identification. Environ Int 2021 Jan;146:106208. [doi: 10.1016/j.envint.2020.106208] [Medline: 33129003]
15. Noghanibehambari H, Bagheri H, Salari M, Tavassoli N, Javid R, Toranji M. Breathing in the future: prenatal exposure to air pollution and infants' health outcomes in the USA. Public Health 2023 Dec;225:198-205. [doi: 10.1016/j.puhe.2023.10.012] [Medline: 37939461]
16. Zhu L, Yuan Y, Mayvaneh F, Sun H, Zhang Y, Hu C. Maternal ozone exposure lowers infant's birthweight: a nationwide cohort of over 4 million livebirths in Iran. Ecotoxicol Environ Saf 2024 Sep 15;283:116840. [doi: 10.1016/j.ecoenv.2024.116840] [Medline: 39126814]
17. Moore JP, Damasceno da Silva RM, Dias MA, Castelhano FJ, Hoinaski L, Requia WJ. Ambient air pollution and low birth weight in Brazil: a nationwide study of more than 10 million births between 2001 and 2018. Chemosphere 2024 Oct;366:143469. [doi: 10.1016/j.chemosphere.2024.143469] [Medline: 39384135]
18. Chen J, Fang J, Zhang Y, et al. Associations of adverse pregnancy outcomes with high ambient air pollution exposure: results from the Project ELEFANT. Sci Total Environ 2021 Mar;761:143218. [doi: 10.1016/j.scitotenv.2020.143218]
19. Chen J, Guo L, Liu H, et al. Modification effects of ambient temperature on associations of ambient ozone exposure before and during pregnancy with adverse birth outcomes: a multicity study in China. Environ Int 2023 Feb;172:107791. [doi: 10.1016/j.envint.2023.107791]
20. Guo P, Chen Y, Wu H, et al. Ambient air pollution and markers of fetal growth: a retrospective population-based cohort study of 2.57 million term singleton births in China. Environ Int 2020 Feb;135:105410. [doi: 10.1016/j.envint.2019.105410] [Medline: 31884132]
21. Zhang C, Yang J, Wei J, et al. Individual ambient ozone exposure during pregnancy and adverse birth outcomes: exploration of the potentially vulnerable windows. J Hazard Mater 2024 Feb;464:132945. [doi: 10.1016/j.jhazmat.2023.132945]
22. do Nascimento FP, de Almeida MF, Gouveia N. Individual and contextual socioeconomic status as effect modifier in the air pollution-birth outcome association. Sci Total Environ 2022 Jan 10;803:149790. [doi: 10.1016/j.scitotenv.2021.149790] [Medline: 34481165]

XSL•FO
RenderX

23. Hunter R, Baird B, Garcia M, et al. Gestational ozone inhalation elicits maternal cardiac dysfunction and transcriptional changes to placental pericytes and endothelial cells. Toxicol Sci 2023 Nov 28;196(2):238-249. [doi: 10.1093/toxsci/kfad092] [Medline: 37695302]
24. Sanidas E, Papadopoulos DP, Grassos H, et al. Air pollution and arterial hypertension. a new risk factor is in the air. J Am Soc Hypertens 2017 Nov;11(11):709-715. [doi: 10.1016/j.jash.2017.09.008] [Medline: 28989071]
25. Wong MCS, Tam WWS, Wang HHX, et al. Exposure to air pollutants and mortality in hypertensive patients according to demography: a 10 year case-crossover study. Environ Pollut 2014 Sep;192:179-185. [doi: 10.1016/j.envpol.2014.05.025] [Medline: 24953346]
26. Cheng Y, Wang P, Zhang L, et al. Ozone exposure during pregnancy and risk of gestational hypertension or preeclampsia in China. JAMA Netw Open 2023 Apr 3;6(4):e236347. [doi: 10.1001/jamanetworkopen.2023.6347] [Medline: 37010870]
27. Hu H, Ha S, Xu X. Ozone and hypertensive disorders of pregnancy in Florida: identifying critical windows of exposure. Environ Res 2017 Feb;153:120-125. [doi: 10.1016/j.envres.2016.12.002] [Medline: 27940104]
28. Yuan K, Sun F, Zhang Y, et al. Maternal exposure to ozone and risk of gestational hypertension and eclampsia in the United States. Sci Total Environ 2023 May 10;872:162292. [doi: 10.1016/j.scitotenv.2023.162292] [Medline: 36801312]
29. Everett TR, Lees CC. Beyond the placental bed: Placental and systemic determinants of the uterine artery Doppler waveform. Placenta 2012 Nov;33(11):893-901. [doi: 10.1016/j.placenta.2012.07.011] [Medline: 22902007]
30. Kingdom JCP, Kaufmann P. Oxygen and placental villous development: origins of fetal hypoxia. Placenta 1997 Nov;18(8):613-621. [doi: 10.1016/S0143-4004(97)90000-X]
31. Thorp JM. Placental vascular compromise: unifying the etiologic pathways of perinatal compromise. Curr Probl Obstet Gynecol Fertil 2001 Dec;24(6):202-220. [doi: 10.1067/mog.2001.119782]
32. Ye C, Ruan Y, Zou L, et al. The 2011 survey on hypertensive disorders of pregnancy (HDP) in China: prevalence, risk factors, complications, pregnancy and perinatal outcomes. PLoS ONE 2014;9(6):e100180. [doi: 10.1371/journal.pone.0100180] [Medline: 24937406]
33. Kroener L, Wang ET, Pisarska MD. Predisposing factors to abnormal first trimester placentation and the impact on fetal outcomes. Semin Reprod Med 2016 Jan;34(1):27-35. [doi: 10.1055/s-0035-1570029] [Medline: 26696276]
34. Villar J, Cheikh Ismail L, Victora CG, et al. International standards for newborn weight, length, and head circumference by gestational age and sex: the newborn cross-sectional study of the INTERGROWTH-21st Project. Lancet 2014 Sep 6;384(9946):857-868. [doi: 10.1016/S0140-6736(14)60932-6] [Medline: 25209487]
35. Wei J, Li Z, Li K, et al. Full-coverage mapping and spatiotemporal variations of ground-level ozone (O3) pollution from 2013 to 2020 across China. Remote Sens Environ 2022 Mar;270:112775. [doi: 10.1016/j.rse.2021.112775]
36. International Statistical Classification of Diseases and Related Health Problems, 10th revision (ICD-10). : World Health Organization; 2019 URL: https://icd.who.int/browse10/Content/statichtml/ICD10Volume2_en_2019.pdf [accessed 2026-03-21]
37. Ji N, Johnson M, Eckel SP, et al. Prenatal ambient air pollution exposure and child weight trajectories from the 3rd trimester of pregnancy to 2 years of age: a cohort study. BMC Med 2023 Sep 7;21(1):341. [doi: 10.1186/s12916-023-03050-y] [Medline: 37674158]
38. Textor J, Hardt J, Knüppel S. DAGitty: A graphical tool for analyzing causal diagrams. Epidemiology (Sunnyvale) 2011 Sep;22(5):745. [doi: 10.1097/EDE.0b013e318225c2be] [Medline: 21811114]
39. Products. National Centers for Environmental Information. URL: https://www.ncdc.noaa.gov/data-access [accessed 2026-03-14]
40. Cummings P. The relative merits of risk ratios and odds ratios. Arch Pediatr Adolesc Med 2009 May;163(5):438-445. [doi: 10.1001/archpediatrics.2009.31] [Medline: 19414690]
41. Deville JC, Särndal CE, Sautory O. Generalized raking procedures in survey sampling. J Am Stat Assoc 1993 Sep;88(423):1013-1020. [doi: 10.1080/01621459.1993.10476369]
42. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. Stat Med 2015 Dec 10;34(28):3661-3679. [doi: 10.1002/sim.6607] [Medline: 26238958]
43. van Noord-Zaadstra BM, Looman CW, Alsbach H, Habbema JD, te Velde ER, Karbaat J. Delaying childbearing: effect of age on fecundity and outcome of pregnancy. BMJ 1991 Jun 8;302(6789):1361-1365. [doi: 10.1136/bmj.302.6789.1361] [Medline: 2059713]
44. Lakshmanan A, Chiu YHM, Coull BA, et al. Associations between prenatal traffic-related air pollution exposure and birth weight: modification by sex and maternal pre-pregnancy body mass index. Environ Res 2015 Feb;137:268-277. [doi: 10.1016/j.envres.2014.10.035] [Medline: 25601728]
45. Zheng H, Zheng H. Assessment and prediction of carbon storage based on land use/land cover dynamics in the coastal area of Shandong Province. Ecol Indic 2023 Sep;153:110474. [doi: 10.1016/j.ecolind.2023.110474]
46. Ng C, Malig B, Hasheminassab S, Sioutas C, Basu R, Ebisu K. Source apportionment of fine particulate matter and risk of term low birth weight in California: exploring modification by region and maternal characteristics. Sci Total Environ 2017 Dec 15;605-606:647-654. [doi: 10.1016/j.scitotenv.2017.06.053] [Medline: 28675874]

XSL•FO
RenderX

47. Buis ML. Direct and indirect effects in a logit model. Stata J 2010;10(1):11-29. [doi: 10.1177/1536867x1001000104] [Medline: 22468140]
48. Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. Epidemiology (Sunnyvale) 1992 Mar;3(2):143-155. [doi: 10.1097/00001648-199203000-00013] [Medline: 1576220]
49. Valeri L, Vanderweele TJ. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. Psychol Methods 2013 Jun;18(2):137-150. [doi: 10.1037/a0031034] [Medline: 23379553]
50. Pearl J. Direct and indirect effects. In: Probabilistic and Causal Inference: The Works of Judea Pearl: Association for Computing Machinery; 2022:373-392.
51. Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. Psychol Methods 2010 Dec;15(4):309-334. [doi: 10.1037/a0020761] [Medline: 20954780]
52. Joo EH, Kim YR, Kim N, Jung JE, Han SH, Cho HY. Effect of endogenic and exogenic oxidative stress triggers on adverse pregnancy outcomes: preeclampsia, fetal growth restriction, gestational diabetes mellitus and preterm birth. Int J Mol Sci 2021 Sep 19;22(18):10122. [doi: 10.3390/ijms221810122] [Medline: 34576285]
53. Clayton CA, Perritt RL, Pellizzari ED, et al. Particle Total Exposure Assessment Methodology (PTEAM) study: Distributions of aerosol and elemental concentrations in personal, indoor, and outdoor air samples in a southern California community. J Expo Anal Environ Epidemiol 1993;3(2):227-250. [Medline: 7694700]
54. Yuan Z, Li Q, Su T, et al. Effects of fine ambient particulate matters on de novo hypertensive disorders of pregnancy and blood pressure before 20 weeks. Environ Res 2023 Feb 1;218:115023. [doi: 10.1016/j.envres.2022.115023] [Medline: 36502896]
55. Liu Y, Li N, An H, et al. Impact of gestational hypertension and preeclampsia on low birthweight and small‐for‐gestational‐age infants in China: a large prospective cohort study. J of Clin Hypertens 2021 Apr;23(4):835-842. [doi: 10.1111/jch.14176]
56. Liu Z, Nawsherwan N, Le Z, et al. The adverse effect of gestational diabetes mellitus and hypertensive disorders of pregnancy on maternal–perinatal outcomes among singleton and twin pregnancies: a retrospective cohort study (2011–2019). Front Endocrinol 2011;14. [doi: 10.3389/fendo.2023.1267338]
57. Hunter R, Wilson T, Lucas S, et al. Characterization of mild delayed gestational hypertension in rats following ozone exposure. Cardiovasc Toxicol 2024 Sep;24(9):843-851. [doi: 10.1007/s12012-024-09887-w] [Medline: 38963633]
58. Xia Y, Niu Y, Cai J, et al. Effects of personal short-term exposure to ambient ozone on blood pressure and vascular endothelial function: a mechanistic study based on DNA methylation and metabolomics. Environ Sci Technol 2018 Nov 6;52(21):12774-12782. [doi: 10.1021/acs.est.8b03044] [Medline: 30259740]
59. Stepan H, Hund M, Andraczek T. Combining biomarkers to predict pregnancy complications and redefine preeclampsia: the angiogenic-placental syndrome. Hypertension 2020 Apr;75(4):918-926. [doi: 10.1161/HYPERTENSIONAHA.119.13763] [Medline: 32063058]
60. Li W, Cui N, Mazzuca MQ, Mata KM, Khalil RA. Increased vascular and uteroplacental matrix metalloproteinase-1 and -7 levels and collagen type i deposition in hypertension in pregnancy: role of TNF-α. Am J Physiol Heart Circ Physiol 2017 Sep 1;313(3):H491-H507. [doi: 10.1152/ajpheart.00207.2017] [Medline: 28626073]
61. Garcia M, Salazar R, Wilson T, et al. Early gestational exposure to inhaled ozone impairs maternal uterine artery and cardiac function. Toxicol Sci 2021 Jan 6;179(1):121-134. [doi: 10.1093/toxsci/kfaa164] [Medline: 33146391]
62. Andruszewski D, Uhlfelder DC, Desiato G, et al. Embryo-restricted responses to maternal IL-17A promote neurodevelopmental disorders in mouse offspring. Mol Psychiatry 2025 Apr;30(4):1585-1593. [doi: 10.1038/s41380-024-02772-6] [Medline: 39384965]
63. Xiao Q, Chen H, Strickland MJ, et al. Associations between birth outcomes and maternal PM2.5 exposure in Shanghai: a comparison of three exposure assessment approaches. Environ Int 2018 Aug;117:226-236. [doi: 10.1016/j.envint.2018.04.050] [Medline: 29763818]

Abbreviations

CHAP: China High Air Pollutants
EMR: electronic medical record
ICD-10: *International Classification of Diseases, Tenth Revision*
INTERGROWTH-21st: International Fetal and Newborn Growth Consortium for the 21st Century
IPSW: inverse probability of selection weight
LBW: low birth weight
NIE: natural indirect effect
NOAA: National Oceanic and Atmospheric Administration
OR: odds ratio
PM: particulate matter
RR: relative risk

XSL•FO
RenderX

SGA: small for gestational age
TE: total effect

Edited by A Mavragani, T Sanchez; submitted 28.Jul.2025; peer-reviewed by F Adjei, LZ Lin; revised version received 04.Mar.2026; accepted 05.Mar.2026; published 08.Apr.2026.

Please cite as:
Zhang Y, Liu Q, Song M, Huang Q, Zhao Q, Tao Q, Zhang C, Li Q, Wang Q
Gestational Hypertension as a Mediator of Prenatal Ozone Exposure and Term Low Birth Weight: Birth Cohort Study
JMIR Public Health Surveill 2026;12:e81412
URL: https://publichealth.jmir.org/2026/1/e81412
doi:10.2196/81412

© Yuxin Zhang, Quanjun Liu, Mengwei Song, Qiulin Huang, Qingbo Zhao, Qingqing Tao, Chenchen Zhang, Qin Li, Qing Wang. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 8.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Projected Health and Economic Benefits of Air Quality Targets in China: Modeling Study

Siyuan Wang[1], MCom; Zhiwei Xu[2], PhD; Gian Luca Di Tanna[3], PhD; Yawen Jiang[4], PhD; Mingsheng Chen[5,6], PhD; Laura Downey[1,7], PhD; Stephen Jan[1,7], PhD; Lei Si[8,9], PhD

[1]The George Institute for Global Health, UNSW, Sydney, New South Wales, Australia
[2]School of Medicine and Dentistry, Griffith University, Gold Coast, Queensland, Australia
[3]Department of Business Economics, Health and Social Care, University of Applied Sciences and Arts of Southern Switzerland, Lugano, Switzerland
[4]School of Public Health (Shenzhen), Sun Yat-Sen University, Shenzhen, China
[5]School of Health Policy and Management, Nanjing Medical University, No. 101 Longmian Avenue, Nanjing, China
[6]Jiangsu Health Vocational College, Nanjing, China
[7]Faculty of Medicine, Imperial College, London, United Kingdom
[8]School of Health Sciences, Faculty of Health, Western Sydney University, Campbelltown, New South Wales, Australia
[9]School of Health Management, Anhui Medical University, Hefei, China

Corresponding Author:
Mingsheng Chen, PhD
School of Health Policy and Management, Nanjing Medical University, No. 101 Longmian Avenue, Nanjing, China

Abstract

Background: Air pollution continues to impose a substantial health and economic burden in China. Despite recent improvements, national annual average PM2.5 (fine particulate matter) concentrations remain substantially above the levels deemed safe by the World Health Organization (WHO), underscoring the need for more stringent air quality control.

Objective: This study aims to quantify the projected health and economic benefits of reducing PM2.5 concentrations in China under the Healthy China 2030 plan.

Methods: Using the 2020 ground-level PM2.5 data as a baseline, we projected cause-specific mortality and morbidity outcomes for 337 prefecture-level cities in China from 2020 to 2030 under four policy scenarios: (1) Healthy China 2030 (10% reduction by 2025; 25 μg/m³ by 2030) and the WHO targets of (2) 15 μg/m³, (3) 10 μg/m³, and (4) 5 μg/m³ by 2030. Mortality for noncommunicable diseases, lower respiratory infections, stroke, ischemic heart disease, lung cancer, and chronic obstructive pulmonary disease was estimated using the Global Exposure Mortality Model. Hospitalizations were modeled using log-linear models based on national evidence.

Results: In 2020, PM2.5 concentrations across the 337 cities ranged from 7 to 63 μg/m³, with a national annual mean of 32.6 μg/m³. Higher concentration levels were observed in eastern China, particularly in the eastern and southeastern regions. Overall, our analysis accounted for nearly 70% of the total population of China in 2020. Our analysis shows that maintaining PM2.5 at 2020 levels was projected to result in 9.04 million (95% CI 7.70 - 10.67 million) attributable deaths, compared with 8.63 million (95% CI 7.42 - 10.30 million) under the policy scenario, corresponding to 0.41 million premature deaths averted under the 14th Five-Year Clean Air Plan. Over 2021 - 2030, PM2.5-attributable deaths declined from 17.76 million (95% CI 14.21 - 20.95 million) under baseline conditions to 15.96 million (95% CI 12.85 - 19.15 million) under the policy scenario. Achieving WHO targets would further reduce attributable deaths to 13.99 million (95% CI 11.25 - 15.25 million) at 15 μg/m³, 12.86 million (95% CI 10.85 - 14.85 million) at 10 μg/m³, and 11.49 million (95% CI 8.96 - 13.45 million) at 5 μg/m³. The annual average hospitalizations declined by 21,422 cardiovascular and 26,545 respiratory admissions under the policy scenario, increasing to 41,690 and 51,681 at 15 μg/m³, 51,884 and 64,333 at 10 μg/m³, and 62,146 and 77,073 at 5 μg/m³, respectively. Subsequently, total economic gains reached US $123.7 billion under the policy scenario and increased to US $185.7 billion, US $240.7 billion, and US $306.5 billion under the 15, 10, and 5 μg/m³ scenarios, respectively.

Conclusions: Our findings suggest that while the Healthy China 2030 Plan offered substantial health gains, achieving stricter WHO air quality targets could yield 2-3 times greater benefits. These findings will support the future development of stricter evidence-based national air quality standards.

(JMIR Public Health Surveill 2026;12:e84809) doi:10.2196/84809

KEYWORDS

China; health impact modeling; particulate matter; economic benefits; air pollution

XSL•FO
RenderX

Introduction

Ambient air pollution represents a major global environmental health challenge. In 2019, the World Health Organization (WHO) estimated that exposure to outdoor air pollution was responsible for more than 4.2 million premature deaths worldwide [1]. In particular, there is a growing body of epidemiological evidence linking exposure to fine particulate matter (PM) with increased risks of cardiovascular and respiratory diseases, lung cancer, and premature mortality [2-6]. In addition to the substantial health burden, exposure to PM2.5 also imposes considerable economic costs, including direct health care expenditures from air pollution–induced illnesses, increased hospitalizations, and indirect losses from reduced workforce productivity [7]. In particular, the burden falls disproportionately on low- and middle-income countries, where around 80% of individuals exposed to unsafe annual PM2.5 concentrations reside and nearly 90% of air pollution–related premature deaths occur [1,8]. Recognizing the increasing health and economic burden associated with exposure to elevated levels of PM, many low- and middle-income countries have introduced national policies aimed at reducing PM levels. For example, countries in Southeast Asia, such as Indonesia, Vietnam, and Thailand, have implemented mitigation strategies to varying degrees, including the establishment of ambient air quality standards, the introduction of emission controls across key sectors, such as transport, industry, and energy, and the strengthening of air quality monitoring systems, with the aim of reducing both emission levels and population exposure [9-11]. In India, the National Clean Air Program has been associated with substantial improvements in urban air quality, with declines in PM10 concentrations across 20 major cities, and is estimated to have averted approximately 62,219 excess mortalities in 2021 compared to 2018 [12].

In China, ambient air pollution remains a critical public health issue, contributing to an estimated 1.67 million premature deaths nationwide in 2020 [13]. Exposure to PM is the nation's fourth leading risk factor for overall mortality and disability-adjusted life years [14]. Overall, this contributes substantially to both the health and economic burden, with several studies estimating associated annual economic losses ranging from 0.92% to 3.6% of the national gross domestic product (GDP) within the past decade [13,15,16]. The Chinese government introduced the Air Pollution Prevention and Control Action Plan in 2013, setting ambitious targets to reduce pollution levels [17]. The plan aimed to lower PM concentrations in prefecture-level cities and above by at least 10% from 2012 levels by 2017. Consequently, the annual average PM2.5 concentration in 74 major cities declined by 33.3%, from 72.2 μg/m³ in 2013 to 47.0 μg/m³ in 2017 [18]. This improvement has resulted in substantial health benefits, including the reduction of 47,240 deaths and the prevention of 710,020 years of life lost [18].

In 2021, the WHO updated its air quality guidelines, lowering the recommended annual average PM2.5 concentration to 5 μg/m³ [19]. This revision reflects a growing body of epidemiological evidence demonstrating that both short-term and long-term exposure at low threshold levels to PM can still result in adverse health effects [3,4,20]. Despite significant improvements in air quality, PM2.5 levels in 2020 remained 7 times higher than the updated recommended limit. Han et al [21] estimated that long-term exposure to ambient PM2.5 above the new WHO guideline was associated with approximately 1.4 million premature deaths and US $1006.9 billion in the total value of statistical life annually. As part of its commitment to building a "Beautiful China," the government introduced a new set of air quality standards under the latest Clean Air Action Plan, aiming to reduce the annual mean PM2.5 concentration by 10% between 2020 and 2025 and achieving the interim target of 25 μg/m³ by 2030 [22,23]. This initiative underscores China's ongoing efforts to mitigate the health and economic burdens associated with air pollution.

Currently, much is known about the national health and economic disease burden associated with PM2.5 exposure. For example, Maji et al [13] estimated the disease burden associated with PM2.5 across 338 cities, reporting that PM2.5-attributable mortality reached 0.964 million in 2016, accounting for 9.98% of total reported deaths in China that year. Other national-level estimates since 2010 have ranged from 1.02 to 2.20 million deaths [24-28]. Additionally, several studies have also retrospectively examined the benefits of previous air quality guidelines or policies, either nationally or at a regional level. Xu et al [29] assessed the health and economic impacts of clean air policies in the Beijing-Tianjin-Hebei region, estimating a reduction of 24,000 premature deaths and associated economic benefits of US $60 billion. Zhang et al [30] compared the benefits of the previous 5-year Clean Air Action Plan with its implementation costs, reporting that the public health benefits of air quality improvements were 1.5 times greater than the mitigation costs. However, the understanding of the potential benefits of future air quality improvements, particularly those projected under the latest clean air policies, remains limited. Given the significant gap between current national air quality standards and WHO recommendations, evaluating the health and economic impacts of national air policies is essential for supporting evidence-based decision-making as China advances toward its Healthy China 2030 goals. To bridge this dearth of evidence, we modeled the attributable health and economic outcomes associated with achieving the latest national regulations from 2021 to 2030, as well as the additional benefits associated with reaching the more stringent WHO targets of 15, 10, and 5 μg/m³ over the same period from a health system perspective.

Methods

Policy Scenarios and Model Period

We projected cause-specific mortality and morbidity outcomes for 337 prefecture-level cities in China from 2020 to 2030 under several air quality improvement scenarios. These include (1) a 10% reduction in PM2.5 by 2025, targeting 25 μg/m³ by 2030, representing the policy scenario under the current clean air action plan; (2-3) a stepwise reduction to WHO interim targets, with annual PM2.5 concentration levels reaching 15 μg/m³ (Interim Target III), and 10 μg/m³ (Interim Target IV) by 2030, respectively; and (4) achieving the WHO's recommended air quality guideline of 5 μg/m³ by 2030. To account for population

XSL•FO
RenderX

changes, age-specific and sex-specific baseline population data were projected for the modeled period under shared socioeconomic pathways for each of the air quality scenarios. Table S1 in Multimedia Appendix 1 presents the matrix summarizing the air quality scenarios and population projections considered in this analysis.

Statistical Analysis

We used the widely used health impact function to model mortality and morbidity outcomes under the policy scenarios [31]. Specifically, the health impact function is given as follows:

$$\Delta y = x0 \times (RR-1)RR \times Pop$$

where the health end point attributable to PM2.5 Δy) is a function of the baseline incidence rate x0, the population attributable risk [(RR-1)/RR], and the population at risk (Pop). Mortality outcomes included the combined category of noncommunicable diseases and lower respiratory infections (NCD + LRI), alongside 5 specific causes: stroke, ischemic heart disease (IHD), chronic obstructive pulmonary disease (COPD), lung cancer (LC), and LRIs. Morbidity outcomes assessed included hospital admissions related to cardiovascular and respiratory conditions.

Mortality Attributable to PM2.5 Exposure

Mortality attributable to PM2.5 exposure under baseline and policy scenarios was estimated using the global exposure mortality model (GEMM), which provides age-specific and cause-specific hazard ratios for nonaccidental mortality [32]. The GEMM is a multivariate random effects model that synthesizes evidence from 41 cohort studies globally to characterize the exposure-response relationship between ambient air pollution and nonaccidental mortality and has been widely applied in health assessment studies of ambient air pollution internationally. Details of the model have been described elsewhere [32]. In brief, the GEMM characterizes the exposure-response relationship as:

$$GEMM(z) = \exp\{\theta \log(z\alpha+1)1+\exp\{-z-\mu\nu\}\}, \text{where } z = \max(0, PM2.5-2.4)$$

where θ, α, μ, and ν are parameters that define the shape of the relationship. We selected the GEMM as it provides the most up-to-date risk estimates for the relationship between ambient air pollution and mortality globally, incorporating risk information from high-pollution regions, such as China. Additionally, the GEMM has demonstrated that mortality estimates attributable to PM exposure may be several times higher compared to those derived from other previously used risk functions, which integrate exposure data from multiple sources, including both indoor and outdoor environments [32,33]. Table S2 in Multimedia Appendix 1 provides details on the parameters of the GEMM.

Morbidity Attributable to PM2.5 Exposure

Cardiovascular and respiratory disease–specific hospital admissions attributable to PM2.5 exposure were estimated using log-linear models. Log-linear models have previously been used in the literature for the assessment of PM2.5-related mortality and morbidity outcomes, particularly in high pollution regions, such as China [13,34,35]. Specifically, the function is defined as follows:

$$RR(z) = \exp[\beta(z-z0)]$$

where β denotes the exposure-response coefficient, representing the change in the incidence of the disease-specific outcome for each μg/m³ increase in PM2.5 concentration, and z0 is the threshold concentration. Consistent with recent epidemiological evidence demonstrating that PM2.5 exposure can have adverse health effects even at low concentrations, no threshold concentration was applied in this analysis. Table S3 in Multimedia Appendix 1 provides an overview of the disease-specific mortality and morbidity outcomes assessed in this study, along with the corresponding exposure-response coefficients derived from national and international studies.

Estimating Health and Economic Benefits

To estimate the economic impact of reductions in mortality and morbidity, we used the value-of-statistical life (VSL) and cost-of-illness (COI) methods. Briefly, VSL represents the monetary value assigned to reducing the risk of mortality, offering a quantitative basis for informing government decisions on environmental and health policies [36]. The estimate is derived from survey-based assessments of individuals' stated willingness to pay for incremental reductions in mortality risk [37]. To ensure the estimates adequately reflected the risk and income specific to the Chinese context, we used VSL values derived from a recent national-level study, estimated at US $689,659 [38]. To quantify morbidity benefits, we used the COI method, which estimates the economic burden of disease [39]. Reductions in direct costs were estimated by multiplying the average hospitalization costs for cardiovascular and respiratory diseases by the projected decrease in hospital admissions under improved air quality scenarios. Baseline figures in 2020 were projected using time series regression, assuming that historical trends continued over the modeled period. We use a 5% discount rate, consistent with China's economic evaluation guidelines [40].

In our reporting, we adhered to the GATHER (Guidelines for Accurate and Transparent Health Estimates Reporting) guidelines (Checklist 1).

Data Source

Air quality monitoring data between January 1 and December 31, 2020, were collected from the National Environmental Monitoring Center, which contain hourly PM2.5 concentrations collected from 1497 monitoring sites in 338 Chinese cities at the prefectural level and above. In our analysis, the annual mean PM2.5 concentrations at the city level were obtained by aggregating hourly monitoring data to annual averages at each station and subsequently assigning these values to the corresponding administrative city unit by averaging across stations within each city. For baseline age-specific and cause-specific mortality in 2020, national-level estimates disaggregated by region (western, central, and eastern China) were utilized in the absence of city-level data. These estimates were derived from the Centers for Disease Control and Prevention's mortality surveillance system, a national death registry covering over 605 counties and districts (21.1% of all counties and districts in China) [41]. Baseline mortalities under the business-as-usual scenario were projected for the modeled period using estimates from an existing study, which assessed

XSL•FO
RenderX

risk factor exposures and associated noncommunicable disease mortality trends based on Global Burden of Disease data through 2030 [42]. For morbidity, baseline cardiovascular and respiratory hospitalization incidences and hospitalization costs were sourced from national and provincial statistical yearbooks and projected forward based on historical trends [43]. Moreover, age-specific and sex-specific baseline demographic data for 337 cities in 2020 were also extracted from statistical yearbooks. Table S4 in Multimedia Appendix 1 presents the primary data sources used in our health prediction model, while Table S5 in Multimedia Appendix 1 summarizes the key assumptions underlying our analyses.

Ethical Considerations

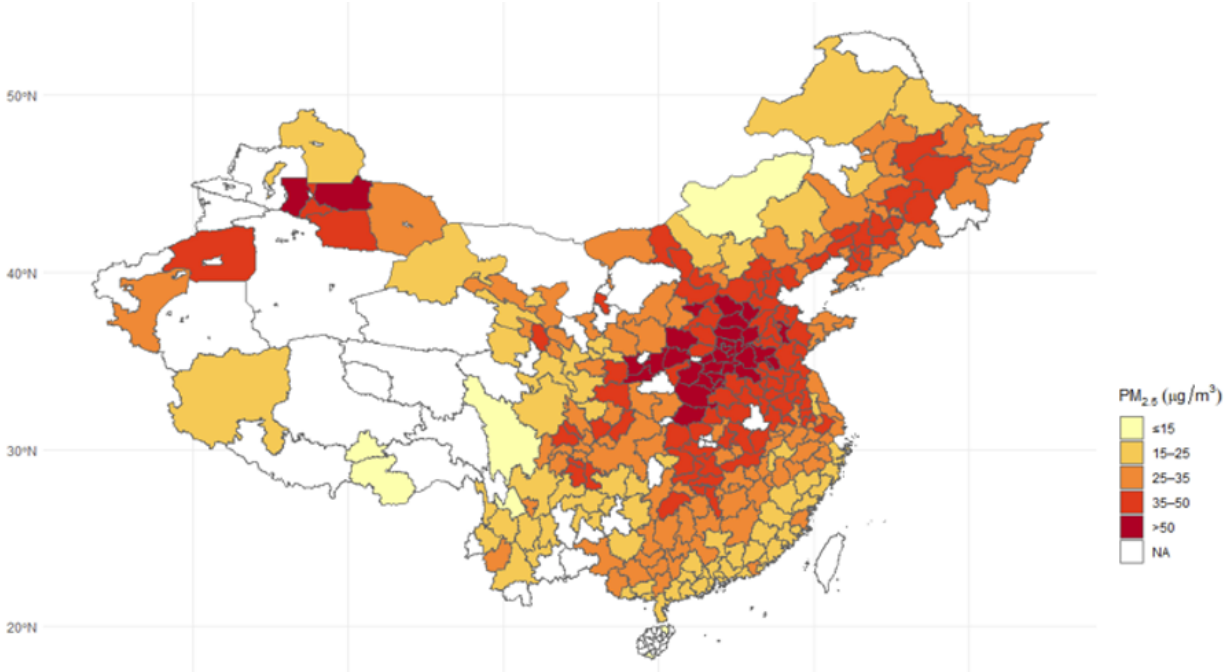
This study is a secondary analysis of preexisting, aggregate-level data obtained from publicly available sources, including the *China Statistical Yearbook* [43] and the *China Health Statistical Yearbook* [44]. These datasets are compiled and released by the National Bureau of Statistics of China as part of routine national statistical and health surveillance. The data are reported at an aggregated city-level and do not contain individual-level data or personal identifiers. Therefore, ethics approval and informed consent were not required as it did not involve direct human participation. Permission to use the data was granted through public access, and all data were used in compliance with the terms and conditions specified by the respective data providers [44].

Results

PM2.5 Concentrations

Figure 1 illustrates the spatial distribution of PM2.5 levels across 337 cities. In 2020, PM2.5 concentrations across 337 cities ranged from 7 to 63 μg/m³, with a national annual average of 32.6 μg/m³, falling below the 35 μg/m³ national guidelines but exceeding the WHO-recommended target of 5 μg/m³ by more than 6-fold. Higher concentrations were observed in the east compared to the west, with densely populated regions such as the Beijing-Tianjin-Hebei (BTH), Fenwei Plain (FP), and Yangtze Delta (YDR) reporting PM2.5 levels of 50.7, 47.8, and 36.7 μg/m³, respectively. Overall, our analysis accounted for nearly 70% of the total population of China in 2020.

Figure 1. Annual average PM2.5 concentrations across Chinese prefecture-level cities in 2020. PM: particulate matter.

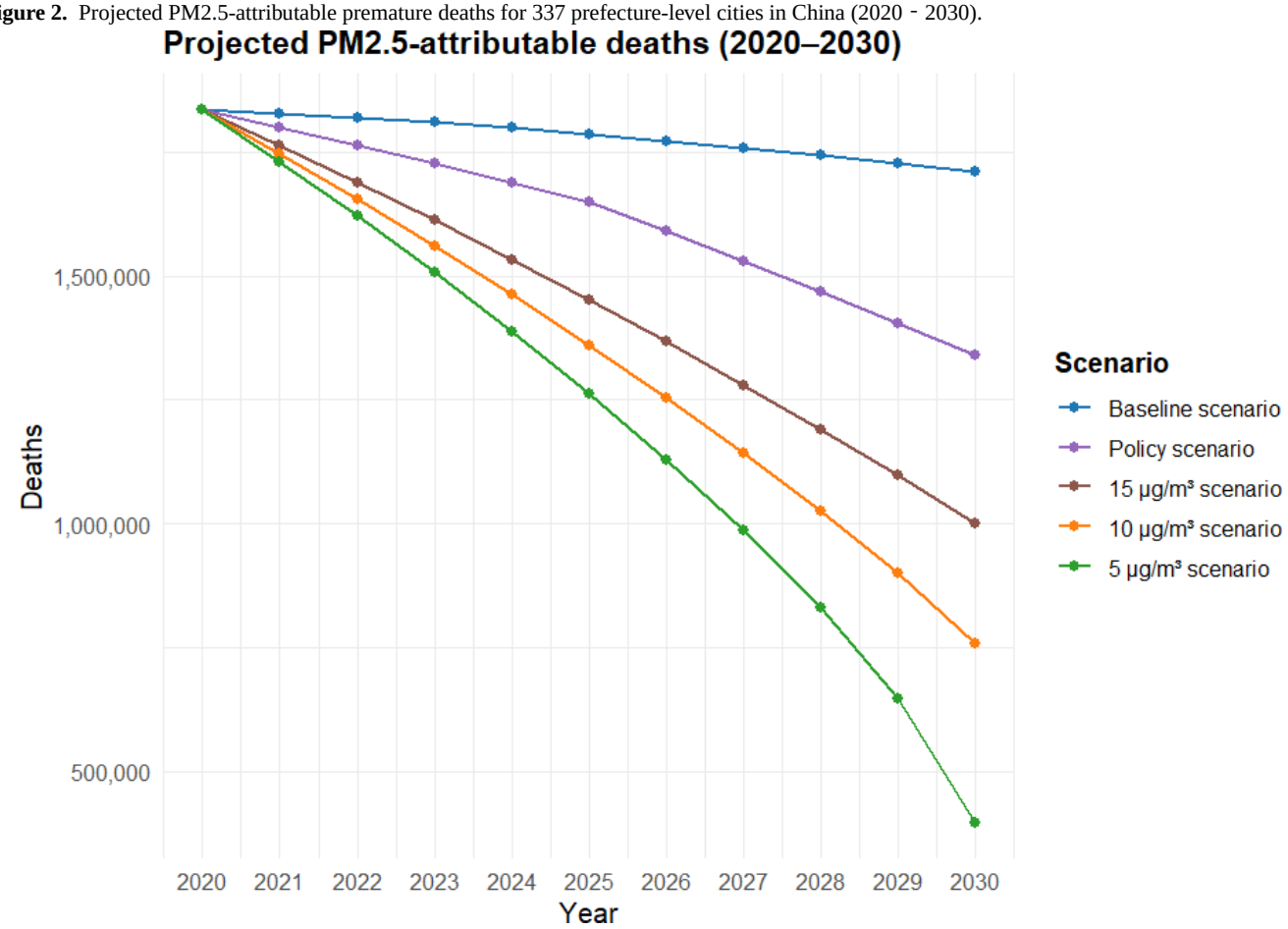


Mortality

Figure 2 presents the estimated burden of premature mortality attributable to PM2.5 exposure across the modeled scenarios, based on the Shared Socioeconomic Pathway 2 (SSP2) population trajectory. Between 2021 and 2025, under the baseline scenario in which PM2.5 concentrations remained at 2020 levels, exposure was projected to contribute to 9.04 million (95% CI 7.70 - 10.67 million) attributable deaths. Under the policy scenario, where concentrations were reduced by 10% from 2020 to 2025, this estimate declined to 8.63 million (95% CI 7.42 - 10.30 million), corresponding to an additional 0.41 million premature deaths averted under the 14th Five-Year Clean Air Plan. Over the full modeled period from 2021 to 2030, the total number of PM2.5-attributable deaths was estimated at 17.76 million (95% CI 14.21 - 20.95 million) under the baseline scenario, while the policy scenario resulted in a lower estimate of 15.96 million (95% CI 12.85 - 19.15 million). If PM2.5 concentrations were further reduced to the WHO's more stringent targets, the estimated number of attributable deaths would decline correspondingly to 13.99 million (95% CI 11.25 - 15.25 million) under a 15 μg/m³ scenario, 12.86 million (95% CI 10.85 - 14.85 million) under a 10 μg/m³ scenario, and 11.49 million (95% CI 8.96 - 13.45 million) under a 5 μg/m³ scenario, respectively.

XSL•FO
RenderX

Figure 2. Projected PM2.5-attributable premature deaths for 337 prefecture-level cities in China (2020 - 2030).



The spatial distribution of the annual average reduction in total premature deaths across provinces for each modeled scenario is shown in Figure 3. The greatest reductions in mortality were observed in the eastern and southeastern regions, particularly in highly populated and heavily polluted areas such as the BTH, FP, and YDR regions, compared to the west. Under the policy scenario, the estimated reduction in average annual premature deaths over the 10-year period was 13,879 for the BTH region, 7,009 for the FP region, and 6,895 for the YDR region. Similar spatial patterns of avoided premature deaths were observed under the WHO scenarios. In the most polluted regions (BTH, FP, and YDR), the estimated annual average benefits of reaching 15 μg/m³ corresponded to 237,580, 165,989, and 142,120 deaths averted, respectively. For the 10 μg/m³ scenario, these estimates increased to 283,878, 217,406, and 171,727, while for the 5 μg/m³ scenario, the reductions in premature deaths reached 338,916, 279,313, and 206,993, respectively.

XSL•FO
RenderX

Figure 3. Spatial distribution of reductions in PM2.5-attributable premature deaths across Chinese prefecture-level cities under improved air quality scenarios for 2021 - 2030: (A) policy scenario, (B) 15 μg/m³ target, (C) 10 μg/m³ target, and (D) 5 μg/m³ target. PM: particulate matter.

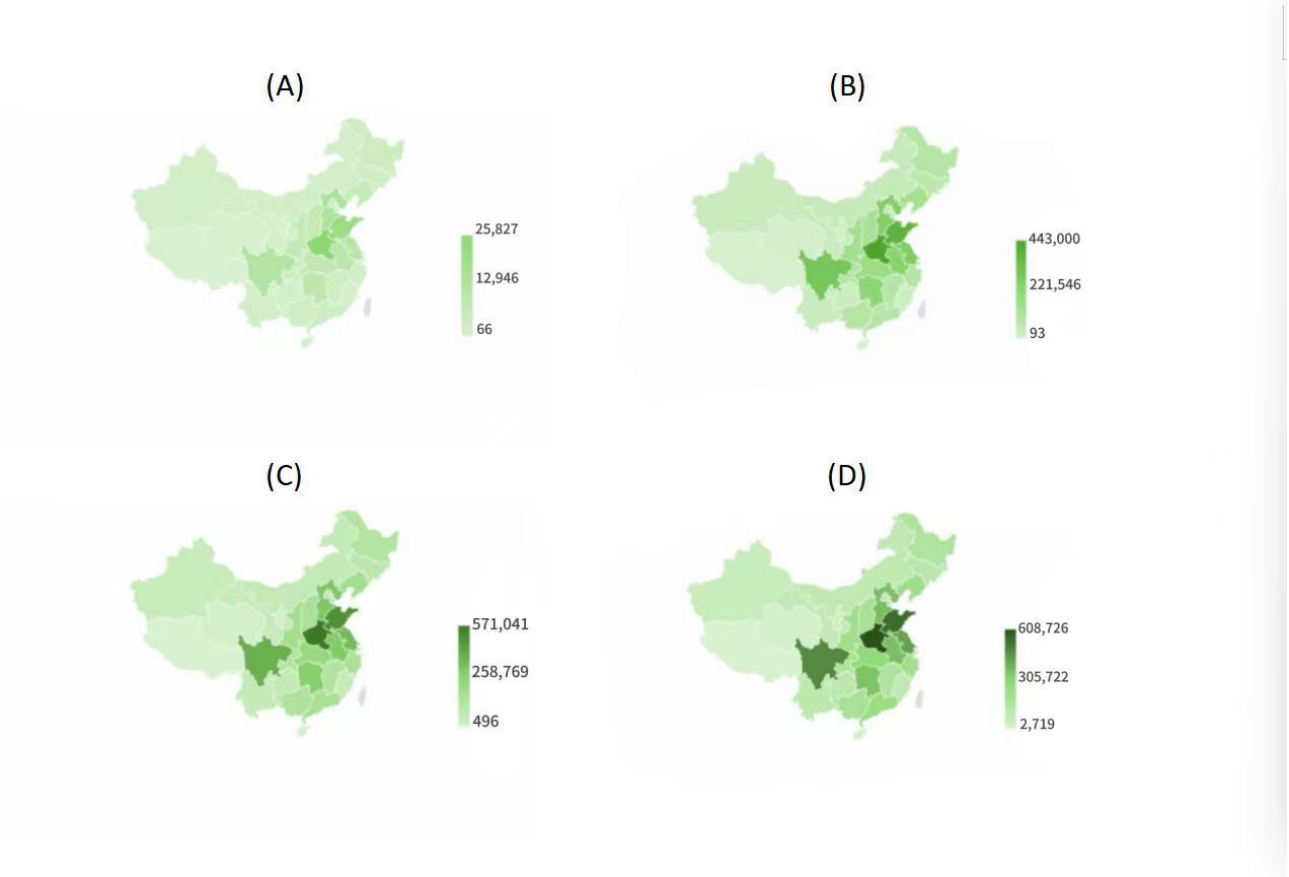
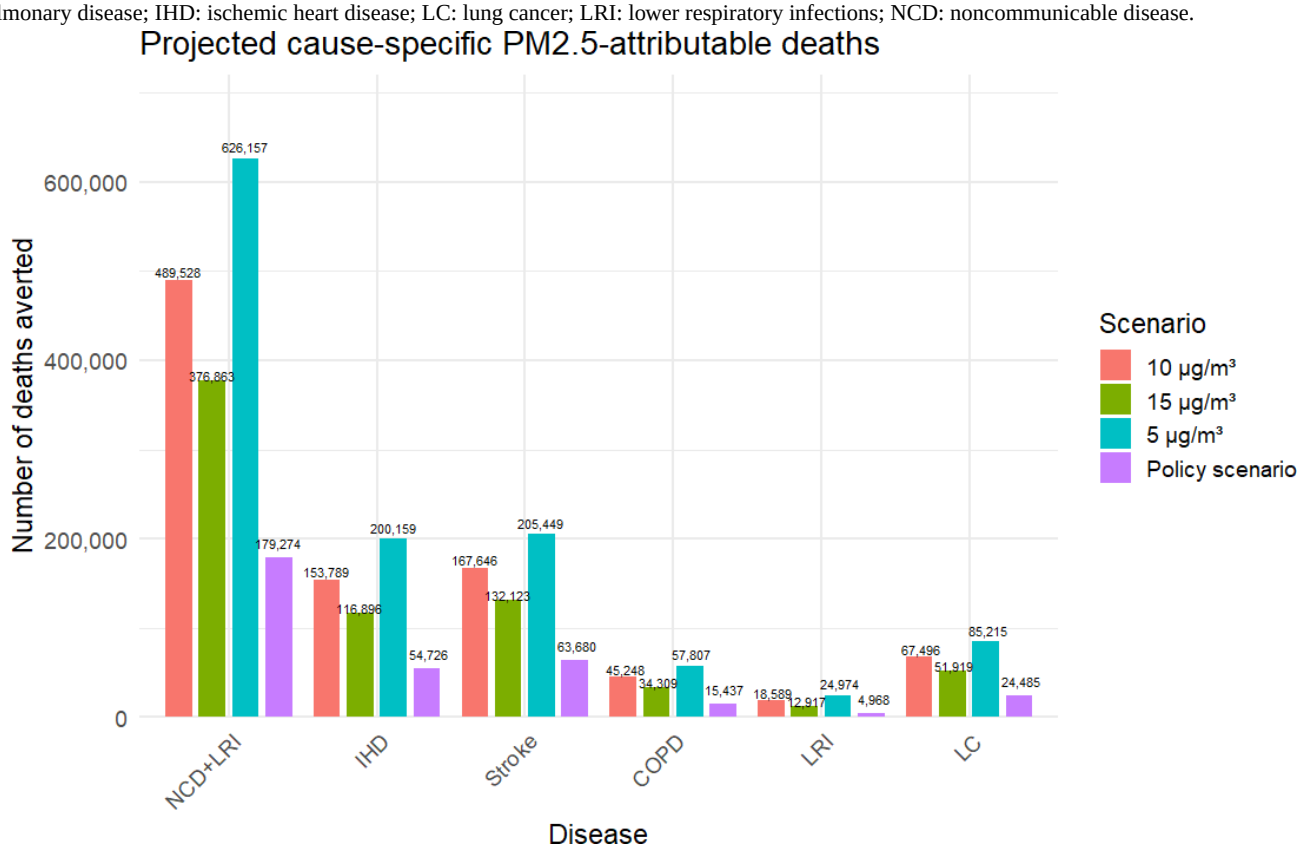


Figure 4 presents the annual average reduction in cause-specific premature deaths attributable to improved air quality for each modeled scenario under the SSP2 pathway. Under the policy scenario, a 10% reduction in PM2.5 concentrations by 2025, followed by a further reduction to 25 μg/m³ by 2030, is projected to result in an average annual reduction of 179,274 premature deaths. This includes 54,726 deaths from IHD, 63,680 from stroke, 15,437 from COPD, 4,968 from LRI, and 24,485 from LC. If PM2.5 concentrations were further reduced to the WHO's interim and final targets (15.10 and 5 μg/m³), the estimated annual reductions in total premature deaths would increase to 376,863, 489,528, and 626,157, respectively. The corresponding reductions in cause-specific mortality would range from 116,896 to 200,159 for IHD, 132,123 to 205,449 for stroke, 34,309 to 57,807 for COPD, 12,917 to 24,974 for LRI, and 51,919 to 85,215 for LC.

XSL•FO
RenderX

Figure 4. Projected annual average premature deaths by cause in China under improved air quality scenarios (2021 - 2030). COPD: chronic obstructive pulmonary disease; IHD: ischemic heart disease; LC: lung cancer; LRI: lower respiratory infections; NCD: noncommunicable disease.



Morbidity

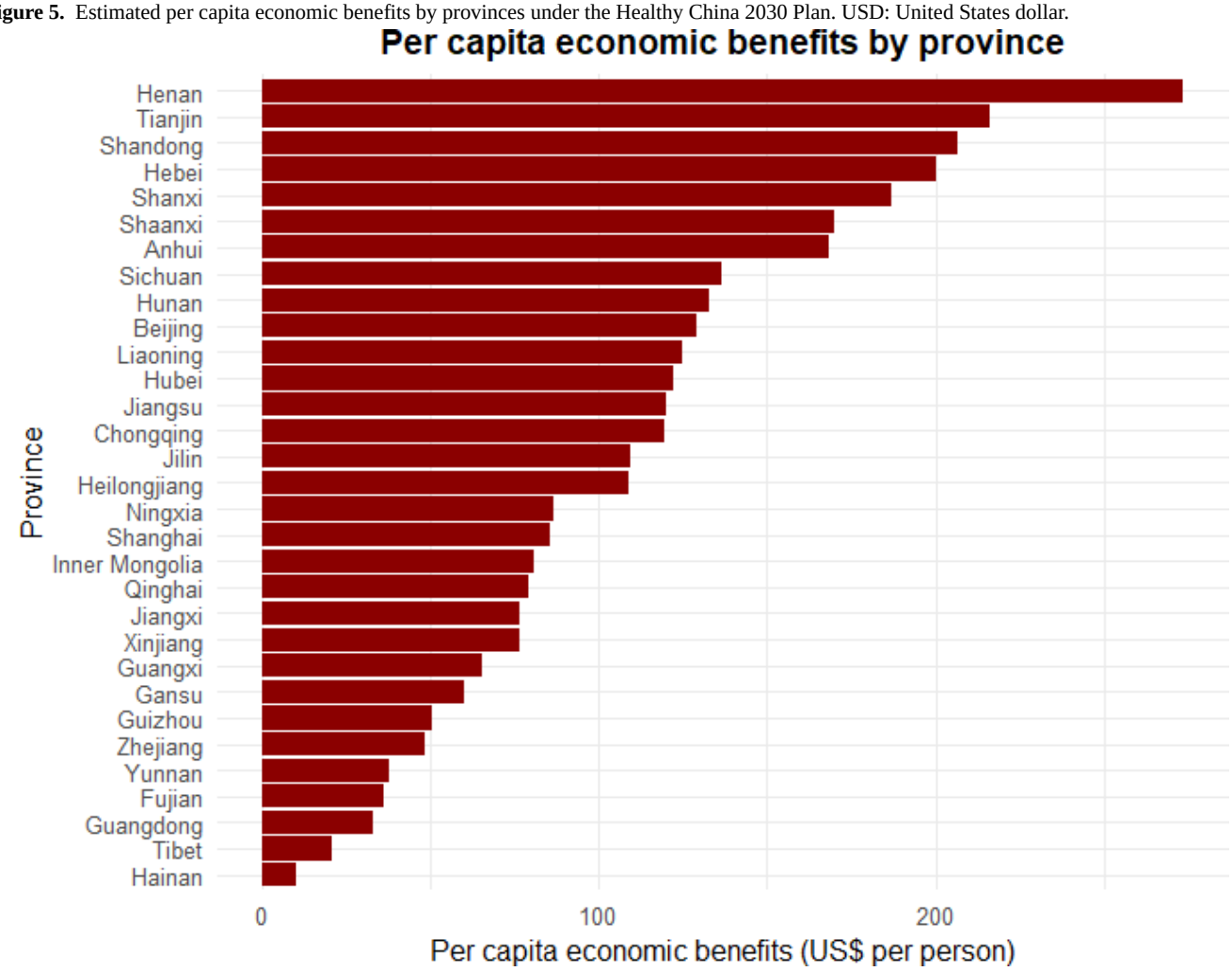
Table S6 in Multimedia Appendix 1 presents the annual average estimates of PM2.5-associated morbidity outcomes for the baseline scenario and the projected benefits under improved air quality scenarios. Under the policy scenario, the annual average morbidity benefits included 21,422 fewer cardiovascular disease admissions and 26,545 fewer respiratory disease admissions. These estimates increased to 41,690 and 51,681 cases under the 15 μg/m³ scenario, 51,884 and 64,333 cases under the 10 μg/m³ scenario, and 62,146 and 77,073 cases under the 5 μg/m³ scenario, respectively.

Economic Benefits

Figure 5 illustrates the estimated per capita annual average economic benefits associated with the modeled scenarios under SSP2 at the provincial level. Under the policy scenario, the annual average economic benefits were estimated at US $123.7 billion, corresponding to approximately US $122.7 per person per year, comprising US $123.5 billion from reductions in premature mortality and US $175.5 million in annual savings from hospitalization costs. With the adoption of more stringent air quality targets of 15, 10, and 5 μg/m³, total annual savings were projected to increase to US $185.7 billion, US $240.7 billion, and US $306.5 billion, respectively. Further, the spatial distribution of economic benefits closely followed PM2.5 concentration patterns, with eastern and southeastern provinces experiencing the most substantial gains, while western regions observed relatively lower benefits. Specifically, under the policy scenario, the per capita annual average economic benefits ranged from US $10.1 in Hainan to US $273.5 in Henan, with other provinces such as Shandong (US $206.4) and Hebei (US $200.5) also realizing substantial gains. Under the more stringent targets, the average economic benefits across all 31 provinces increased progressively under similar distribution patterns, with projected benefits ranging from US $51.1 per person in Hainan to US $644.6 per person in Henan under the most stringent 5 μg/m³ target.

XSL•FO
RenderX

Figure 5. Estimated per capita economic benefits by provinces under the Healthy China 2030 Plan. USD: United States dollar.



Scenario Analysis

Tables S7-S10 in Multimedia Appendix 1 provide summary estimates of the mortality and morbidity outcomes under different population trajectories. Annual reductions in total premature deaths ranged from 177,857 (under SSP4) to 180,500 (under SSP3) under the policy scenario, aligning closely with the SSP2 "middle of the road" estimates. The average annual reduction in morbidity was the highest under SSP3 (48,284 cases) and lowest under SSP5 (47,588 cases). Further reductions in premature deaths were observed when achieving lower PM2.5 targets of 15, 10, and 5 μg/m³, with estimated reductions ranging from 373,985 - 379,377 deaths, 485,767 - 492,843 deaths, and 621,288 - 630,487 deaths, respectively. Annual reductions in hospitalizations corresponded to 92,661 - 93,967 cases for 15 μg/m³, 115,333 - 116,965 cases for 10 μg/m³, and 138,159 - 140,119 cases for 5 μg/m³. The estimated annual economic benefits under the policy scenario ranged from US $123.1 billion to US $124.9 billion, further increasing to US $184.3 to US $186.8 billion for 15 μg/m³, US $238.9 to US $242.2 billion for 10 μg/m³, and US $304.2 to US $308.5 billion for 5 μg/m³.

Discussion

Principal Findings

In this study, we modeled the health and economic benefits of air quality improvements in China under several population and PM2.5 concentration scenarios. Our analysis suggests that attaining China's current national policy, which aims to reduce PM2.5 levels by 10% from 2021 to 2025 and reach 25 μg/m³ by 2030, would prevent an annual average of approximately 179,274 premature deaths and 47,967 hospitalizations, translating to economic benefits totaling US $123.7 billion (US $122.7 per capita). This would represent approximately 0.6% of China's national GDP in 2020. Under this policy scenario, the annual average reduction in premature deaths accounted for roughly 10% of the total PM2.5-attributable deaths estimated in 2020. Reductions in deaths from IHD and stroke accounted for most of these benefits, comprising 31.4% and 34.2% of the total annual average averted deaths, respectively. Under the VSL of US $689,659, monetized mortality benefits were overwhelmingly the largest source of economic gain, comprising over 99% of the total. The greatest benefits were observed in regions with high PM2.5 concentrations and population densities, such as the BTH, FP, and YDR region. At the provincial level, Shandong, Henan, and Hebei yielded the highest estimated annual benefits, equating to 1.0%, 1.6%, and 1.4% of their respective 2020 GDPs. Moreover, our model

XSL•FO
RenderX

estimates suggested that substantially greater health and economic benefits could be realized under more stringent air quality targets. Specifically, the additional annual average economic benefits were projected to be approximately 2.1, 2.7, and 3.5 times higher under the 5, 10, and 15 μg/m³ targets, respectively, compared to the national policy scenario. These results underscore the substantial potential for improved health and economic outcomes through more stringent air quality control, supporting the national government's ongoing efforts to enhance air quality and emphasizing the necessity to set more ambitious future guidelines.

Our study contributes to the literature in several ways. First, it provides an updated analysis to support identifying at-risk populations and future policy design by projecting health and economic outcomes in the future at the national level. Limited evidence exists in this area of research. Previously, Wang et al [45] quantified the health benefits for 100% and 50% improvements in air quality based on 2010 levels, reporting a reduction in the number of PM2.5-related premature deaths in China by approximately 129,278 by 2020 and 217,988 by 2030, compared with 2010. Since then, China has significantly reduced PM2.5 concentrations, with an annual average of 32.6 μg/m³ in 2020. With the introduction of its latest 14th 5-Year Action Plan and Healthy 2030 Plan, China is set to reach the target of 25 ug/m^3 by 2030 [22]. Additionally, in 2021, the WHO updated its air quality guidelines, recommending several updated interim and final targets to reflect the growing body of updated epidemiology evidence [19]. These updates in policies and air quality guidelines both nationally and internationally necessitate an updated analysis to support identifying at-risk populations and informing future policy design. Additionally, in this analysis, we use the GEMM to estimate mortality. Evidence shows that GEMM predicts mortality estimates over 120% higher than other established exposure-response functions, such as the WHO's Integrated Exposure-Response model or the log-normal model by considering for epidemiological evidence specific to ambient air pollution, as well as including evidence from high pollution regions [32]. Third, our findings further contribute to the global evidence demonstrating the substantial health and economic benefits of significant reductions in ambient air pollution, particularly in highly polluted countries. For example, research from India has shown significant reductions in excess mortality for both urban and rural populations following national actions to reduce PM [12,46]. Collectively, this body of evidence encourages governments in South and South-East Asia, which already bear a disproportionate burden from high PM exposure, to establish and strengthen effective national air quality regulatory standards. Fourth, we provide an economic perspective by quantifying the potential economic benefits of air quality improvements. Using both willingness to pay and COI estimates, our analysis translates health gains into monetary terms to better inform policy and facilitate decision-making.

Several limitations and future research questions should be acknowledged. As with all modeling analyses, the accuracy of our estimates is constrained by simplifying assumptions. First, we assumed that annual average PM2.5 concentrations serve as a proxy for long-term exposure, despite individual exposure being influenced by spatial-temporal variability, socioeconomic factors, and personal health behaviors [47,48]. Second, our economic valuation heavily relied on the VSL estimate, despite the lack of a nationally recognized figure. Previous VSL studies have reported values ranging from US $23,745 to US $614,805, highlighting the significant variability [38]. In our analysis, we opted for a recent VSL figure estimated in multiple cities representative of the national population. Third, due to the absence of city-level data on age-specific population trajectories, baseline mortality, morbidity, and health care utilization, we relied on provincial and regional data from statistical yearbooks and national death registries. Fourth, future advancements in health care services could significantly alter hospitalization rates and costs, yet our projections are based on historical trends. Further, a broader societal perspective on the benefits of reducing PM2.5 could be considered. For example, this study did not incorporate indirect benefits, such as avoided productivity losses. Additionally, previous studies have indicated that air pollution mitigation can involve substantial costs, including direct abatement costs and broader social and economic impacts [30,49]. Although these were beyond the scope of this study, they may affect the overall net economic benefit of such policies and warrant further investigation. Fifth, in our analysis, although population growth was projected using the shared socioeconomic pathways, we assumed that the population age structure remained unchanged over time. This is a notable limitation in the context of China's rapidly aging population, as epidemiological evidence indicates that older adults are more susceptible to the adverse health effects of PM [32]. Consequently, our estimates may underestimate the future disease burden attributable to air pollution and the corresponding health benefits achievable through air quality improvements. Finally, we assumed a steady decline in PM2.5 concentrations over our modeled period, while in reality, PM2.5 reductions occur through a nonlinear process.

Conclusion

We modeled the potential health and economic benefits of improving air quality in China from 2021 to 2030. Achieving a 10% reduction in national PM2.5 concentrations by 2025, with a target of 25 μg/m³ by 2030, was estimated to prevent an average of 179,274 premature deaths and 47,967 hospital admissions annually, corresponding to US $123.7 billion in annual average economic benefits. However, more ambitious targets could yield substantially greater benefits, with annual death reductions of 376,863, 489,528, and 626,157 and hospitalization reductions of 93,371, 116,217, and 139,219, under 15, 10, and 5 μg/m³ targets, respectively, representing a 2-fold to 3-fold increase compared to the national target. Our findings underscore the substantial additional health and economic benefits of stricter air quality policies and support the development of future evidence–based national air quality guidelines.

XSL•FO
RenderX

Acknowledgments

The authors attest that no generative artificial intelligence tools were used in the preparation of this manuscript.

Funding

This study was funded by the National Natural Science Foundation of China (grant numbers: 71874086, 72174093, and 72474109).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: LS, MC, SJ, SW
Data curation: LS, SW
Formal analysis: LS, SW
Methodology: LS, SJ, SW, ZX
Resources: LS, MC
Supervision: LS
Writing – original draft: SW
Writing – reviewing & editing: GLDT, LD, LS, MC, SJ, SW, YJ, ZX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Summary of model assumptions, parameters, and results of scenario analyses.
[DOCX File, 53 KB - publichealth_v12i1e84809_app1.docx]

Checklist 1
GATHER checklist.
[DOCX File, 22 KB - publichealth_v12i1e84809_app2.docx]

References

1. Ambient (outdoor) air pollution. World Health Organisation. URL: https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health [accessed 2026-01-22]
2. Pun VC, Kazemiparkouhi F, Manjourides J, Suh HH. Long-term PM2.5 exposure and respiratory, cancer, and cardiovascular mortality in older US adults. Am J Epidemiol 2017 Oct 15;186(8):961-969. [doi: 10.1093/aje/kwx166] [Medline: 28541385]
3. Chen J, Hoek G. Long-term exposure to PM and all-cause and cause-specific mortality: a systematic review and meta-analysis. Environ Int 2020 Oct;143:105974. [doi: 10.1016/j.envint.2020.105974]
4. Sun Y, Milando CW, Spangler KR, et al. Short term exposure to low level ambient fine particulate matter and natural cause, cardiovascular, and respiratory morbidity among US adults with health insurance: case time series study. BMJ 2024 Feb 21;384:e076322. [doi: 10.1136/bmj-2023-076322] [Medline: 38383039]
5. Karimi B, Samadi S. Long-term exposure to air pollution on cardio-respiratory, and lung cancer mortality: a systematic review and meta-analysis. J Environ Health Sci Eng 2024 Jun;22(1):75-95. [doi: 10.1007/s40201-024-00900-6] [Medline: 38887768]
6. Alexeeff SE, Deosaransingh K, Van Den Eeden S, Schwartz J, Liao NS, Sidney S. Association of long-term exposure to particulate air pollution with cardiovascular events in California. JAMA Netw Open 2023 Feb 1;6(2):e230561. [doi: 10.1001/jamanetworkopen.2023.0561] [Medline: 36826819]
7. Awe YA, Larsen BK, Sanchez-Triana E. The global health cost of PM2.5 air pollution a case for action beyond 2021. : World Bank Group; 2022 URL: https://documents1.worldbank.org/curated/en/455211643691938459/pdf/The-Global-Health-Cost-of-PM-2-5-Air-Pollution-A-Case-for-Action-Beyond-2021.pdf [accessed 2026-03-13]
8. Rentschler J, Leonova N. Global air pollution exposure and poverty. Nat Commun 2023 Jul 22;14(1):4432. [doi: 10.1038/s41467-023-39797-4] [Medline: 37481598]
9. Lukman KM, Hastuti LP, Oktavia D, Pratiwi D, Uchiyama Y, Harding D. Towards blue skies: a comprehensive review and regional mapping of ambient air quality in Indonesian cities. J Environ Manage 2025 Aug;389:126132. [doi: 10.1016/j.jenvman.2025.126132] [Medline: 40505559]
10. Nhung NTT, Duc VT, Ngoc VD, et al. Mortality benefits of reduction fine particulate matter in Vietnam, 2019. Front Public Health 2022;10:1056370. [doi: 10.3389/fpubh.2022.1056370] [Medline: 36466445]

XSL•FO
RenderX

11. Jainontee K, Pongkiatkul P, Wang YL, et al. Strategy design of PM2.5 controlling for Northern Thailand. Aerosol Air Qual Res 2023;23(6):220432. [doi: 10.4209/aaqr.220432]
12. Gopikrishnan GS, Kuttippurath J. Impact of the National Clean Air Programme (NCAP) on the particulate matter pollution and associated reduction in human mortalities in Indian cities. Sci Total Environ 2025 Mar 10;968:178787. [doi: 10.1016/j.scitotenv.2025.178787] [Medline: 39970552]
13. Maji KJ, Ye WF, Arora M, Shiva Nagendra SM. PM2.5-related health and economic loss assessment for 338 Chinese cities. Environ Int 2018 Dec;121(Pt 1):392-403. [doi: 10.1016/j.envint.2018.09.024] [Medline: 30245362]
14. Zhou M, Wang H, Zeng X, et al. Mortality, morbidity, and risk factors in China and its provinces, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet 2019 Sep;394(10204):1145-1158. [doi: 10.1016/S0140-6736(19)30427-1]
15. Zhang X, Ou X, Yang X, et al. Socioeconomic burden of air pollution in China: province-level analysis based on energy economic model. Energy Econ 2017 Oct;68:478-489. [doi: 10.1016/j.eneco.2017.10.013]
16. Xie Y, Dai H, Zhang Y, Wu Y, Hanaoka T, Masui T. Comparison of health and economic impacts of PM2.5 and ozone pollution in China. Environ Int 2019 Sep;130:104881. [doi: 10.1016/j.envint.2019.05.075] [Medline: 31200152]
17. The State Council issued the "Action Plan for Continuous Improvement of Air Quality" [Website in Chinese]. Central People's Government of the People's Republic of China. URL: https://www.gov.cn/yaowen/liebiao/202312/content_6919218.htm [accessed 2025-03-08]
18. Huang J, Pan X, Guo X, Li G. Health impact of China's Air Pollution Prevention and Control Action Plan: an analysis of national air quality monitoring and mortality data. Lancet Planet Health 2018 Jul;2(7):e313-e323. [doi: 10.1016/S2542-5196(18)30141-4] [Medline: 30074894]
19. WHO global air quality guidelines: particulate matter (PM25 and PM10) , ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. : World Health Organisation; 2021 URL: https://iris.who.int/server/api/core/bitstreams/551b515e-2a32-4e1a-a58c-cdaecd395b19/content [accessed 2026-03-13]
20. Sangkham S, Phairuang W, Sherchan SP, et al. An update on adverse health effects from exposure to PM2.5. Environ Adv 2024 Dec;18:100603. [doi: 10.1016/j.envadv.2024.100603]
21. Han C, Xu R, Ye T, et al. Mortality burden due to long-term exposure to ambient PM2.5 above the new WHO air quality guideline based on 296 cities in China. Environ Int 2022 Aug;166:107331. [doi: 10.1016/j.envint.2022.107331] [Medline: 35728411]
22. The CPC central committee and the state council issued the "Outline of the 'Healthy China 2030' plan" [Article in Chinese]. Central People's Government of the People's Republic of China. URL: https://www.gov.cn/xinwen/2016-10/25/content_5124174.htm [accessed 2025-03-08]
23. Zheng Y, Xue T, Zhao H, Lei Y. Increasing life expectancy in China by achieving its 2025 air quality target. Environ Sci Ecotechnol 2022 Oct;12:100203. [doi: 10.1016/j.ese.2022.100203] [Medline: 36157339]
24. Lim SS, Vos T, Flaxman AD, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2224-2260. [doi: 10.1016/S0140-6736(12)61766-8] [Medline: 23245609]
25. Zheng S, Schlink U, Ho KF, Singh RP, Pozzer A. Spatial distribution of PM(2.5)-related premature mortality in China. Geohealth 2021;5(12):e2021GH000532. [doi: 10.1029/2021GH000532]
26. Liang F, Xiao Q, Huang K, et al. The 17-y spatiotemporal trend of PM2.5 and its mortality burden in China. Proc Natl Acad Sci U S A 2020 Oct 13;117(41):25601-25608. [doi: 10.1073/pnas.1919641117] [Medline: 32958653]
27. Xie R, Sabel CE, Lu X, et al. Long-term trend and spatial pattern of PM2.5 induced premature mortality in China. Environ Int 2016 Dec;97:180-186. [doi: 10.1016/j.envint.2016.09.003] [Medline: 27614532]
28. Song J, Li C, Hu Y, Zhang J, Kang J, Zhao L. Premature deaths and economic losses attributable to long- and short-term PM 2.5 exposure in China. Ecosyst Health Sustain 2023 Jan;9:0111. [doi: 10.34133/ehs.0111]
29. Xu M, Qin Z, Zhang S, Xie Y. Health and economic benefits of clean air policies in China: a case study for Beijing-Tianjin-Hebei region. Environ Pollut 2021 Sep 15;285:117525. [doi: 10.1016/j.envpol.2021.117525] [Medline: 34380223]
30. Zhang J, Jiang H, Zhang W, et al. Cost-benefit analysis of China's Action Plan for Air Pollution Prevention and Control. Front Eng Manag 2019 Dec;6(4):524-537. [doi: 10.1007/s42524-019-0074-8]
31. Pozzer A, Anenberg SC, Dey S, Haines A, Lelieveld J, Chowdhury S. Mortality attributable to ambient air pollution: a review of global estimates. Geohealth 2023 Jan;7(1):e2022GH000711. [doi: 10.1029/2022GH000711] [Medline: 36636746]
32. Burnett R, Chen H, Szyszkowicz M, et al. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. Proc Natl Acad Sci U S A 2018 Sep 18;115(38):9592-9597. [doi: 10.1073/pnas.1803222115] [Medline: 30181279]
33. Burnett RT, Pope CA 3rd, Ezzati M, et al. An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. Environ Health Perspect 2014 Apr;122(4):397-403. [doi: 10.1289/ehp.1307049] [Medline: 24518036]

XSL•FO
RenderX

34. Fang D, Wang Q, Li H, Yu Y, Lu Y, Qian X. Mortality effects assessment of ambient PM2.5 pollution in the 74 leading cities of China. Sci Total Environ 2016 Nov 1;569-570:1545-1552. [doi: 10.1016/j.scitotenv.2016.06.248] [Medline: 27395080]
35. Wyatt LH, Xi Y, Kshirsagar A, et al. Association of short-term exposure to ambient PM2.5 with hospital admissions and 30-day readmissions in end-stage renal disease patients: population-based retrospective cohort study. BMJ Open 2020 Dec 15;10(12):e041177. [doi: 10.1136/bmjopen-2020-041177] [Medline: 33323443]
36. Mortality risk valuation. US Environmental Protection Agency. URL: https://www.epa.gov/environmental-economics/mortality-risk-valuation [accessed 2025-03-03]
37. Keller E, Newman JE, Ortmann A, Jorm LR, Chambers GM. How much is a human life worth? A systematic review. Value Health 2021 Oct;24(10):1531-1541. [doi: 10.1016/j.jval.2021.04.003] [Medline: 34593177]
38. Cao C, Song X, Cai W, et al. Estimating the value of a statistical life in China: a contingent valuation study in six representative cities. Chin J Popul Resour Environ 2023 Dec;21(4):269-278. [doi: 10.1016/j.cjpre.2023.11.009]
39. Jo C. Cost-of-illness studies: concepts, scopes, and methods. Clin Mol Hepatol 2014 Dec;20(4):327-337. [doi: 10.3350/cmh.2014.20.4.327] [Medline: 25548737]
40. China guidelines for pharmacoeconomic evaluations 2020 edition. : Chinese Pharmaceutical Association, China Society for Pharmacoeconomics and Outcomes Research, and ISPOR Beijing Chapter; 2020 URL: https://consultorsalud.com/wp-content/uploads/2021/06/China-Guidelines-for-Pharmacoeconomic-Evaluations-2020.pdf [accessed 2026-03-13]
41. Liu S, Wu X, Lopez AD, et al. An integrated national mortality surveillance system for death registration and mortality surveillance, China. Bull World Health Organ 2015 Jan 1;94(1):46-57. [doi: 10.2471/BLT.15.153148] [Medline: 26769996]
42. Li Y, Zeng X, Liu J, et al. Can China achieve a one-third reduction in premature mortality from non-communicable diseases by 2030? BMC Med 2017 Jul 11;15(1):132. [doi: 10.1186/s12916-017-0894-5] [Medline: 28693510]
43. China statistical yearbook. National Bureau of Statistics of China. URL: https://www.stats.gov.cn/english/Statisticaldata/yearbook/ [accessed 2025-03-03]
44. China Health Statistical Yearbook. National Bureau of Statistics of China. URL: https://www.nhc.gov.cn/mohwsbwstjxxzx/tjzxtjcbw/tjsj_list.shtml [accessed 2026-01-31]
45. Wang Q, Wang J, Zhou J, Ban J, Li T. Estimation of PM2·5-associated disease burden in China in 2020 and 2030 using population and air quality scenarios: a modelling study. Lancet Planet Health 2019 Feb;3(2):e71-e80. [doi: 10.1016/S2542-5196(18)30277-8] [Medline: 30797415]
46. Pathak M, Kuttippurath J. Elucidating the changing particulate matter pollution and associated health effects in rural India during 2000-2019. Environ Pollut 2024 May 1;348:123830. [doi: 10.1016/j.envpol.2024.123830] [Medline: 38518972]
47. Lim S, Bassey E, Bos B, et al. Comparing human exposure to fine particulate matter in low and high-income countries: a systematic review of studies measuring personal PM2.5 exposure. Sci Total Environ 2022 Aug 10;833:155207. [doi: 10.1016/j.scitotenv.2022.155207] [Medline: 35421472]
48. McCarron A, Semple S, Braban CF, Gillespie C, Swanson V, Price HD. Personal exposure to fine particulate matter (PM2.5) and self-reported asthma-related health. Soc Sci Med 2023 Nov;337:116293. [doi: 10.1016/j.socscimed.2023.116293] [Medline: 37837949]
49. Zhang F, Xing J, Zhou Y, et al. Estimation of abatement potentials and costs of air pollution emissions in China. J Environ Manage 2020 Apr 15;260:110069. [doi: 10.1016/j.jenvman.2020.110069] [Medline: 32090813]

Abbreviations

BTH: Beijing-Tianjin-Hebei
COI: cost-of-illness
COPD: chronic obstructive pulmonary disease
FP: Fenwei Plain
GATHER: Guidelines for Accurate and Transparent Health Estimates Reporting
GDP: gross domestic product
GEMM: global exposure mortality model
IHD: ischemic heart disease
LC: lung cancer
LRI: lower respiratory infection
NCD: noncommunicable disease
PM: particulate matter
SSP: Shared Socioeconomic Pathway
VSL: value-of-statistical life
WHO: World Health Organization
YDR: Yangtze Delta

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 25.Sep.2025; peer-reviewed by J Kuttippurath, Y Ding; revised version received 06.Feb.2026; accepted 25.Feb.2026; published 01.Apr.2026.

Please cite as:
Wang S, Xu Z, Di Tanna GL, Jiang Y, Chen M, Downey L, Jan S, Si L
Projected Health and Economic Benefits of Air Quality Targets in China: Modeling Study
JMIR Public Health Surveill 2026;12:e84809
URL: https://publichealth.jmir.org/2026/1/e84809
doi:10.2196/84809

© Siyuan Wang, Zhiwei Xu, Gian Luca Di Tanna, Yawen Jiang, Mingsheng Chen, Laura Downey, Stephen Jan, Lei Si. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 1.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Assessing the Effectiveness of Crowdsourced Data to Detect Established Tick Populations in Quebec, Canada: Retrospective Ecological Study

Ariane Dumas[1,2,3], PhD; Jade Savage[4], PhD; Kirsten E Crandall[5], PhD; Jules K Koffi[2,6], PhD; Patrick A Leighton[2,3], PhD; Nicholas H Ogden[1,2], BVSc, DPhil, FCAHS; Erin E Rees[1,2,3], PhD; Catherine Bouchard[1,2,3], PhD, DVM

[1]Modelling Hub Division, Applied Public Health Sciences Directorate, Science and Policy Integration Branch, Public Health Agency of Canada, 200, boulevard René-Lévesque Ouest, Tour Est, Montréal, QC, Canada

[2]Groupe de Recherche en Épidémiologie des Zoonoses et Santé Publique, St-Hyacinthe, QC, Canada

[3]Département de pathologie et microbiologie, Faculté de médecine vétérinaire, Université de Montréal, St-Hyacinthe, QC, Canada

[4]Department of Biology and Biochemistry, Bishop's University, Sherbrooke, QC, Canada

[5]Unité Maladies infectieuses en communauté, Direction des risques biologiques, Institut national de santé publique du Québec, Montreal, QC, Canada

[6]Zoonoses Division, Centre for Foodborne, Environmental and Zoonotic Infectious Diseases, Public Health Agency of Canada, Saint-Hyacinthe, QC, Canada

Corresponding Author:
Ariane Dumas, PhD
Modelling Hub Division, Applied Public Health Sciences Directorate, Science and Policy Integration Branch, Public Health Agency of Canada, 200, boulevard René-Lévesque Ouest, Tour Est, Montréal, QC, Canada

Abstract

Background: The range expansion of ticks transmitting zoonotic diseases in Canada poses challenges for public health surveillance. Active monitoring of where tick populations are emerging helps inform public health responses, but is resource intensive and logistically complex. Crowdsourced data may provide a cost-effective approach to augment surveillance. In this study, we assessed the value of crowdsourced data from eTick, a platform whereby the public submits tick photographs for expert identification.

Objective: We aimed to (1) characterize spatial patterns of eTick submissions and identify socioecological factors associated with their occurrence and frequency, and (2) assess the ability of indicators derived from eTick to predict census subdivisions (CSDs) with established tick populations.

Methods: We conducted a retrospective ecological assessment (2020 - 2023) for the effectiveness of crowdsourced data to detect established blacklegged tick (*Ixodes scapularis*) populations in Quebec, Canada. First, we applied a spatial scan approach to assess significantly high submission rates in human- and animal-origin data separately. Then, using zero-inflated regression models, we assessed socioecological factors associated with the occurrence and counts of tick submissions. Next, we used logistic regression models to assess the value of using eTick surveillance data to predict the presence of established tick populations, as documented by active surveillance. Finally, we calculated thresholds for eTick submissions to maximize sensitivity, specificity, or a balance of both metrics for detecting established tick populations, and developed scenarios to interpret linear predictors of submission counts across different socioecological contexts.

Results: No clusters in submissions were detected for human and animal-origin data. However, high concentrations of ticks found on humans corresponded with areas reporting a high risk of Lyme disease. Tick submissions of animal origin extended beyond endemic areas, highlighting their complementary surveillance value. Zero-inflated models identified socioeconomic, climatic, and land cover factors significantly associated with submission rates. Logistic regression models, including significant socioecological factors, performed best to identify CSDs with established tick populations. Predictive performance was similar for ticks from humans (area under the curve [AUC] 0.86), animals (AUC 0.83), and all submissions combined (AUC 0.85). Using the best model adjusted for population size, median income, median age, and cumulative degree days above 0 °C, an optimal threshold of 3.32 crowdsourced tick submissions allowed the detection of CSDs with established tick populations while maximizing sensitivity and specificity (both 0.76).

Conclusions: This work introduces a framework for analyzing crowdsourced surveillance data by correcting for participation bias and benchmarking against active surveillance. Unlike approaches relying on traditional or unadjusted data, it incorporates socioecological factors associated with reporting. We show that crowdsourced data can yield reliable inferences about vector distributions and provide a methodological approach that could be applied across species and regions. Applied to eTick, it supports

XSL•FO
RenderX

integration into surveillance systems to improve early detection, targeting interventions, and cost-effective monitoring of emerging risks by public health.

(JMIR Public Health Surveill 2026;12:e90734) doi:10.2196/90734

KEYWORDS

surveillance; tick-borne disease; Lyme disease; crowdsourced data; tick exposure; tick digital images; participatory science; socioecological factors; climate suitability; landcover

Introduction

Tick-borne diseases represent a growing public health concern in Canada. In recent decades, several tick species have expanded their geographic ranges and population sizes, largely driven by climate and land use changes [1,2]. Notably, *Ixodes scapularis*, the primary vector of the agent of Lyme disease (LD), the most commonly reported vector-borne illness in Canada, has seen significant range expansion [3-5]. Geographic ranges of other species transmitting various pathogens, such as *Ixodes cookei* and *Dermacentor* spp., are also expanding, thereby increasing current and future risks of tick-borne diseases in Canada [1,6,7]. In this context, national and provincial public health authorities across Canada have implemented surveillance programs to identify where tick populations are spreading and to monitor disease incidence trends. These efforts identify where tick-borne disease risk is occurring, allowing targeted promotion of public awareness of risk and prevention methods and providing guidance to health care professionals [8]. However, tick presence and disease risk vary regionally due to ecological and environmental factors, including the availability of suitable hosts for ticks and pathogens, climate and habitat suitability for tick survival, and human exposure [1,9]. Canada's vast territory and diverse environments present major challenges for conducting consistent nationwide tick surveillance [10]. Tick-borne diseases demonstrate the importance of a One Health approach, as disease risk emerges from interactions among environmental conditions, animal hosts, and human behavior. While tick occurrence depends on climate, habitat, and wildlife hosts, exposure of humans and domestic animals is shaped by patterns of outdoor activity and contact with suitable tick habitat [11]. Domestic animals, which are frequently exposed to ticks, can serve as sentinels of local tick presence, underscoring the value of integrated surveillance approaches that integrate data from multiple sources (ie, humans, pets, and wildlife) and the environment [11-14].

Traditional methods for tick surveillance include active and passive surveillance conducted by public health organizations. In active surveillance, field teams perform standardized environmental sampling, generally involving drag sampling to collect questing ticks [10]. This approach provides strong scientific rigor because it allows the collection of tick abundance data, ensuring repeatability and comparability across sites due to standardized protocols for sampling and site selection [15,16]. However, this approach is also labor-intensive, and its inherent high costs limit its use geographically and temporally [17]. In passive surveillance, wider geographic coverage is possible at a lower cost, as ticks are voluntarily submitted to public health laboratories by the general population via health care professionals, medical clinics, or veterinary clinics. These laboratories undertake tick species identification and molecular screening for targeted pathogens [18]. Historically, passive surveillance has played a central role in detecting the emergence of new tick populations in Canada, providing early warning signals [18-20]. However, it has been gradually phased out in most regions across Canada due to insufficient laboratory resources to handle the increasing submission volumes associated with both higher numbers of ticks in endemic regions and expanding tick species distributions [16,21,22]. Consequently, an alternative passive surveillance framework is needed to ensure continued monitoring of emerging tick species and tick-borne diseases and to enable long-term data collection critical for understanding the spread of tick-borne diseases [23].

Citizen science approaches such as crowdsourcing-based data are increasingly used to complement field observations to estimate species distributions [24]. Crowdsourcing-based tick data collection is an emerging approach that engages the public to contribute to monitoring efforts via web platforms and applications [25]. Numerous examples from Europe and North America demonstrate the utility of such participatory approaches in complementing traditional tick surveillance methods [26-30]. The benefits of these approaches for public health include access to large volumes of data, broad geographic coverage, timely observations, and increased public engagement and risk awareness [31]. In 2017, eTick was launched as the first crowdsourcing platform dedicated to tick data collection in Canada. The general public, veterinarians, and health care professionals are invited to submit digital images of ticks found on people, on animals, or in the environment for free professional identification using a validated method [32]. When submitting a tick, users mark the location on a map to identify the presumed location of acquisition and provide additional information about recent travel. Submitters receive tick species identification results along with health prevention recommendations within 2 business days, and the municipalities where ticks were found are displayed on a publicly accessible map.

Multiple factors are thought to influence the data structure in citizen science and passive tick surveillance initiatives and must be identified and accounted for if the data are to be unbiased and useful for driving public health actions. Participation varies with age and sex [33-36], personal values, intrinsic interest, and a sense of belonging to a community [31,37-39], as well as education, income, and ethnicity, which may either encourage participation through increased awareness and risk perception or represent barriers to participation [31,34,36,37,39-41]. In tick-borne disease initiatives, pet owners are typically the most

XSL•FO
RenderX

represented group, along with people living in houses with yards and those regularly doing outdoor recreational activities [35].

Geographical biases in sampling are a common issue with crowdsourced data. Areas that are more accessible and that offer higher recreational interest, such as public parks and protected areas, may yield more observations [42,43]. Most tick submissions tend to be associated with peridomestic exposures [44], but these are influenced by the degree with which human settlement is juxtaposed with wildland [45]. The geographic scope of tick surveillance is limited by the geographic scope of the human populations. Tick submission numbers are proportional to the size of the participating human population and must be accounted for in data analysis [18-20,40,41]. Furthermore, expectations of spatial accuracy of tick encounter locations identified in passive surveillance must be realistic, as ticks can remain attached to their hosts for several days and often go unnoticed [21,46]. If possible, tick data from submitters with a recent history of long-distance travel should be eliminated from analysis [47].

A significant potential weakness of crowdsourced data is that, for some species, individual specimens may be found outside their geographic range due to their dispersion by migratory animals [48]. Crowdsourced data from these species require careful analysis and processing to avoid overestimating geographic ranges and ecological niches. For many tick species, ticks found by the public do not necessarily indicate locations where annually reproducing tick populations have become established. Many tick species are dispersed over long distances from their source populations by hosts (particularly migratory birds), and these ticks can survive to molt and feed as the next instar on hosts, including humans and their pets [48]. Locations where submitted ticks comprise only these "adventitious ticks" must be distinguished from locations with established tick populations [20]. The primary method for making this distinction involves determining a threshold number of submitted ticks per geographic unit based on statistical comparisons of the submission rates from locations with and without established populations identified through "gold standard" active surveillance [20]. Including environmental data that determine where tick populations can persist (habitat, host densities, and climate [1,9,49,50]) in analyses of passive surveillance data can help discriminate ticks coming from established populations versus adventitious ticks [20].

This study aimed to develop robust, evidence-based indicators for assessing the risk of tick population establishment using crowdsourced tick surveillance data. Specifically, our objectives were to (1) characterize spatial patterns in eTick submission rates and identify socioecological factors influencing their occurrence and frequency and (2) assess the ability of indicators derived from eTick data to predict census subdivisions (CSDs) with established tick populations. We focused on *I scapularis* in Quebec, Canada, where long-term active surveillance provides a reliable benchmark for validation. By leveraging Quebec's extensive dataset, this study generated practical, scalable insights to improve tick surveillance and risk assessment frameworks across Canada and in comparable international contexts.

Methods

Study Design and Setting

In this study, we conducted a retrospective ecological assessment for the effectiveness of crowdsourced data to detect established tick populations in Quebec, Canada, from 2020 to 2023. Here, an established tick population refers to a locally reproducing population considered persistent over time. The criteria used to infer establishment followed the *Institut national de santé publique du Québec's* (INSPQ) risk classification approach [51], itself based on [52] (see section Active Tick Surveillance for details). The unit of analysis was the municipality, defined as CSDs, totaling 1282 across the province [53], representing the finest spatial scale at which all relevant data were available. Approximately 80% of the province's population of 8 million lives in metropolitan areas, concentrated in the south, with the most populated being Montréal, Québec City, and Sherbrooke [54]. The province has seen ticks and LD emerging gradually over the last two decades. The most endemic regions, with the longest history of emergence, are the southernmost regions of Montérégie and Estrie, near the US border [55,56]. Since initial emergence, endemic regions have progressively expanded across southern Quebec, and human LD cases have been constantly rising. The dominant forest types are deciduous and mixed forests, and the potential exposure to ticks varies depending on land use and human activities, ranging from peridomestic and rural environments to recreational areas, such as natural parks in urban settings.

Data Sources and Variable Construction

Active Tick Surveillance

Active surveillance data were obtained from the INSPQ for 2014 to 2023. The INSPQ is Quebec's public health expertise and reference center, playing a key role in the surveillance and prevention of LD and other tick-borne diseases. Active surveillance was conducted at sampling sites located in public nature parks or on private lands with owner authorization. Site selection was based on habitat suitability for ticks (eg, deciduous or mixed forests) and signals of emerging risk, such as reported human LD cases or *I scapularis* collected through active and passive surveillance. Sampling took place between May and August under conditions without rain and with dry soil. At each site, a standardized drag sampling protocol was used. A team of 2 field technicians sampled each site, using a 1-m² flannel cloth dragged horizontally along 2 transects: one along the trail edge and one 25 m into the forest, parallel to the trail. Each technician sampled 1000 m², for a total sampled area of 2000 m² per site. The flannel cloths were inspected every 25 m for ticks. Ticks were removed with tweezers, placed in tubes containing 70% ethanol, and sent to the Quebec Public Health Laboratory (*Laboratoire de santé publique du Québec*) for species identification using taxonomic keys and pathogen testing [57]. Counts of the 3 life stages of *I scapularis* (larvae, nymphs, and adults) were calculated for each site visit and used to classify CSDs as having established tick populations or not. A CSD was considered established if at least six ticks of any life stage or at least two different life stages were collected within the same year in the municipality. If the indicator was met for at least 1

XSL•FO
RenderX

year during our study period, the municipality was considered as having an established *I scapularis* population [52,58].

Crowdsourced Tick Surveillance

Crowdsourced tick surveillance data were obtained from the eTick platform. We extracted all submissions identified as *I scapularis* for 2020 to 2023 in Quebec. Although eTick was launched in 2017, this time frame was considered representative of a stable period for analysis, following an initial phase of increasing familiarity and adoption of the eTick platform by the public. Retrospective analysis of the number of tick submissions suggested that the COVID-19 pandemic did not have a large effect on submissions, possibly due to limited restrictions on peridomestic activity and outdoor pursuits. Furthermore, the launch of the mobile eTick application in early 2020, just prior to lockdown, likely facilitated engagement and sustained user participation.

Data reported by users for each submission included the geographic location where the tick was found (ie, users marked the presumed location of acquisition on a zoomable map, from which the research team extracted coordinates), the date it was found, and whether the tick was found attached to an animal, attached to a human, or unattached (eg, crawling on a host or free in the environment), which is referred to as the origin of the tick throughout the paper. Additionally, for cases where the tick was attached to a human, users reported the age and gender of the individual. Users also reported whether they had traveled outside their municipality within 2 weeks prior to finding the tick. Starting in 2023, data collection also included information from the submitting individual regarding their profession (health care professional, veterinarian, or general public participant). Finally, additional details were provided by eTick's professional entomologist team: the tick life stage and sex, and unique unidentifiable user IDs. A public map is available on the platform, displaying tick submission locations, along with the tick species, locality, date of observation, origin, and tick pictures. All other information provided by users remains confidential.

The total number of individual submissions per user was calculated, and the distribution of the number of submissions per user was inspected graphically to detect potential outliers (ie, superusers). Superusers were defined a priori as users with submission counts exceeding the 99.9th percentile of the distribution.

Total tick submissions by origin type and overall (including unknown origins) were summed at the CSD level, given submission locations found within CSD boundaries. These counts were calculated both with and without the superuser outliers. Ticks that could have been acquired during travel were excluded from the analysis to reduce potential bias from submissions not representing the geographic location of tick exposure [19,20].

Socioecological Factors

Here, socioecological factors refer to demographic and socioeconomic characteristics that may influence community participation and environmental conditions known to affect tick populations. Corresponding variables were constructed and considered as predictors or confounders in the statistical models described below (see section Analysis for Details).

CSD statistics for human population size, median age, and median household income were obtained from the 2021 census [53,59]. All monetary values are reported in Canadian dollars. The conversion rate at the time of the study was CAD $1=US $0.72. Missing data occurred for these variables at the CSD level, primarily due to census data suppression or nonrelease resulting from confidentiality rules, low population counts, or high nonresponse rates [60]. These missing data largely occurred in remote or Indigenous CSDs, which represent subpopulations that differ from the population relevant to this study (ie, potential users of the eTick platform). CSDs with missing values for any of these variables were therefore not included in the analysis (194 CSDs excluded, representing 15% of Quebec's 1282 CSDs).

A variable derived from Google Trends was used to account for CSD-level participation toward eTick [61,62]. Google Trends is an open-access tool to extract information about Google search trends [63]. We assumed that Google searches that included keywords "ticks" and "Lyme disease," in either English or French, represented interest in ticks and LD, and thus served as a proxy to contribute to eTick. Google Trends data for the study variable were extracted at the municipal level for 2020 to 2023. Municipalities were then matched to their corresponding CSD based on name. The returned data use a relative scale from 0 to 100, where 100 represents the highest proportion of keyword searches relative to the total searches in a given region and period, and 0 represents the lowest level of interest [64].

A variable representing human exposure to tick habitats was derived as the proportion of interface zones between human settlements and wildlands. Data from the global wildland-urban interface (WUI) [65] were used to compute the WUI proportion per CSD. The calculation included all WUI categories dominated by forest, shrub, and wetland types, including both intermix and interface zones. The full methodology used to produce this layer has been published [65]. In brief, the WUI classifies areas into 2 categories: intermix and interface. The intermix WUI has ≥6.17 buildings per km² and ≥50% wildland vegetation. The interface WUI has ≥6.17 buildings per km² and <50% wildland vegetation but is located within 2.4 km of a large patch (≥5 km²) of wildland with >75% vegetation cover.

Finally, to compute annual values of cumulative degree days (DD) >0 °C, we used the Daymet dataset [66], which provides daily meteorological data at a 1 km spatial resolution across North America by integrating data from weather stations and other supporting sources through interpolation methods. This variable served as a proxy for 2 factors. First, the yearly exposure duration to host-seeking ticks [41], and second, a determinant of the potential for tick population establishment, given an approximate constraint of 2843 DD >0 °C to maintain populations [67,68]. Yearly CSD values were calculated as the mean 1-km resolution of pixel values whose centroids fell within CSD boundaries.

XSL•FO
RenderX

Ethical Considerations

The eTick platform collects routine passive surveillance data on ticks in Canada. On the basis of the nature of the data collected and the collection process, and acknowledging that no formal research question was specifically formulated to acquire these data, the Bishop's University Research Ethics Board stated that no ethical review is required for data collection through the eTick platform (Multimedia Appendix 1). Here, we make a secondary use of nonidentifiable data; individuals cannot be identified from any images or materials in the paper or supplementary files, a case defined in the Canadian research ethics guidelines for research involving humans (TCPS 2) as not requiring ethical review [69].

Before submission, all eTick users must consent to the use of their deidentified data for surveillance purposes and for future unspecified research [70]. Personal information (name, username, email address) is collected exclusively for communication purposes and visible only to site administrators. Data extracts used for surveillance or research do not include personal information in their data fields. No compensation was provided to participants.

Analysis

All analyses were conducted in R version 4.5.1 (R Foundation for Statistical Computing) [71].

Objective 1: Spatial Patterns and Socioecological Factors Associated With eTick Submissions

Assessing for Clusters in Tick Submissions

Spatial patterns in tick submissions were analyzed separately for tick submissions originating from humans and animals. The total CSD-level submissions for both types of submissions over the study period were visualized as rates normalized by CSD population size. Rates were calculated as: (number of submissions in the CSD/CSD human population)×100,000 and then mapped using QGIS (version 3.44.1).

The spatial scan statistic assessed for significantly high rates in the number of submissions at the CSD level [72]. CSD locations were represented by the centroid coordinates. The method used a circular scanning window that moved across the study area, varying in size from 0% to 50% of the study population, to identify potential windows in which submission rates are significantly higher than in surrounding areas [73]. Prior to the scan, overdispersion in submission counts was assessed using the Dean test [74]. As significant overdispersion was present, spatial scan analyses were conducted under a negative binomial distribution, implemented in the *DCluster* package in R [75]. As shown by Loh and Zhu [76], this approach is appropriate for overdispersed count data and reduces the risk of detecting false-positive clusters driven by extravariability rather than true spatial clustering [73]. The statistical significance of each potential cluster was evaluated using Monte Carlo simulation with 999 random permutations. Clusters were considered statistically significant at $P \geq .05$. The spatial scan analysis was performed separately for human- and animal-origin data.

Identifying Socioecological Factors Associated With Tick Submission Occurrence and Counts

We assessed for socioecological factors associated with the occurrence and counts of tick submissions, at the CSD-year levels, using zero-inflated regression models.

Logarithmic transformations were first applied to normalize skewed predictor variable distributions, then all predictors were centered and scaled to ensure appropriate model behavior (ie, reduce multicollinearity, help produce reliable model estimates, and facilitate interpretation of model estimates). Collinearity among variables was verified using Pearson correlation coefficients and the variance inflation factor (using maximal thresholds of 0.7 and 3, respectively [77]). Significant overdispersion was detected based on the dispersion ratio, and the presence of excess zeros was confirmed through the Vuong test.

A zero-inflated negative binomial model was fitted using the *pscl* package, version 1.5.9 [78]. These models have 2 distinct components: a logistic component that models the probability of having structural zeros (ie, populations not at risk; in our case, regions unsuitable for the presence of tick populations) and a count component that models the number of events (including zero counts) in the at-risk population (in our case, regions suitable for tick presence, where tick submissions may or may not occur). A backward stepwise selection procedure was used for final variable selection [79].

Model fit was assessed through graphical inspection of the residuals. Cook distance was calculated to identify potential outliers with a threshold of 1. When outliers were detected, models were refitted excluding the outlier CSDs. Observations were considered influential if their associated difference in beta exceeded the conventional threshold ($2/\sqrt{n}$) for any key coefficient or if their removal altered the statistical significance of predictors [80]; such CSDs were excluded; otherwise, they were retained. The presence of residual spatial autocorrelation, which could bias the results, was assessed using the global Moran I test on standardized Pearson residuals with the *ape* package in R.

Objective 2: Ability of eTick Indicators for Detecting Established Tick Populations

Detecting Established Tick Populations

Logistic regression models were used to assess the value of using eTick surveillance data to predict the presence of an established tick population in a CSD. Observed tick establishment status was determined using INSPQ active surveillance data, following the criteria described in section Active Tick Surveillance, and including CSDs with at least one active surveillance visit since 2014 (n=343). Model building was conducted separately for 3 types of eTick indicators: number of submissions of (A) human origin, (B) animal origin, and (C) all origins combined. Three model structures were assessed for each indicator: (1) a global model including the respective eTick indicator and all socioecological factors identified as significant in objective 1 (models A1, B1, and C1), (2) a reduced model including the eTick indicator and CSD population size (models

XSL•FO
RenderX

A2, B2, and C2), and (3) a basic model including only the eTick indicator (models A3, B3, and C3).

We used the same variable data transformation and model building approaches as described for objective 1. Backward stepwise selection based on Akaike information criterion was used to identify the best model for the global model structures. We compared the explanatory power of the nested models using the likelihood ratio test, and assessed model performance using area under the curve (AUC) and Akaike information criterion. Finally, model fit was evaluated using the Hosmer-Lemeshow test, along with graphical inspection and outlier detection as described earlier in objective 1.

Identifying Thresholds of Submissions Optimizing Tick Population Detection Accuracy

Thresholds for eTick submission data were calculated to maximize sensitivity, specificity, or a balance between both metrics for detecting established tick populations. Analyses focused on submissions from all origins, with thresholds derived from the linear predictors of models C1, C2, and C3. Optimal cutoffs for the linear predictors were determined using the *OptimalCutpoints* package in R [81]. The first 2 methods maximize sensitivity and specificity separately, while maintaining the other parameter at a minimum threshold [82-84]. For instance, in the method optimizing sensitivity, we set specificity to 0.5 and maximized sensitivity under this condition. The third method sought a balance between sensitivity and specificity by minimizing the absolute difference between them [85-87].

To facilitate interpretation of the resulting linear predictor thresholds, illustrative scenarios were subsequently developed. These scenarios were not intended for inference, but rather to translate the model-derived thresholds into interpretable submission counts, by fixing covariates at representative values reflecting different socioecological contexts within the study area. To this end, CSDs were grouped according to the 2 most influential covariates affecting the relationship between submissions and tick population presence: climate suitability and human population size. CSDs were first classified according to their climatic suitability for tick establishment, using the threshold of 2843 DD>0 °C [67]. Within climatically suitable CSDs, a population threshold of 100,000 inhabitants was applied to distinguish large population centers from medium and small population centers [88]. As climatically unsuitable CSDs generally have smaller populations, we used the small population center threshold (30,000 inhabitants [88]) in this group to improve descriptive stratification. In each group, the number of submissions corresponding to each optimized cutoff of the linear predictor was calculated while holding model covariates at their mean values.

Results

Objective 1: Spatial Patterns and Socioecological Factors Associated With eTick Submissions

Characteristics and Spatial Patterns of Tick Submissions

A total of 2668 photographic submissions to eTick were identified as *I scapularis* collected in Quebec during the study period, including 429 (16.1%) submissions in 2020, 722 (27.1%) in 2021, 855 (32.1%) in 2022, and 662 (24.8%) in 2023.

Of these, 313 (11.7%) submissions were excluded from the analysis because users reported travel outside their municipality of residence within the 2 weeks prior to submission.

Among the 2355 included submissions, most (n=1594, 67.7%) submitted ticks were found attached to animals, whereas 28.6% (n=673) were found attached to humans and 3.7% (n=88) were found from the environment (Table 1). Most submitted ticks were adult females (1639/1955, 83.8%), followed by adult males (270/1955, 13.8%) and nymphs (46/1955, 2.4%). The life stage could not be distinguished for 400 (17%) of 2355 submissions. Among submissions with reported submitter profession (this question was introduced in 2023 and was therefore missing for 1,524/2355, 64.7%, submissions), most were submitted by members of the public (719/831, 86.5%), with smaller proportions of submissions from veterinary professionals and human health care professionals (Table 1).

XSL•FO
RenderX

Table . Distribution of blacklegged tick (*Ixodes scapularis*) submissions reported through the eTick crowdsourcing passive surveillance platform in Quebec, Canada, between 2020 and 2023, categorized by tick, host, and submitter characteristics[a].

Variables and characteristics	Submissions, n (%)
Tick stage and sex[b] (n=1955)	
Nymph	46 (2.4)
Adult (male)	270 (13.8)
Adult (female)	1639 (83.8)
Origin of the tick (n=2355)	
On a human	673 (28.6)
On an animal	1594 (67.7)
In the environment	88 (3.7)
Age of the human host (y; n=671)[c]	
0 - 20	185 (25.6)
21 - 40	131 (19.5)
41 - 60	229 (34.1)
≥61	126 (18.8)
Submitter profession[d] (n=831)	
Member of the public	719 (86.5)
Veterinary professional	101 (12.2)
Human health care professional	11 (1.3)

[a]Only considering the submissions for which the user did not report travel history within the previous 2 weeks (n=2355).

[b]Was missing in 400 (17%) of 2355 cases because not all ticks could be identified to a life stage from the submitted picture.

[c]Was missing in 1684 (71.5%) of 2355 cases because not all users reported this information.

[d]Was missing in 1524 (64.7%) of 2355 cases because it has been only added to the eTick questionnaire in 2023. For users who submitted before 2023 and later returned, their profession was retrieved and added to their earlier submissions.

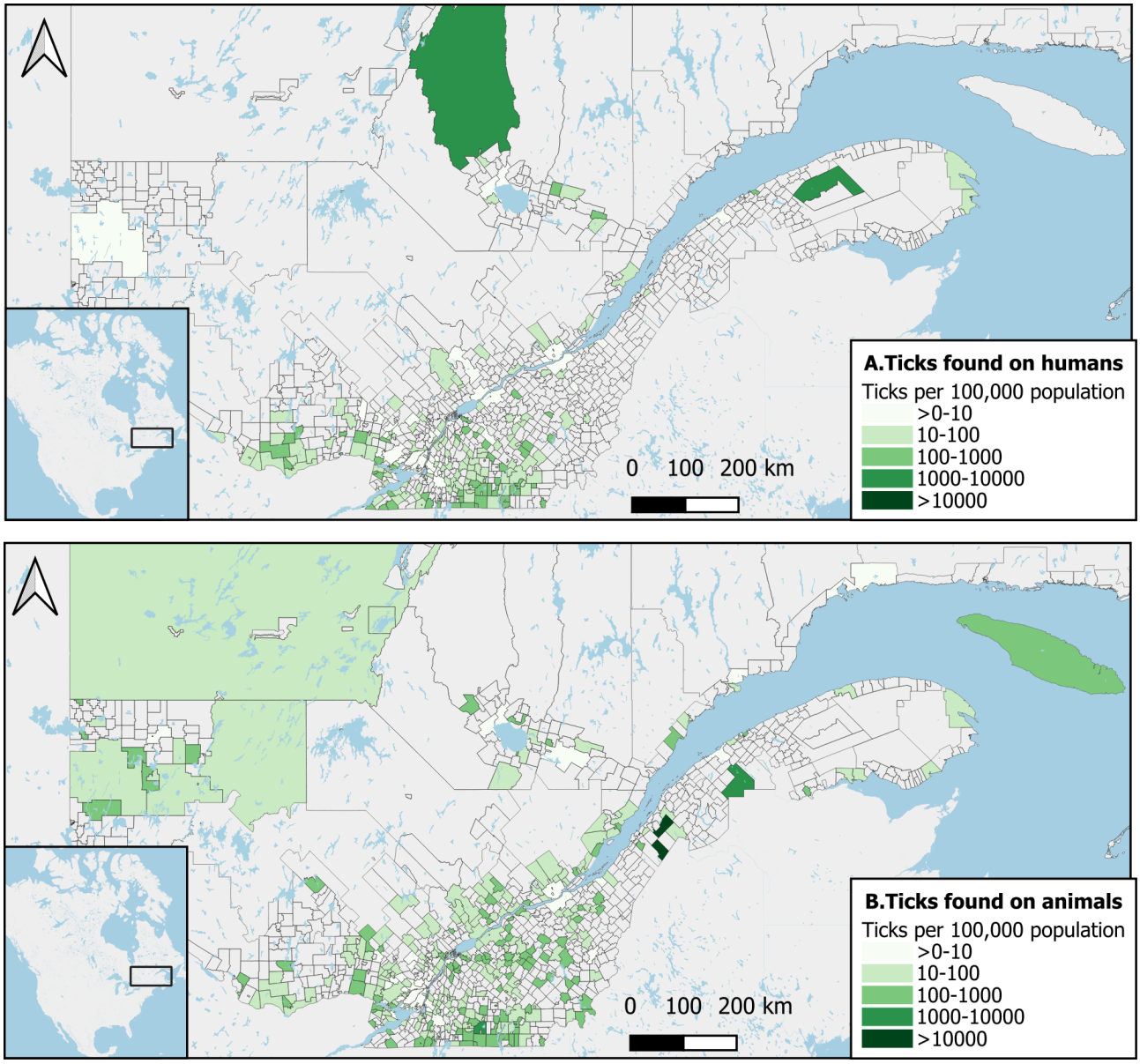
Most (1600/1789, 89.4%) users made only one submission to the platform, with an average of 1.49 (SD 7.38) submissions per user. Among users who made multiple submissions, 2 were outliers regarding submission frequency (submitting 137 and 264 times) and were therefore excluded from the subsequent analysis. After removing these outliers, the average number of submissions per user was 1.27 (SD 2.38), and the number of submissions ranged from 0 to 148 per CSD, with a mean of 1.56 (SD 6.82) submissions per CSD.

Of the 1282 CSDs, 454 (35.4%) had at least one tick submission. Among these tick-positive CSDs, 210 (46.3%) were identified exclusively through submissions of ticks found attached to animals, 85 (18.7%) were identified through submissions of ticks found attached to humans, and 150 (33%) were identified through both categories of submissions. Strong overdispersion was detected in CSD-level submission counts for both human and animal-origin ticks (Dean test, $P<.001$). Spatial scan analyses accounting for overdispersion did not identify any statistically significant spatial clusters for either human- or animal-origin submissions. Spatial patterns are therefore described qualitatively.

For both categories, the submissions were geographically distributed across municipalities in southern Quebec (Figure 1). Submissions of ticks found attached to humans accounted for 29% of all submissions and were predominantly reported from the southern regions of Quebec, particularly Estrie, Outaouais, and Montérégie. Additional areas with higher submission rates for human-origin ticks were observed in Centre-du-Québec and Chaudière-Appalaches. Ticks found attached to animals represented 68% of all submissions and showed a broader geographic distribution. Additionally, higher submission rates for animal-origin ticks were observed in more northerly regions, including Abitibi-Témiscamingue, Mauricie, Capitale-Nationale, and Centre-du-Québec.

XSL•FO
RenderX

Figure 1. Blacklegged tick (*Ixodes scapularis*) submission rates reported through the eTick crowdsourcing passive surveillance platform, in Quebec, Canada, from 2020 to 2023, showing the spatial distribution of ticks found attached to (A) humans and (B) animals at the census subdivision level.



Socioecological Factors Associated With Tick Submissions

The human population size of CSDs was the strongest predictor of both the likelihood (Table 2) of tick submissions and their count (Table 3). Each unit increase in the log-transformed population was associated with a 2.04-fold increase in submission counts (incidence risk ratio [IRR] 2.04, 95% CI 1.86 - 2.25; $P<.001$) and a reduction in the odds of structural zeros (odds ratio [OR] 0.13, 95% CI 0.05 - 0.31; $P<.001$). The log-transformed proportion of WUI zones significantly influenced both components of the model. CSDs with higher WUI proportions had more submissions (IRR 1.43, 95% CI 1.26 - 1.63; $P<.001$), but they were also associated with greater odds of being a structural zero (OR 2.65, 95% CI 1.21 - 5.79; $P=.01$). The number of DD>0 °C significantly reduced the odds of structural zeros (OR 0.10, 95% CI 0.04 - 0.24; $P<.001$). Socioeconomic factors also positively influenced submission counts. Higher median household income and median age were associated with increases in submissions (IRR 1.35, 95% CI 1.20 - 1.52; $P<.001$; and IRR 1.45, 95% CI 1.26 - 1.67; $P<.001$, respectively). One CSD, Lac-Brome in the Estrie region, was identified as an outlier for submission counts (n=148) and excluded from the model fit. No residual spatial autocorrelation was detected ($P=.60$), supporting the robustness of the model.

XSL•FO
RenderX

Table . Factors associated with the occurrence of blacklegged tick (*Ixodes scapularis*) submissions, reported through the eTick crowdsourcing passive surveillance platform across census subdivisions (CSDs) in Quebec, Canada, from 2020 to 2023, according to the zero-inflation component of a zero-inflated negative binomial model

Effect[a]	Estimate	SE	*P* value	OR[b] (95% CI)
Human population[c]	−2.08	0.50	<.001	0.13 (0.05 - 0.31)
Proportion of WUI[c,d]	0.97	0.40	.02	2.65 (1.21 - 5.79)
Annual cumulated Degree Days >0 °C	−2.39	0.48	<.001	0.10 (0.04 - 0.24)

[a]All the variables were centered and scaled before the analysis.

[b]OR: odds ratio.

[c]Logged variable.

[d]WUI: wildland-urban interface.

Table . Factors associated with the abundance of blacklegged tick (*Ixodes scapularis*) submissions, reported through the eTick crowdsourcing passive surveillance platform across census subdivisions (CSDs) in Quebec, Canada, from 2020 to 2023, according to the count component of a zero-inflated negative binomial model (n=1067).

Effect[a]	Estimate	SE	*P* value	IRR[b] (95% CI)
Google Trends index	0.08	0.04	.03	1.08 (1.01 - 1.16)
Median income	0.30	0.06	<.001	1.35 (1.20 - 1.52)
Median age	0.37	0.07	<.001	1.45 (1.26 - 1.67)
Human population[c]	0.71	0.05	<.001	2.04 (1.86 - 2.25)
Proportion of WUI[c,d]	0.36	0.07	<.001	1.43 (1.26 - 1.63)

[a]All the variables were centered and scaled before the analysis.

[b]IRR: incidence risk ratio.

[c]Logged variable.

[d]WUI: wildland-urban interface.

Objective 2: Ability of eTick Indicators for Detecting Established Tick Populations

The Detection of Established Blacklegged Tick Populations

In all 3 sets of models, the more complex models that incorporated socioecological factors outperformed the simpler models, which included submissions with population size or submissions alone (Table 4). The model using submissions of ticks found attached to humans and socioecological covariates (model A1) performed best among all models (AUC 0.86). The model using all types of submissions (model C1) had lower performance (AUC 0.85) compared to the model using submissions of ticks found attached to humans only (model A1), followed by the model using submissions of ticks found attached to animals only (model B1, AUC 0.83).

XSL•FO
RenderX

Table . Performance of logistic regression models predicting the presence of established blacklegged tick (*Ixodes scapularis*) populations at the census subdivision (CSD) level in Quebec, Canada, between 2020 and 2023, using eTick crowdsourced surveillance indicators[a].

Model[b]	Set of covariables included for each candidate model	AIC[c]	ΔAIC[d]	AUC[e]
A1	Submissions of ticks found attached to humans, population size, income, age, proportion of WUI[f], and DD[g] >0 °C	289.24	0.00	0.86
C1	All tick submissions, population size, income, age, and DD >0 °C	295.82	6.58	0.85
B1	Submissions of ticks found attached to animals, population size, income, age, and DD >0 °C	312.26	23.02	0.83
A2	Submissions of ticks found attached to humans and population size	366.99	77.75	0.70
A3	Submissions of ticks found attached to humans	370.23	80.99	0.67
C2	All tick submissions and population size	379.54	90.31	0.68
C3	All tick submissions	384.70	95.46	0.64
B2	Submissions of ticks found attached to animals and population size	396.15	106.91	0.61
B3	Submissions of ticks found attached to animals	396.17	106.93	0.59
0	Intercept only	409	119.77	0.50

[a]Accounting or not for confounding variables identified in the objective.

[b]Models are built separately for 3 types of eTick indicators: the number of submissions of (A) human origin, (B) animal origin, and (C) all origins combined and are compared across 3 levels of model complexity.

[c]AIC: Akaike information criterion.

[d]Differences in AIC (ΔAIC) represent the comparison for all models against the top model, A1.

[e]AUC: area under the curve.

[f]WUI: wildland-urban interface.

[g]DD: degree days.

The number of human-attached submissions to eTick was significantly associated (OR 3.34, 95% CI 2.30 - 4.86; $P<.001$; Table 5) with an increased likelihood of finding established tick populations through active surveillance. Socioecological factors influencing tick presence or population participation in crowdsourced surveillance improved model performance. Positive effects were observed for the log DD >0 °C (OR 4.73, 95% CI 2.97 - 7.52; $P<.001$) and for the proportion of WUI (OR 1.37, 95% CI 1.03 - 1.83; $P=.03$). Negative effects were observed for the log of population size (OR 0.37, 95% CI 0.24 - 0.56; $P<.001$), median income (OR 0.42, 95% CI 0.28 - 0.63; $P<.001$), and median age (OR 0.62, 95% CI 0.42 - 0.92; $P=.02$).

XSL•FO
RenderX

Table . Outputs of the best logistic regression models predicting the presence of established blacklegged tick (*Ixodes scapularis*) populations at the census subdivision (CSD) level in Quebec, Canada, between 2020 and 2023, using eTick crowdsourced surveillance indicators and socioecological factors[a].

Variable[b]	OR[c] (95% CI)	*P* value
A1. Best model using submissions of ticks found attached to humans and socioecological covariables		
Number of submissions found attached to humans[d]	3.34 (2.30 - 4.86)	<.001
Population size[d]	0.37 (0.24 - 0.55)	<.001
Median income	0.42 (0.28 - 0.63)	<.001
Median age	0.62 (0.42 - 0.92)	.02
Proportion of WUI[e]	1.37 (1.03 - 1.83)	.03
DD[f] >0 °C	4.73 (2.97 - 7.52)	<.001
B1. Best model using submissions of ticks found attached to animals and socioecological covariables		
Number of submissions found attached to animals[d]	2.90 (1.93 - 4.36)	<.001
Population size[d]	0.36 (0.23 - 0.56)	<.001
Median income	0.50 (0.35 - 0.73)	<.001
Median age	0.67 (0.46 - 0.97)	.03
DD >0 °C	4.67 (3.04 - 7.18)	<.001
C1. Best model using all tick submissions and socioecological covariables		
Number of submissions[d]	3.86 (2.51 - 5.93)	<.001
Population size[d]	0.31 (0.20 - 0.48)	<.001
Median income	0.45 (0.30 - 0.66)	<.001
Median age	0.63 (0.43 - 0.92)	.017
DD >0 °C	4.78 (3.05 - 7.50)	<.001

[a]Models are built separately for 3 types of eTick indicators: the number of submissions of (A) human origin, (B) animal origin, and (C) all origins combined.

[b]All variables were scaled prior to the analysis.

[c]OR: odds ratio.

[d]Log-transformed variable.

[e]WUI: wildland-urban interface.

[f]DD: degree days.

Thresholds of Submissions Optimizing Tick Population Detection Accuracy

Thresholds of tick submissions calculated using the linear predictor from the most complex model (C1) outperformed those from the simpler models across all optimization methods. A cutoff of −2.30 optimized sensitivity at 0.96, while a cutoff of 0.26 optimized specificity at 0.91 (Table 6). These optimizations were achieved while ensuring that the opposite parameter did not fall below 0.50. When optimizing both parameters simultaneously, sensitivity and specificity were balanced at 0.76 with a cutoff of −0.66. The method that minimized the overall error rate was the one optimizing specificity, resulting in an error rate of 0.22.

XSL•FO
RenderX

Table . Cutpoints of linear predictors optimizing sensitivity, specificity, or the equilibrium between both for detecting established tick populations at the census subdivision level in Quebec, Canada, between 2020 and 2023, according to the best-performing model using the eTick crowdsourced surveillance indicator, including all origins combined.

Model and method of cut-point optimization	Sensitivity (%)	Specificity (%)	Error rate (%)	Linear predictor cutpoint (submissions)
Optimization of sensitivity[a]				
C3. Model using all submissions	0.55	0.61	0.41	−0.87
C2. Model using all submissions and population size	0.72	0.53	0.42	−1.18
C1. Model using all submissions, population size, income, age, and DD[b] >0 °C	0.96	0.51	0.36	−2.30
Optimization of specificity[c]				
C3. Model using all submissions	0.55	0.61	0.41	−0.81
C2. Model using all submissions and population size	0.5	0.86	0.25	−0.50
C1. Model using all submissions, population size, income, age, and DD >0 °C	0.5	0.91	0.22	0.26
Optimization of the balance between sensitivity and specificity				
C3. Model using all submissions	0.55	0.61	0.41	−0.81
C2. Model using all submissions and population size	0.63	0.63	0.37	−0.90
C1. Model using all submissions, population size, income, age, and DD >0 °C	0.76	0.76	0.24	−0.66

[a]The minimum value tolerance for specificity is 0.50.

[b]DD: degree days.

[c]The minimum value tolerance for sensitivity is 0.50.

The CSDs grouped by climatic suitability and human population showed diverse environmental and demographic characteristics (Table 7). Across the study region, for a typical area with an average DD >0 °C of 3199 and an average human population of 19,838, the optimal submission threshold for detecting an established tick population—balancing sensitivity and specificity—is estimated at 3.32. This threshold increases to 7.54 when optimizing specificity and decreases to 0.51 (rounded to 1 submission in practice) when optimizing sensitivity. When analyzing CSDs by group, climatically unsuitable medium population centers had the highest threshold values, followed by climatically suitable large population centers, climatically unsuitable small population centers, and, finally, climatically suitable small and medium population centers.

XSL•FO
RenderX

Table . Optimal number of submissions (cutpoints) for optimizing sensitivity, specificity, or both, based on the top-performing model (C1) predicting established blacklegged tick (*Ixodes scapularis*) populations at the census subdivision (CSD) level in Quebec, Canada, between 2020 and 2023, using the eTick crowdsourced surveillance indicator including all origins combined and socioecological factors[a].

CSD group	Average covariate values across the CSD group				Optimal number of submissions (tick count) according to each cutpoint optimization method		
	DD >0 °C	Population (n)	Median income (CAD $)	Median age (y)	Sensitivity≈specificity	↑ Specificity	↑ Sensitivity
Overall average	3199	19,838	75,384	47.5	3.32	7.54	0.51
Climatically suitable large population centers (>2843 DD[b] >0 °C; ≥100,000 population; n=9)	3603	430,893	74,933	42.2	12	24.67	3.55
Climatically suitable small and medium population centers (>2843 DD >0 °C; <10,000 population; n=247)	3344	9202	77,717	46.4	1.92	4.76	0.02
Climatically unsuitable medium population centers (<2843 DD >0 °C; ≥30,000 population; n=3)	2769	97,172	61,300	50.2	64.40	128.16	21.90
Climatically unsuitable small population centers (<2843 DD>0 °C; <30,000 population; n=74)	2684	3478	68,129	51.7	11.87	24.41	3.51

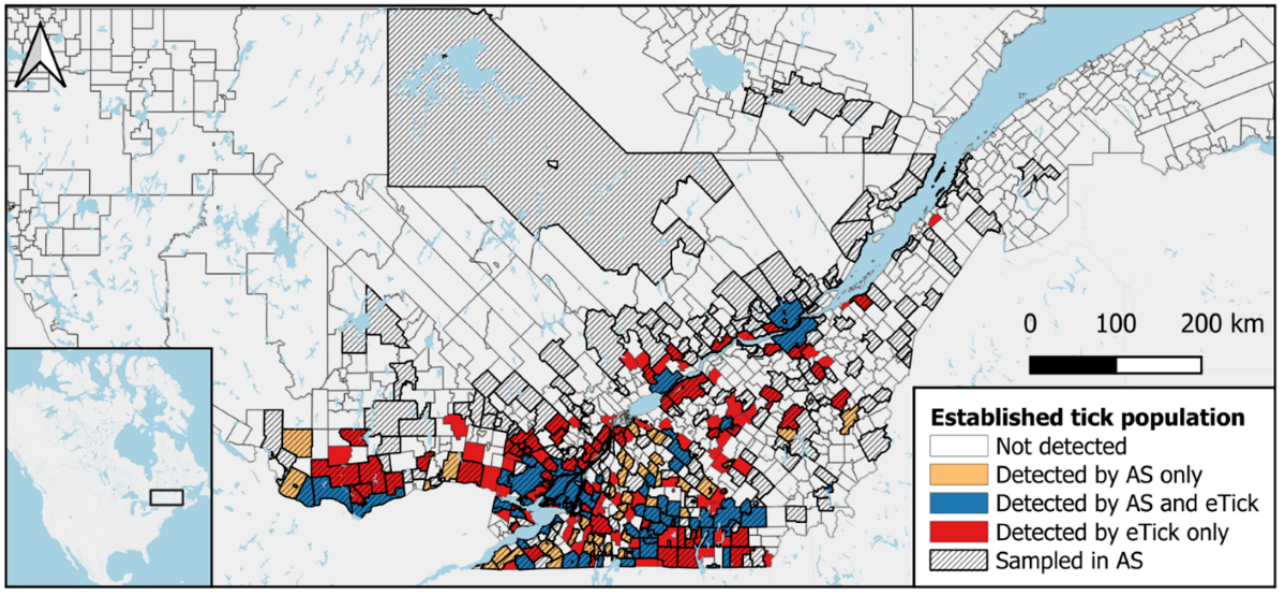
[a] For ease of interpretation, the cutpoints are calculated for typical CSDs grouped by climate suitability and population size, with demographics and environmental covariates fixed at their mean values within each CSD group.

[b]DD: degree days.

According to the active surveillance criteria for the presence of established blacklegged tick populations and the eTick indicator based on submission thresholds, all CSDs with established populations are located in the southern part of the province, up to the approximate 47˚N latitude of Quebec City (Figure 2). Among the 343 CSDs sampled through active surveillance, 248 (72.3%) were identified as positive for the presence of established tick populations by at least one of the 2 indicators. Of these 248 positive CSDs, 45 (18.1%) were identified by the active surveillance criterion only, 148 (59.7%) were identified by the eTick indicator only (using thresholds ranging from 2 to 64 submissions, as detailed in Table 7), and 55 (22.2%) were identified by both indicators (Figure 2).

XSL•FO
RenderX

Figure 2. Comparison of established blacklegged tick (*Ixodes scapularis*) populations at the census subdivision (CSD) level in Quebec, Canada, between 2020 and 2023, based on active surveillance (AS) criteria (n=343 CSD sampled) and the eTick thresholds. Based on AS data, an *I scapularis* population is considered established if either 6 or more ticks have been found or at least two different life stages are found within the same year in a municipality. The eTick thresholds represent the number of submissions optimizing the balance between sensitivity and specificity and are derived from the top-performing model (C1) predicting established blacklegged tick populations using the eTick crowdsourced surveillance indicator, including all origins combined and socioecological factors. Thresholds are applied to CSDs according to climate suitability and population size, as detailed in Table 7. Depending on the CSD, thresholds range from 2 to 64 submissions to be considered indicative of an established population.



Discussion

Principal Findings

This study demonstrates that crowdsourced eTick data can be used to develop robust, evidence-based indicators of blacklegged tick population establishment in Quebec. Although we did not detect spatial clusters in submission rates, we found that the rates were associated with socioecological factors, including climate suitability, population size, landscape configuration, and population socioeconomic characteristics. The results also highlighted key sources of variation that must be accounted for when deriving indicators from crowdsourced surveillance data. Building on these findings, we developed and tested indicators based on eTick submissions that were effective in detecting CSDs with established tick populations while accounting for confounding factors. Overall, the results support integration of crowdsourced surveillance data into the Canadian tick-borne disease surveillance framework as a complementary tool to traditional active surveillance. Previous studies using eTick data have applied different approaches to predict tick distribution or public health risk, providing useful points of comparison for the present findings. Bowser et al [89] assessed whether early eTick data (2018 - 2020) could predict LD risk at the CSD level, in an endemic region of Quebec, Canada. They reported weaker associations and poorer model fit when eTick submissions were used as LD risk predictors compared to other indicators, attributing these results to the novelty of the platform at the time and variable levels of public adoption and participation. More recently, Westcott et al [90] used eTick submissions with other citizen science datasets to develop niche models predicting the current and future distribution of *I scapularis*. Both studies relied on raw submission data and did not explicitly account for socioecological factors influencing participation and reporting intensity, which may have limited risk signal detection in Bowser et al [89] and biased habitat suitability estimates in Westcott et al [90]. While these studies illustrate the potential utility of crowdsourced surveillance data, the lack of consideration of these factors may represent an important limitation. Indeed, results from this study indicate that eTick submission counts are influenced by both human and environmental factors, including population size, socioeconomic characteristics, and habitat suitability affecting tick occurrence and abundance. Submission rates likelihood reflect exposure risk among humans and companion animals as driven by behavior and tick presence and density, as well as factors influencing the decision to submit samples to eTick. Accordingly, the value of these data for contributing to surveillance programs depends on accounting for biases affecting submission rates.

Submissions to eTick were distributed across southwestern and central Quebec, corresponding to the known distribution range of *I scapularis* in the province [91]. Although no statistically significant spatial clusters were detected, there was geographic heterogeneity in submission rates for both submission types. Submission of ticks found attached to humans were mostly reported in the southernmost regions of the province, which are also the highest risk regions for LD (Outaouais, Montérégie, and Estrie [92]). In contrast, submissions of ticks found attached to animals were distributed across a wider geographic area, including regions located outside the currently defined endemic zones. Several of these regions include municipalities with lower risk levels for LD or are not considered endemic for LD according to the provincial risk classification [91]. Under this classification, municipalities are considered endemic when they meet one or more criteria regarding the number of declared human LD cases, the number of ticks collected through passive

XSL•FO
RenderX

surveillance, or the number and stages of *Borrelia burgdorferi*–infected ticks collected in active surveillance [93]. Taken together, these patterns highlight the broader geographic scope of detections obtained through animal-origin submissions and suggest that combining submissions from both origins may provide complementary information for surveillance of tick presence beyond established endemic areas. As eTick data continue to accumulate over time, there may be sufficient statistical power to detect higher than expected submission rates, if present, which could help identify and test for currently unknown local factors (human and pet population, or tick) affecting exposure risk.

Socioecological Factors Associated With Tick Submissions

In zero-inflated negative binomial models analyzing the factors associated with tick submissions, the total human population of CSDs was the strongest predictor of both submission likelihood and count, indicating that more populated areas are more likely to report tick presence. This result aligns with previous studies, which have shown that submission counts are generally proportionate to population size, where the odds of submission increase with the number of potential contributors [18-20,40,41]. These findings reinforce the importance of accounting for population size when analyzing passive surveillance data [18,19]. Additional socioecological factors also influenced submission occurrence and counts.

First, municipalities that accumulated more annual DD >0 °C were more likely to have tick detections. Ticks require ambient temperatures above 0 °C to quest for hosts, with warmer conditions accelerating interstadial development [94]. Globally warmer conditions throughout the year are thought to accelerate the completion of their life cycle and facilitate their establishment. Indeed, previous modeling and field studies have shown that a threshold of approximately 2800 DD >0 °C is needed for ticks to establish viable populations [67,95]. While DD >0 °C was a determinant of the presence of tick populations, it did not significantly affect submission counts once populations were established. This suggests that once tick populations are well established, other factors such as host densities, human-tick encounters, or participation behavior may drive the numbers of submissions to eTick.

The proportion of WUI in CSDs was positively associated with tick submission counts. This variable measures the degree to which human settlements (ie, neighborhoods where people live, infrastructures where people work, or roads providing access to the territory) are surrounded by natural areas, including woodlands, ecotones, and prairies, which may be suitable habitats for ticks and their hosts [65]. Results indicate that CSDs with higher WUI represent higher tick exposure risk areas [45]. Conversely, the WUI indicator was negatively associated with the occurrence of tick submissions (as denoted by a significant positive relationship with the likelihood of structural zeros), which is a surprising result that warrants further analysis. Here, we studied a region that may not be entirely suitable for tick establishment, with very different biomes across the territory [5]. It is unknown, for example, if boreal forest provides a habitat where *I scapularis* populations can become established [96]. Consequently, this result may reflect another underlying relationship not captured by our current indicators or may relate to the coarse ecological nature of the indicator, which is aggregated at the municipal level and includes all types of natural habitats. Moreover, ticks require specific ecological conditions regarding habitat type, host availability, and landscape connectivity at fine spatial scales, which likely could not be detected at the municipal scale used in this work [45,97,98].

Finally, CSDs with higher median household income and age were associated with more submissions. This aligns with other findings that citizen science initiatives tend to predominantly attract middle-aged and older groups and those with greater access to educational and financial resources [35,37]. This is a common point of critique [37,99] and an area where improvement is needed through initiatives aimed at reaching more diverse groups. Doing so would help ensure that the benefits of participation are more equitably distributed across the population [31] and that the indicators derived from these data are more representative of all areas and population profiles. Our proxy of general population awareness of ticks and LD, constructed using Google Trends data, showed a weak but positive relationship with submission counts in CSDs. This aligns with expectations that public awareness of the problematics under focus is generally associated with higher participation in citizen science initiatives [31,37,100]. However, several limitations inherent to Google Trends may have contributed to the weak association observed. Keyword-based indicators rely on specific search terms and may not capture the broader range of queries related to tick information seeking [101]. Additionally, Google search activity can be strongly influenced by short-term, media-driven attention, which may not reflect sustained awareness or a durable willingness to engage with surveillance initiatives such as eTick [101]. More broadly, public health intelligence derived from open-access, user-generated sources, such as Google Trends, is subject to participation bias, with contributions disproportionately originating from populations with greater internet access and digital literacy. As a result, these data may not accurately represent the population at risk to the public health threat [102,103]. Furthermore, Google Trends data can be unstable or imprecise when analyzed at small spatial scales or in areas with low population sizes, such as many CSDs, in the study area [104]. Collectively, these limitations may reduce the validity of this indicator as a proxy for awareness, as noted in previous studies [101]. Integrating complementary data from diverse sources, ranging from traditional public health surveillance systems to nontraditional data, such as open-access social media, may help strengthen signal detection and mitigate participation biases [102].

The Detection of Established Blacklegged Tick Populations

To evaluate the ability to detect established *I scapularis* populations, we built and compared models using 1 of 3 crowdsourced surveillance indicators based on tick origin: ticks found attached to humans, animals, and from all origins combined. For each indicator, we compared simple models with more complex ones that accounted for socioecological factors

XSL•FO
RenderX

associated with submissions in objective 1. In all cases, the complex models significantly outperformed simpler alternatives, including those accounting for population size or only using submission counts. While simple models can be useful for ease of interpretation and indicator development for public health use [18], the loss in performance here would be too high to justify this approach. This result highlights the relevance of analytically considering socioecological covariates when using passive or crowdsourced data, consistent with findings and approaches from other studies [20,41,44].

The 3 top-performing models yielded similar results: model A1, based on ticks found attached to humans and socioecological factors; model B1, based on ticks found on animals and socioecological factors; and model C1, based on all tick origins combined with socioecological factors. Given these similar performances and the complementarity observed in the geographic distributions of both submission types, we pursued subsequent analyses using the submissions from all origins to identify the threshold optimizing the detection of established tick populations. By doing so, we also ensured that all data collected by users were leveraged.

These analyses showcase a methodology for indicator development that can be adapted according to the objectives of the surveillance or research program and the epidemiological context. Previous Canadian passive surveillance studies have shown that ticks recovered from animals, particularly dogs, were effective at providing early signals and wider geographical extent of emerging tick populations [18-20,105]. This relationship has been attributed to differences in exposure, as dogs are generally more directly in contact with vegetation and leaf litter, where they are more likely to be bitten by questing ticks, even with low tick densities [13,14]. However, ticks recovered from animals may be less specific indicators than those from humans, as very low tick densities detected through animals may reflect adventitious ticks transported by migratory birds rather than established tick populations [18,19]. On the other hand, our analyses suggest that ticks from animals can be efficient at detecting emerging tick populations where active surveillance may not be possible or may be insufficient. Although active surveillance remains the most standardized method available for tick surveys, its efficacy depends on multiple factors, including vegetation type and weather conditions before and during sampling [106-108]. Drag sampling may be less effective for detecting emerging tick populations in northern Quebec, where coniferous forests and dense understory vegetation dominate, compared to deciduous forests in southern Quebec [106-109]. This highlights that both active and crowdsourced surveillance indicators may have strengths and limitations and that we would benefit from both by developing integrated indicators [15,100,110].

Thresholds of Submissions Optimizing Tick Population Detection Accuracy

To identify submission thresholds that optimize tick population detection, we searched for cutpoints that maximized sensitivity, specificity, or both, using models based on submission types. In the best-performing model, a threshold of 3 submissions (rounded from 3.32) maximized both sensitivity and specificity for detecting tick populations. At this threshold, both sensitivity and specificity reached 76%. These levels are comparable to those of a previous study conducted in Quebec, which used passive surveillance to predict the presence of established *I scapularis* populations, as measured by active surveillance carried out from 2007 to 2008 [20]. That study identified an optimal threshold of 0.073 (Predicted Linear Index), which predicted the presence of ticks in the environment with 74.19% sensitivity and 73.68% specificity, using a similar methodology that maximized both indicators. Their model included passive surveillance data, an index of environmental suitability based on cumulative annual DD >0 °C, and an index estimating the number of ticks arriving on migratory birds. Overall, our results of crowdsourced surveillance shared similar characteristics and limitations with comparable traditional passive surveillance datasets from other Canadian studies: more adult female ticks were submitted than immature ticks, more ticks were recovered from animals than from humans, and similar factors affected the submission data [14,18-20]. The predictive performance of this dataset in identifying established tick populations, as confirmed by active surveillance, was consistent with that of traditional passive surveillance. Therefore, our analysis supports the validity of crowdsourced data as a complement to traditional active and passive surveillance. We believe these approaches should remain complementary, as each provides distinct information. Active surveillance provides standardized measures of tick density and confirmation of locally reproducing populations, while crowdsourced data offer broader geographic coverage and rapid detection of emerging areas [111], and traditional passive surveillance also more readily enables pathogen testing, which is generally not feasible with crowdsourced submissions. Future work should include an assessment of how these data should be used to build indicators that are adaptive to evolving epidemiological context, such as the one prevailing in Canada [11,105].

The optimal thresholds of submissions maximizing tick population detection accuracy were highly variable across the study area. This variability reflects the province's pronounced heterogeneity regarding population size, land cover, climate, and socioeconomic characteristics. To facilitate interpretation, we exemplified what these thresholds would be across different groups of CSDs, regrouped by population size and climate suitability. Climatically suitable regions for ticks are concentrated in the southern parts of Quebec. These areas are well-known endemic zones for *I scapularis* and LD, including the highest incidence regions of Estrie and Montérégie [56]. Most CSDs in these regions have small population sizes and are suburban or rural. These areas also have a high proportion of WUI, meaning there is significant overlap between tick-suitable natural environments and human habitation. In such contexts, only a very small number of tick submissions was sufficient to indicate the likely presence of an established tick population. This suggests that even limited crowdsourced surveillance signals in these areas should raise awareness about possible reproducing tick populations and emerging risk of tick-borne disease [112]. In contrast, large population centers in climatically suitable areas required a higher threshold of submissions to balance sensitivity and specificity. In climatically unsuitable regions, a similar threshold was identified for small

XSL•FO
RenderX

population centers, while the highest one was identified for medium population centers. These findings suggest that in emerging areas, a high number of ticks must be submitted before suspecting an established population, which may indicate that adventitious ticks are detected well before actual establishment occurs. Alternatively, it could reflect a mismatch between active and passive surveillance data in these regions. For example, active surveillance may fail to detect ticks in more northern environments due to limitations in sampling sensitivity, especially in the dense understory vegetation of coniferous forests [106,108], or because of lower sampling effort deployed in these areas. In contrast, passive surveillance may be less affected by these limitations, as it relies on ticks encountered by the public or professionals across a wider range of habitats and conditions.

This study has several limitations that should be considered when interpreting its findings. First, the spatial accuracy of eTick submissions implies an inherent level of uncertainty, as reported locations reflect the user's presumed site of tick acquisition and cannot be assumed to be exact. These locations remain subject to recall and information bias, as users may have incomplete recollection of where exposure occurred or may infer locations based on routine activities. Spatial accuracy is expected to be higher when ticks are detected promptly after exposure or when individuals perform regular tick checks and are aware of their recent movements [28]. Excluding submissions from users reporting recent travel history mitigated the impact of this source of uncertainty; however, it was not possible to fully control for this limitation with the available data. Second, uncertainty in indicator precision may arise from spatial aggregation at the CSD level and from the use of active surveillance as an imperfect reference standard. Ecological processes relevant to tick establishment operate at finer spatial scales than those captured here, and active surveillance may fail to detect established populations under certain environmental or sampling conditions. Together, these limitations reflect the challenge of relying on relatively coarse spatial indicators. Likewise, although we attempted to account for participation biases to which crowdsourced initiatives are subject by including sociodemographic covariates in our modeling framework, residual bias likely remains. Finally, the generalizability of the proposed indicators beyond Quebec and over time will require further evaluation. Differences in surveillance systems, population engagement, ecological contexts, and public awareness across provinces or countries may affect indicator performance. Over time, the usefulness of a crowdsourced approach such as eTick depends on sustained participation from both the general public and professionals [113]. Since future engagement is unknown and may not remain stable over time, the continued usefulness of this surveillance method cannot be guaranteed. Regular reassessment will therefore be necessary to ensure data quality and long-term relevance.

Conclusions

In this study, we analyzed the patterns in blacklegged tick submissions to the platform eTick in Quebec, Canada. This work introduces a rigorous analytical framework for interpreting crowdsourced surveillance data by comparing submission rates with long-term active surveillance. A novel contribution is using a socioecological approach to adjust for potential participation bias in eTick submission rates. Unlike prior Canadian tick passive surveillance studies, which rely primarily on traditional surveillance methods or that use unadjusted crowdsourced data, our approach explicitly incorporates socioecological factors associated with reporting into the modeling framework. This work demonstrates that crowdsourced surveillance data, when appropriately modeled, can yield reliable inferences about disease vector distributions. It provides a methodological approach applicable to other medically important vector species and geographic contexts, and contributes to improving the robustness of ecological inference of vector distributions. Specifically, our findings show that eTick can effectively complement traditional passive surveillance methods for detecting established blacklegged tick populations. Although this analysis focused on *I scapularis*, eTick allows the collection of all tick species and could therefore support monitoring of other medically important species in Canada, such as *Ixodes cookei*, *Dermacentor variabilis*, *Rhipicephalus sanguineus*, or *Amblyomma americanum*. These contributions have practical implications for public health practice. More confident identification of tick distributions supports targeted public health actions and improved surveillance strategies. The broad geographic coverage of eTick makes it well suited for early detection of emerging tick-borne diseases in undersurveyed regions. Integrating eTick data into formal surveillance systems could significantly enhance Canada's capacity to detect those emerging risks in a cost-effective way. Future work should focus on expanding the analytical framework to other tick species and pathogens, conducting field validation of predicted tick populations using active surveillance, and improving the representativeness of reporting populations across diverse and historically excluded communities.

Acknowledgments

We thank Jeremie Bouffard for assistance in managing the eTick program and all participants who submitted specimens to eTick. We also thank Ariane Adam-Poupart for facilitating the data agreement between Institut national de santé publique du Québec and Public Health Agency of Canada. The authors confirm that no generative AI tools were used in the preparation of this manuscript.

Funding

This study was funded by the Public Health Agency of Canada through internal operational research funds.

XSL•FO
RenderX

Data Availability

The datasets analyzed during this study are not publicly available due to legal and data sharing restrictions but can be requested from the data owners. Requests for access to Québec tick surveillance data should be directed to the Infocentre de santé publique of the INSPQ (infocentre@inspq.qc.ca), and requests for eTick data should be addressed to Jade Savage, Bishop's University (jsavage@ubishops.ca).

Authors' Contributions

Conceptualization: AD, JS, KEC, PAL, NHO, EER, CB
Data curation: AD
Formal analysis: AD
Investigation: AD, JS
Methodology: AD, EER, CB
Project administration: AD, CB
Supervision: EER, CB
Validation: AD
Visualization & writing – original draft: AD
Writing – review & editing: JS, KEC, JKK, PAL, NHO, EER, CB

Conflicts of Interest

None declared.

Multimedia Appendix 1
Ethics review exemption from Bishop's University Research Ethics Board.
[PDF File, 174 KB - publichealth_v12i1e90734_app1.pdf]

References

1. Bouchard C, Dibernardo A, Koffi J, Wood H, Leighton PA, Lindsay LR. N Increased risk of tick-borne diseases with climate and environmental changes. Can Commun Dis Rep 2019 Apr 4;45(4):83-89. [doi: 10.14745/ccdr.v45i04a02] [Medline: 31285697]
2. Leighton PA, Lindsay LR, Ogden NH. Climate and the emergence of tick-borne disease in canada. In: Nuttal P, editor. Climate, Ticks and Disease: CABI; 2021:507-512. [doi: 10.1079/9781789249637.0073]
3. Clow KM, Leighton PA, Ogden NH, et al. Northward range expansion of Ixodes scapularis evident over a short timescale in Ontario, Canada. PLoS One 2017;12(12):e0189393. [doi: 10.1371/journal.pone.0189393] [Medline: 29281675]
4. Ogden NH, Lindsay LR, Morshed M, Sockett PN, Artsob H. The emergence of Lyme disease in Canada. Can Med Assoc J 2009 Jun 9;180(12):1221-1224. [doi: 10.1503/cmaj.080148]
5. Ripoche M, Bouchard C, Irace-Cima A, Leighton P, Thivierge K. Current and future distribution of Ixodes scapularis ticks in Québec: field validation of a predictive model. PLoS One 2022;17(2):e0263243. [doi: 10.1371/journal.pone.0263243] [Medline: 35113941]
6. Dergousoff SJ, Galloway TD, Lindsay LR, Curry PS, Chilton NB. Range expansion of dermacentor variabilis and dermacentor andersoni (Acari: Ixodidae) near their northern distributional limits. J Med Entomol 2013 May;50(3):510-520. [doi: 10.1603/me12193] [Medline: 23802445]
7. Gasmi S, Bouchard C, Ogden NH, et al. Evidence for increasing densities and geographic ranges of tick species of public health significance other than Ixodes scapularis in Québec, Canada. PLoS ONE 2018;13(8):e0201924. [doi: 10.1371/journal.pone.0201924] [Medline: 30133502]
8. Tam T. Lyme Disease in Canada: A Federal Framework: Public Health Agency of Canada; 2017.
9. Ghanbari H, Siebels K, Dumas A, et al. A machine learning framework for estimating the probability of blacklegged tick population establishment in eastern Canada using Earth observation data. PLoS One 2025;20(9):e0332582. [doi: 10.1371/journal.pone.0332582] [Medline: 40982485]
10. Guillot C, Bouchard C, Buhler K, Pelletier R, Milord F, Leighton P. Quality over quantity in active tick surveillance: sentinel surveillance outperforms risk-based surveillance for tracking tick-borne disease emergence in southern Canada. Can Commun Dis Rep 2023 Feb 1;49(2-3):50-58. [doi: 10.14745/ccdr.v49i23a04] [Medline: 38090724]
11. Bouchard C, Dumas A, Baron G, et al. Integrated human behavior and tick risk maps to prioritize Lyme disease interventions using a "One Health" approach. Ticks Tick Borne Dis 2023 Mar;14(2):102083. [doi: 10.1016/j.ttbdis.2022.102083]
12. Pelletier J, Guillot C, Rocheleau JP, et al. The added value of one health surveillance: data from questing ticks can provide an early signal for anaplasmosis outbreaks in animals and humans. Can J Public Health 2023 Apr;114(2):317-324. [doi: 10.17269/s41997-022-00723-8] [Medline: 36471231]

XSL•FO
RenderX

13. Hamer SA, Tsao JI, Walker ED, Mansfield LS, Foster ES, Hickling GJ. Use of tick surveys and serosurveys to evaluate pet dogs as a sentinel species for emerging Lyme disease. Am J Vet Res 2009 Jan;70(1):49-56. [doi: 10.2460/ajvr.70.1.49] [Medline: 19119948]
14. Ogden NH, Trudel L, Artsob H, et al. Ixodes scapularis ticks collected by passive surveillance in Canada: analysis of geographic distribution and infection with Lyme borreliosis agent Borrelia burgdorferi. J Med Entomol 2006 May 1;43(3):600-609. [doi: 10.1093/jmedent/43.3.600]
15. Lyons LA, Brand ME, Gronemeyer P, et al. Comparing contributions of passive and active tick collection methods to determine establishment of ticks of public health concern within Illinois. In: Rich S, editor. J Med Entomol 2021 Jul 16;58(4):1849-1864. [doi: 10.1093/jme/tjab031] [Medline: 33855433]
16. Nelder MP, Russell C, Lindsay LR, et al. Population-based passive tick surveillance and detection of expanding foci of blacklegged ticks Ixodes scapularis and the Lyme disease agent Borrelia burgdorferi in Ontario, Canada. PLoS ONE 2014;9(8):e105358. [doi: 10.1371/journal.pone.0105358] [Medline: 25171252]
17. Guillot C, Bouchard C, Buhler K, et al. Sentinel surveillance contributes to tracking Lyme disease spatiotemporal risk trends in Southern Quebec, Canada. Pathogens 2022 May 2;11(5):531. [doi: 10.3390/pathogens11050531] [Medline: 35631052]
18. Ripoche M, Gasmi S, Adam-Poupart A, et al. Passive tick surveillance provides an accurate early signal of emerging Lyme disease risk and human cases in Southern Canada. J Med Entomol 2018 Jun 28;55(4):1016-1026. [doi: 10.1093/jme/tjy030] [Medline: 29522180]
19. Gasmi S, Ogden NH, Ripoche M, et al. Detection of municipalities at-risk of Lyme disease using passive surveillance of Ixodes scapularis as an early signal: a province-specific indicator in Canada. PLoS ONE 2019;14(2):e0212637. [doi: 10.1371/journal.pone.0212637] [Medline: 30779789]
20. Koffi JK, Leighton PA, Pelcat Y, et al. Passive surveillance for I. scapularis ticks: enhanced analysis for early detection of emerging Lyme disease risk. J Med Entomol 2012 Mar;49(2):400-409. [doi: 10.1603/me11210] [Medline: 22493860]
21. Nelder MP, Russell CB, Dibernardo A, et al. Monitoring the patterns of submission and presence of tick-borne pathogens in Ixodes scapularis collected from humans and companion animals in Ontario, Canada (2011-2017). Parasit Vectors 2021 May 17;14(1):260. [doi: 10.1186/s13071-021-04750-1] [Medline: 34001256]
22. Wilson C, Gasmi S, Bourgeois AC, et al. Surveillance for Ixodes scapularis and Ixodes pacificus ticks and their associated pathogens in Canada, 2019. CCDR 2022;48(5):208-218. [doi: 10.14745/ccdr.v48i05a04]
23. Ebi KL, Ogden NH, Semenza JC, Woodward A. Detecting and attributing health burdens to climate change. Environ Health Perspect 2017 Aug 7;125(8):085004. [doi: 10.1289/EHP1509] [Medline: 28796635]
24. Lusk D, Wolf S, Svidzinska D, et al. Crowdsourced biodiversity monitoring fills gaps in global plant trait mapping. Nat Commun 2026;17(1):1203. [doi: 10.1038/s41467-026-68996-y]
25. Ranard BL, Ha YP, Meisel ZF, et al. Crowdsourcing--harnessing the masses to advance health and medicine, a systematic review. J Gen Intern Med 2014 Jan;29(1):187-203. [doi: 10.1007/s11606-013-2536-8] [Medline: 23843021]
26. Bald L, Ratnaweera N, Hengl T, et al. Assessing tick attachments to humans with citizen science data: spatio-temporal mapping in Switzerland from 2015 to 2021 using spatialMaxent. Parasit Vectors 2025 Jan 23;18(1):22. [doi: 10.1186/s13071-024-06636-4] [Medline: 39849565]
27. Hart CE, Bhaskar JR, Reynolds E, et al. Community engaged tick surveillance and tickMAP as a public health tool to track the emergence of ticks and tick-borne diseases in New York. PLOS Glob Public Health 2022;2(6):e0000215. [doi: 10.1371/journal.pgph.0000215]
28. Kache PA, Bron GM, Zapata-Ramirez S, et al. Evaluating spatial and temporal patterns of tick exposure in the United States using community science data submitted through a smartphone application. Ticks Tick Borne Dis 2023 Jul;14(4):102163. [doi: 10.1016/j.ttbdis.2023.102163] [Medline: 37001417]
29. Kopsco HL, Duhaime RJ, Mather TN. Crowdsourced tick image-informed updates to U.S. county records of three medically important tick species. J Med Entomol 2021 Nov 9;58(6):2412-2424. [doi: 10.1093/jme/tjab082] [Medline: 33973636]
30. Laaksonen M, Sajanti E, Sormunen JJ, et al. Crowdsourcing-based nationwide tick collection reveals the distribution of Ixodes ricinus and I. persulcatus and associated pathogens in Finland. Emerg Microbes Infect 2017 May 10;6(5):e31. [doi: 10.1038/emi.2017.17] [Medline: 28487561]
31. Den Broeder L, Devilee J, Van Oers H, Schuit AJ, Wagemakers A. Citizen science for public health. Health Promot Int 2018 Jun 1;33(3):505-514. [doi: 10.1093/heapro/daw086] [Medline: 28011657]
32. Koffi JK, Savage J, Thivierge K, et al. Evaluating the submission of digital images as a method of surveillance for Ixodes scapularis ticks. Parasitology 2017 Jun;144(7):877-883. [doi: 10.1017/S0031182017000117] [Medline: 28345501]
33. Brown P. Popular epidemiology revisited. Curr Sociol 1997 Jul;45(3):137-156. [doi: 10.1177/001139297045003008]
34. Merenlender AM, Crall AW, Drill S, Prysby M, Ballard H. Evaluating environmental education, citizen science, and stewardship through naturalist programs. Conserv Biol 2016 Dec;30(6):1255-1265. [doi: 10.1111/cobi.12737] [Medline: 27109290]
35. Fernandez MP, Bron GM, Kache PA, et al. Usability and feasibility of a smartphone app to assess human behavioral factors associated with tick exposure (The Tick App): quantitative and qualitative study. JMIR Mhealth Uhealth 2019 Oct 24;7(10):e14769. [doi: 10.2196/14769] [Medline: 31651409]

XSL•FO
RenderX

36. Mac Domhnaill C, Lyons S, Nolan A. The citizens in citizen science: demographic, socioeconomic, and health characteristics of biodiversity recorders in Ireland. Citiz Sci Theory Pract 2020 Aug 6;5(1):16. [doi: 10.5334/cstp.283]
37. Hobbs SJ, White PCL. Motivations and barriers in relation to community participation in biodiversity recording. J Nat Conserv 2012 Dec;20(6):364-373. [doi: 10.1016/j.jnc.2012.08.002]
38. Lotfian M, Ingensand J, Brovelli MA. A framework for classifying participant motivation that considers the typology of citizen science projects. ISPRS Int J Geo Inf 2020;9(12):704. [doi: 10.3390/ijgi9120704]
39. West S, Pateman R. Recruiting and retaining participants in citizen science: what can be learned from the volunteering literature? Citiz Sci Theory Pract 2016;1(2):15. [doi: 10.5334/cstp.8]
40. Pernat N, Kampen H, Ruland F, Jeschke JM, Werner D. Drivers of spatio-temporal variation in mosquito submissions to the citizen science project “Mückenatlas”. Sci Rep 2021 Jan 14;11(1):1356. [doi: 10.1038/s41598-020-80365-3] [Medline: 33446753]
41. Tran T, Porter WT, Salkeld DJ, Prusinski MA, Jensen ST, Brisson D. Estimating disease vector population size from citizen science data. J R Soc Interface 2021 Nov;18(184):20210610. [doi: 10.1098/rsif.2021.0610] [Medline: 34814732]
42. Bradter U, Mair L, Jönsson M, Knape J, Singer A, Snäll T. Can opportunistically collected Citizen Science data fill a data gap for habitat suitability models of less common species? Methods Ecol Evol 2018 Jul;9(7):1667-1678. [doi: 10.1111/2041-210X.13012]
43. Fei S, Yu F. Quality of presence data determines species distribution model performance: a novel index to evaluate data quality. Landscape Ecol 2016 Jan;31(1):31-42. [doi: 10.1007/s10980-015-0272-7]
44. Porter WT, Motyka PJ, Wachara J, et al. Citizen science informs human-tick exposure in the Northeastern United States. Int J Health Geogr 2019 May 7;18(1):9. [doi: 10.1186/s12942-019-0173-0] [Medline: 31064416]
45. Diuk-Wasser MA, VanAcker MC, Fernandez MP. Impact of land use changes and habitat fragmentation on the eco-epidemiology of tick-borne diseases. In: Reisen W, editor. J Med Entomol 2021 Jul 16;58(4):1546-1564. [doi: 10.1093/jme/tjaa209] [Medline: 33095859]
46. Eisen L, Eisen RJ. Critical evaluation of the linkage between tick-based risk measures and the occurrence of Lyme disease cases. J Med Entomol 2016 Sep 1;53(5):1050-1062. [doi: 10.1093/jme/tjw092] [Medline: 27330093]
47. Eisen L, Eisen RJ. Benefits and drawbacks of citizen science to complement traditional data gathering approaches for medically important hard ticks (Acari: Ixodidae) in the United States. J Med Entomol 2021 Jan 12;58(1):1-9. [doi: 10.1093/jme/tjaa165] [Medline: 32772108]
48. Ogden NH, Lindsay LR, Hanincová K, et al. Role of migratory birds in introduction and range expansion of Ixodes scapularis ticks and of Borrelia burgdorferi and Anaplasma phagocytophilum in Canada. Appl Environ Microbiol 2008 Mar;74(6):1780-1790. [doi: 10.1128/AEM.01982-07] [Medline: 18245258]
49. Clow KM, Ogden NH, Lindsay LR, Michel P, Pearl DL, Jardine CM. The influence of abiotic and biotic factors on the invasion of Ixodes scapularis in Ontario, Canada. Ticks Tick Borne Dis 2017 Jun;8(4):554-563. [doi: 10.1016/j.ttbdis.2017.03.003] [Medline: 28359751]
50. Sonenshine DE. Range expansion of tick disease vectors in North America: implications for spread of tick-borne disease. Int J Environ Res Public Health 2018 Mar 9;15(3):478. [doi: 10.3390/ijerph15030478] [Medline: 29522469]
51. Cartographie de la répartition actuelle et future des tiques ixodes scapularis au québec en contexte de changements climatiques - rapport méthodologique. INSPQ. 2024. URL: https://www.inspq.qc.ca/publications/3483 [accessed 2024-05-23]
52. Dennis DT, Nekomoto TS, Victor JC, Paul WS, Piesman J. Reported distribution of Ixodes scapularis and Ixodes pacificus (Acari: Ixodidae) in the United States. J Med Entomol 1998 Sep;35(5):629-638. [doi: 10.1093/jmedent/35.5.629] [Medline: 9775584]
53. Population and dwelling counts: canada and census subdivisions (municipalities). Government of Canada SC. 2022. URL: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=9810000201 [accessed 2022-12-22]
54. Focus on geography series, 2021 census - Quebec, province. Government of Canada SC. 2022. URL: https://www12.statcan.gc.ca/census-recensement/2021/as-sa/fogs-spg/page.cfm?lang=E&topic=1&dguid=2021A000224 [accessed 2025-01-28]
55. Rapport de surveillance de la maladie de lyme et des autres maladies transmises par la tique ixodes scapularis au québec entre 2004 et 2013. INSPQ.: Institut national de santé publique du Québec; 2016. URL: https://www.inspq.qc.ca/publications/2160 [accessed 2023-08-16]
56. INSPQ. Portrait de la maladie de Lyme au Québec : 2006-2019 [Article in French]. : Gouvernement du Québec; 2022 URL: https://www.inspq.qc.ca/publications/2844 [accessed 2026-03-03]
57. Plan d’analyse de la surveillance intégrée de la maladie de lyme. INSPQ. 2017. URL: https://www.inspq.qc.ca/publications/2257 [accessed 2023-08-16]
58. Cartographie de la répartition actuelle et future des tiques Ixodes scapularis au québec en contexte de changements climatiques - rapport méthodologique. INSPQ. 2024. URL: https://www.inspq.qc.ca/publications/3483 [accessed 2025-09-02]
59. Data products, 2021 census. Government of Canada. 2021. URL: https://www12.statcan.gc.ca/census-recensement/2021/dp-pd/index-eng.cfm [accessed 2026-07-31]
60. 2021 census data quality guidelines. Government of Canada. 2021. URL: https://www12.statcan.gc.ca/census-recensement/2021/ref/98-26-0006/982600062021001-eng.cfm [accessed 2026-04-01]

XSL•FO
RenderX

61. Kornellia E, Syakurah RA. Use of Google Trends database during the COVID-19 pandemic: systematic review. Multidiscip Rev 2023;6(2):2023017. [doi: 10.31893/multirev.2023017]
62. Wisnieski L, Gruszynski K, Faulkner V, Shock B. Challenges and opportunities in One Health: Google Trends search data. Pathogens 2023 Nov 9;12(11):1332. [Medline: 38003796]
63. Carneiro HA, Mylonakis E. Google Trends: a web-based tool for real-time surveillance of disease outbreaks. Clin Infect Dis 2009 Nov 15;49(10):1557-1564. [doi: 10.1086/630200] [Medline: 19845471]
64. Trends help. Google. 2025. URL: https://support.google.com/trends/ [accessed 2025-07-28]
65. Schug F, Bar-Massada A, Carlson AR, et al. The global wildland-urban interface. Nature 2023 Sep;621(7977):94-99. [doi: 10.1038/s41586-023-06320-0] [Medline: 37468636]
66. Thornton MM, Shrestha R, Wei Y, Thornton PE, Kao SC. Daymet: daily surface weather data on a 1-km grid for North America, version 4 R1. ORNL Distributed Active Archive Center. Preprint posted online on Nov 1, 2022. [doi: 10.3334/ORNLDAAC/2129]
67. Ogden NH, Bigras-Poulin M, O'Callaghan CJ, et al. A dynamic population model to investigate effects of climate on geographic range and seasonality of the tick Ixodes scapularis. Int J Parasitol 2005 Apr 1;35(4):375-389. [doi: 10.1016/j.ijpara.2004.12.013] [Medline: 15777914]
68. Ogden NH, Ben Beard C, Ginsberg HS, Tsao JI. Possible effects of climate change on ixodid ticks and the pathogens they transmit: predictions and observations. J Med Entomol 2021 Jul 16;58(4):1536-1545. [doi: 10.1093/jme/tjaa220] [Medline: 33112403]
69. TCPS 2 (2022) – chapter 2: scope and approach. Government of Canada. 2022. URL: https://ethics.gc.ca/eng/tcps2-eptc2_2022_chapter2-chapitre2.html [accessed 2026-06-11]
70. Privacy policy – etick.ca and etick mobile application. eTick. URL: https://etick.ca/en/content/privacy-policy [accessed 2026-08-05]
71. R: a language and environment for statistical computing. The R Foundation. 2025. URL: https://www.R-project.org/ [accessed 2026-07-31]
72. Kulldorff M, Nagarwalla N. Spatial disease clusters: detection and inference. Stat Med 1995 Apr 30;14(8):799-810. [doi: 10.1002/sim.4780140809] [Medline: 7644860]
73. Bivand RS, Pebesma EJ, Gómez-Rubio V. Applied Spatial Data Analysis with R: Springer; 2008. [doi: 10.1007/978-0-387-78171-6]
74. Dean CB. Testing for overdispersion in poisson and binomial regression models. J Am Stat Assoc 1992 Jun;87(418):451-457. [doi: 10.1080/01621459.1992.10475225]
75. Gómez-Rubio V, Ferrándiz-Ferragud J, López-Quílez A, Bivand R. DCluster: functions for the detection of spatial clusters of diseases. 2024. URL: https://cran.r-project.org/web/packages/DCluster/DCluster.pdf [accessed 2026-03-26]
76. Loh JM, Zhu Z. Accounting for spatial correlation in the scan statistic. Ann Appl Stat 2007;1(2):560-584. [doi: 10.1214/07-AOAS129]
77. Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM. Mixed Effects Models and Extensions in Ecology with R: Springer; 2009. [doi: 10.1007/978-0-387-87458-6]
78. Zeileis A, Kleiber C, Jackman S. Regression models for count data in R. J Stat Softw 2008;27:1-25. [doi: 10.18637/jss.v027.i08]
79. Wang Z, Ma S, Wang CY. Variable selection for zero‐inflated and overdispersed data with application to health care demand in Germany. Biometrical J 2015 Sep;57(5):867-884 [FREE Full text] [doi: 10.1002/bimj.201400143]
80. Kutner M, Nachtsheim C, Neter J, Li W. Applied Linear Statistical Models: McGraw; 2005.
81. López-Ratón M, Rodríguez-Álvarez MX, Cadarso-Suárez C, Gude-Sampedro F. OptimalCutpoints: an r package for selecting optimal cutpoints in diagnostic tests. J Stat Softw 2014(61):1-36. [doi: 10.18637/jss.v061.i08]
82. Gallop RJ, Crits-Christoph P, Muenz LR, Tu XM. Determination and interpretation of the optimal operating point for ROC curves derived through generalized linear models. Understanding Stat 2003 Nov;2(4):219-242. [doi: 10.1207/S15328031US0204_01]
83. Vermont J, Bosson JL, François P, Robert C, Rueff A, Demongeot J. Strategies for graphical threshold determination. Comput Methods Programs Biomed 1991 Jun;35(2):141-150. [doi: 10.1016/0169-2607(91)90072-2] [Medline: 1914452]
84. Schäfer H. Constructing a cut-off point for a quantitative diagnostic test. Stat Med 1989 Nov;8(11):1381-1391. [doi: 10.1002/sim.4780081110] [Medline: 2692111]
85. Greiner M. Two-graph receiver operating characteristic (TG-ROC): a Microsoft-EXCEL template for the selection of cut-off values in diagnostic tests. J Immunol Methods 1995 Sep 11;185(1):145-146. [doi: 10.1016/0022-1759(95)00078-o] [Medline: 7665897]
86. Peng CYJ, So TSH. Logistic regression analysis and reporting: a primer. Understanding Stat 2002 Feb 2;1(1):31-70. [doi: 10.1207/S15328031US0101_04]
87. Hosmer DW, Lemeshow S. Applied Logistic Regression, 1st edition: Wiley; 2000. [doi: 10.1002/0471722146]
88. Census of population, 2021 – population centre (POPCTR). Government of Canada SC. 2021. URL: https://www12.statcan.gc.ca/census-recensement/2021/ref/dict/az/definition-eng.cfm?ID=geo049a [accessed 2026-04-02]

XSL•FO
RenderX

89. Bowser N, Bouchard C, Sautié Castellanos M, et al. Self-reported tick exposure as an indicator of Lyme disease risk in an endemic region of Quebec, Canada. Ticks Tick Borne Dis 2024 Jan;15(1):102271. [doi: 10.1016/j.ttbdis.2023.102271] [Medline: 37866213]
90. Westcott JR, Bowden JJ, Savage J, Doody KM. Rapid northward expansion of the blacklegged tick, Ixodes scapularis, in response to climate change. Glob Chang Biol 2025 Nov;31(11):e70591. [doi: 10.1111/gcb.70591] [Medline: 41195696]
91. Maladies transmises par les tiques. INSPQ. 2025. URL: https://www.inspq.qc.ca/zoonoses/tiques [accessed 2025-10-16]
92. Résultats annuels de surveillance des maladies transmises par les tiques: 2023. INSPQ. 2024. URL: https://www.inspq.qc.ca/zoonoses/tiques/surveillance/2023 [accessed 2025-04-09]
93. Mise à jour de la définition de la zone endémique de la maladie de Lyme au Québec en vue de son élargissement. INSPQ. 2024. URL: https://www.inspq.qc.ca/publications/3499 [accessed 2025-10-16]
94. Ogden NH, Lindsay LR. Effects of climate and climate change on vectors and vector-borne diseases: ticks are different. Trends Parasitol 2016 Aug;32(8):646-656. [doi: 10.1016/j.pt.2016.04.015] [Medline: 27260548]
95. Gabriele-Rivet V, Arsenault J, Badcock J, et al. Different ecological niches for ticks of public health significance in Canada. PLoS One 2015;10(7):e0131282. [doi: 10.1371/journal.pone.0131282] [Medline: 26131550]
96. Ogden NH, Dumas A, Gachon P, Rafferty E. Estimating the incidence and economic cost of Lyme disease cases in Canada in the 21st century with projected climate change. Environ Health Perspect 2024 Feb;132(2):27005. [doi: 10.1289/EHP13759] [Medline: 38349724]
97. Baldwin H, Landesman WJ, Borgmann-Winter B, Allen D. A geographic information system approach to map tick exposure risk at a scale for public health intervention. J Med Entomol 2022 Jan 12;59(1):162-172. [doi: 10.1093/jme/tjab169] [Medline: 34642748]
98. Dumas A, Bouchard C, Lindsay LR, Ogden NH, Leighton PA. Fine-scale determinants of the spatiotemporal distribution of Ixodes scapularis in Quebec (Canada). Ticks Tick Borne Dis 2022 Jan;13(1):101833. [doi: 10.1016/j.ttbdis.2021.101833]
99. Palmer JRB, Oltra A, Collantes F, et al. Citizen science provides a reliable and scalable tool to track disease-carrying mosquitoes. Nat Commun 2017;8(1):916. [doi: 10.1038/s41467-017-00914-9]
100. Aiello AE, Renson A, Zivich PN. Social media– and internet-based disease surveillance for public health. Annu Rev Public Health 2020 Apr 2;41(1):101-118. [doi: 10.1146/annurev-publhealth-040119-094402] [Medline: 31905322]
101. Nuti SV, Wayda B, Ranasinghe I, et al. The use of Google trends in health care research: a systematic review. In: Voracek M, editor. PLoS One 2014;9(10):e109583. [doi: 10.1371/journal.pone.0109583] [Medline: 25337815]
102. Rees EE, Sotoodeh M, Denis-Robichaud J, Carabin H, de Montigny SD. Predicting local COVID-19 emergences: a time-series classification approach and value of data from social media, search engines, and neighbouring regions. Epidemics 2026 Mar;54:100891. [doi: 10.1016/j.epidem.2026.100891] [Medline: 41604747]
103. Denis-Robichaud J, Rees EE, Daley P, et al. Linking opinions shared on social media about COVID-19 public health measures to adherence: repeated cross-sectional surveys of twitter use in Canada. J Med Internet Res 2024 Aug 13;26:e51325. [doi: 10.2196/51325] [Medline: 39137009]
104. Djorno C, Santillana M, Yang S. Restoring the forecasting power of Google Trends with statistical preprocessing. arXiv. Preprint posted online on Apr 9, 2025. [doi: 10.48550/arXiv.2504.07032]
105. Bouchard C, Leonard E, Koffi JK, et al. The increasing risk of Lyme disease in Canada. Can Vet J 2015 Jul;56(7):693-699. [Medline: 26130829]
106. Schulze TL, Jordan RA, Hung RW. Biases associated with several sampling methods used to estimate abundance of Ixodes scapularis and Amblyomma americanum (Acari: Ixodidae). J Med Entomol 1997 Nov;34(6):615-623. [doi: 10.1093/jmedent/34.6.615] [Medline: 9439115]
107. Schulze TL, Jordan RA. Daily variation in sampled densities of Ixodes scapularis and Amblyomma americanum (Acari: Ixodidae) nymphs at a single site—implications for assessing acarological risk. J Med Entomol 2022 Mar 16;59(2):741-751. [doi: 10.1093/jme/tjab213] [Medline: 34994380]
108. Tack W, Madder M, De Frenne P, Vanhellemont M, Gruwez R, Verheyen K. The effects of sampling method and vegetation type on the estimated abundance of Ixodes ricinus ticks in forests. Exp Appl Acarol 2011 Jul;54(3):285-292. [doi: 10.1007/s10493-011-9444-6] [Medline: 21431925]
109. Classification écologique du territoire québécois. Ministère des Forêts, de la Faune et des Parcs. 2021. URL: https://mrnf.gouv.qc.ca/nos-publications/classification-ecologique-territoire-quebecois/ [accessed 2026-02-16]
110. Isaac NJB, Jarzyna MA, Keil P, et al. Data integration for large-scale models of species distributions. Trends Ecol Evol 2020 Jan;35(1):56-67. [doi: 10.1016/j.tree.2019.08.006] [Medline: 31676190]
111. Clow KM, Leighton PA, Pearl DL, Jardine CM. A framework for adaptive surveillance of emerging tick-borne zoonoses. One Health 2019 Jun;7:100083. [doi: 10.1016/j.onehlt.2019.100083] [Medline: 30809583]
112. Ogden NH, Lindsay LR, Leighton PA. Predicting the rate of invasion of the agent of Lyme disease Borrelia burgdorferi. J Appl Ecol 2013 Apr;50(2):510-518. [doi: 10.1111/1365-2664.12050]
113. Gisondi S, Giardullo P, Lenzi A, Di Buccio E, Roversi PF, Campanaro A. Engaging me softly: comparing social drivers for continuative citizens' participation in a long-term citizen science initiative on protected species monitoring. PLoS One 2025;20(6):e0324701. [doi: 10.1371/journal.pone.0324701] [Medline: 40493641]

XSL•FO
RenderX

Abbreviations

AUC: area under the curve
CSD: census subdivision
DD: degree days
INSPQ: Institut National de Santé Publique du Québec
IRR: incidence risk ratio
LD: Lyme disease
OR: odds ratio
WUI: wildland-urban interface

Edited by S Brini; submitted 02.Jan.2026; peer-reviewed by P Chuard, S Maxwell, S Cork; revised version received 12.Jun.2026; accepted 16.Jun.2026; published 19.Aug.2026.

Please cite as:
Dumas A, Savage J, Crandall KE, Koffi JK, Leighton PA, Ogden NH, Rees EE, Bouchard C
Assessing the Effectiveness of Crowdsourced Data to Detect Established Tick Populations in Quebec, Canada: Retrospective Ecological Study
JMIR Public Health Surveill 2026;12:e90734
URL: https://publichealth.jmir.org/2026/1/e90734
doi:10.2196/90734

© Ariane Dumas, Jade Savage, Kirsten E Crandall, Jules K Koffi, Patrick A Leighton, Nicholas H Ogden, Erin E Rees, Catherine Bouchard. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 19.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Prevalence of Early Rheumatic Heart Disease Among Asymptomatic Students in Underserved Communities in Ethiopia: Cross-Sectional Observational Study

Amsalu Tomas Chuma[1,2], MTech, MSc; Desalew Mekonnen Kassie[3], MD; Jens-Uwe Voigt[4], MD, PhD; Melkamu Hunegnaw Asmare[1,5], PhD; Ahmed Saeed Youssef[4,6], MD, PhD; Carolina Varon[7], PhD; Rik Willems[4], MD, PhD; Michelle Yates[8], MD; Adane Petros Sikamo[3], MD; Yidnekachew Asrat Birhan[3], MD; Chunzhuo Wang[1], PhD; Bart Vanrumste[1], PhD

[1]Department of Electrical Engineering, eMedia Research Lab/ Center for Dynamical Systems, Signal Processing, and Data Analytics, Katholieke Universiteit Leuven, Leuven, Belgium

[2]College of Engineering, Addis Ababa Science and Technology University, Addis Ababa, Ethiopia

[3]College of Health Sciences, Addis Ababa University, Addis Ababa, Ethiopia

[4]Department of Cardiovascular Sciences, Katholieke Universiteit Leuven, Leuven, Belgium

[5]Institute of Technology, Center of Biomedical Engineering, Addis Ababa University, Addis Ababa, Ethiopia

[6]Department of Cardiovascular Medicine, Suez Canal University, Ismailia, Egypt

[7]Department of Electrical Engineering, STADIUS, Katholieke Universiteit Leuven, Leuven, Belgium

[8]Department of Pediatrics, Soddo Christian General Hospital, Wolaita Soddo, Ethiopia

Corresponding Author:
Amsalu Tomas Chuma, MTech, MSc
Department of Electrical Engineering
eMedia Research Lab/ Center for Dynamical Systems, Signal Processing, and Data Analytics
Katholieke Universiteit Leuven
Andreas Veisaliusstraat 13
Leuven, 3000
Belgium
Phone: 32 016 30 10 30
Email: amsaluthomas.chuma@kuleuven.be

Abstract

Background: Rheumatic heart disease (RHD) is a sequela of recurrent, untreated group A *Streptococcus* infections. RHD disproportionately affects children and young adults in the Global South. Intermittent mass screening of early RHD by using affordable tools in these disease-endemic regions is essential for effective prevention.

Objective: This study examined multimodal physiological data for assessing the prevalence of early RHD in a cohort of asymptomatic, at-risk students in rural Ethiopia.

Methods: A total of 584 asymptomatic children aged 10 to 20 years were randomly selected for screening and stratified into 2 groups (≤14 and >14 years). Electrocardiogram, phonocardiogram, and echocardiography screening were performed, with diagnoses based on the 2012 World Heart Federation criteria.

Results: After excluding 1.4% (8/584) of the children, who had nonrheumatic findings, 576 participants were analyzed, including 334 (58.0%) female and 242 (42.0%) male children. The mean age was 16.1 (SD 2.4) years (95% CI 15.9-16.3). Echocardiographic screening identified 19 cases of RHD (n=10, 52.6% borderline and n=9, 47.4% definite). Female children accounted for 68.4% (13/19) of cases, and the association between female sex and RHD was not statistically significant (odds ratio 1.59, 95% CI 0.60-4.25; *P*=.35). The estimated prevalence of RHD was 32.5 per 1000 population (95% CI 18.1-46.9; SE 7.4; 19/576, 3.3%), which was significantly higher than the most recent multicenter prevalence estimate of 19 per 1000 population (95% CI 13.9-23.4; odds ratio 2.12; *P*=.03). Mitral regurgitation was the predominant lesion (16/19, 84.2%), followed by mitral stenosis (2/19, 10.5%) and aortic regurgitation (1/19, 5.3%). Phonocardiogram analysis showed mitral regurgitation (10/19, 52.6%), mitral stenosis (2/19, 10.5%), and subclinical findings in the rest of patients with RHD. Prolonged PR intervals were observed in 10.5% (2/19) of the RHD-positive participants.

XSL•FO
RenderX

Conclusions: This study confirms persistent high prevalence of asymptomatic RHD among students in rural regions of Ethiopia. Although there was a female predominance in RHD incidence, the difference between the sexes was not statistically significant.

(JMIR Public Health Surveill 2026;12:e87039) doi:10.2196/87039

KEYWORDS

rheumatic heart disease; echocardiography screening; valvular heart diseases; electrocardiogram; phonocardiogram; prevalence

Introduction

Rheumatic heart disease (RHD) arises from a systemic immune response to β-hemolytic streptococcal throat infections, particularly affecting children younger than 15 years [1]. Untreated infections can lead to acute rheumatic fever (ARF) and subsequent cardiac damage resulting in RHD [2], which often remains asymptomatic in the early stages [3]. RHD has been a global health priority since 2018. It disproportionately affects low- and middle-income countries (LMICs) and socioeconomically marginalized populations despite being a preventable disease [4-6]. A recent World Health Organization report [7] stated that approximately 55 million people worldwide are affected by the disease, resulting in over 360,000 deaths each year. While the burden of RHD has decreased significantly to almost none in high-income countries over recent decades, it remains a critical public health issue in the Global South, particularly in regions with high rates of ARF. The prevalence of RHD in LMICs has increased by nearly 50% since 1990, with projections suggesting that this trend will continue until 2030 [8].

Epidemiological studies conducted in Ethiopia indicate that chronic RHD ranks among the most commonly diagnosed cardiovascular conditions, with valvular heart disease constituting 40% of annual cardiac cases—80% of which are linked to RHD [9]. However, accurate estimates of prevalence of RHD remain challenging due to variations in screening methods, demographic characteristics of the participants, and sampling techniques. An echocardiography-based overall pooled national prevalence estimate in students suggests approximately 27 cases per 1000 individuals [4]. A multicenter study involving 3238 students in regional cities reported definite RHD cases being 2.9 times more prevalent than borderline RHD (1.4% vs 0.5%, respectively) [10]. Subclinical RHD was reported 7 to 8 times more commonly than symptomatic RHD in at-risk populations (21.1 per 1000 individuals vs 2.7 per 1000 individuals, respectively) [11].

Furthermore, the reported prevalence of RHD among students in Ethiopia varies widely across studies (ranging from 3.2 to 56.7 per 1000 individuals), and most of these reports are over a decade old. For instance, echocardiography-based studies have reported a prevalence rate of 30.5 per 1000 individuals in a study including 2000 children [12], 56.7 per 1000 individuals in another study of 987 participants [13], 19 per 1000 individuals in 3238 participants [10], and 3.2 per 1000 individuals in 1874 students [14]. Socioeconomic and environmental factors such as overcrowding, income, and proximity to health centers have been associated with an increased risk of RHD [15]. These factors pose challenges to accurately estimating the current health burden of RHD and hinder the development of prophylactic interventions aimed at preventing recurrent ARF, which are critical to reduce the progression to subclinical RHD.

Subclinical RHD is characterized by the presence of RHD identified via echocardiography in the absence of detectable heart murmurs on auscultation and without accompanying clinical symptoms [16]. The mitral valve (MV) is the most involved valve in patients with RHD with subsequent regurgitation, stenosis, or combined lesion [1,2,17]. In the acute phase, valvulitis may present as mitral regurgitation (MR) due to annular dilation and chordal elongation. Chronic RHD often leads to mitral stenosis (MS) resulting from leaflet thickening, commissural fusion, and chordal shortening [2,17]. Most patients with chronic MR, particularly in the pediatric population, remain asymptomatic for many years [18]. However, over time, the condition can lead to left ventricular dysfunction, and patients typically present with clinical signs in later years of their life. Notably, physiological valvular regurgitations may occur in healthy children as well [19].

While echocardiography remains the gold standard for RHD detection, offering superior sensitivity and specificity compared with auscultation or electrocardiogram (ECG) screening [20], its implementation in resource-limited settings is often constrained by a shortage of skilled personnel and appropriate equipment [21,22]. Nevertheless, structural and functional hemodynamic abnormalities of rheumatic cardiac valves are often subtle, particularly during the early or subclinical stages of the disease. These abnormalities are undetectable through auscultation or ECG alone, resulting in poor sensitivity for identifying asymptomatic RHD. Furthermore, the cardiac hemodynamics of asymptomatic patients with RHD are frequently within normal limits [23]. ECG patterns are also inconsistent in RHD. The most common are prolonged PR intervals and QTc intervals [24-26]. Handheld ultrasound devices, on the other hand, demonstrate high diagnostic performance, with sensitivities reported to be between 92% and 96% when used by expert sonographers [20,27]. Although evidence on the diagnostic accuracy of phonocardiogram (PCG) and ECG remains limited, they can be regarded as potential prescreening tools to identify high-risk individuals who would benefit the most from confirmatory echocardiographic evaluation. Moreover, advances in machine learning provide a promising strategy to enhance their diagnostic utility, potentially improving automated screening sensitivity. Assessment of such scalable and cost-effective RHD screening strategies requires further validation in endemic populations.

Therefore, the aim of this study was to determine the current prevalence of RHD among at-risk, asymptomatic students using the 2012 World Heart Federation (WHF) echocardiographic criteria [2]. We also performed PCG and ECG assessments in students with echocardiography-confirmed RHD to estimate

XSL•FO
RenderX

the proportion of cases that would be missed in routine clinical screening. The PR, QTc, and QRS interval duration was measured for ECG analysis, whereas PCG recordings were assessed for the presence of any pathological systolic or diastolic murmurs. This highlights gaps in early detection for population-level surveillance.

Methods

Study Setting

This study was conducted on students recruited from 4 public schools in rural areas of southern Ethiopia (Multimedia Appendix 1). The schools were selected using a systematic random sampling method based on previous demographic surveys conducted at a nearby hospital. In this sample, participants were randomly selected from each grade level, ensuring an even distribution of age and sex. Data were collected by a team of 4 local health care workers, 2 cardiologists and 2 general practitioners.

Selection of Participants and Randomization

This study used a cross-sectional observational design to assess the prevalence of early RHD among asymptomatic underserved students at schools in rural Ethiopia. The students who met all the inclusion criteria, shown in Multimedia Appendix 1, were included in the study. The inclusion criteria ensured that only medically stable participants without interfering clinical conditions such as severe anemia, chronic lung or renal disease, and malnutrition were enrolled in the study. Children and adolescents outside the 10- to 20-year age range; who had any history of cardiac surgery, replaced cardiac valve, or implanted devices; and who were taking cardiac medications were excluded from screening. Subsequently, the registry of this target population was encoded and anonymized. A systematic random sampling method was used to select participants from each classroom in the study period (March 2023 to September 2023). First, the students in the classrooms were stratified by age groups of 14 years or below and above 14 years. Each student in the classroom was assigned a unique alphanumeric code, and approximately 30% of the numbered list was selected at random. The selected students and, in case of being minors, their parents were then informed about the research protocol via a participant information sheet.

Sample Size Determination

We based our assumptions on a literature review indicating a disease prevalence of at least 2.7% among students at schools [28]. To estimate such a prevalence with a precision of up to 2% at a 95% confidence level, we determined that a minimum sample size of 504 was necessary. Considering a potential dropout rate of 10%, we aimed to enroll 560 students. The sample size was divided proportionally among the participating schools.

Ethical Considerations

This study was approved by the ethics committees of University Hospitals Leuven (B3222022001075) and Soddo Christian Hospital (SCH1941015). The study protocol conformed to the ethical guidelines of the 1975 Declaration of Helsinki as reflected in a prior approval from the institutions' human research ethics committees. Written informed consent was obtained from all participants or their guardians. Data were anonymized to protect participant confidentiality. No compensation or tokens of appreciation were provided for participation in the study. To maintain confidentiality, unique identification codes were assigned to the study participants. The principal investigators securely stored and controlled access to the data, limiting it to authorized personnel only.

Statistical Analysis

Statistical analyses were conducted using the R software (version 4.1.0; R Foundation for Statistical Computing). Continuous variables were summarized as means and SDs for normally distributed data and medians for nonnormally distributed data. Normality was assessed using the Shapiro-Wilk test. For comparison of continuous variables between 2 groups, the Welch *t* test was used for normally distributed data, and the Mann-Whitney *U* test was used for nonnormally distributed data. Categorical variables were presented as counts and percentages, and comparisons were made using the chi-square test. *P* values of less than .05 were considered statistically significant. The prevalence of RHD was estimated per 1000 population. We calculated the sum of RHD-positive cases from the total number of participants examined and extrapolated the result to an estimate per 1000 population.

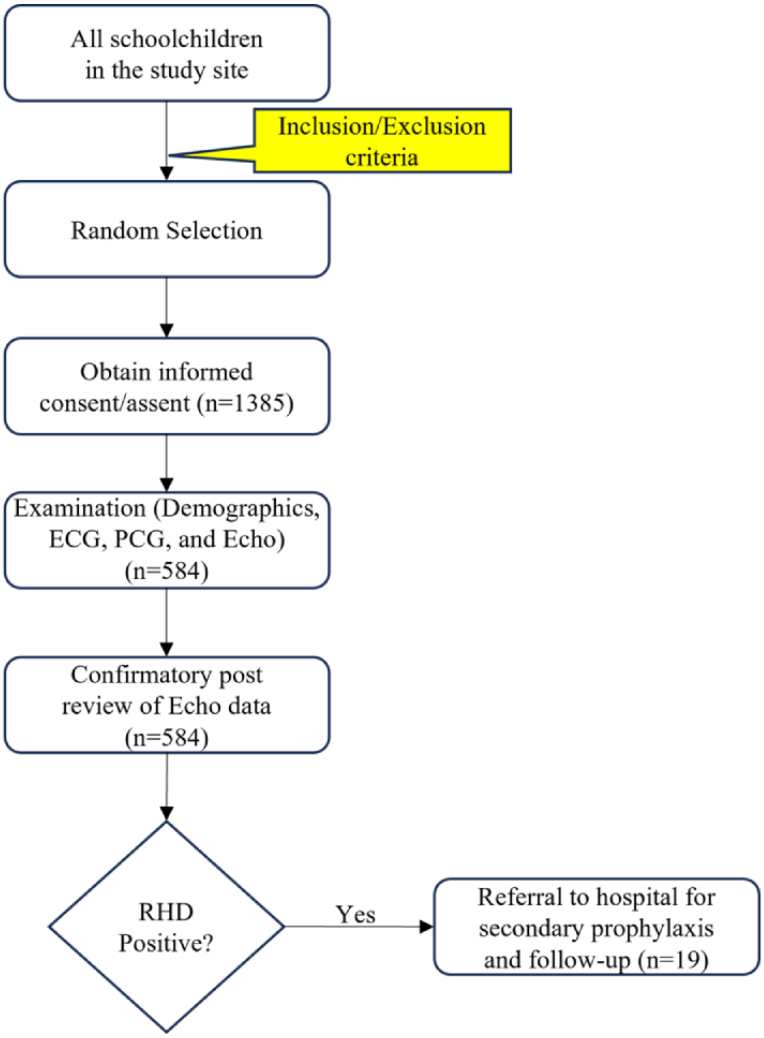
Data Collection Procedures

Overview

Study participants underwent a screening examination procedure, as shown in Figure 1. From each study participant, demographic data and ECG, PCG, and echocardiographic data were collected according to International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use good clinical practice. This is a widely adopted internationally agreed upon standard for designing, conducting, recording, and reporting clinical trials involving human participants. No clinical evaluation, including blood tests or swab samples related to ARF, were performed. A total of 1385 participants gave consent participants, of whom only 584 (42.2%) completed the ECG, PCG, and echocardiographic screening procedures, whereas the remainder missed one or more of the examinations mainly due to conflicting schedules, anxiety about echocardiographic screening, peer pressure, or unspecified personal reasons.

XSL•FO
RenderX

Figure 1. Screening examination flow diagram outlining participant inclusion and recruitment; the sequence of electrocardiogram (ECG), phonocardiogram (PCG), and echocardiographic (Echo) examinations; and the final diagnostic decision. RHD: rheumatic heart disease.



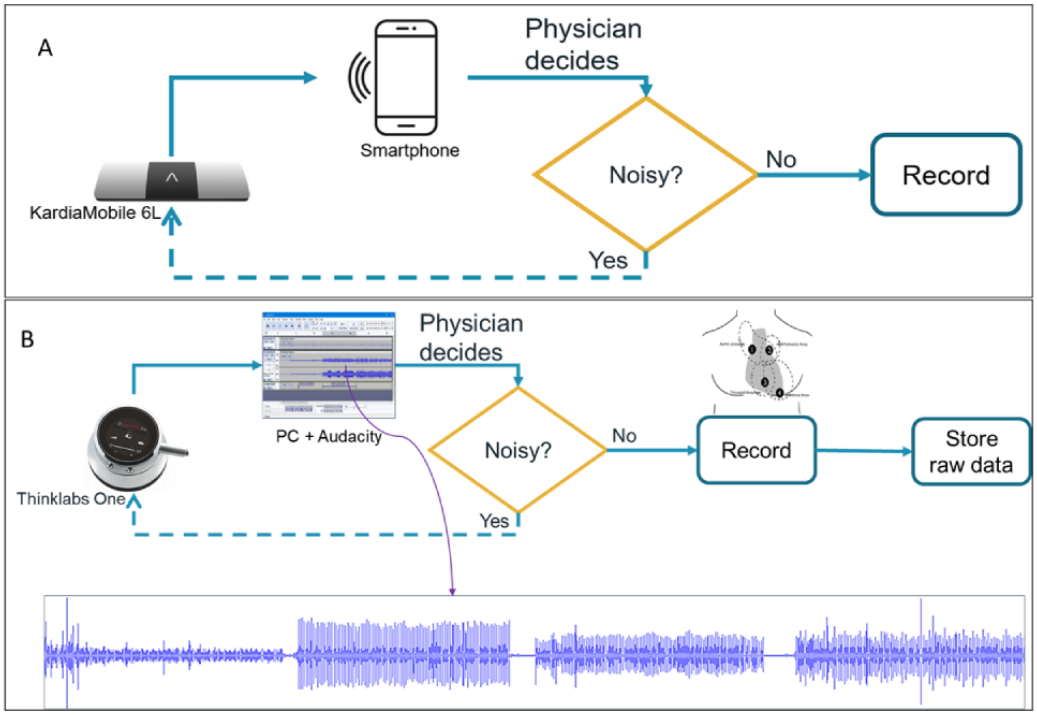
ECG Data

ECG data were recorded using a commercial CE-marked device, KardiaMobile 6L (AliveCore, Inc). KardiaMobile is small pocket device that is connected via Bluetooth to a compatible smartphone that has the Kardia app from AliveCor, Inc, installed, shown in Figure 2A. The device has 3 electrodes with a sampling frequency of 300 Hz. For obtaining an ECG recording, participants must touch 3 electrodes on the device, 1 on the back with their left leg and 2 on the front with the fingers of the left and right hand. ECG recordings were obtained at least twice, both from the knee or ankle positions, and repeated if deemed necessary (Figure 3C-D). Each recording had a duration of 30 seconds.

The temporal evaluations on ECG waveforms were performed manually by 2 physicians. After performing a 50-Hz powerline interference and Butterworth zero-phase band-pass filter from 1 to 100 Hz, an averaged ECG was computed by aligning and averaging the cardiac cycles in the lead II waveform using the R wave as a reference point. Three ECG intervals, PR, QRSd, and QTc, were then evaluated for RHD-positive participants from an averaged ECG. Prolonged changes in these intervals have been reported in patients with RHD [24-26]. The intervals were normalized using the heart rate (HR) and calculated considering a threshold margin of –10 to +10 ms for the QRSd interval and –15 to +15 ms for the PR and QTc intervals. HR-adjusted PR intervals [29] were used to determine PR prolongation. QTc values for each heartbeat were calculated using the formula by Bazett [30]. The QRS morphology was classified as normal if it was within the range of less than 105 ms considering the age of the cohort [31].

XSL•FO
RenderX

Figure 2. Connectivity block diagrams of the devices used during screening for rheumatic heart disease: (A) KardiaMobile device (AliveCor, Inc) and smartphone to record electrocardiogram signals and (B) Thinklabs One stethoscope with a computer to record phonocardiogram signals. In cases of noisy recordings, the physician performed a repeat acquisition to ensure signal quality.

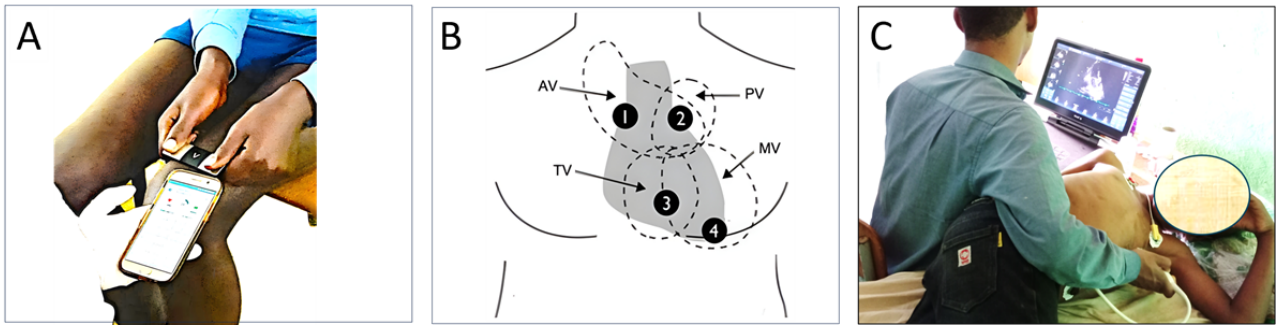


PCG Data

PCG data were recorded using a Thinklabs One stethoscope, a CE-marked medical device for auscultation. The stethoscope was connected to a PC, and input cardiac sounds were recorded using the Audacity open-source software (version 3.7.2; Muse Group) [32], as illustrated in Figure 2B. A wideband (20-2000 Hz) device mode audio filtering was set to acquire the heart sounds sampled at 44.1 kHz. The recording was taken from 4 positions: aortic, pulmonic, tricuspid, and mitral areas, as shown in Figure 3B.

The PCG recordings of mild regurgitations from participants with a BMI above 24 kg/m^2 were not audible, so different positioning strategies, such as left lateral decubitus, sit-up-lean-forward, and supine, were used to obtain a better signal. A total of 60 seconds of cardiac sound data were recorded from the aortic, pulmonic, tricuspid, and mitral areas. The recorded auscultations were played back and listened to by 2 physicians. Any pathological murmur during systole over the apex was considered an MR. Diastolic murmurs were labeled as aortic regurgitation (AR) or MS based on the intensity and duration of the murmur. Any other murmur was labeled as "other."

Figure 3. Data acquisition from the participants screened for rheumatic heart disease: (A) electrocardiogram data acquisition from the knee position, (B) 4 locations from where phonocardiogram data acquisition was performed (1=aortic area, 2=pulmonic area, 3=tricuspid area, and 4=mitral area), and (C) transthoracic echocardiography acquisition. AV: aortic valve; MV: mitral valve; PV: pulmonic valve; TV: tricuspid valve.



Echocardiography

A comprehensive transthoracic echocardiogram was performed by trained cardiologists using a portable Vivid iq (GE Vingmed Ultrasound) equipped with an M5Sc 2.2-MHz transducer (Figure 3C). Image acquisition was performed with the patient in a left decubitus position. Cardiac size and function, as well as the site and severity of the valvular lesion, were assessed using B-mode, spectral Doppler, and color Doppler imaging. We recorded clips with 3 to 5 cardiac cycles or 10 cycles in the case of atrial fibrillation. Afterward, the recorded images were independently reviewed by 2 cardiologists, and decisions were made using the WHF 2012 criteria.

XSL•FO
RenderX

The WHF 2012 criteria [2] (Multimedia Appendix 1) detail the diagnosis for definite RHD, borderline RHD, and normal echocardiographic findings in A, B, C, and D evidence groupings. Definite RHD requires either pathological regurgitation (MR or AR) combined with at least 2 morphological features of RHD in the affected valve, MS with a mean gradient of 4 mm Hg or higher, or concurrent borderline features in both the MV and aortic valve (AV). Borderline RHD is characterized by isolated morphological valve abnormalities or isolated pathological MR or AR without meeting the full criteria for definite RHD. Borderline RHD diagnosis cannot be based on a single abnormality; it requires either mild pathological regurgitation accompanied by at least one morphological feature or multiple morphological features of valvular involvement. Normal findings may include physiological regurgitation and isolated morphological changes without associated pathological stenosis or regurgitation. Participants with other nonrheumatic findings or comorbidities were excluded from the final analysis. Participants and/or their legal representatives were notified about findings indicative of RHD. Furthermore, all RHD-positive patients were referred to a nearby hospital for further guideline-recommended clinical management.

Results

Prevalence of RHD

A total of 584 study participants (n=340, 58.2% female and n=244, 41.8% male) were included in the study, of whom 8 (1.4%) were found to have nonrheumatic findings and excluded. Thus, 576 participants (n=334, 58% female and n=242, 42% male) were analyzed, as shown in Table 1. The nonrheumatic cases were observed mostly in female individuals (n=6) than in male individuals (n=2); 4 participants (n=3 female and n=1 male) had a first-degree MV prolapse with no rheumatic etiology, and 1 male participant had aortic sclerosis, whereas a female participant had atrial septal defect. Two additional participants exhibited physiological regurgitations associated with congenital AV anomalies: a bicuspid valve with right-left coronary cusp fusion (female individual) and a quadricuspid valve with a smaller accessory cusp (male individual).

Study participant characteristics are shown in Table 2, and the distribution of confirmatory results among the cohort overall and among RHD-positive participants is shown in Figure 4 and Figure 5, respectively. The overall mean age was 16.1 (SD 2.4; 95% CI 15.9-16.3) years. Participants younger than 13 years accounted for only 15.5% (89/576) of the total cohort due to challenges encountered in obtaining consent from legal guardians. A total of 5.2% (30/576) had a BMI above 24 kg/m^2, and 86.7% (26/30) of these participants were female. By applying the WHF 2012 guidelines mentioned in the Echocardiography section, we identified 19 cases of RHD (n=10, 52.6% borderline RHD and n=9, 47.4% definite RHD). Extrapolating from our data, the suggested prevalence of RHD in the age group of 10 to 20 years was 32.5 per 1000 persons (95% CI 18.1-46.9; SE 7.4). The mean age of the RHD-positive participants was 17.0 (SD 2.2) years, with the earliest incidence of RHD being at 13 years. The RHD-positive participants were predominantly female (13/19, 68.4% vs 6/19, 31.6% male), although the difference was not statistically significant (odds ratio 1.59, 95% CI 0.60-4.25; P=.35). Overall, 3.9% (13/334) of the female participants were RHD positive compared to 2.5% (6/242) of the male participants. There was no statistically significant correlation between RHD and household size (P=.49; t=0.705), age (P=.09; t=1.771), or sex (N=576, χ^2_1=0.8; P=.36).

Table 1. Summary of demographic data.

Variables	Primary school participants (n=296)	Secondary school participants (n=280)
Age (y), mean (SD)	13.5 (1.8)	18.0 (1.2)
Sex, n (%)		
Female	202 (68.2)	132 (47.1)
Male	94 (31.8)	148 (52.9)
Height (cm), mean (SD)	156 (12.0)	167 (8.1)
Weight (kg), mean (SD)	47.3 (9.8)	55 (7.7)
BMI (kg/m^2), mean (SD)	19.2 (2.9)	19.8 (2.5)
Family size, mean (SD)	6 (2)	7 (2)

XSL•FO
RenderX

Table 2. Summary of demographic data of rheumatic heart disease (RHD)–positive and healthy participants.

Variables	Healthy (n=557)	RHD-positive (n=19)
Age (y), mean (SD)	16.1 (2.4)	17.0 (2.2)
10-13, n (%)	86 (15.4)	1 (5.3)
14-16, n (%)	207 (37.2)	7 (36.8)
17-20, n (%)	264 (47.4)	11 (57.9)
Sex, n (%)		
Female	321 (57.6)	13 (68.4)
Male	236 (42.4)	6 (31.6)
BMI (kg/m^2), mean (SD)	19.5 (2.7)	20.0 (3.2)
HR[a] (bpm[b]), mean (SD)	77 (14)	80 (13)

[a]HR: heart rate.
[b]bpm: beats per minute.

Figure 4. Distribution of confirmatory results for rheumatic heart disease (RHD)–negative and RHD-positive (both borderline and definite) participants. F: female; M: male.

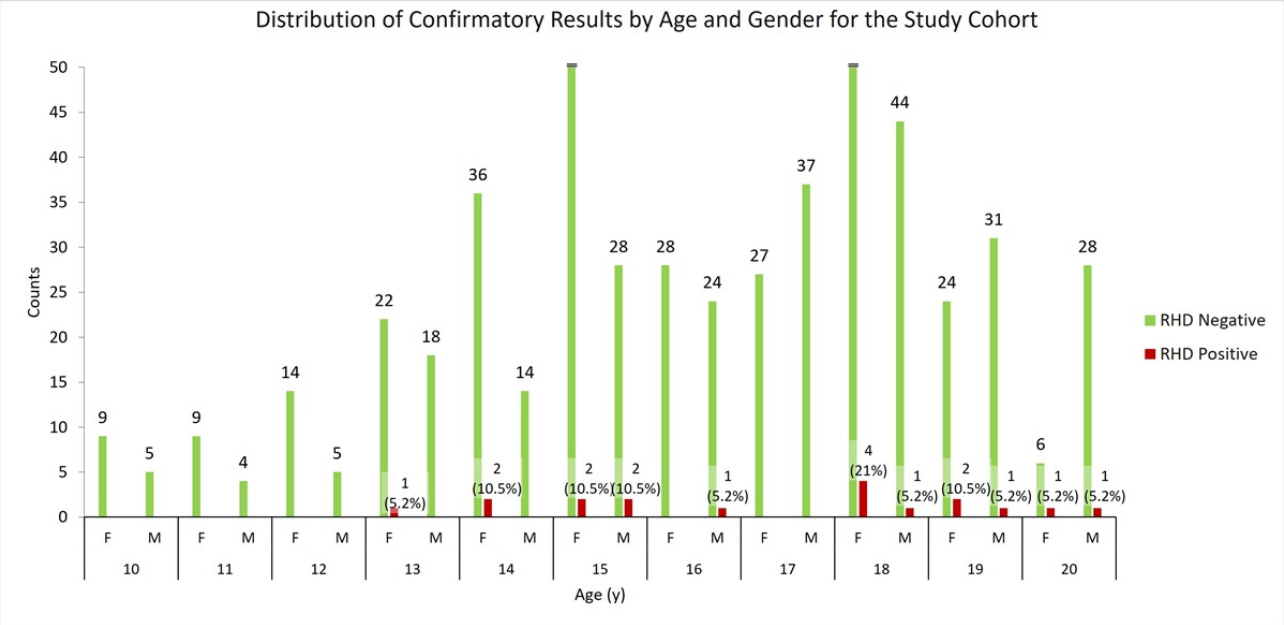
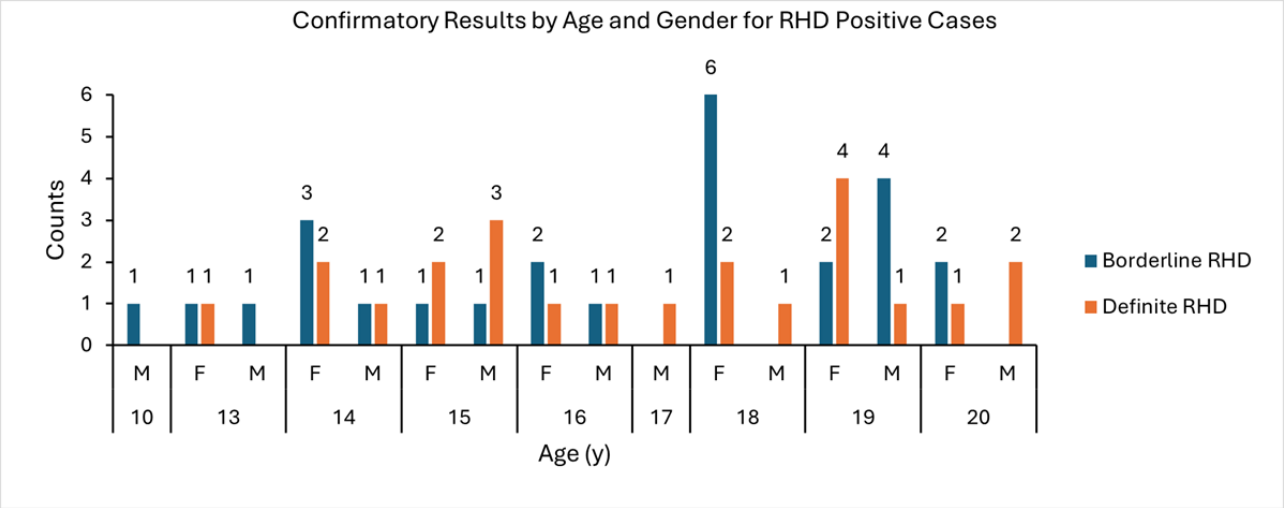


Figure 5. Distribution of confirmatory results for participants with borderline and definite rheumatic heart disease (RHD) only. F: female; M: male.

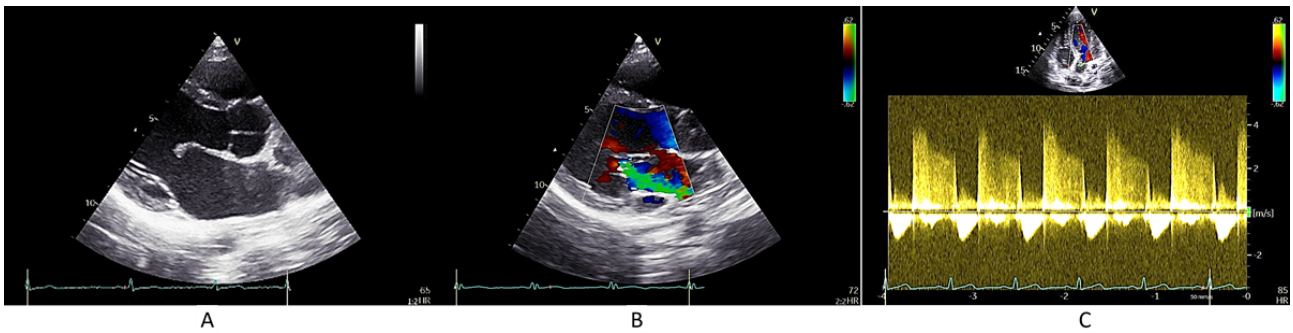


XSL•FO
RenderX

Echocardiographic Findings

We analyzed the involvement of the valves and their structures according to the WHF 2012 A, B, C, and D groupings discussed in the Echocardiography section and Multimedia Appendix 1. Illustrative echocardiographic images from an RHD-positive participant are shown in Figure 6 [3]. We observed that thickening of the MV leaflets and restricted leaflet motions were dominant features. Specifically, 73.7% (14/19; 9/14, 64.3% female; and 5/14, 35.7% male) of RHD-positive participants exhibited restricted MV leaflet motion. Excessive anterior leaflet motion during systole was observed in 47.4% (9/19; n=7 female and n=2 male) of the participants, whereas posterior leaflet annulus thickening occurred in 21.1% (4/19; all female). Additionally, irregular or focal thickening of the AV was noted in 31.6% (6/19; n=4 female and n=2 male) of the participants. While pathological MR and AR were the most common findings, observed in 84.2% (16/19; 10/16, 62.5% female and 6/16, 37.5% male) of the participants, MS was observed in 10.5% (2/19; all female) of the participants, and isolated AR was observed in only 5.3% (1/19) of the participants. MS coexisted with MR in both cases (Multimedia Appendix 1).

Figure 6. Pathological findings based on World Heart Federation 2012 criteria [3] for a patient with rheumatic heart disease: (A) parasternal long axis (PLAX) view showing doming of anterior mitral leaflet and restricted posterior mitral leaflet, (B) mitral regurgitation via color Doppler imaging in PLAX view, and (C) aortic regurgitation in continuous-wave Doppler imaging. The mitral valve is shown using an arrow.



ECG Findings

The average HR values showed a monotonically decreasing trend as age increased. We observed a nonsignificant tendency toward higher HR in female individuals than in male individuals across all ages (mean 86, SD 9 beats per minute vs mean 67, SD 10 beats per minute in RHD-positive participants and mean 81, SD 12 beats per minute vs mean 71, SD 12 beats per minute in RHD-negative participants, respectively; both not significant with the corresponding P=.08 and P=.79. Abnormal rhythm irregularities were not observed in the RHD-positive participants. In total, 10.5% (2/19) of the RHD-positive participants (both female) had prolonged QTc above the normal range of 450 ms. A QRS duration above 100 ms was observed in 10.5% (2/19) of the RHD-positive participants (both female). A prolonged PR interval was also observed in 10.5% (2/19; n=1 female and n=1 male) of the RHD-positive participants.

PCG Findings

PCG evaluation results of the RHD-positive participants are shown in Table 3. The findings showed MR in 52.6% (10/19) of the participants and MS in 10.5% (2/19) of the participants. The murmur was faint in 15.8% (3/19) of the participants, whereas it was determined as “other” in 21.1% (4/19) of the participants. Systolic murmurs were mostly isolated MRs without MS. In students with a BMI above 24 kg/m^2, S1 was often muffled due to an increased amount of tissue that affects the resonance of sound waves. However, there were no findings of pleural effusion or pericardial effusion that may have attenuated the heart sounds in these participants.

Such a muffled S1 heart sound with an audible S4 gallop was observed in 5.3% (1/19) of the RHD-positive participants. Moreover, a split S1 sound, suggestive of delayed left ventricular conduction, was observed in 10.5% (2/19) of the participants. For the first participant in the latter case, a 2D echocardiograph image and M-mode ultrasound showed asynchronous myocardial wall motion where contraction of the left ventricular septal wall succeeded the right ventricular contraction. Early closure of the MV preceding tricuspid valve closure, combined with thickened anterior mitral leaflets and excessive leaflet tip motion, contributed to low-velocity regurgitant flow into the left atria during systole. In the second participant, synchronous conduction with progressively prolonged contraction intervals were observed.

Table 3. Phonocardiogram evaluation results on confirmed rheumatic heart disease–positive participants (n=19).

Sex	MR[a], n (%)	MS[b], n (%)	Other, n (%)	Normal, n (%)
Female	8 (42.1)	1 (5.3)	2 (10.5)	1 (5.3)
Male	2 (10.5)	1 (5.3)	2 (10.5)	2 (10.5)

[a]MR: mitral regurgitation.

[b]MS: mitral stenosis.

XSL•FO
RenderX

Discussion

Principal Findings

Overview

This study investigated the current prevalence of asymptomatic RHD among students aged 10 to 20 years in underserved areas in Ethiopia. The potential use of affordable clinical examination for periodic surveillance via PCG and ECG was assessed on echocardiography-confirmed cases. Our findings showed that there is a persistent high prevalence of RHD in at-risk communities, with female predominance. Echocardiography revealed MV involvement as the most common manifestation of RHD, typically presenting with leaflet thickening and restricted tip motion. Definite participants exhibited characteristic murmurs, whereas borderline or asymptomatic participants often showed nonspecific auscultatory findings. Low-cost ECG and PCG were insufficient for early detection as PR interval prolongation did not reliably indicate early-stage disease, although this prolongation was more prominent in advanced chronic participants [24-26]. These results underscore the importance of echocardiography-based screening for timely identification and management of subclinical RHD.

Prevalence

RHD is highly prevalent among students in endemic regions, affecting female individuals more often [1,2]. This study also found an asymptomatic RHD prevalence rate of 32.5 per 1000 population among students, which is consistent with reports from Ethiopia (Multimedia Appendix 1). However, our figures are slightly higher than the pooled prevalence figures from other similar studies in LMICs [10,12]. The findings provide further evidence of a nondecreasing trend in RHD burden in high-risk areas, with persistent transmission of group A *Streptococcus* infections and gaps in primary prevention.

Distribution of RHD

As can be observed in Table 2, the RHD burden was 60% higher in female than in male individuals. Most of the RHD-positive participants (16/19, 84.2%) were older than 15 years, which aligns with trends observed in previous studies [10]. Overall, the asymptomatic presence of early-stage RHD in school-aged children and adolescents was not strongly correlated with household size, age, or sex in this cohort. While elevated prevalence in female individuals has historically been linked to childbearing age susceptibility [21], our data suggest elevated rates among even younger school-aged girls, which prompts further investigation into the potential role of biological, environmental, or sociocultural risk factors. Disease severity also correlated with advancing age, likely reflecting cumulative valvular damage from untreated ARF recurrences, as described in longitudinal cohorts [11]. However, the pathophysiological basis for sex disparities remains unclear. Importantly, while overcrowding was reported to have a positive and strong association with the prevalence of ARF and subsequently RHD, no strong associations were found for socioeconomic factors such as household income, parental occupation, or literacy levels [15]. This reinforces the necessity of population-wide screening irrespective of demographic proxies in disease-prone regions for early diagnosis and timely treatment of group A *Streptococcus* infections.

Valvular Involvement

Valvular involvement patterns in our cohort align with global epidemiological profiles, where the MV is most frequently affected, followed by the AV, consistent with reports from Australia, New Zealand, and Uganda [19,20]. This sequence of valve involvement is likely attributed to the MV's hemodynamic vulnerability during ARF, where inflammatory insults preferentially target valve leaflets and chordae [17]. Leaflet thickening and restricted motion dominated MV pathology, confirming the revised 2023 WHF RHD diagnostic criteria. As reported in the study by Webb et al [19], isolated MR was more prevalent than MS in asymptomatic participants, thereby supporting the paradigm of early-stage rheumatic damage characterized by inflammatory retraction and chordal elongation rather than fibrotic fusion. However, combined MR and AR occurred in 84.2% (16/19) of the participants in this study, which is consistent with previously reported findings [11]. Conversely, isolated AR was uncommon in this study cohort, with a single case reported. This is inconsistent with the results reported in the study by Rothenbühler et al [11], who found AR prevalence to be 21%. This discrepancy may reflect regional variations in ARF severity, delayed diagnosis, or genetic susceptibility to multivalvular involvement [4-6].

When these patterns of structural and functional involvement are interpreted within the framework of echocardiographic staging of rheumatic valvular disease in the 2012 WHF criteria [2], early inflammatory valvular changes can be distinguished from chronic advanced RHD. Borderline RHD is suggested by isolated rheumatic morphological abnormalities without pathological regurgitation, such as mitral leaflet thickening, posterior mitral annular thickening, restricted leaflet motion, or irregular or focal thickening of the AV. Definite RHD is suggested when these characteristic morphological changes are accompanied by pathological MR and/or AR, indicating established RHD with functional hemodynamic consequences. However, prolonged PR intervals were only observed in participants with borderline RHD (2/19, 10.5%), not in those with definite RHD.

Contribution of Screening Modalities

It is noteworthy that PCG screening detected pathogenic audible murmurs in approximately two-thirds of the participants (12/19, 63.2%), which surpassed the levels previously reported in auscultation-based studies [28]. This supports the hypothesis that subclinical RHD generates subtle murmurs associated with early valvular changes without a clinically pathological murmur [2]. However, in endemic regions, clinically silent RHD is 7 to 8 times more prevalent than symptomatic RHD [11]. This highlights the limitations of murmur-dependent screening and underscores the need for multimodal data evaluation and echocardiography as the gold standard.

The ECG findings revealed QRSd and QTc interval prolongation in RHD-positive participants, consistent with conduction abnormalities attributed to myocardial fibrosis or inflammatory blockages [24-26]. These changes exhibited a correlation with

XSL•FO
RenderX

the timing of ARF and disease severity. However, these results differed from those of studies that emphasized arrhythmias due to atrial ectopy and heart block in advanced RHD [24,26]. Atrial ectopy increases the risk of atrial fibrillation, whereas heart block typically produces bradyarrhythmia due to impaired electrical conduction. This is because rhythmic disturbances were minimal in our early-stage cohort. This observation suggests that arrhythmic manifestations may emerge at a later stage in the disease progression, although longitudinal data are necessary to substantiate this claim.

In summary, our findings advocate for integrated, technology-enhanced screening frameworks to combat RHD in resource-limited settings. Recent advancements in wearable ultrasounds for cardiac imaging have demonstrated potential to make imaging more cost-effective while maintaining comparable performance to that of standard portable echocardiography devices [33]. Hu et al [33] proposed an automated approach that uses deep learning algorithms to estimate the volume of the left ventricle and interpret B-mode ultrasound images with an intersection over union score above 85%, thereby facilitating more precise and efficient analysis.

Furthermore, the evaluation of MR jet length as a single criterion to detect RHD has shown sensitivity between 73% and 78.9% and specificity of 82.4% to 87.3% [4]. The application of such wearable devices with simplified criteria in combination with low-cost PCG or ECG sensors could improve diagnosis and reduce global RHD-associated burden. The development of such a cost-effective and scalable approach relies on the availability of large, diverse, and multimodal datasets capable of capturing the complex and heterogeneous features of RHD to support accurate and timely diagnosis. Future studies should explore using predictive modeling of multimodal ECG, PCG, and echocardiogram data to stratify the risk of RHD progression and evaluate cost-effective interventions in real-world settings.

Limitations

The study had limitations. First, a high dropout rate was observed, with 57.8% (801/1385) of enrolled students not completing all the screening examinations, introducing a skewed age distribution toward older students. Consequently, the generalizability of the findings to the broader pediatric population, particularly those under 13 years of age, is limited, and the statistical power to detect significant effects was reduced. Second, the results indicate a high RHD burden among students in a rural, disease-endemic area. The study sample's demographic profile reflects a low socioeconomic status in crowded living conditions; the prevalence among urban, relatively healthier students is unclear. Third, although recurrent ARF is a major precipitating factor for RHD, structured clinical data on ARF episodes were unavailable for analysis. Future studies should include students from diverse socioeconomic backgrounds to provide a more comprehensive understanding of RHD prevalence and risk factors.

Conclusions

The findings of this study indicate a persistently high prevalence of early RHD, predominantly involving the MV, among asymptomatic students (32.5 per 1000 population). The RHD incidence was higher in female individuals, but the difference between the sexes did not reach statistical significance. In resource-limited settings where echocardiographic facilities are unavailable, clinical screening based on persistent mitral or aortic regurgitant murmurs remains an important approach for large-scale case identification. However, clinical assessment of ECG and PCG alone is insufficient for holistic detection of RHD as a substantial proportion of cases may remain clinically silent. Therefore, persistent screening programs, particularly in school-based settings, are essential to facilitate timely identification and a prompt initiation of secondary prophylaxis. Nevertheless, the use of ECG and PCG assisted with artificial intelligence as a prescreening tool is worth further exploration as it may enable rapid, scalable community screening for early detection of RHD.

Acknowledgments

The authors are grateful to all the participants and study team members, the collaborating hospitals (University Hospitals Leuven, Tikur Anbessa Specialized Hospital, and Soddo Christian Hospital), and Flanders AI Research Program for their unwavering support. The authors also acknowledge the use of the generative artificial intelligence tool ChatGPT from OpenAI for improving the readability of the manuscript.

Funding

This research was funded by KU Leuven with reference number B/22/032 and the 5 E – Group T Fund with reference number REF23123123 under Leuven Centre for Affordable Health Technology. RW is supported as a fundamental clinical researcher by the Research Foundation – Flanders. The researchers received no financial compensation from the device manufacturers (Thinklabs; AliveCor, Inc; or GE Vingmed Ultrasound) to conduct this study. None of the companies were involved in the planning of the study, data analysis, or reporting of the data.

Data Availability

The Rheumatic Heart Disease database analyzed in this study is not currently available publicly due to the GDPR (General Data Protection Regulation) policies. However, it can be shared from the corresponding author on reasonable request.

XSL•FO
RenderX

Authors' Contributions

ATC conceptualized the study, performed computer programming, collected the dataset, carried out the experiments, and prepared the first draft of the manuscript. JUV, MHA, DMK, CV, RW, and MY designed dataset acquisition, analyzed the experiments and results, and reviewed the manuscript. APS and YAB collected the dataset, analyzed the results, and reviewed the manuscript draft. CW reviewed the experiments and the manuscript. BV conceptualized the study, designed the methodology, analyzed the experiments and results, and reviewed the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The echocardiographic diagnosis criteria used for diagnosis is given in A1. The map of the study site and inclusion exclusion criteria used for recruiting the participants is provided in A2 and A3, respectively. The overall echocardiograph-based prevalence of rheumatic heart disease among students in Ethiopia from previous studies is given in A4 and echocardiograph measurements of the rheumatic heart disease cases of this study cohort is given in A5.

[DOCX File , 2775 KB - publichealth_v12i1e87039_app1.docx]

References

1. Dass C, Kanmanthareddy A. Rheumatic heart disease. In: StatPearls [Internet]. Treasure Island, FL: StatPearls Publishing; 2023.
2. Reményi B, Wilson N, Steer A, Ferreira B, Kado J, Kumar K, et al. World Heart Federation criteria for echocardiographic diagnosis of rheumatic heart disease--an evidence-based guideline. Nat Rev Cardiol 2012 Feb 28;9(5):297-309 [FREE Full text] [doi: 10.1038/nrcardio.2012.7] [Medline: 22371105]
3. Ou Z, Yu D, Liang Y, Wu J, He H, Li Y, et al. Global burden of rheumatic heart disease: trends from 1990 to 2019. Arthritis Res Ther 2022 Jun 11;24(1):138 [FREE Full text] [doi: 10.1186/s13075-022-02829-3] [Medline: 35690787]
4. Passos LS, Nunes MC, Aikawa E. Rheumatic heart valve disease pathophysiology and underlying mechanisms. Front Cardiovasc Med 2021 Jan 18;7:612716. [doi: 10.3389/fcvm.2020.612716] [Medline: 33537348]
5. Dougherty S, Khorsandi M, Herbst P. Rheumatic heart disease screening: current concepts and challenges. Ann Pediatr Cardiol 2017;10(1):39-49 [FREE Full text] [doi: 10.4103/0974-2069.197051] [Medline: 28163427]
6. Baker MG, Gurney J, Oliver J, Moreland NJ, Williamson DA, Pierse N, et al. Risk factors for acute rheumatic fever: literature review and protocol for a case-control study in New Zealand. Int J Environ Res Public Health 2019 Nov 15;16(22):4515 [FREE Full text] [doi: 10.3390/ijerph16224515] [Medline: 31731673]
7. Rheumatic heart disease. World Health Organization. 2025 Jan 24. URL: https://www.who.int/news-room/fact-sheets/detail/rheumatic-heart-disease [accessed 2026-03-26]
8. Hu Y, Tong Z, Huang X, Qin JJ, Lin L, Lei F, et al. The projections of global and regional rheumatic heart disease burden from 2020 to 2030. Front Cardiovasc Med 2022 Oct 18;9:941917 [FREE Full text] [doi: 10.3389/fcvm.2022.941917] [Medline: 36330016]
9. Yadeta D, Guteta S, Alemayehu B, Mekonnen D, Gedlu E, Benti H, et al. Spectrum of cardiovascular diseases in six main referral hospitals of Ethiopia. Heart Asia 2017 Jun 19;9(2):e010829 [FREE Full text] [doi: 10.1136/heartasia-2016-010829] [Medline: 29492110]
10. Yadeta D, Hailu A, Haileamlak A, Gedlu E, Guteta S, Tefera E, et al. Prevalence of rheumatic heart disease among school children in Ethiopia: a multisite echocardiography-based screening. Int J Cardiol 2016 Oct 15;221:260-263. [doi: 10.1016/j.ijcard.2016.06.232] [Medline: 27404686]
11. Rothenbühler M, O'Sullivan CJ, Stortecky S, Stefanini GG, Spitzer E, Estill J, et al. Active surveillance for rheumatic heart disease in endemic regions: a systematic review and meta-analysis of prevalence among children and adolescents. Lancet Glob Health 2014 Dec;2(12):e717-e726 [FREE Full text] [doi: 10.1016/S2214-109X(14)70310-9] [Medline: 25433627]
12. Engel ME, Haileamlak A, Zühlke L, Lemmer CE, Nkepu S, van de Wall M, et al. Prevalence of rheumatic heart disease in 4720 asymptomatic scholars from South Africa and Ethiopia. Heart 2015 Sep 15;101(17):1389-1394. [doi: 10.1136/heartjnl-2015-307444] [Medline: 26076935]
13. Gemechu T, Mahmoud H, Parry EH, Phillips DI, Yacoub MH. Community-based prevalence study of rheumatic heart disease in rural Ethiopia. Eur J Prev Cardiol 2017 May;24(7):717-723. [doi: 10.1177/2047487316687104] [Medline: 28071960]
14. Mulatu HA, Woldemichael M, Aberra M. Prevalence of rheumatic heart disease among primary school students in mid-eastern Ethiopia. Biol Syst Open Access 2015;05(01). [doi: 10.4172/2329-6577.1000149]
15. Okello E, Kakande B, Sebatta E, Kayima J, Kuteesa M, Mutatina B, et al. Socioeconomic and environmental risk factors among rheumatic heart disease patients in Uganda. PLoS One 2012;7(8):e43917 [FREE Full text] [doi: 10.1371/journal.pone.0043917] [Medline: 22952810]

XSL•FO
RenderX

16. Godown J, Lu JC, Beaton A, Sable C, Mirembe G, Sanya R, et al. Handheld echocardiography versus auscultation for detection of rheumatic heart disease. Pediatrics 2015 Apr;135(4):e939-e944. [doi: 10.1542/peds.2014-2774] [Medline: 25780068]
17. Reményi B, Steer A, Cheung M. Rheumatic fever and rheumatic heart disease. In: Lai WW, Mertens LL, Cohen MS, editors. Echocardiography in Pediatric and Congenital Heart Disease: From Fetus to Adult. Hoboken, NJ: John Wiley & Sons; 2016:750-762.
18. Nishimura RA, Otto CM, Bonow RO, Carabello BA, Erwin JP3, Fleisher LA, et al. 2017 AHA/ACC focused update of the 2014 AHA/ACC guideline for the management of patients with valvular heart disease: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Circulation 2017 Jun 20;135(25):e1159-e1195. [doi: 10.1161/CIR.0000000000000503] [Medline: 28298458]
19. Webb RH, Gentles TL, Stirling JW, Lee M, O'Donnell C, Wilson NJ. Valvular regurgitation using portable echocardiography in a healthy student population: implications for rheumatic heart disease screening. J Am Soc Echocardiogr 2015 Aug;28(8):981-988. [doi: 10.1016/j.echo.2015.03.012] [Medline: 25959548]
20. Beaton A, Lu JC, Aliku T, Dean P, Gaur L, Weinberg J, et al. The utility of handheld echocardiography for early rheumatic heart disease diagnosis: a field study. Eur Heart J Cardiovasc Imaging 2015 May;16(5):475-482 [FREE Full text] [doi: 10.1093/ehjci/jeu296] [Medline: 25564396]
21. Aliyu IA, Bala JA, Yusuf I, Amole TG, Musa BM, Yahaya G, et al. Rheumatic heart disease burden in Africa and the need to build robust infrastructure. JACC Adv 2024 Dec 30;3(12):101347 [FREE Full text] [doi: 10.1016/j.jacadv.2024.101347] [Medline: 39817077]
22. Rwebembera J, Marangou J, Mwita JC, Mocumbi AO, Mota C, Okello E, et al. 2023 World Heart Federation guidelines for the echocardiographic diagnosis of rheumatic heart disease. Nat Rev Cardiol 2024 Apr;21(4):250-263. [doi: 10.1038/s41569-023-00940-9] [Medline: 37914787]
23. Ferrer MI, Harvey RM, Cathcart RT, Cournand A, Richards DWJ. Hemodynamic studies in rheumatic heart disease. Circulation 1952 Nov;6(5). [doi: 10.1161/01.cir.6.5.688] [Medline: 12988349]
24. Asmare MH, Chuma AT, Varon C, Woldehanna F, Janssens L, Vanrumste B. Characterization of rheumatic heart disease from electrocardiogram recordings. Physiol Meas 2023 Mar 01;44(2). [doi: 10.1088/1361-6579/aca6cb] [Medline: 36595302]
25. Sokolow M. Significance of electrocardiographic changes in rheumatic fever. Am J Med 1948 Sep;5(3):365-378. [doi: 10.1016/0002-9343(48)90087-4] [Medline: 18875467]
26. Alp H, Baysal T, Altın H, Karataş Z, Karaarslan S. QT and P-wave dispersions in rheumatic heart disease: prospective long-term follow up. Pediatr Int 2014 Oct;56(5):681-688. [doi: 10.1111/ped.12328] [Medline: 24628726]
27. Diamantino A, Beaton A, Aliku T, Oliveira K, Oliveira C, Xavier L, et al. A focused single-view hand-held echocardiography protocol for the detection of rheumatic heart disease. Cardiol Young 2018 Jan;28(1):108-117. [doi: 10.1017/S1047951117001676] [Medline: 28889812]
28. Asfaw T. Prevalence of rheumatic heart disease in Ethiopia: a systematic review and meta-analysis. SAGE Open Med 2023 Apr 29;11:20503121231169389. [doi: 10.1177/20503121231169389] [Medline: 37152842]
29. Park MK. Park's Pediatric Cardiology for Practitioners, Sixth Edition. Maryland Heights, MO: Mosby; 2014.
30. Bazett HC. An analysis of the time-relations of electrocardiograms. Heart 1920;7:353-370 [FREE Full text]
31. Palhares DM, Marcolino MS, Santos TM, da Silva JL, Gomes PR, Ribeiro LB, et al. Normal limits of the electrocardiogram derived from a large database of Brazilian primary care patients. BMC Cardiovasc Disord 2017 Jun 13;17(1):152 [FREE Full text] [doi: 10.1186/s12872-017-0572-8] [Medline: 28610624]
32. Audacity 3.4.2. Audacity. URL: https://support.audacityteam.org/additional-resources/changelog/older-versions/audacity-3.4/audacity-3.4.2 [accessed 2026-03-30]
33. Hu H, Huang H, Li M, Gao X, Yin L, Qi R, et al. A wearable cardiac ultrasound imager. Nature 2023 Jan;613(7945):667-675 [FREE Full text] [doi: 10.1038/s41586-022-05498-z] [Medline: 36697864]

Abbreviations

AR: aortic regurgitation
ARF: acute rheumatic fever
AV: aortic valve
ECG: electrocardiogram
GDPR: General Data Protection Regulation
HR: heart rate
LMIC: low- or middle-income country
MR: mitral regurgitation
MS: mitral stenosis
MV: mitral valve
PCG: phonocardiogram
RHD: rheumatic heart disease

XSL•FO
RenderX

WHF: World Heart Federation

Edited by A Mavragani, T Sanchez; submitted 03.Nov.2025; peer-reviewed by O Ahaiwe, E Abiy; comments to author 12.Jan.2026; revised version received 01.Feb.2026; accepted 25.Feb.2026; published 17.Apr.2026.

Please cite as:
Chuma AT, Kassie DM, Voigt JU, Asmare MH, Youssef AS, Varon C, Willems R, Yates M, Sikamo AP, Birhan YA, Wang C, Vanrumste B
Prevalence of Early Rheumatic Heart Disease Among Asymptomatic Students in Underserved Communities in Ethiopia: Cross-Sectional Observational Study
JMIR Public Health Surveill 2026;12:e87039
URL: https://publichealth.jmir.org/2026/1/e87039
doi:10.2196/87039
PMID:

©Amsalu Tomas Chuma, Desalew Mekonnen Kassie, Jens-Uwe Voigt, Melkamu Hunegnaw Asmare, Ahmed Saeed Youssef, Carolina Varon, Rik Willems, Michelle Yates, Adane Petros Sikamo, Yidnekachew Asrat Birhan, Chunzhuo Wang, Bart Vanrumste. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 17.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Population Estimates and Hypertension and Diabetes Prevalence: Cross-Sectional Quantitative Study Comparing Electronic Health Record–Derived Counts, Census, and Centers for Disease Control and Prevention Population Level Analysis and Community Estimates

Tyler NA Winkelman[1], MSc, MD; Kelly Bergmann[2], DO; Peter Bodurtha[1], MPP; Alanna M Chamberlain[3], PhD; R Adams Dudley[4], MD; David Haynes[4], PhD; Steven G Johnson[4], PhD; Thomas E Kottke[5], MSPH, MD; Karen L Margolis[5], MPH, MD; Gabriela Vazquez Benitez[5], MS, PhD; Devon Nerstad[4], BS; Patrick Olson[6], MPH, MSHI; James M Peacock[7], PhD; Nayanjot Kaur Rai[4], MPH; Stephen C Waring[8], PhD; Bjorn Westgard[5]; Paul Englund Drawz[4], MS, MHS, MD[‡]

[1]Hennepin Healthcare Research Institute, Minneapolis, MN, United States
[2]Children's Minnesota, Minneapolis, MN, United States
[3]Mayo Clinic, Rochester, MN, United States
[4]University of Minnesota, Minneapolis, MN, United States
[5]HealthPartners, St Paul, MN, United States
[6]Healthy Quin Counties, Thief River Falls, MN, United States
[7]Minnesota Department of Health, St Paul, MN, United States
[8]Essentia Health, Duluth, MN, United States
[‡]on behalf of the MN EHR Consortium

Corresponding Author:
Paul Englund Drawz, MS, MHS, MD
University of Minnesota
717 Delaware St SE
Minneapolis, MN, 55414
United States
Phone: 1 6126249444
Email: pedrawz00@yahoo.com

Abstract

Background: Accurate small-area estimates of vaccination rates and disease burden can inform public health interventions.

Objective: This study aimed to compare population denominators derived from census data and electronic health record (EHR) data from a statewide collaboration in Minnesota and examine concordance between Centers for Disease Control and Prevention and EHR-based estimates of diabetes and hypertension prevalence at the census tract level.

Methods: A retrospective study was conducted using EHR data from 2018 to 2022 from the Minnesota EHR Consortium (MNEHRC), population estimates from the 2020 census, and disease prevalence estimates among adults from the Centers for Disease Control and Prevention Population Level Analysis and Community Estimates (PLACES) project. Patients were included if they had a Minnesota address and a clinic visit in the last 3 years. Patients with hypertension and diabetes were identified based on the presence of at least 1 diagnosis code in the Observational Medical Outcomes Partnership condition occurrence table in the last 5 years or an elevated outpatient blood pressure (systolic blood pressure ≥140 mm Hg or diastolic blood pressure ≥90 mm Hg) on 2 or more days in the last 3 years for hypertension or at least 1 hemoglobin A_{1c} value of ≥6.5 in the last 3 years for diabetes.

Results: The 2020 census estimate for the population of Minnesota was 5,707,254. A total of 5,271,191 (92.4% of the census estimate) unique individuals visited 1 of the 11 MNEHRC health care systems in 3 years (2018-2020). The ratio of MNEHRC patients to the Minnesota statewide 2020 census estimate was higher for female individuals (0.97) than for male individuals (0.88)

XSL•FO
RenderX

and higher for older age groups (individuals aged 65 years and older: 1.05) than for younger age groups (individuals aged 0-17 years: 0.83). The MNEHRC patient-to-census ratio also differed by race—the ratio was the highest for Black Minnesotans (1.17) and the lowest for American Indian and Alaska Native Minnesotans (0.68). According to MNEHRC data, the percentage of adults in Minnesota with diabetes in 2022 was 9.5% (415,914/4,376,805), and the percentage of adults in Minnesota with hypertension in 2021 was 32.2% (1,365,413/4,234,000). Estimates from PLACES for diabetes were 9.9% (435,481/4,389,028) and for hypertension were 29.9% (1,311,459/4,389,028). The percentage of census tracts where the MNEHRC estimate was within 10% of the PLACES estimate was 40.3% (605/1500) for diabetes and 42.3% (635/1500) for hypertension; 77.9% (1168/1500) of census tracts for diabetes and 79.7% (1195/1500) for hypertension were within 25% agreement.

Conclusions: Our analysis suggests that there are both similarities and important differences between small-area estimates derived from EHR and survey data. Such differences suggest that further research is needed to determine the optimal collection method for local estimates of health conditions.

(JMIR Public Health Surveill 2026;12:e86337) doi:10.2196/86337

KEYWORDS

prevalence; population; hypertension; diabetes; census

Introduction

Accurate population-level denominators are essential for providing reliable estimates of disease burden and evaluating public health interventions [1]. Many governmental and academic entities rely on census-derived population estimates to determine the rate of procedures and prevalence of chronic diseases [2,3]. These estimates of condition prevalence can guide intervention efforts to improve health at the neighborhood, state, and national levels [4-6]. However, inaccuracies in census data and discrepancies between census data and health data measures—particularly among underrepresented populations—have at times led to anomalies in reporting [7]. For example, during the COVID-19 pandemic, vaccination rates exceeding 100% were reported in some groups in both California and Massachusetts [8,9]. In Minnesota, we developed a statewide distributed data network that captures electronic health record (EHR) data for more than 90% of the population, offering an alternative source for deriving population-level denominators for public health data [10-13].

In addition to syndromic surveillance, population estimates play a critical role in allocating public health resources and designing interventions to address inequities [14]. The Centers for Disease Control and Prevention (CDC), through the Behavioral Risk Factor Surveillance System (BRFSS) and its Population Level Analysis and Community Estimates (PLACES) project, reports estimates of chronic disease prevalence for adults aged 18 years and older at the census tract level using small-area modeling that incorporates survey and census data [15]. While these modeled estimates guide federal and state funding to address health disparities, their accuracy may be limited by sparse sampling at the neighborhood level. On average, BRFSS includes only approximately 5 individuals per census tract per year [16]. By contrast, our EHR-based system enables direct observation of diagnosed chronic disease prevalence in nearly every census tract across Minnesota, offering an opportunity to compare EHR-based and model-based small-area estimates [10].

In this study, we compared population denominators derived from census data and Minnesota-based EHRs and examined concordance between CDC and EHR-based estimates of hypertension and diabetes prevalence at the census tract level in the adult population. We assessed concordance between CDC-modeled and EHR-derived prevalence estimates across Minnesota and explored the area-level sociodemographic factors associated with prevalence estimates from EHR data and CDC PLACES. These findings have implications for improving the accuracy of population health surveillance and ensuring equitable distribution of public health resources.

Methods

Minnesota EHR Consortium

The Minnesota EHR Consortium (MNEHRC) is a collaboration among individuals from the 11 largest health care systems in Minnesota, the Minnesota Department of Health, local public health departments, the University of Minnesota, and Minnesota Community Measurement. Each of the 11 health care systems has mapped its EHR data to the Observational Medical Outcomes Partnership (OMOP) common data model [17]. Patient addresses have been geocoded to census tracts to facilitate area-level analyses. Individuals who have sought care at multiple health systems are deduplicated using a secure, privacy-preserving record linkage process with hashed identifiers from each health care system [18,19]. An algorithm based on frequency of care in the last 3 years assigns individuals to 1 health care system for reporting purposes.

Data are from the Health Trends Across Communities (HTAC) in Minnesota project, which includes EHR-derived prevalence estimates for several chronic conditions at the census tract level. Patients are eligible for inclusion for a given year if they are alive and aged 18 years or older on December 31 of that year and have had an encounter in their assigned health care system in the past 3 years. Therefore, the EHR data represent a health care–seeking population. Patients with hypertension were identified based on the presence of at least 1 hypertension diagnosis code in the OMOP condition occurrence table in the last 5 years or an elevated outpatient blood pressure (systolic blood pressure ≥140 mm Hg or diastolic blood pressure ≥90 mm Hg) on 2 or more days in the last 3 years. Patients with diabetes were identified based on the presence of at least 1 diabetes diagnosis code in the last 5 years or at least 1

XSL•FO
RenderX

hemoglobin A_{1c} value of ≥6.5 in the last 3 years. To compare with census data, we used 2020 counts for population estimates (individuals alive on December 31, 2020, with an encounter in 2018-2020). To evaluate concordance with the most recent data available from CDC, 2022 data were used for counts of diabetes, and 2021 data were used for counts of hypertension.

Census and CDC PLACES Estimates

Census data at the county level were obtained from the 2020 Census Demographic and Housing Characteristics Person and Housing tables [20]. Race and ethnicity data were grouped to be consistent with mutually exclusive categories in the EHR data: Hispanic and then non-Hispanic American Indian; Asian or Native Hawaiian or Other Pacific Islander; Black; White; and other (more than 1 race or other).

Census tract estimates of diabetes and hypertension prevalence among adults were obtained from the CDC PLACES data portal [21]. These estimates were based on results from BRFSS questions that ask whether individuals have ever been told by a physician or a health professional that they have diabetes or high blood pressure [22]. PLACES generates estimates at the census tract level using multilevel regression and poststratification that incorporates BRFSS results and the US Census Bureau's American Community Survey data [23].

Census Tract–Level Variables

Patient addresses were geocoded to a census tract at each of the MNEHRC health care systems. Census tracts for addresses that did not successfully geocode were assigned based on population ratios using a zip code to a census tract crosswalk from the US Department of Housing and Urban Development [24]. Census tract–level sociodemographic variables were obtained from multiple sources. Social Vulnerability Index (SVI), a measure "designed to identify and quantify communities experiencing social vulnerability," was obtained from the CDC [25]. For analytic purposes, the continuous SVI was categorized based on quartiles, with SVI 1 referring to those in the lowest quartile of vulnerability. Rurality was assigned based on the US Department of Agriculture's Economic Research Service rural-urban commuting area codes and the percentage of individuals living in rural areas from the US Census American Community Survey [19]. Census data were used to determine the age, race distribution, and the percentage of the population with a high school degree or General Educational Development [20].

Analyses

Statewide population counts for various demographic subgroups were tabulated from MNEHRC's 2020 HTAC data and compared with corresponding estimates from the 2020 census. County population counts from HTAC data were also compared with 2020 Census counts to assess coverage at the county level.

MNEHRC estimates of census tract hypertension and diabetes prevalence were calculated using 2021 and 2022 HTAC data, respectively. The obtained estimates were compared with CDC PLACES prevalence estimates, both of which were scaled 0 to 100. To assess agreement, the relative difference between the MNEHRC and PLACES estimates was computed. Bland-Altman plots were created to visualize agreement and systematic bias across all prevalence values.

Associations between tract demographic and socioeconomic characteristics and the prevalence of diabetes and hypertension were estimated separately for the 2 sources using linear regression. Predictors measured as proportions (eg, age >65 years, race, and poverty level) were modeled as a percentage of census tract population with a range from 0 to 100. Analyses were conducted using R (version 4.4.2; R Foundation for Statistical Computing).

Ethical Considerations

HTAC is a public health surveillance project conducted in partnership with the Minnesota Department of Health and local public health agencies and, therefore, was not considered research per 45 CFR 46.102(l)(2). Therefore, it was not subject to institutional review board approval [26].

Results

The 2020 census estimate for the population of Minnesota was 5,707,254. A total of 5,271,191 (92.4% of the census estimate) unique individuals visited 1 of the 11 MNEHRC health care systems in 3 years (2018-2020; Table 1). The ratio of MNEHRC patients to the Minnesota statewide census estimate was higher for female individuals (0.97) than for male individuals (0.88) and higher for the older age groups (65 years and older: 1.05) than for the younger age groups (0-17 years: 0.83). The MNEHRC patient-to-census ratio also differed by race—the ratio was highest for Black Minnesotans (1.17) and lowest for American Indian and Alaska Native Minnesotans (0.68). The MNEHRC data also had a greater number of individuals with missing, other, or unknown race (MNEHRC-to-census ratio 1.31). Coverage by county was generally consistent across the state, with a median MNEHRC patient-to-census estimate ratio of 0.89 (IQR 0.73-0.96; Figure 1).

XSL•FO
RenderX

Table 1. The 2020 Minnesota Electronic Health Record Consortium (MNEHRC) and census population estimates for Minnesota.

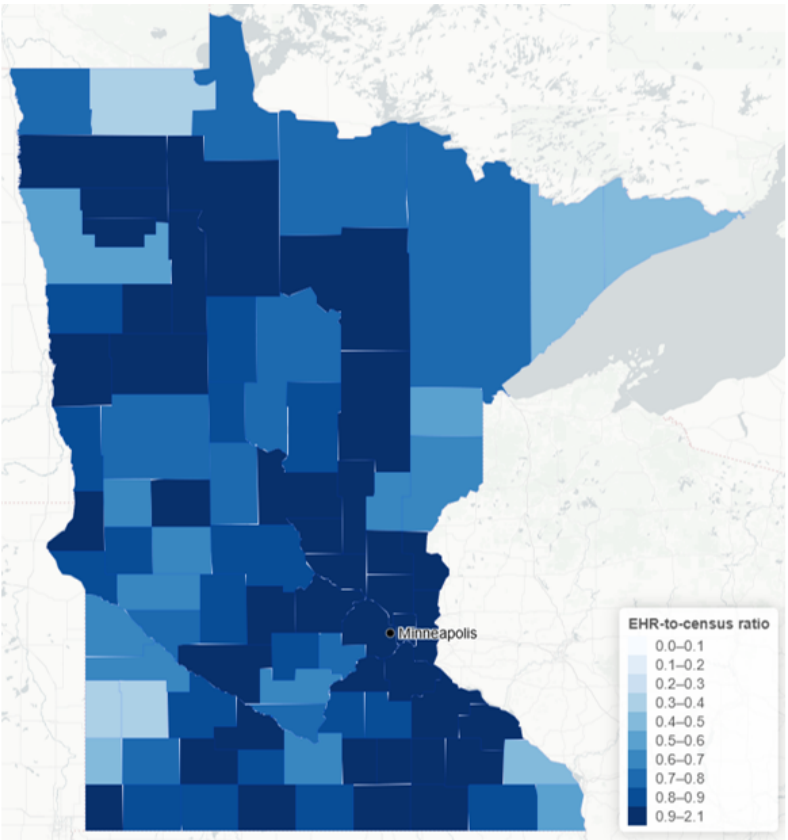
Characteristic	MNEHRC[a], n; census[b], n; ratio				
	All ages[c]	Aged 0-17 years[c]	Aged 18-44 years[c]	Aged 45-64 years[c]	Aged >65 years[c]
Sex					
Male	2,482,165; 2,835,448; 0.88	557,640; 673,827; 0.83	857,548; 1,012,726; 0.85	612,529; 714,245; 0.86	454,448; 434,650; 1.05
Female	2,789,026; 2,871,046; 0.97	531,812; 643,634; 0.83	1,027,682; 992,022; 1.04	688,909; 720,747; 0.96	540,623; 514,643; 1.05
Race or ethnicity					
American Indian or Alaska Native	38,614; 57,046; 0.68	9944; 17,486; 0.57	16,504; 21,439; 0.77	8745; 12,859; 0.68	3421; 5262; 0.65
Asian or Pacific Islander	259,865; 300,081; 0.87	74,740; 83,539; 0.89	113,407; 140,585; 0.81	50,488; 56,245; 0.90	21,230; 19,712; 1.08
Black	460,302; 392,850; 1.17	148,679; 135,497; 1.10	200,991; 161,303; 1.25	83,392; 72,834; 1.14	27,240; 23,216; 1.17
Hispanic	256,188; 345,640; 0.74	81,608; 126,873; 0.64	116,815; 150,957; 0.77	46,029; 54,280; 0.85	11,736; 13,530; 0.87
White	3,923,405; 4,353,880; 0.90	690,332; 844,827; 0.82	1,311,314; 1,438,283; 0.91	1,048,115; 1,197,790; 0.88	873,644; 872,980; 1.00
Other	336,761; 256,997; 1.31	85,216; 109,239; 0.78	128,012; 92,181; 1.39	65,072; 40,984; 1.59	58,461; 14,593; 4.01

[a]MNEHRC: Minnesota electronic health record consortium.

[b]Population estimates from the Minnesota statewide 2020 census.

[c]The values in all the columns indicate MNEHRC patients, Minnesota statewide 2020 census estimate, and the ratio of MNEHRC patients to the Minnesota statewide census estimate.

Figure 1. Ratio of 2020 Minnesota Electronic Health Record Consortium patient count to census population estimate, by county. EHR: electronic health record.



XSL•FO
RenderX

According to the MNEHRC data, the percentage of adults in Minnesota with diabetes in 2022 was 9.5% (415,914/4,376,805), and the percentage with hypertension in 2021 was 32.2% (1,365,413/4,243,000). Crude prevalence estimates from PLACES data for diabetes were 9.9% (435,481/4,389,028) and for hypertension were 29.9% (1,311,459/4,389,028). There was a moderate level of concordance between the prevalence of diabetes and hypertension at the census tract level among adults comparing MNEHRC and PLACES estimates (Figure 2). Relative to the PLACES estimates, MNEHRC estimates tended to be higher in high-prevalence census tracts. This was particularly true for hypertension. The percentage of census tracts where the MNEHRC estimate was within 10% of the PLACES estimate was 40.3% (605/1500) for diabetes and 42.3% (635/1500) for hypertension. The percentage of census tracts where the MNEHRC estimate was within 25% of the PLACES estimate was 77.9% (1168/1500) for diabetes and 79.7% (1195/1500) for hypertension (Figure 3).

Figure 2. Bland-Altman plots comparing diabetes and hypertension prevalence estimates from the Centers for Disease Control and Prevention (CDC) Population Level Analysis and Community Estimates and cross-sectional electronic health record data from the Minnesota Electronic Health Record Consortium (MNEHRC) using census tract–level data.

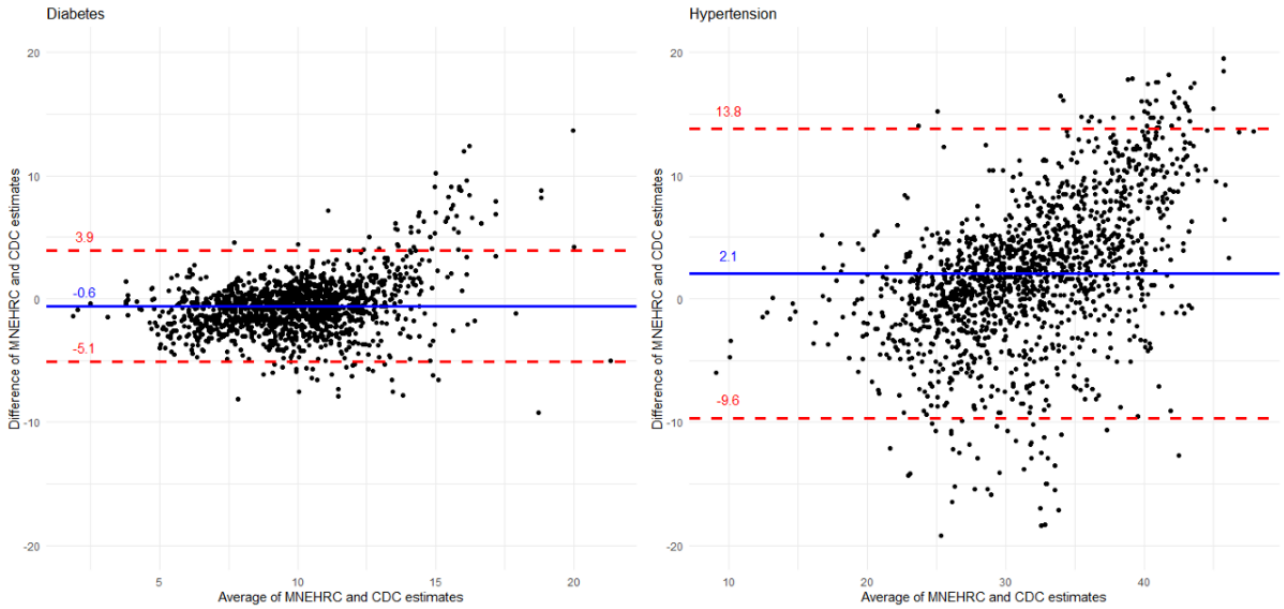
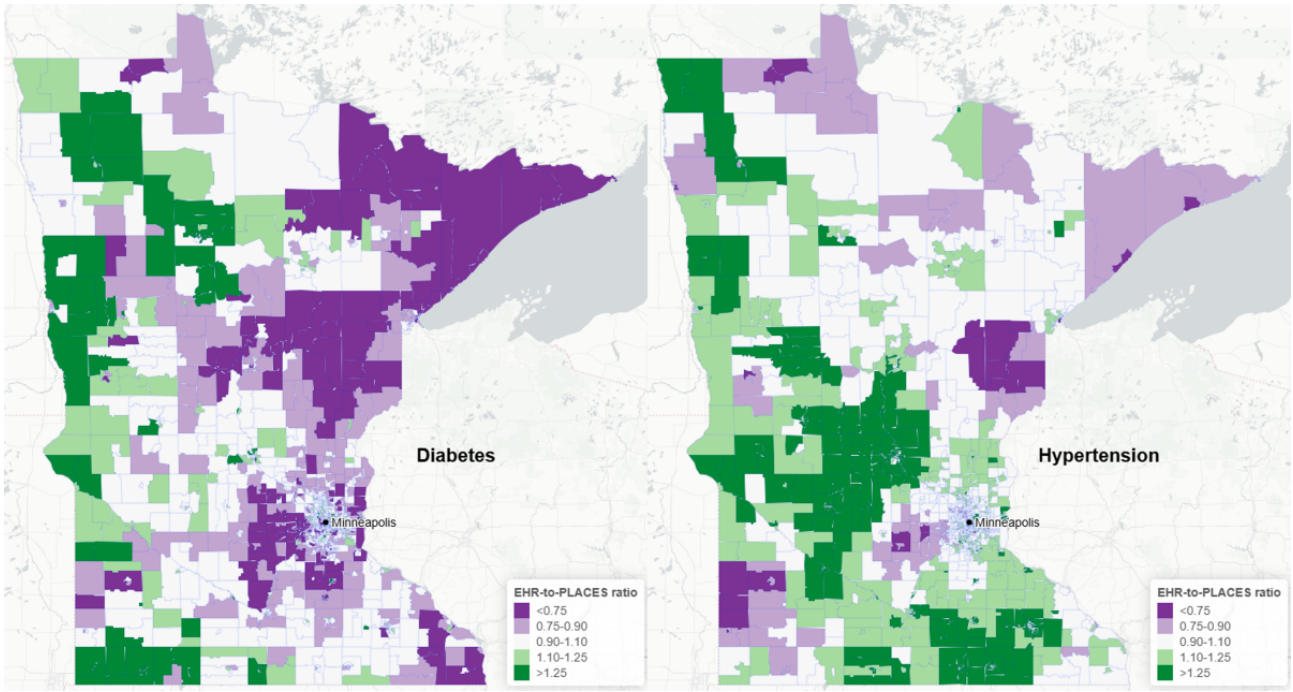


Figure 3. Ratio of Minnesota Electronic Health Record Consortium (MNEHRC) to Centers for Disease Control and Prevention Population Level Analysis and Community Estimates (PLACES) prevalence estimates for diabetes and hypertension at the census tract level. Blue represents census tracts where the MNEHRC estimate is lower than the PLACES estimate. Green represents census tracts where the MNEHRC estimate is higher than the PLACES estimate. White represents census tracts where the 2 estimates are within 10% of each other. EHR: electronic health record.



XSL•FO
RenderX

Area-level factors associated with diabetes and hypertension prevalence were generally similar in linear regression models for MNEHRC and PLACES (Table 2). Factors associated with higher prevalence in all models included a higher proportion of persons aged 65 years and older, a higher SVI, nonurban census tracts, and a lower proportion of individuals with at least a high school degree or General Educational Development. A higher proportion of Black individuals in a census tract was associated with lower hypertension prevalence in MNEHRC models but higher diabetes and hypertension prevalence in PLACES models. SVI was a stronger predictor of diabetes and hypertension in MNEHRC models compared with PLACES models. Model R^2 for diabetes was 0.46 in the MNEHRC model and 0.75 in the PLACES model; for hypertension, model R^2 was 0.52 in the MNEHRC model and 0.81 in the PLACES model.

Table 2. Census tract factors associated with hypertension and diabetes prevalence estimates from the Centers for Disease Control and Prevention Population Level Analysis and Community Estimates (PLACES) and a cross-sectional study using electronic health record data from the Minnesota Electronic Health Record Consortium (MNEHRC).

	MNEHRC, coefficients (95% CI)[a]	MNEHRC, *P* value	PLACES, coefficients (95% CI)[b]	PLACES, *P* value
Hypertension				
Individuals aged >65 years (%)	0.54 (0.48 to 0.60)	<.001	0.61 (0.59 to 0.64)	<.001
Black (%)	–0.10 (–0.14 to –0.06)	<.001	0.07 (0.05 to 0.09)	<.001
American Indian or Alaska Native (%)	–0.12 (–0.18 to –0.05)	<.001	0.05 (0.03 to 0.08)	<.001
Asian or HPI[c] (%)	–0.08 (–0.13 to –0.03)	.001	–0.05 (–0.07 to –0.03)	<.001
SVI[d] 2 (vs SVI 1)	1.62 (0.80 to 2.43)	<.001	0.07 (–0.28 to 0.41)	.70
SVI 3 (vs SVI 1)	2.19 (1.31 to 3.06)	<.001	0.55 (0.18 to 0.92)	.004
SVI 4 (vs SVI 1)	2.53 (1.43 to 3.63)	<.001	1.23 (0.76 to 1.70)	<.001
Exurban (vs urban)	4.16 (3.13 to 5.19)	<.001	3.00 (2.56 to 3.43)	<.001
Rural (vs urban)	4.25 (3.28 to 5.22)	<.001	3.31 (2.89 to 3.72)	<.001
Small town (vs urban)	5.98 (4.80 to 7.16)	<.001	2.03 (1.52 to 2.53)	<.001
Below poverty (%)	–0.05 (–0.08 to –0.03)	<.001	0.00 (–0.01 to 0.01)	.95
High school or General Educational Development (%)	–0.23 (–0.31 to –0.15)	<.001	–0.31 (–0.34 to –0.28)	<.001
Diabetes				
Individuals aged >65 years (%)	0.14 (0.12 to 0.16)	<.001	0.22 (0.21 to 0.24)	<.001
Black (%)	0.01 (–0.01 to 0.03)	.21	0.03 (0.02 to 0.04)	<.001
American Indian or Alaska Native (%)	0.07 (0.04 to 0.10)	<.001	0.07 (0.05 to 0.08)	<.001
Asian or HPI (%)	0.02 (–0.00 to 0.04)	.07	0.00 (–0.01 to 0.02)	.42
SVI 2 (vs SVI 1)	0.86 (0.52 to 1.20)	<.001	0.03 (–0.14 to 0.20)	.72
SVI 3 (vs SVI 1)	1.51 (1.14 to 1.87)	<.001	0.29 (0.11 to 0.48)	.002
SVI 4 (vs SVI 1)	1.88 (1.42 to 2.34)	<.001	0.46 (0.23 to 0.70)	<.001
Exurban (vs urban)	1.00 (0.57 to 1.43)	<.001	1.23 (1.01 to 1.45)	<.001
Rural (vs urban)	2.29 (1.89 to 2.70)	<.001	1.42 (1.22 to 1.63)	<.001
Small town (vs urban)	1.93 (1.44 to 2.43)	<.001	0.93 (0.68 to 1.18)	<.001
Below poverty (%)	–0.01 (–0.02 to –0.00)	.008	0.01 (0.00 to 0.01)	<.001
High school or General Educational Development (%)	–0.16 (–0.19 to –0.13)	<.001	–0.19 (–0.21 to –0.18)	<.001

[a]Model R^2 for hypertension = 0.52; Model R^2 for diabetes = 0.46.

[b]Model R^2 for hypertension = 0.81; Model R^2 for diabetes = 0.75.

[c]HPI: Hawaiian or Pacific Islander.

[d]SVI: Social Vulnerability Index.

XSL•FO
RenderX

Discussion

In this analysis comparing Minnesota statewide EHR data with census and CDC PLACES data, we demonstrated that population estimates from EHR data and the census are generally similar, though there are important differences by age, sex, and race. This finding has important implications for public health data; when possible, it is preferable to obtain numerators and denominators from the same source to avoid overestimating or underestimating prevalence or uptake of services. We also demonstrated moderate concordance between small-area prevalence estimates from EHR data and CDC PLACES. However, we did observe a greater than 25% relative difference in estimates for 22% of census tracts for diabetes and 20% of census tracts for hypertension. These results indicate the need for multiple data sources to accurately estimate small-area prevalence for chronic conditions.

Our results add to the literature evaluating the accuracy of BRFSS-based estimates. In the Boston validation study, BRFSS estimates were compared with estimates from a survey of 7340 individuals in 15 neighborhoods [27]. The BRFSS model estimates of diabetes prevalence were within the 95% CIs of the Boston survey for 80% of the neighborhoods. In the Missouri County-Level Study, 50,690 adults were surveyed across 115 counties [28]. The correlation between the BRFSS estimate for diabetes prevalence at the county level and the Missouri County-Level Study estimate was only 0.51. These prior studies were limited in that they compared estimates between 2 surveys. The results of this study allow a comparison between BRFSS and an estimate from a different data source, namely, EHR data.

Differences in population estimates between MNEHRC and the census may be due to multiple factors. Census undercounting may lead to a lower estimate of the number of Black Minnesotans. American Indian and Alaska Native Minnesotans may receive care at facilities that are not included in the MNEHRC, such as Indian Health Service clinics. The causes of differences in disease prevalence estimates between MNEHRC and PLACES are less apparent. PLACES may be inaccurate in certain census tracts due to their low survey rate, which necessitates modeling-based prevalence estimates. This may have led to some inaccuracies in prevalence estimates, which may be more apparent for the high-prevalence census tracts, as we observed with hypertension. Notably, the relationship between census tract characteristics and disease prevalence differed between MNEHRC and PLACES. A higher proportion of Black and American Indian or Alaska Native residents was associated with lower hypertension prevalence in MNEHRC data but higher hypertension prevalence in PLACES. This could represent differential access to care, lower rates of diagnoses, or PLACES model assumptions—such as age distribution—that are inaccurate in Minnesota. SVI was more strongly associated with hypertension and diabetes prevalence in MNEHRC models than in PLACES models, which may be due to a stronger association between SVI and prevalence in Minnesota compared with national estimates. The higher overall R^2 in the PLACES model is expected as those estimates are inherently model based using census tract characteristics.

Estimating small-area prevalence using EHR data has several potential advantages. First, EHR data can capture nearly the entire population. In our MNEHRC data, more than 90% of the census-estimated population had an encounter at a participating health care system within a 3-year period. Comparatively, surveys collect data from a small number of individuals within each census tract, and most census tracts are not represented in the National Health and Nutrition Examination Survey at all. This study included the availability of statewide data from a large, multi–health care system distributed EHR network that is scalable in other states. The number of adults in the EHR data for most census tracts is >75% of the census estimate for that population, which makes the EHR-based estimates robust. This allowed for, what we believe is, the first comparison of CDC PLACES– and EHR-based small-area estimates of diabetes and hypertension prevalence. Second, established data collection methods reduce the need for expensive and time-consuming survey collection and field work. Third, EHR data can be produced in near real time. While our analysis cannot comment on which approach (EHR based vs survey) is more accurate, it appears that using EHR data to assess area-level trends would provide similar information to survey-based data and reduce the risk of errors related to discrepancies between numerator and denominator sources. Both EHR data and survey data are prone to some misclassification. While inaccuracies in EHR data may be a particular concern for people who have infrequent health care use, a common concern with survey data is recall bias. Factors that would favor use of EHR data include (1) high percentage of population included in EHR data sources, (2) ability to use objective data to identify conditions (eg, blood pressure for hypertension and hemoglobin A_{1c} values for diabetes), (3) conditions where EHR data have a high sensitivity, (4) conditions with low prevalence that would require surveying thousands of individuals, (5) conditions that patients may be unlikely to report on surveys either due to stigma or because they are unaware they have the disease, (6) when disease prevalence estimates for key subpopulations would be informative, and (7) when assessing local temporal trends in disease prevalence.

Our study has important limitations. First, the EHR patient count may overestimate the population due to incomplete deduplication, though this is likely to be small. Second, EHR data are unable to account for out-migration, which might represent 1% to 2% of the effect on the population in any given year [29]. By contrast, the EHR patient count may underestimate the population due to individuals being seen by health care providers not included in MNEHRC and individuals who have not been seen by a health care provider in the last 3 years. Third, patients who never seek care may vary in important ways from health care–seeking patients, resulting in EHR-based population and prevalence estimates that do not reflect the entire population. By contrast, the census and PLACES data represent individuals who respond to surveys. Relatedly, both EHR estimates and survey estimates likely underestimate disease prevalence due to undiagnosed disease and lack of disease awareness, but EHR estimates may underestimate disease prevalence less than self-report. Fourth, our study does not allow us to identify a threshold of EHR data coverage above which it becomes more

XSL•FO
RenderX

appropriate to use EHR data rather than survey data for public health decision-making. The results of this study are relevant to distributed data networks with robust geographic coverage, particularly those that use OMOP. Finally, our models were estimated via ordinary least squares regression, which assumes independence across census tracts. Because tracts are spatially correlated, uncertainty surrounding coefficient estimates is likely understated, although the estimates themselves are unbiased.

In conclusion, we identified important differences between EHR and census data by age, sex, and race. These differences should be considered in public health reporting when 2 different data sources are used for the numerator and denominator, as was frequently done during the COVID-19 pandemic, where vaccination rates were estimated using census denominators. The prevalence of diabetes and hypertension at the census tract level was generally concordant between MNEHRC and PLACES, but differences suggest that EHR data may help inform targeting of public health interventions to reduce prevalence and adverse outcomes from chronic conditions. Our analysis suggests that there are both similarities and important differences between small-area estimates derived from EHR and survey data. Such differences suggest that further research with validation against an assessment of all individuals using clinical gold standards is needed to determine the optimal collection method for local estimates of health conditions.

Data Availability

The datasets generated or analyzed during this study are not publicly available as they include patient data from multiple health systems. Only aggregate results were shared centrally.

Funding

The study was funded by the Minnesota Department of Health as part of the Health Trends Across Communities project. Individuals from the Minnesota Department of Health and local public health agencies participated in the design of the study and interpretation of the results.

Authors' Contributions

Conceptualization: TNAW, PED
Data curation: PED, PB, DN
Formal analysis: PED, DN
Funding acquisition: TNAW
Methodology: TNAW, PB, AMC
Project administration: TNAW, NKR
Supervision: TNAW
Validation: DN, PED
Writing—original draft: PED
Writing—review & editing: TNAW, KB, PB, AMC, AD, DH, SGJ, TEK, KLM, GVB, DN, PO, JMP, NKR, SCW, BW

Conflicts of Interest

GVB reports receiving nonrelated research support from AbbVie and Sanofi.

References

1. Zhang X, Holt JB, Lu H, Wheaton AG, Ford ES, Greenlund KJ, et al. Multilevel regression and poststratification for small-area estimation of population health outcomes: a case study of chronic obstructive pulmonary disease prevalence using the behavioral risk factor surveillance system. Am J Epidemiol 2014 Apr 15;179(8):1025-1033. [doi: 10.1093/aje/kwu018] [Medline: 24598867]
2. Ramer S, Nguyen AT, Hollier LM, Rodenhizer J, Warner L, Whiteman MK. Abortion surveillance - United States, 2022. MMWR Surveill Summ 2024 Nov 28;73(7):1-28. [doi: 10.15585/mmwr.ss7307a1] [Medline: 39602470]
3. Razzaghi H, Wang Y, Lu H, Marshall KE, Dowling NF, Paz-Bailey G, et al. Estimated county-level prevalence of selected underlying medical conditions associated with increased risk for severe COVID-19 illness - United States, 2018. MMWR Morb Mortal Wkly Rep 2020 Jul 24;69(29):945-950 [FREE Full text] [doi: 10.15585/mmwr.mm6929a1] [Medline: 32701937]
4. Casper M, Kramer MR, Peacock JM, Vaughan AS. Population health, place, and space: spatial perspectives in chronic disease research and practice. Prev Chronic Dis 2019 Sep 05;16:E123 [FREE Full text] [doi: 10.5888/pcd16.190237] [Medline: 31489834]
5. Lu H, Wang Y, Liu Y, Holt JB, Okoro CA, Zhang X, et al. County-level geographic disparities in disabilities among US adults, 2018. Prev Chronic Dis 2023 May 11;20:E37 [FREE Full text] [doi: 10.5888/pcd20.230004] [Medline: 37167553]
6. Vaughan AS, Coronado F, Casper M, Loustalot F, Wright JS. County-level trends in hypertension-related cardiovascular disease mortality-United States, 2000 to 2019. J Am Heart Assoc 2022 Apr 05;11(7):e024785 [FREE Full text] [doi: 10.1161/JAHA.121.024785] [Medline: 35301870]

XSL•FO
RenderX

7. O'Hare WP. Differential Undercounts in the U.S. Census: Who is Missed?. Cham, Switzerland: Springer; 2019.
8. COVID-19 vaccine progress dashboard data by ZIP code. California Health and Human Services. URL: https://data.chhs.ca.gov/dataset/covid-19-vaccine-progress-dashboard-data-by-zip-code [accessed 2025-10-13]
9. Bor J, Assoumou SA, Lane K, Diaz Y, Ojikutu BO, Raifman J, et al. Inequities in COVID-19 vaccine and booster coverage across Massachusetts ZIP codes after the emergence of Omicron: a population-based cross-sectional study. PLoS Med 2023 Jan 31;20(1):e1004167 [FREE Full text] [doi: 10.1371/journal.pmed.1004167] [Medline: 36719864]
10. Zellmer L, Van Siclen R, Bodurtha P, Drawz PE, Waring SC, Chamberlain AM, et al. Estimating health condition prevalence among a statewide cohort with recent homelessness or incarceration. J Gen Intern Med 2025 Nov;40(15):3733-3742. [doi: 10.1007/s11606-025-09814-x] [Medline: 40897968]
11. Contag S, Dufort EM, Lim S, Winkelman T, Zipprich J, Erickson L, et al. Notes from the field: human parvovirus B19 infections among pregnant persons - Minnesota, January-September 2024. MMWR Morb Mortal Wkly Rep 2024 Nov 28;73(47):1087-1089 [FREE Full text] [doi: 10.15585/mmwr.mm7347a4] [Medline: 39602427]
12. DeSilva MB, Knowlton G, Rai NK, Bodurtha P, Essien I, Riddles J, et al. Vaccine effectiveness against SARS-CoV-2 related hospitalizations in people who had experienced homelessness or incarceration - findings from the Minnesota EHR Consortium. J Community Health 2024 Jun;49(3):448-457 [FREE Full text] [doi: 10.1007/s10900-023-01308-3] [Medline: 38066221]
13. Bozio CH, Butterfield KA, Briggs Hagen M, Grannis S, Drawz P, Hartmann E, et al. Protection from COVID-19 mRNA vaccination and prior SARS-CoV-2 infection against COVID-19-associated encounters in adults during Delta and Omicron predominance. J Infect Dis 2023 Jun 15;227(12):1348-1363. [doi: 10.1093/infdis/jiad040] [Medline: 36806690]
14. Crews DC, Patzer RE, Cervantes L, Knight R, Purnell TS, Powe NR, et al. Designing interventions addressing structural racism to reduce kidney health disparities: a report from a National Institute of Diabetes and Digestive and Kidney Diseases workshop. J Am Soc Nephrol 2022 Dec;33(12):2141-2152 [FREE Full text] [doi: 10.1681/ASN.2022080890] [Medline: 36261301]
15. Interactive PLACES map. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/places/tools/interactive-map-tool.html [accessed 2025-10-13]
16. Zhu L, Zhou W, Yang Y, Rao H, Liu Y, Rao Y. Association between cancer and cardiovascular disease risk: a cross-sectional study of 241,064 individuals. Front Oncol 2026 Feb 4;16:1734601 [FREE Full text] [doi: 10.3389/fonc.2026.1734601] [Medline: 41717419]
17. Standardized data: the OMOP Common Data Model. Observational Health Data Sciences and Informatics. URL: https://www.ohdsi.org/data-standardization/ [accessed 2025-10-13]
18. Vatsalan D, Sehili Z, Christen P, Rahm E. Privacy-preserving record linkage for big data: current approaches and research challenges. In: Zomaya A, Sakr S, editors. Handbook of Big Data Technologies. Cham, Switzerland: Springer; 2017.
19. Winkelman TN, Rai NK, Bodurtha PJ, Chamberlain AM, DeSilva M, Jeruzal J, et al. Trends in COVID-19 vaccine administration and effectiveness through October 2021. JAMA Netw Open 2022 Mar 01;5(3):e225018 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.5018] [Medline: 35357452]
20. Manson S, Schroeder J, Van Riper D, Knowles K, Kugler T, Roberts F, et al. IPUMS national historical geographic information system: version 18.0. IPUMS. 2023. URL: http://doi.org/10.18128/D050.V18.0 [accessed 2026-04-03]
21. PLACES Data Portal. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/places/tools/data-portal.html [accessed 2025-10-13]
22. Health outcomes. Centers for Disease Control and Prevention. 2024. URL: https://www.cdc.gov/places/measure-definitions/health-outcomes.html [accessed 2025-10-13]
23. Methodology. Centers for Disease Control and Prevention. 2025. URL: https://www.cdc.gov/places/methodology/index.html [accessed 2025-10-13]
24. HUD-USPS ZIP Code Crosswalk files. HUD User. URL: https://www.huduser.gov/portal/datasets/usps_crosswalk.html [accessed 2025-12-23]
25. Social vulnerability index. Agency for Toxic Substances and Disease Registry. URL: https://www.atsdr.cdc.gov/place-health/php/svi/index.html [accessed 2025-10-13]
26. eCFR: 45 CFR 46.102 -- definitions for purposes of this policy. National Archives and Records Administration. URL: https://www.ecfr.gov/current/title-45/subtitle-A/subchapter-A/part-46/subpart-A/section-46.102#p-46.102(l [accessed 2025-10-22]
27. Wang Y, Holt JB, Zhang X, Lu H, Shah SN, Dooley DP, et al. Comparison of methods for estimating prevalence of chronic diseases and health behaviors for small geographic areas: Boston validation study, 2013. Prev Chronic Dis 2017 Oct 19;14:E99 [FREE Full text] [doi: 10.5888/pcd14.170281] [Medline: 29049020]
28. Zhang X, Holt JB, Yun S, Lu H, Greenlund KJ, Croft JB. Validation of multilevel regression and poststratification methodology for small area estimation of health indicators from the behavioral risk factor surveillance system. Am J Epidemiol 2015 Jul 15;182(2):127-137 [FREE Full text] [doi: 10.1093/aje/kwv002] [Medline: 25957312]
29. An analysis of migration trends and patterns in Minnesota. Minnesota State Demographic Center. URL: https://mn.gov/admin/assets/Migration%20Report_FINAL_tcm36-620018.pdf [accessed 2025-10-13]

XSL•FO
RenderX

Abbreviations

BRFSS: Behavioral Risk Factor Surveillance System
CDC: Centers for Disease Control and Prevention
EHR: electronic health record
HTAC: Health Trends Across Communities
MNEHRC: Minnesota Electronic Health Record Consortium
OMOP: Observational Medical Outcomes Partnership
PLACES: Population Level Analysis and Community Estimates
SVI: Social Vulnerability Index

Edited by A Mavragani, T Sanchez; submitted 22.Oct.2025; peer-reviewed by O Ajileye, O Rosales, ME Wandai; comments to author 10.Dec.2025; revised version received 16.Mar.2026; accepted 25.Mar.2026; published 07.May.2026.

Please cite as:
Winkelman TNA, Bergmann K, Bodurtha P, Chamberlain AM, Dudley RA, Haynes D, Johnson SG, Kottke TE, Margolis KL, Benitez GV, Nerstad D, Olson P, Peacock JM, Rai NK, Waring SC, Westgard B, Drawz PE
Population Estimates and Hypertension and Diabetes Prevalence: Cross-Sectional Quantitative Study Comparing Electronic Health Record–Derived Counts, Census, and Centers for Disease Control and Prevention Population Level Analysis and Community Estimates
JMIR Public Health Surveill 2026;12:e86337
URL: https://publichealth.jmir.org/2026/1/e86337
doi:10.2196/86337
PMID:42097616

©Tyler NA Winkelman, Kelly Bergmann, Peter Bodurtha, Alanna M Chamberlain, R Adams Dudley, David Haynes, Steven G Johnson, Thomas E Kottke, Karen L Margolis, Gabriela Vazquez Benitez, Devon Nerstad, Patrick Olson, James M Peacock, Nayanjot Kaur Rai, Stephen C Waring, Bjorn Westgard, Paul Englund Drawz. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 07.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Correction: Awareness of Sexual Partner's HIV Status Among Men Who Have Sex With Men in China: Cross-Sectional Survey Study

Qian Ma[1], MA; Tingting Jiang[2], MA; Wanjun Chen[2], MA; Shaoqiang Jiang[3], BA; Jinlei Zheng[2], MA; Hui Wang[2], MA; Lin He[2], MA

[1]Department of Rheumatology Immunology & Allergy, The Children's Hospital, Zhejiang University School of Medicine, National Clinical Research Center for Child Health, Hangzhou, China

[2]Zhejiang Provincial Center for Disease Control and Prevention, 3399 Bin Sheng Road, Binjiang District, Hangzhou, China

[3]Coastal Service Center, Qiantang District, Hangzhou, Zhejiang, China

Corresponding Author:
Lin He, MA
Zhejiang Provincial Center for Disease Control and Prevention, 3399 Bin Sheng Road, Binjiang District, Hangzhou, China

Related Article:

Correction of: https://publichealth.jmir.org/2025/1/e66139/

(JMIR Public Health Surveill 2026;12:e101200) doi:10.2196/101200

In "Awareness of Sexual Partner's HIV Status Among Men Who Have Sex With Men in China: Cross-Sectional Survey Study" [1], the authors made one correction.

Affiliation 1 has been changed from the following:

> *Department of Rheumatology Immunology & Allergy, Zhejiang University School of Medicine, Hangzhou, China*

The affiliation now reads:

> *Department of Rheumatology Immunology & Allergy, The Children's Hospital, Zhejiang University School of Medicine, National Clinical Research Center for Child Health, Hangzhou, China*

The correction will appear in the online version of the paper on the JMIR Publications website together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Reference

1. Ma Q, Jiang T, Chen W, et al. Awareness of sexual partner's HIV status among men who have sex with men in China: cross-sectional survey study. JMIR Public Health Surveill 2025 Jan 13;11:e66139. [doi: 10.2196/66139] [Medline: 39810334]

Submitted 12.May.2026; this is a non–peer-reviewed article;accepted 19.May.2026; published 31.Jul.2026.

Please cite as:
Ma Q, Jiang T, Chen W, Jiang S, Zheng J, Wang H, He L
Correction: Awareness of Sexual Partner's HIV Status Among Men Who Have Sex With Men in China: Cross-Sectional Survey Study
JMIR Public Health Surveill 2026;12:e101200
URL: https://publichealth.jmir.org/2026/1/e101200
doi:10.2196/101200

© Qian Ma, Tingting Jiang, Wanjun Chen, Shaoqiang Jiang, Jinlei Zheng, Hui Wang, Lin He. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 31.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Correction: Cross-Sectional Study on Oral Nicotine Product Sales Trends in Scandinavia From 2018 to 2025

Marina A Murphy[1], BSc, MSc, PhD; Diane Henenberg[2], BBus; Lindsay Reese[2], BS, PhD

[1]HAYPP Limited, 33 Clarke Road, Mount Farm, Milton Keynes, United Kingdom
[2]Snusbolaget AB, Stockholm, Sweden

Corresponding Author:
Marina A Murphy, BSc, MSc, PhD
HAYPP Limited, 33 Clarke Road, Mount Farm, Milton Keynes, United Kingdom

Related Article:

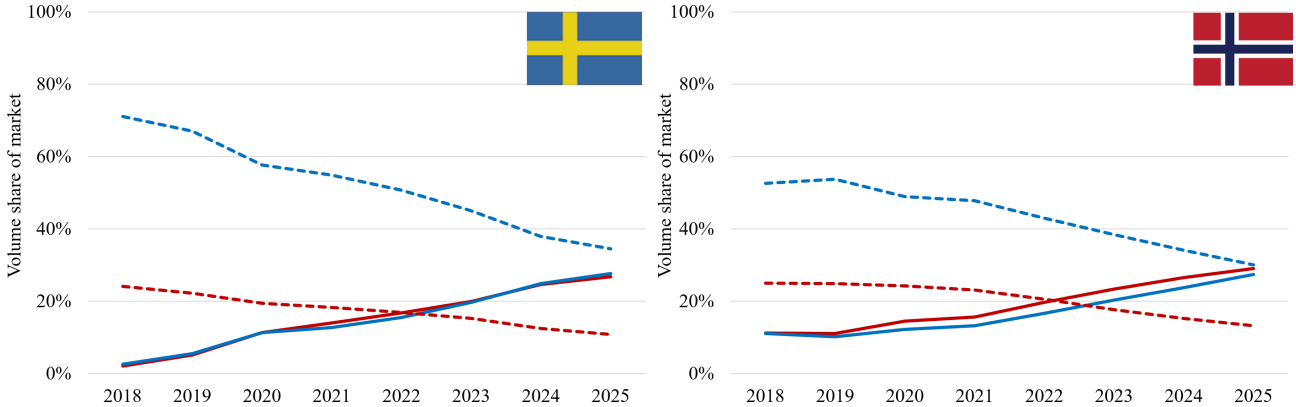
Correction of: https://publichealth.jmir.org/2026/1/e85490/

(JMIR Public Health Surveill 2026;12:e94157) doi:10.2196/94157

In "Cross-Sectional Study on Oral Nicotine Product Sales Trends in Scandinavia From 2018 to 2025" [1], the authors noted one error.

In the original published paper, the Figure 1 image was erroneously duplicated for Figure 2. Figure 2 has been updated to the following image (attached here as Figure 1):

Figure 1. New Figure 2: Cross-sectional volume shares of the oral nicotine product markets (2018 - 2025) for men (blue) and women (red) for nicotine pouches (NP, solid lines) and snus (dotted lines) in Sweden (left) and Norway (right).



The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Reference

1. Murphy MA, Henenberg D, Reese L. Cross-sectional study on oral nicotine product sales trends in Scandinavia from 2018 to 2025. JMIR Public Health Surveill 2016;12:e85490. [doi: 10.2196/85490] [Medline: 41604316]

Submitted 25.Feb.2026; this is a non–peer-reviewed article;accepted 25.Feb.2026; published 05.Mar.2026.

Please cite as:
Murphy MA, Henenberg D, Reese L
Correction: Cross-Sectional Study on Oral Nicotine Product Sales Trends in Scandinavia From 2018 to 2025
JMIR Public Health Surveill 2026;12:e94157
URL: https://publichealth.jmir.org/2026/1/e94157
doi:10.2196/94157

XSL•FO
RenderX

© Marina A Murphy, Diane Henenberg, Lindsay Reese. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 5.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Correction: Diagnostic Testing Preferences in Rural and Vulnerable Populations During a Pandemic: Discrete Choice Experiment

Eline van den Broek-Altenburg[1*], MA, MSci, PhD; Jamie Benson[2*], BA; Yvonne Jonk[3], PhD; Abimbola Leslie[1], MD; Jan Carney[1,4], MD; Gary Stein[1,5], MD

[1]Department of Radiology, College of Medicine, University of Vermont, Burlington, VT, United States
[2]Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States
[3]Maine Rural Health Research Center, University of Southern Maine, Portland, OR, United States
[4]Department of Medicine, College of Medicine, University of Vermont, Burlington, VT, United States
[5]Department of Biochemistry, College of Medicine, University of Vermont, Burlington, VT, United States
[*]these authors contributed equally

Corresponding Author:
Eline van den Broek-Altenburg, MA, MSci, PhD
Department of Radiology, College of Medicine
University of Vermont
89 Beaumont Avenue
Burlington, VT, 05405
United States
Phone: 1 8024956029
Email: eline.altenburg@med.uvm.edu

Related Article:

Correction of: https://publichealth.jmir.org/2025/1/e68734

(JMIR Public Health Surveill 2026;12:e87714) doi:10.2196/87714

In “Diagnostic Testing Preferences in Rural and Vulnerable Populations During a Pandemic: Discrete Choice Experiment” [1], the authors noted one error.

The Ethical Considerations section has been changed from the following:

> *This study (STUDY00002116) has been reviewed by the Institutional Review Board of the University of Vermont and the Institutional Official has authorized the publication of the research data.. Focus group protocols were reviewed by the IRB using the exempt procedures set forth under 45 CFR 46.104, specifically, under Exemption Category: (2)(ii) tests, surveys, interviews, or observation (low risk), with waiver of documentation of consent under 46.117(c)(1). The Discrete Choice Experiment (DCE) was originally deemed not human subjects research by the IRB’s assessment tool; a quality review later determined that the study should have been reviewed and would likely have been approved with a waiver of written consent. In the DCE, survey data were collected without identifiers. Respondents received $5 compensation for their time from the survey company Centiment. Additionally, no identification of individual participants in any images of the manuscript or supplementary material is possible.*

This section now reads:

> *Focus group protocols were reviewed by the IRB using the exempt procedures set forth under 45 CFR 46.104, specifically, under Exemption Category: (2)(ii) tests, surveys, interviews, or observation (low risk), with waiver of documentation of consent under 46.117(c)(1). The Discrete Choice Experiment (DCE) was originally deemed not human subjects research by the IRB’s assessment tool; a quality review later determined that the study should have been reviewed and would likely have been approved with a waiver of written consent. In the DCE, survey data were collected without identifiers. Respondents received $5 compensation for their time from the survey company Centiment. Additionally, no identification of individual participants in any images of the manuscript or supplementary material is possible. The IRB director at UVM gave permission to publish this manuscript.*

The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication

XSL•FO
RenderX

of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Reference

1. van den Broek-Altenburg E, Benson J, Jonk Y, Leslie A, Carney J, Stein G. Diagnostic testing preferences in rural and vulnerable populations during a pandemic: discrete choice experiment. JMIR Public Health Surveill 2025 Oct 21;11:e68734 [FREE Full text] [doi: 10.2196/68734] [Medline: 41118676]

Submitted 13.Nov.2025; this is a non–peer-reviewed article;accepted 13.Nov.2025; published 13.Mar.2026.

Please cite as:
van den Broek-Altenburg E, Benson J, Jonk Y, Leslie A, Carney J, Stein G
Correction: Diagnostic Testing Preferences in Rural and Vulnerable Populations During a Pandemic: Discrete Choice Experiment
JMIR Public Health Surveill 2026;12:e87714
URL: https://publichealth.jmir.org/2026/1/e87714
doi:10.2196/87714
PMID:41824945

©Eline van den Broek-Altenburg, Jamie Benson, Yvonne Jonk, Abimbola Leslie, Jan Carney, Gary Stein. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 13.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Building Public Health Data Dashboards: Tutorial Playbook

Elwin Wu[1], PhD; Raymond Balise[2], PhD; Benjamin Katz[3], PhD; Daniel Harris[4], PhD; Matthew Bullard[5], MPH; Naleef Fareed[6], PhD; Marc Larochelle[7], MD, MPH; Jennifer Villani[8], PhD

[1]Social Intervention Group, Columbia University School of Social Work, 1255 Amsterdam Avenue, New York, NY, United States

[2]Division of Biostatistics and Bioinformatics, Department of Public Health Sciences, University of Miami, Miami, FL, United States

[3]Institute for Implementation Science in Population Health, City University of New York (CUNY), New York, NY, United States

[4]Institute for Biomedical Informatics, University of Kentucky, Lexington, KY, United States

[5]Center for Health Data Science, Boston University School of Public Health, Boston, MA, United States

[6]CATALYST – The Center for the Advancement of Team Science, Analytics, and Systems Thinking, College of Medicine, The Ohio State University, Columbus, OH, United States

[7]Clinical Addiction Research and Education Unit, Section of General Internal Medicine, Department of Medicine, Boston University School of Medicine and Boston Medical Center, Boston, MA, United States

[8]National Institute of Drug Abuse, National Institutes of Health, Bethesda, MD, United States

Corresponding Author:
Elwin Wu, PhD
Social Intervention Group, Columbia University School of Social Work, 1255 Amsterdam Avenue, New York, NY, United States

Abstract

Public health data dashboards have substantial potential to improve transparency, understanding, and decision-making at multiple levels, from individuals to public health practitioners and policymakers. However, creating effective dashboards presents many challenges. In this case-based tutorial on public health dashboard development, we share lessons learned from our experience developing data dashboards for the HEALing Communities Study (HCS), a National Institutes of Health (NIH)–funded, community-engaged intervention to deploy evidence-based practices to reduce opioid overdose deaths in 67 communities across 4 states. We present key decision points dashboard teams must address, along with the major considerations and trade-offs that shaped our approach. First, we describe core considerations of the who, what, why, where, when, and how of data dashboard development. Second, we outline steps in data curation, including the identification of key metrics and potential data sources and developing processes to acquire the data. Third, we discuss practical aspects of developing data visualizations that can effectively communicate key messages to the end users of interest. Fourth, we describe the infrastructure considerations to host and publish data dashboards. And finally, we discuss maintenance and sustainability of the dashboard. While the material can be read sequentially as a step-by-step guide, we refer to this resource as a "playbook" because readers may engage with specific domains in a random-access fashion, that is, based on their specific needs and/or starting point rather than a fixed sequence. The information, supplemental materials, and resources will assist individuals and organizations seeking to build data dashboards by fostering context-sensitive evaluation of design and implementation choices to realize the promise of data-driven decision-making.

(JMIR Public Health Surveill 2026;12:e83157) doi:10.2196/83157

KEYWORDS

dashboard systems; data visualization; public health; opioid epidemic; drug overdose

Introduction

Automobile instrument panels are perhaps the most familiar type of dashboard. They combine analog and digital visualizations—for example, gauges, counters, and icons—to help drivers monitor key aspects of their vehicle, including fuel levels, speed, engine temperature, and mileage. Warning indicators alert drivers to critical issues like low oil, low tire pressure, or an unfastened seatbelt. Designed for instant readability, these dashboards prioritize essential information, with standardized warning lights and speedometers placed centrally to capture the driver's immediate attention.

Public health data dashboards function much like automobile dashboards; both use data visualizations to communicate critical information and support decision-making. By consolidating valid and useful epidemiological metrics and other public health data, these dashboards enhance awareness, facilitate monitoring, and drive informed action. Their value has been especially evident during public health crises, such as the COVID-19 pandemic and the opioid overdose epidemic, where real-time data have played a crucial role in response efforts [1-3].

Despite an influx of public health data, challenges and barriers have emerged with regard to the effective and efficient communication of these data to a broad range of individuals. Issues critical to the uptake and use of public health data

XSL•FO
RenderX

dashboards include a design that contains the most useful data, usability, cognitive load, implementation and evaluation, and sustainability [4]. Studies on data visualizations and dashboards have discussed notable challenges that pertain to these issues, including lack of rigorous scientific methodology, superficial or incorrect application of design principles, data literacy of end users, and limited scope in the representation of end user input [5-7].

Within this context, our research team engaged in the development of data-driven dashboards for a large, multisite cluster-randomized pragmatic trial. The HEALing Communities Study (HCS) was a community-based cluster-randomized trial aimed at addressing the opioid crisis in communities with high opioid overdose death rates. The National Institute on Drug Abuse (NIDA), in partnership with the Substance Abuse and Mental Health Services Administration (SAMSHA), funded the study in April 2019 to test the immediate impact of the implementation of an integrated set of evidence-based practices across health care, behavioral health, justice, and other community-based settings to prevent opioid overdose deaths and treat opioid use disorder [8]. Among the 4 states ("study sites") involved in the HCS, communities (n=67) were randomized to receive the intervention first or be in the wait-list control group. In each community, a community coalition was established to provide local leadership and data-driven decision-making for the intervention [9]. Data dashboards—to support data-driven decision-making [10]—were co-created with coalition members to provide visualizations of key metrics on opioid overdoses and treatment for opioid use disorder within a community [3]. The coalition stakeholders (1) used data to define their local opioid situation, including prevention and treatment gaps and resources; (2) developed a data-driven action plan for implementing evidence-based practices linked to local needs; and (3) after deploying the action plan, they used data dashboards to monitor progress and adjust as needed.

Eighteen months after deployment, we conducted a usability evaluation of the HCS dashboards using a mixed methods approach that involved quantitative feedback from the System Usability Survey (SUS) [11] and qualitative interviews guided by the Technology Adoption Framework [12]. The SUS scores indicated the end users found HCS dashboards to be acceptable for usability; for example, the overall grand mean SUS score from 62 users of the HCS dashboards (39% response rate) was 73 (SD 4.6), which falls between the 70th and 80th percentile of SUS responses and is comparable to scores reported for widely used tools such as Microsoft Word and PowerPoint (Microsoft Corp) [13]. The interview results indicated positive feedback around dashboard usability and intention to use from the end users sampled.

The considerations described in this case-based tutorial playbook are grounded in our prior work in the development and evaluation of HCS data dashboards. Although some aspects of HCS dashboard development reflected consensus decisions or externally imposed requirements (eg, minimum required metrics and National Institutes of Health [NIH] funding conditions), the purpose of this tutorial playbook is not to present consensus guidance or a prescriptive methodology. Rather, we focus on identifying key decision points that arise during public health dashboard development and the attendant major considerations and trade-offs that shape how those decisions are approached. By emphasizing decision points rather than prescriptive solutions, this tutorial playbook supports adaptation across different public health settings. The manuscript is organized along a typical dashboard development lifecycle, from conceptual considerations through postlaunch revision. While this can be used as a step-by-step guide, dashboard development is often nonlinear; thus, we refer to this document as a "playbook" to indicate that readers can refer to specific sections as dictated by their context, starting point, and immediate needs at a given time.

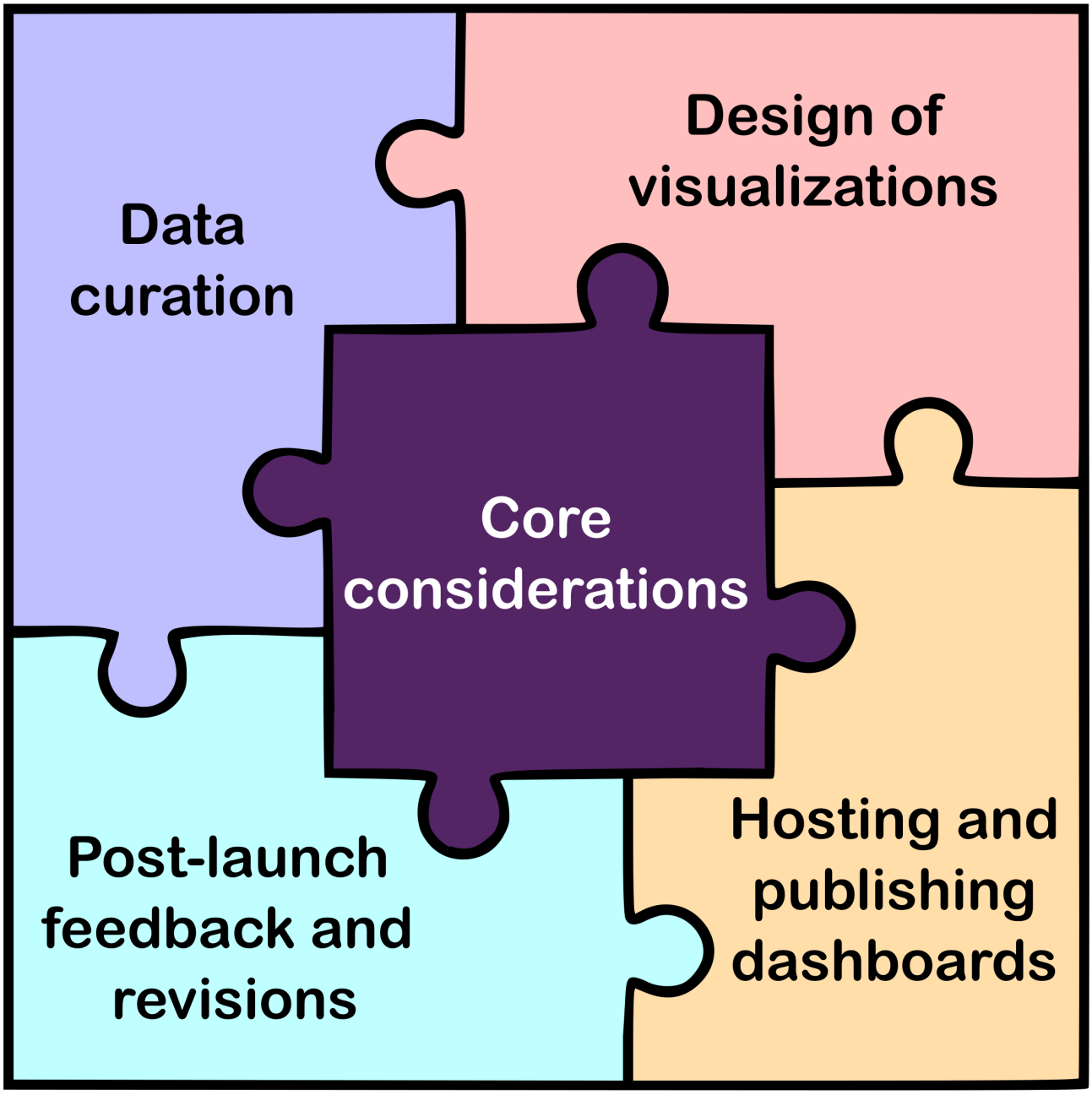
Public Health Dashboards: Key Tasks

Based on lessons learned from the HCS, we provide this playbook for creating public health data dashboards. While HCS dashboards focused exclusively on opioid use disorder and overdose-related metrics for each community, the information and considerations provided here can be applied to myriad public health issues. Consistent with the case-based foundation of this manuscript, the core considerations and domains are presented as guiding questions and decision points rather than as required artifacts (ie, tools or steps that are applicable across all contexts, such as accessibility review checklists, compliance sign-off documentation, and articulated user personas), success criteria, or approval gates that are likely to vary substantially both qualitatively and quantitatively across contexts.

Each section is a conceptual domain of questions that help guide dashboard developers as they design and deploy a dashboard. For purposes of this tutorial playbook, we organize these along a prototypical dashboard lifecycle, starting with (1) core considerations and then progressing through (2) data curation, (3) design of visualizations, (4) hosting and publishing dashboards, and (5) postlaunch feedback and revision. However, as depicted in Figure 1, the process is neither linear nor discrete. Each section in the playbook is equally important, and the interrelationships and/or interplay among domains require developers to revisit decisions across domains over time.

XSL•FO
RenderX

Figure 1. Overview of the playbook.



Core Considerations

A useful framework for guiding dashboard development at every stage is based on the classic journalistic "Five Ws and One H" approach: who, what, why, where, when, and how. These questions help ensure a structured and comprehensive approach to designing, deploying, and maintaining public health data dashboards (Textbox 1). By addressing these key decision-making topics, dashboard developers can ensure a holistic approach to planning, designing, and deploying effective dashboards. The core decision-making topics manifest throughout the key domains involved in dashboard development and publishing that are described in the ensuing sections of this playbook.

XSL•FO
RenderX

Textbox 1. Core considerations.

The "Five Ws and One H" approach:

- Who is the target audience (ie, end users)?
- What do the end users want to be able to learn or infer from the data?
- Why are dashboards a useful or the optimal tool to meet the need?
- Where will the dashboard be published and/or made available?
- When do the dashboards need to be published, updated, etc?
- How will we actually build and publish the dashboards?

Who: Understanding the End Users

Identifying the target audience—or end users—is essential. Dashboard developers must assess and tailor to their users' technical expertise, data literacy, professional roles, and specific needs. Workflows also shape dashboard requirements; for instance, some users may need real-time monitoring, while others might only check a dashboard weekly, monthly, or quarterly.

What: Defining the Dashboard's Purpose

Dashboards should be designed to answer key questions for end users efficiently. Developers must determine, "What information do users need? How can the data be presented to support decision-making?" Ideally, visualizations should allow users to grasp insights quickly with minimal cognitive effort, delivering essential information "at a glance."

Why: Establishing the Dashboard's Core Objectives

This question prompts developers to define the overarching goals of the dashboard. What key messages or insights should be communicated? What operational or policy decisions should the dashboard support? Design choices—such as whether to use graphical summaries or textual explanations—should align with these objectives to maximize clarity and usability.

When: Considering the Timeline for Development and Updates

Developers must plan for the time required to acquire and integrate data, design and test the dashboard, and manage its lifecycle post launch. This includes establishing a schedule for updates, revisions, and long-term sustainability.

Where: Ensuring Availability and Visibility

The dashboard's availability is a critical consideration. Developers must decide whether it should be hosted on a public website or restricted to an internal platform. If public, should it be optimized for search engines, or should access be controlled? Additionally, dashboards should be designed to function seamlessly across different devices—desktops, laptops, tablets, and smartphones—each of which has unique constraints related to screen size, processing power, and internet connectivity. In some cases, a dedicated app or embedded dashboards within an existing platform may enhance usability.

How: Assessing Resources and Infrastructure

Successful dashboards require technological and human resources. Developers must evaluate factors such as cybersecurity, data storage, processing capabilities, and integration with external data sources. Managing these elements effectively ensures the dashboard remains functional, secure, and scalable over time.

In addition, governance is an important consideration, as decisions about data access, processing, and publication are often shaped by legal, contractual, and institutional constraints. In practice, dashboard teams need to address how data use agreements (DUAs), applicable privacy regulations (eg, the Health Insurance Portability and Accountability Act [HIPAA]), and data suppression rules affect what data can be ingested, how it can or needs to be transformed and displayed, and with whom it can be shared. These constraints often extend beyond technical implementation to influence workflows, timelines, and responsibilities across the data lifecycle, from data acquisition and quality assurance through publication, revision, and eventual retirement.

An information architecture diagram can help document how data move through a system, from source data through transformation and storage to visualization, access control, and monitoring. Such diagrams may be useful for identifying or clarifying the human and technical resources involved, how those resources are organized, key dependencies, and where governance and security considerations arise, as well as for supporting communication among technical and nontechnical stakeholders involved in dashboard development.

Data Curation

Data curation involves identifying desired metrics, sourcing available data, and establishing procedures to acquire or "ingest" the data (Textbox 2). This process is foundational to building a reliable and actionable public health dashboard. Desired metrics—shaped primarily by the who, what, and why—typically are the primary outcome and key epidemiological characteristics and/or targets of prevention and intervention efforts (eg, incidence or prevalence of a disease or phenomenon of interest). Other data that are often desired include secondary outcomes, key drivers of the outcomes of interest, and data that allow for subgroup characterization (eg, age, sex, race, ethnicity, geography).

XSL•FO
RenderX

Textbox 2. Decision points for data curation.

The following questions outline key decision points to consider when planning and implementing data curation processes for dashboards and other data-driven systems:

- What are the desired metrics and visualizations?
- What data are available?
- What are the acquisition requirements (eg, data use agreements and security needs) and associated costs for infrastructure and staffing?
- What are ways to address data quality and refresh discipline?
- How do these requirements match available or planned resources and funding?
- Based on all of the above, what are the relative priorities of potential metrics and visualizations?

Many metrics require calculations that integrate multiple data points. For example, prevalence is derived from case counts relative to a denominator—typically population size—which may also need stratification by demographic variables. Similarly, spatial visualizations require both temporal and geographic details for each case. Thus, desired metrics may involve multidimensional databases or combining and calculating across different data sources. Regardless of the exact metric, the most effective dashboards ensure that such data are not only accurate but also actionable (eg, lag between data collection and reporting does not undermine relevance for decision-making).

In identifying available data, several potential sources should be considered. Administrative data may be collected and made available from government agencies and initiatives (eg, vital statistics, Medicaid, or prescription drug monitoring programs) as well as service providers (eg, health care providers and social service agencies). Data may be collected and made available from research studies, program evaluations, population-level surveys (eg, the National Survey on Drug Use and Health and the Behavioral Risk Factor Surveillance System), and contracted third parties (eg, HCS used data from IQVIA and Millenium Health). Many data repositories also make data available (eg, Centers for Disease Control and Prevention [CDC] WONDER and the National Notifiable Diseases Surveillance System). Existing dashboards and other surveillance or monitoring systems may provide data directly or, when permitted, be incorporated into planned dashboards via embed codes that display existing visualizations without transferring the underlying data, or through scraping data from existing visualizations in accordance with applicable terms of use for local processing into a new visualization.

Constraints often govern the use and visualization of data, shaping how dashboards can present and share information. Data providers may impose restrictions on who can access specific datasets, limiting or prohibiting certain audiences from viewing them. These conditions are typically outlined in a DUA, which defines the expectations, requirements, and responsibilities between data providers and those developing and publishing the dashboard. A DUA specifies what data are being shared, the procedures for data transmission, storage, and publication, and any restrictions on visualization. These may include requirements for attribution, contextual definitions, or limitations on further distribution. Additionally, government regulations and institutional policies often impose further constraints, such as suppression rules designed to protect privacy and confidentiality. These rules dictate how data can be displayed, ensuring that sensitive or identifiable information remains secure while maintaining the integrity and utility of the dashboard. Other privacy-preserving computation approaches (eg, secure multiparty computation and federated approaches) may also be relevant, particularly for analyses that span data across multiple organizations or jurisdictions, although such approaches can involve greater technical complexity and coordination.

Within these constraints, data quality and data refresh are ongoing considerations in data curation. HCS teams regularly examined and tracked data quality dimensions such as completeness, timeliness and lag, internal consistency, and stability over time [14]. Dashboard designers may want to display refresh schedules or lag indicators, and the use of tools such as data quality scorecards (eg, to assess whether data meet thresholds indicating sufficient quality or warrant caution) may be useful for standardization, monitoring, and quality assurance. We note, however, that appropriate thresholds and related decisions depend heavily on context, including data source characteristics, suppression requirements, the potential clinical or policy impact of over- or misinterpretation, and intended dashboard use. Those are some of the aspects that allow for tailoring of data quality and refresh discipline (ie, the exact metrics, thresholds, or automation approaches for detecting problems stemming from lower quality or lag in data).

Data ingestion refers to the process of transferring data from its source to the recipient responsible for building the dashboard. This transfer can occur through various methods, including pushing, where the data provider actively transmits the data; pulling, where the recipient requests and retrieves the data; and scraping, where data are electronically extracted from existing documents or publications when permitted and in accordance with publishers' terms of use. In many cases, ingestion also involves preliminary data processing to ensure compatibility with dashboard visualization and storage. This preprocessing follows an extract, transform, and load workflow, which may include data cleaning (such as removing disallowed values or correcting inconsistencies), transforming variables (such as calculating rates from raw counts or aggregating data across time periods), and restructuring datasets to align with the platform used for visualization. The efficiency and reliability of data ingestion improve with greater automation. For example, some data sources provide an application programming interface, a standardized set of protocols that allow software to

XSL•FO
RenderX

directly request and retrieve data from the source, subject to authorization and applicable restrictions on use, storage, or display. However, implementing automation requires investment in human resources for design, programming, and ongoing maintenance as data systems evolve or become obsolete. Before establishing a data ingestion plan, it is essential to assess the capabilities of the existing cyberinfrastructure. If these resources are inadequate, the costs—whether in terms of personnel, hardware, or development time—must be factored into planning. When budgets and timelines are constrained, trade-offs may become necessary, potentially limiting the scope of data sources or the complexity of visualizations that can be supported.

Selection of data to be ingested for visualizations often cycles back and forth among desired versus available data, utility of data, and resources needed to transform the data into visualizations as well as maintain the cyberinfrastructure that hosts and publishes the dashboards. Desired metrics usually drive a targeted search for the relevant data needed to calculate and display the metric. Conversely, an examination of available data may prompt the desire for additional data (eg, disaggregated into subgroups such as race or ethnicity or age group, subtypes of substances, or variants of infectious agents; greater specificity with respect to time or geography). Given that resources do not allow for every possible visualization—nor is it advisable from a health communications perspective to provide vast quantities of information all at once—the full menu of potential visualizations should be prioritized for creation and inclusion in the dashboards. This decision-making process is about the appropriate or, ideally, optimal balance among data quality, utility of the metric, and costs or resources available to curate the data and publish the dashboards.

In public health, researchers, policymakers, and practitioners play a central role in determining key metrics and are typically considered the primary end users of data dashboards. The HCS operated under the principle that community members bring invaluable perspectives on the relevance, utility, and limitations of different data and metrics. Toward this end, HCS researchers developed the Community-Engaged Data Science model to assist data scientists working with communities in identifying and integrating relevant data for making informed decisions [15]. For data dashboard development, this ethos translated into a co-creation approach that sought to ensure that those affected by public health decisions, whether or not they hold formal decision-making roles, were actively involved in identifying and prioritizing the metrics and visualizations that would shape the dashboard. Engaging community members from the outset of the project fostered greater relevance, usability, and impact, making the dashboard a more effective tool for addressing the overdose crisis and other real-world public health challenges.

Design of Visualizations

Ideally, the team-building dashboards include experts in human-centered design and graphical design who can work with the consumers of the displays (Textbox 3). A starting point is understanding the wants and needs of the end users. Interviews of end users can capture both what they want to know at a glance and what inferences and conclusions they would like to support (eg, trends in the data over time). During the initial development, focus groups and public meetings allow design teams to identify the types of people that will be using the dashboard. Understanding their data literacy—the ability to read and understand data and data visualizations—is paramount. For example, knowing that a dashboard will be used by public health practitioners would suggest that terms like “incidence” and “prevalence” can be used unabashedly, but if the same dashboard will be used by people without any background in statistics or epidemiology, avoiding jargon and using plain language is preferred to avoid confusing and alienating the audience. In the context of designing public health dashboards, it is critical to identify all end user groups and then develop an outreach plan to make sure that all constituents are assessed, rather than just the vocal and/or immediately available audience who may first provide information.

Textbox 3. Decision points for the design of visualizations.

The following questions highlight key considerations for designing effective visualizations that support user understanding and informed decision-making:

- How can we engage and elicit feedback from all end users?
- How can the data be transformed and/or presented so the end users can grasp the desired point at a glance?
- What additional visual elements will aid in comprehension or reduce misinterpretation?
- How can specific design elements be tweaked or refined to take advantage of how humans process information?
- What ancillary information is needed or helpful to the user?
- How can the dashboard interface allow the user to access such information without distracting from the main visualization?

A well-designed dashboard allows end users to quickly extract key insights and draw meaningful inferences. Within the design community, a widely used approach is to identify and/or envision different user “personas”—distinct groups of individuals who will interact with the data—and tailor the visualization to meet their specific needs [16]. Effective visual design prioritizes clarity and accessibility. High-level, at-a-glance insights can be embedded in intuitive graphics that communicate key trends immediately, while more detailed or supporting information can be provided separately in text or tables that would take additional time to process [17]. Well-crafted data visualizations enable users to perceive both central tendencies and variability immediately. For instance, a straight trend line clearly indicates whether a pattern is increasing or decreasing over time. However, as illustrated with synthetic data in Figure 2A, a flexible line—such as a curve generated by mathematical smoothing techniques like beta splines or locally estimated scatterplot smoothing—can reveal

XSL•FO
RenderX

nuanced departures from linear trends [18-20]. These features are particularly important for data that could have seasonal trends (eg, drug overdose patterns that vary by time of year) or identifying sudden shocks in data, such as unexpected spikes in mortality.

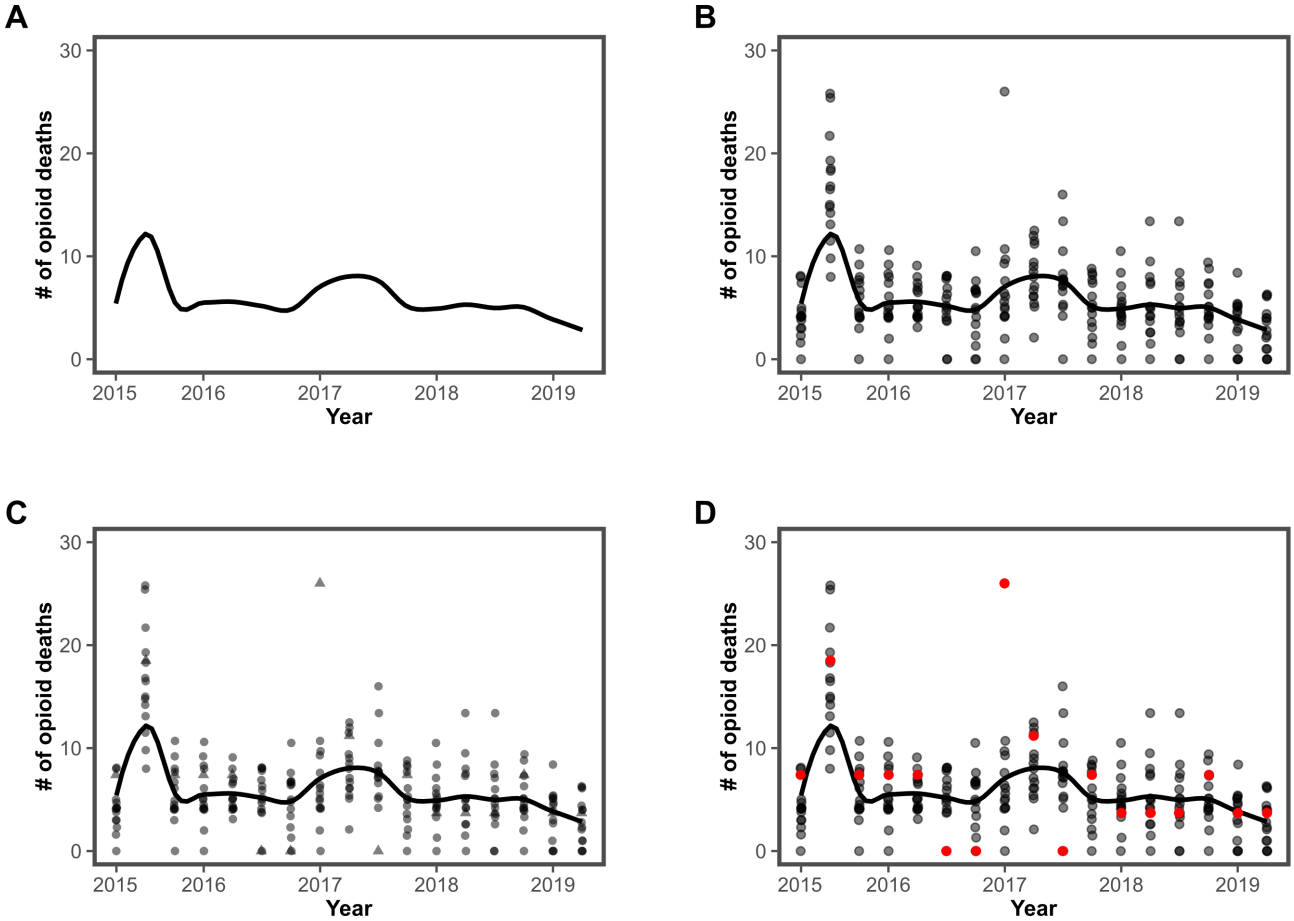
Combining multiple visualization elements enhances interpretability. As shown in Figure 2B, overlaying individual data points on a trend curve allows users to assess both the overall pattern and the underlying variability. This approach is especially valuable for identifying extreme events, such as a sudden surge in mortality due to drug supply contamination. At the same time, it helps distinguish genuine outliers from data anomalies that may stem from reporting errors rather than true shifts in the trend. When relevant, tagging outliers with contextual metadata can prevent misinterpretation (eg, a change in data reporting rather than a change in the underlying trend) and ensure that decision-making is based on accurate, reliable insights. By applying these visualization principles, dashboards can effectively support data-driven decision-making, enabling users to quickly grasp patterns, detect anomalies, and respond to public health challenges with confidence.

Some visualizations are static, meaning they function as fixed images where the information displayed remains unchanged. These are essentially bitmaps; what the user sees is exactly what they get. However, digital dashboards presented in a web browser or app can incorporate interactivity, allowing visualizations to change based on user actions. Interactive features enhance user engagement and data exploration. For example, users can zoom in or out on a graph by clicking buttons, dragging a slider, or pinching on a touchscreen. Filtering options, such as checkboxes, allow users to display or hide specific subgroups based on characteristics like race, age, or sex. Another valuable feature is the hover-over function, which reveals additional details when a user places the cursor over a data element, providing deeper insights without cluttering the initial visualization. While interactivity can significantly improve a dashboard's functionality by allowing users to focus on information most relevant to them, it comes with trade-offs. Implementing interactive features requires additional human and computational resources, both for development and ongoing maintenance. Design challenges also emerge, as excessive interactivity can be distracting or even counterproductive. Too many visual elements, animations, or overly complex interactions may overwhelm the user, while poorly optimized interactivity can slow down rendering, diminishing the user experience. Thoughtful design ensures that interactivity enhances rather than hinders the dashboard's effectiveness, striking a balance between flexibility and usability.

To ensure clarity and prevent misinterpretation, dashboard visualizations must be tested with multiple, diverse audiences. This is especially critical in public health, where data suppression is often necessary to protect anonymity or account for incomplete reporting. Misrepresentation can be particularly problematic when tracking rare events, such as overdoses. In small communities, just a few deaths can cause sharp fluctuations in rates, while in large urban areas, high death counts may appear less pronounced due to the sheer size of the population. In the HCS data dashboards, population-based rates were used to account for these disparate but real-world scenarios. However, even these metrics required careful handling. Overdose death rates, for example, were based on highly sensitive data that had to be suppressed in locations with low case counts to protect individual privacy. Simply omitting data points to comply with suppression rules could create misleading gaps in trend lines, making it difficult for policymakers to assess emerging patterns. Instead, smoothing techniques, such as locally estimated scatterplot smoothing curves (as shown in Figure 2A), allowed users to view approximate trends without revealing suppressed values, preserving both privacy and interpretability.

XSL•FO
RenderX

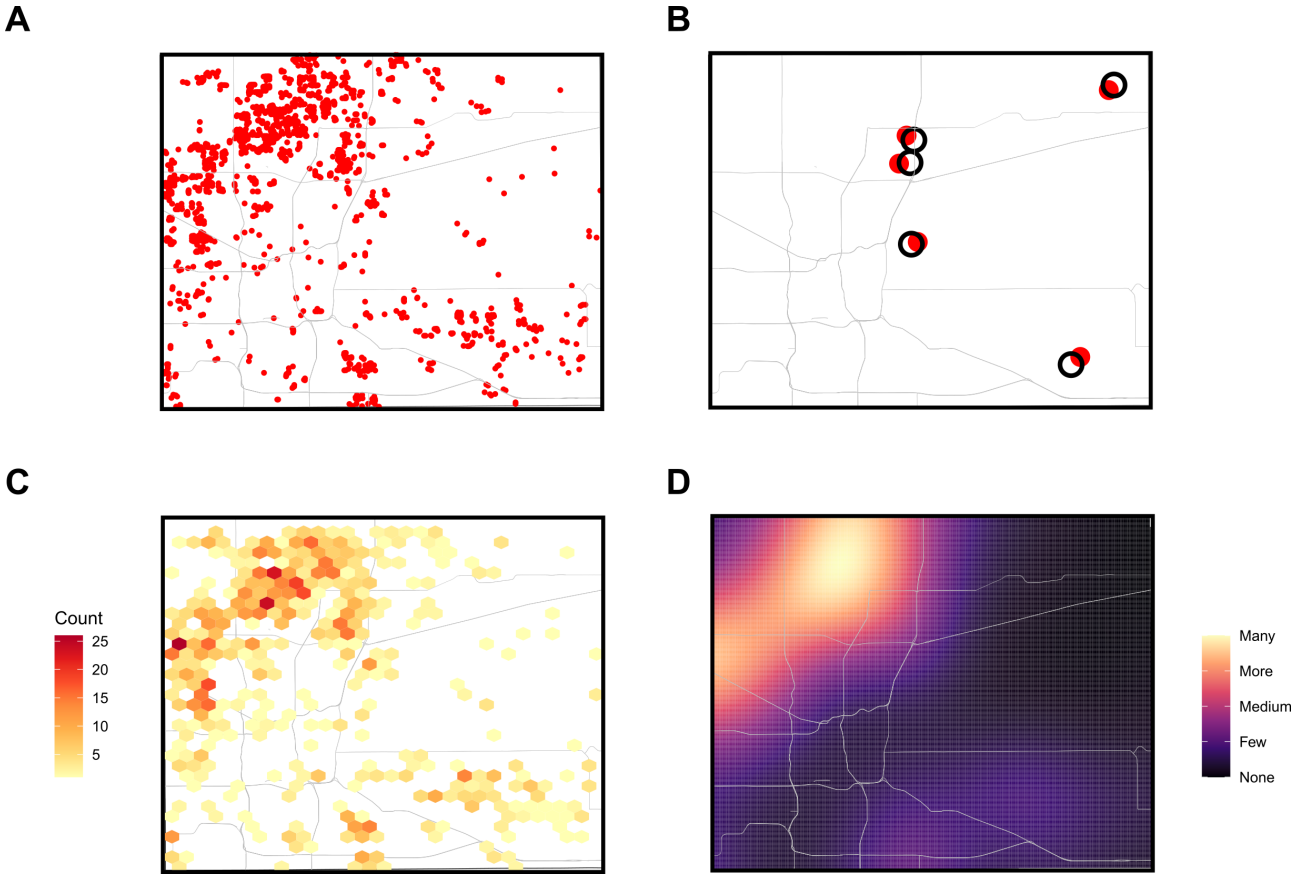
Figure 2. Scatterplot variations of the same synthetic data of opioid deaths versus time. (A) Panel A shows the general trend in opioid-related overdose deaths for all counties in an illustrative state. (B) Panel B adds markers for each county. (C) Panel C distinguishes a county of interest using a different shape. (D) Panel D distinguishes the same county of interest using color.



Similar challenges arise when incorporating maps into dashboards. Policymakers rely on geospatial data to allocate resources—such as overdose reversal kits—to the areas most in need. However, to protect individual privacy, maps cannot display the precise locations of overdoses. Various methods can address this issue, each with trade-offs in accuracy and usability. One approach, illustrated with maps of synthetic data in Figure 3B, involves randomly shifting ("jittering") [21], overdose locations to nearby points on the map; this preserves the overall count information but could foster misinterpretation for users who do not read the legend carefully. Alternative methods include overlaying a grid and shading areas based on case density or using mathematical density functions to blur exact locations. While these approaches help preserve anonymity, they can reduce granularity and make interpretation more challenging. Additionally, choices on the distance to offset or the amount of blur (Figure 3C and 3Ds) are complex [22,23]. A higher number of events in very densely populated areas (eg, urban areas) can safely allow for smaller offsets or blur, but a lower number of events in sparsely populated regions (eg, rural areas) requires larger jitter parameters to hide the exact locations of events.

XSL•FO
RenderX

Figure 3. Map variations of the same geographical data. (A) Panel A shows an illustrative county with its primary roads along with simulated locations of overdoses, which are shown as red dots. (B) Panel B shows the same county with only 5 simulated overdose events displayed as large red dots along with slightly shifted (mean =1.25 km) locations as black circles. (C) Panel C shows the counts as colored hex shapes (hexagon area =1.4 km^2) with deeper red colors indicating many events. (D) Panel D shows a density plot from kernel smoothing (40×40 km, max-min normalization) [21], with brighter and lighter colors indicating more overdoses.



Ultimately, the most effective visualization strategies depend on the needs of the end users. Engaging with stakeholders throughout the design process ensures that visualizations balance privacy, accuracy, and usability, making them both ethically responsible and practically useful for decision-making. Consistent with the community co-creation ethos described earlier, the HCS used the community-engaged data science model [15] as an approach to community co-creation key considerations, such as differences in data familiarity, interpretation, accessibility needs, and perceived relevance across roles, geographies, and communities. Community co-creation inputs informed visualization design decisions, including choices related to language and reading level, use of plain-language labels or multilingual resources, level of detail, and use of interactivity. Community engagement also informed decisions related to data equity (eg, when suppression risked obscuring inequities and how data gaps were annotated) and how uncertainty or data limitations were communicated. In addition, stakeholder input highlighted practical access considerations, such as the need to account for low-bandwidth environments, mobile access, and alternative formats for resource-limited settings. We note that approaches to community engagement are likely to vary across communities and over time, reflecting differences in local context, resources, intended dashboard use, and public health topics.

Consulting with graphic designers when building dashboards can greatly enhance the clarity and usability of data visualizations. Experts in visual design understand how to optimize information processing, making it easier and faster for end users to extract key insights [24,25]. Principles such as the Gestalt laws leverage the human tendency to group related elements based on similarity and proximity, allowing users to quickly interpret patterns and relationships within the data. The ideal visualization enables users to grasp critical information at a glance.

Reducing cognitive load is a core principle of good design. Directly labeling elements within a visualization, rather than relying on a separate key, prevents users from having to remember reference points while analyzing the graphic. Other mnemonic techniques, such as using meaningful symbols (eg, hollow dots to indicate missing or suppressed data), can further streamline comprehension. However, such symbols must be clearly defined to avoid confusion. Effective visual design requires continuous feedback from end users to ensure that all details are easily understood across diverse audiences.

Certain chart types are inherently more effective than others. Research has shown that bar charts and dot plots allow for faster and more accurate comparisons than pie charts, which can be more difficult to interpret [26-28]. In addition to choosing the right type of visualization, graphic designers use visual signals—such as color, shape, and size—to draw attention to

XSL•FO
RenderX

key data points [17,29]. For instance, Figures 2C and 2D compare 2 methods for highlighting a county of interest, one using shape and the other using color. Determining which method is more effective depends on factors such as the overall size of the visualization, the density of the data, and the distribution of key elements. Direct engagement with end users can help refine these choices, ensuring that the most intuitive and efficient design is selected.

The design process should begin with an initial sketch or storyboard of the layout, progress through choices and prototypes, and culminate in a final, refined display. Dashboards are often accessed across devices with different screen sizes and bandwidth constraints (eg, desktop use in a networked office vs use on a personal mobile device with cellular bandwidth). Design decisions related to responsive layouts, content prioritization, and simplification of visual elements may therefore be needed to support use on smaller screens or in lower-bandwidth settings. These decisions may include determining which metrics or visualizations need to be retained across screen sizes and connection types, which elements can be condensed or hidden without loss of meaning, and how interactions are adjusted when space or network capacity is limited. Throughout this process, accessibility must remain a priority. Removing unnecessary clutter helps users focus on essential insights without distraction [30]. Additionally, dashboards should comply with applicable accessibility standards, such as the Americans with Disabilities Act (ADA), to support inclusive use. In practice, this may involve attention to issues such as color contrast and color blindness, keyboard navigation and focus states, the use of alternative text for non-text elements, and accessible labeling or markup to support screen readers. Addressing these considerations can help ensure that dashboards remain usable for individuals with a range of visual, motor, and cognitive needs. Accessibility checklists may be useful tools, including those informed by the Web Content Accessibility Guidelines (WCAG), for reviewing common accessibility considerations. Thoughtful, user-centered design ultimately ensures that dashboards are both functional and accessible, allowing all audiences to interact with data effectively.

In addition to the design decisions for the visualizations themselves, dashboards need to be constructed to allow users to expand their knowledge beyond what is on the page through proper attribution to the data. Data attribution refers to listing the data source for the underlying data used in the visualization. Specifically, the design team needs to decide if those details will always be visible, which increases visual clutter, or available on demand when users interact with the display, increasing the risk of users not identifying the sources of the data. For example, a visualization of drug overdose deaths may include a note that the data originate from the December 2022 mortality report from the state department of health. For the HCS displays, where we had a great deal of information to put on the dashboard, we typically opted to display additional information when the end user hovered over data or clicked on buttons to learn more. Feedback gathered from the community suggested that they were able to easily glean the details for the sources of the data.

Hosting and Publishing Dashboards

Hosting and publishing dashboards are shaped by governance considerations that affect data access, release, and ongoing maintenance (Textbox 4). Rather than prescribing a specific governance model, this playbook emphasizes key governance-related decision points that arise as dashboards are hosted and published, such as who has authority to approve data use and release; how DUAs and privacy requirements (eg, 42 CFR Part 2 obligations) are monitored over time; how suppression rules are implemented within analytic and visualization workflows; how changes to data sources, metrics, or visualizations are reviewed and communicated; and how equity and inclusion goals are incorporated into these processes.

Textbox 4. Decision points for hosting and publishing dashboards.

The following questions highlight key considerations when planning the infrastructure, resources, and processes required to host and publish dashboards:

- What are the cyberinfrastructure requirements for data storage, creation of visualizations, publishing the dashboards, hosting the platform, and access permissions?
- What are the human resource requirements (eg, key roles, levels of expertise needed, and amount of effort)?
- Does this infrastructure meet requirements and regulatory rules for data storage and security?
- What are the existing institutional capacities with regard to servers, data visualization tools, and content management systems?
- What is the schedule for testing and quality assurance that will achieve a smooth publication of the official release version of the dashboard?

Implementing dashboards requires technical resources for both hosting the dashboard software and the data engineering processes necessary to convert raw source data into standardized measures [14]. Infrastructure requires resources to host or purchase commercial web hosting services. Personnel are needed for multiple roles and tasks, back-end database management staff, analytical staff to generate clean, standardized results, embedded testing staff for quality assurance, visual designers to implement effective visualizations, front-end engineers to produce dashboards and visualizations for end users, and a manager to coordinate internally and communicate externally to engage key stakeholders and the dashboard's target audiences. Real-world resource constraints may dictate that more than one role and task be assumed by one individual. By some combination of skills and experience, technical expertise at every level of the process—as well as allocating sufficient effort and time—is just as integral to success as the infrastructure itself.

In addition to staffing and infrastructure needs, hosting and publishing dashboards involve ongoing resource and cost

XSL•FO
RenderX

considerations beyond initial development. These can include the scale and complexity of data pipelines and hosting environments, recurring effort related to system administration, security reviews, software updates, and dependency maintenance, and the level of effort needed to support different refresh cadences or degrees of automation. As in the HCS, decisions about hosting and publishing often require balancing available staffing, technical capacity, desired functionality, and timeline considerations. Hosting and publishing dashboards also involve a range of operational tasks across build-out, launch, and maintenance. In practice, this may include clarifying staff roles and responsibilities for tasks, such as data engineer for extract, transform, and load processes; software engineer for visualization coding; designer for website design; and administration and support staff for hosting and administration, user support, and coordination across technical and programmatic staff. Operational considerations may also include determining appropriate data refresh patterns for different metrics and ensuring that routine maintenance activities—such as monitoring data pipelines, updating dependencies, and addressing security issues—are incorporated into ongoing workflows. Consistent with the case-based nature of this manuscript, these issues are treated as considerations rather than as a prescriptive costing or operational blueprint, with feasible approaches varying by organizational size, mission, context, and the maturity of available technologies.

The data infrastructure supporting raw data, analytical results, and dashboards must meet or exceed the security and protection requirements specified in the associated DUA and, when applicable, comply with broader regulatory frameworks, such as the HIPAA. For instance, storing and processing unsuppressed data may necessitate a secure network server with robust cybersecurity measures before the data are processed and prepared for publication.

Access control is another critical consideration. Depending on the sensitivity of the data, it may be necessary to restrict access to entire dashboards or specific features. Managing user access requires dedicated human resources for tasks such as account creation, access approval, validation, and ongoing monitoring. Access models can range from fully open dashboards, where no account management is needed, to tightly controlled environments in which each user must be individually verified and assigned an account. Intermediate approaches include self-managed accounts, where users can opt-in or sign up for access, balancing security with convenience. Integration with external authentication systems can further influence infrastructure requirements. For example, allowing Open Authorization—which enables users to log in with existing credentials from work accounts, Gmail (Google LLC), or other third-party services—can simplify access management but requires additional technical and security considerations. Ensuring that the cyberinfrastructure can support these access control mechanisms while maintaining compliance with security and privacy regulations is essential for safeguarding sensitive data while maintaining usability.

There is a wide variety of software for dashboarding, database management, analysis, and content management with varying degrees of cost and technical skill needed for implementation. Important considerations for dashboard development teams during evaluation, comparison, and selection of platforms and services include costs for licenses and feature sets; required technical expertise; scalability (eg, based on data volume, types of visualizations, user base size and heterogeneity, and accessibility); and the maintenance burden. Additional considerations include if and how dashboards can be embedded into other web content—for example, by providing application programming interfaces along with attendant restrictions on reproduction and further use—and how well tools integrate with existing infrastructure and workflows. In the HCS, different sites selected tools based on local expertise, resources, and institutional constraints. Open-source solutions, such as Shiny (Posit PBC) or D3, minimized licensing costs and allowed for greater customization but often required more advanced programming skills to implement and maintain. Commercial platforms, such as Tableau (Salesforce) and Microsoft Power BI, provided graphical user interfaces that supported rapid development of common visualization types but were less flexible for customization. In addition, academic and government organizations may already have enterprise licenses for commercial dashboard software, which can strongly influence platform selection.

Prior to publication and dashboard launch, the project implementation team should outline a plan for end user testing and review of all data visualizations, including feedback on navigation, design style, functionality, and processing speeds. Staging or test sites are common and widely available across different hosting platforms and should allow for different user access and permissions during this prelaunch review phase. Unforeseen obstacles are common and should be included in any work plan and accounted for in project launch timelines, as one programming solution will usually lead to a new and separate issue to solve. Often overlooked, testers need to also test for acceptability and performance on different device and browser types for any web-based application during a review phase. Once a thorough and transparent feedback and review process has been completed and final dashboard versions are ready for launch, it is recommended to use a “soft launch” to allow for time to address any bugs not detected on test site development. While every system and dashboard build varies and presents different challenges from development to publication, maintaining flexibility and the ability to adapt to new inputs or requirements is critical.

Postlaunch Feedback and Revision

After the initial launch, live dashboards require routine maintenance and resolution of bugs to maintain trust and prevent burdensome obstacles to stakeholder engagement (Textbox 5). This requires a transparent mechanism in place for users to report bugs, submit feedback, and request new data or visualization formats. This mechanism should be easily accessible to end users either using an embedded form or a contact email address solely used for reporting issues or general feedback. Whichever method is implemented, site administrators would optimally receive automatic alerts to facilitate responsiveness. The dashboard team should also maintain a tracking tool to keep detailed records of each case to respond quicker and prevent similar issues in the future.

XSL•FO
RenderX

Textbox 5. Decision points for postlaunch feedback and revisions of dashboards.

The following questions highlight important considerations for monitoring dashboard performance after launch, collecting user feedback, and implementing revisions to improve functionality and usability over time.

- What is the plan for identifying bugs and issues and then implementing resolutions on an ongoing basis?
- What types and levels of end user feedback are integral to dashboard optimization, revision, and/or sunsetting?
- What are available methods for collecting end user feedback?
- What processes are in place for revising and enhancing dashboard components?
- How are postlaunch changes to metrics and visualizations documented and communicated to users?

Ongoing long-term feedback and revision of the dashboard system are equally necessary processes for assessing and evaluating whether dashboard aims are being met. These processes can iterate over the course of a project timeline but should be given consideration and built into initial scopes of work in early stages. Specific dashboard aims will differ depending on the project criteria; however, a central question in all public health dashboards should center on whether key metrics reach the intended audience and meet their needs. Are end users receiving useful information quick enough to inform decisions, understandable, and easy to apply in their own work or research? The answer to this question can determine whether information or insights gleaned from a data dashboard is in fact actionable in real-world settings (eg, setting program targets and monitoring implementation outcomes).

There are various methods to assess end-user experience. Well-documented and applied methods to collect feedback include usability testing (ie, direct observation of a participant performing a task), user surveys, focus groups, interviews, and website analytics tools that collect real-time site traffic data [31,32]. Each of these methods varies in resources required to implement and the types of data and insights they elicit. Balancing these factors and identifying one or a combination of methods throughout the course of the project can help ensure dashboard sustainability, as this feedback should be a key factor in the dashboard implementation team's decisions on potential revisions and enhancements to the initial platform content and design.

Website analytics tools (eg, Google Analytics) require few resources to set up and provide useful real-time and trend data on user sessions, most visited pages, average time spent on a page, and pathways of navigation. A simple metric, such as which pages are most visited, should not be overlooked, as this can inform how to prioritize certain visualization enhancements and modifications over others. Website analytics also provide data on end user characteristics (eg, where users are located, device type, and language), which are important inputs when prioritizing revisions and new feature development.

User surveys are a common method for gathering website feedback and vary in resource allocation from automated pop-up surveys that require little maintenance to more comprehensive questionnaires that may incorporate validated usability and user satisfaction scales, such as the SUS or its modified version, the Standard User Experience Percentile Rank Questionnaire (SUPR-Q) [33]; these validated questionnaires include prompts that address end users' attitudes about data credibility, appearance, ease of navigation, and willingness to return to the website and recommend to others. Despite the benefits of using these validated questionnaires, they alone will not provide the level of specificity needed to make targeted revisions for any single dashboard audience. While requiring more time in survey design, recruitment, and analysis, a more comprehensive survey that combines targeted prompts based on existing dashboard visualizations could prove to be a valuable source of end user feedback.

Focus groups, live webinars, and in-depth interviews are a more direct method of collecting end user feedback. While resource-intensive, these ensure direct community input and an opportunity to collect more nuanced real-time feedback from key stakeholders. In our prior usability study with the HCS, one important lesson we identified was the need to use an interdisciplinary approach to understand and improve engagement with data visualizations, which involves use of innovative technologies such as eye-tracking and audit log files and use of plan-do-study-act cycles to ensure fidelity with the co-creation process [12]. Such approaches generate multiperspective feedback that can support better design of a data dashboard.

Measuring a dashboard's usefulness and ability to facilitate an action or decision is often a challenge for any dashboard evaluation. While the various feedback collection methods mentioned, combined with website analytics, may provide insight into end-user satisfaction and dashboard traffic, they will not sufficiently capture the breadth of how end users are applying information from the dashboard and insights learned in real-world settings. A routine process of soliciting these examples from key dashboard stakeholders may involve pop-up prompts, requests via email listserv, and invitations to participate in follow-up surveys, webinars, focus groups, or user story-telling events. The process of gathering examples of real-world applications of dashboard content should repeat over the lifetime of a dashboard, and sustained documentation of these occurrences is integral to a comprehensive evaluation at the end of a project cycle.

Evaluation of public health dashboards can occur at multiple levels, ranging from usability and user experience to patterns of use and, in some cases, downstream decision-making. As noted earlier, evaluation in the HCS focused primarily on usability and perceived usefulness, using the SUS and qualitative interviews informed by the Technology Adoption Model [12]. Beyond usability, teams may also benefit from examining how dashboards are used over time (eg, views, click-through rates,

XSL•FO
RenderX

and dashboard item interactions) and how insights from dashboards inform decision-making processes. However, appropriate evaluation metrics, thresholds, and study designs (eg, AB testing) depend heavily on context, including dashboard purpose, data availability, organizational capacity, and intended use.

Once mechanisms for collecting evaluation data and feedback are in place, consistent and organized documentation of evaluation findings, end user feedback, ongoing maintenance issues, and new data and features requests is essential to informing prioritization. A collaborative process between dashboard implementers, key project stakeholders, and end users must weigh factors ranging from feasibility and required resources to potential impact when deciding which dashboard components to revise and enhance [34]. Across the HCS dashboards, feedback from different sources and endeavors informed multiple variations and versions of dashboard content and design. For example, qualitative feedback interviews indicated a need for support in data interpretation; this led to the team revising the dashboards to include "story" panels that contained brief textual summaries of temporal trends in the data. In another instance, community members raised questions during coalition planning meetings about the reach of online antistigma campaigns, which prompted the team to update the dashboard website to include visualizations of page views and click-through rates by campaign. These examples illustrate how user feedback informed dashboard development decisions without relying on a single, standardized feedback-to-change workflow.

As dashboards evolve post launch, changes to data sources, metrics, calculations, or visualizations should be tracked and documented (eg, changelog) in a systematic fashion. This includes situations in which metrics are modified, replaced, deprecated, or no longer supported due to changes or updates in data availability, definitions, reporting, or analytic approaches. Clear documentation of these changes can help maintain transparency, support interpretation of trends, and reduce confusion among end users, particularly when historical values are revised or when metrics are removed. While specific approaches to versioning, changelogs, or handling deprecated metrics will vary by platform and organizational context, the key task is considering how changes are documented and communicated.

Discussion

Summary

Well-designed dashboards serve as powerful tools for communicating public health data to key stakeholders and communities. Just as a carefully crafted automobile instrument panel can ensure a driver can reach their destination safely, an effective public health dashboard empowers community stakeholders to make informed, data-driven decisions. This is particularly crucial when epidemiological trends are rapidly changing or evolving across different locations and time periods. Unlike static data displays, interactive dashboards with real-time data provide a means of tracking and responding to public health challenges as they unfold. This playbook for developing data dashboards outlined steps and key considerations in data curation, design of visualizations, dashboard hosting and publishing, and feedback and revision.

Limitations

While we expect this playbook to be valuable for a wide range of public health applications, our experience with the HCS may not always be generalizable. HCS focused specifically on opioid overdose and opioid use disorder, shaping the selection of dashboard metrics and data to reflect epidemiological trends and modifiable factors related to overdose prevention and treatment, such as safer prescribing practices and naloxone availability. The unique challenges encountered—such as suppression rules for sensitive data—may not directly apply to other public health issues, particularly those with different temporal, geographical, or population-level dynamics that influence the usability of visualizations and dashboard features. The intervention being tested in HCS was highly community-engaged; thus, opportunities for community co-creation and feedback for dashboards may have been more readily available. We also recognized that the underlying basis for this material is a series of case studies rather than a systematic, empirical investigation of dashboard development and deployment. The coordinated, multisite structure of the HCS allowed individual dashboard teams to benefit from cross-site discussion, problem-solving, and deliberation around key considerations. While the lessons presented here reflect consensus developed across multiple teams, similar opportunities for coordination and shared learning may not be available in other dashboard development efforts. Finally, as part of a grant-funded study, HCS dashboards were not required to include formal or coordinated sustainability planning. As a result, issues shaping sustainability (eg, funding, technical debt, knowledge transfer, and ongoing stakeholder engagement) were not systematically addressed. We posit that these considerations remain important for future dashboard sustainability efforts but are also highly context- and resource-dependent, thus prohibiting us from recommending universal sustainability templates, diagrams, or checklists. Altogether, while this case-based tutorial playbook offers practical considerations, further research and evaluation are needed to refine effective and practical practices for designing and implementing public health data dashboards across diverse contexts.

Resources

The HCS dissemination efforts include a brief video entitled "How to Build Opioid Data Dashboards for Community Decision-Making: Best Practices from the HCS" [35]. This is a preview for a more in-depth, 5-session webinar for organizations interested in building public health data dashboards; webinar recordings and downloadable resources can be retrieved from the Data Dashboard Learning Collaborative Portal [36] (access provided upon request). The HCS also created a publicly accessible GitHub repository [37] with materials from a workshop organized along the following topics: (1) data acquisition and quality; (2) co-designing dashboards; (3) implementing and using dashboards in the community; (4) best practices for dashboard use; and (5) supporting the dashboard enterprise. The repository also includes

XSL•FO
RenderX

a DUA template, synthetic data, and sample code to construct metrics (eg, apply suppression rules), and workbooks to visualize data from our HCS dashboards.

Next Steps

Although the examples presented in this playbook are drawn from HCS dashboards developed for opioid overdose surveillance, we aimed to identify and describe decision points and considerations with potential portability to other public health issues. By emphasizing decision points rather than specific or prescriptive solutions, the tutorial playbook is intended to support adaptation across settings where data sources, governance structures, regulatory constraints, and available technologies may differ. We have provided Multimedia Appendix 1 to provide illustrative information and examples of using the playbook for building a dashboard for maternal mortality. However, public health encompasses a vast array of topics, and unforeseen challenges may arise when applying these principles to different issues beyond the scope of this work.

The scientific principles guiding the co-design of public health dashboards are still emerging and most likely will be subject to continual evolution, with many concepts remaining open to interpretation. This playbook provides a structured approach to addressing the challenges of designing rigorous, stakeholder-informed dashboards, offering key insights into the who, what, why, where, when, and how of dashboard development within the context of opioid use disorder and the overdose epidemic. This playbook's design is a representation of first-hand experience and lessons learned while co-creating dashboards to communicate important public health information to community stakeholders, not a representation of scientific research such as an empirical review. While sustainability was mentioned, it remains an area that warrants further exploration; readers and developers are encouraged to seek successful models from government-funded endeavors, public-private collaborations, or community efforts. Additionally, some notable technological advances—such as machine learning (ML) and generative artificial intelligence (AI)—emerged during the period in which these dashboards were developed but were not implemented at the time (hence, outside the scope of this case-based tutorial playbook). Future public health dashboards may explore the value of such technologies for dashboarding, including forecasting trends, identifying emerging patterns, or supporting scenario-based planning [38]. Beyond analytic applications, AI tools have also been described as potential "design collaborators" that may augment the utility, effectiveness, and appeal of visualizations [39]. At the same time, ML and AI raise important challenges, including data privacy and governance considerations, the need for specialized technical expertise, and the risk of misinterpretation or overreliance on model outputs (eg, with spurious findings or hallucinations). Taking such care, ML and AI in public health dashboards remain potentially exciting and valuable areas for further investigation. Altogether, as the allied fields and technologies continue to evolve, dashboard development paradigms should be refined and evolve with allied fields—such as data science, graphic design, psychophysics, and cognitive sciences—to maximize the public health impact that dashboards can bring to the field.

Acknowledgments

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, the Substance Abuse and Mental Health Services Administration, or the NIH HEAL Initiative. The authors wish to acknowledge Guertson Jean-Baptiste for developing the graphics for Figure 3. The authors declare the use of generative AI in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to GAI tools under full human supervision: proofreading and editing. The GAI tool used was: ChatGPT 5.2.

Funding

This research was supported by the National Institutes of Health and the Substance Abuse and Mental Health Services Administration through the NIH HEAL (Helping to End Addiction Long-term) Initiative under award numbers UM1DA049394, UM1DA049406, UM1DA049412, UM1DA049415, and UM1DA049417 (ClinicalTrials.gov Identifier: NCT04111939). Dr. Villani was substantially involved in UM1DA049394, UM1DA049406, UM1DA049412, UM1DA049415, and UM1DA049417, consistent with her role as a Scientific Officer.

Conflicts of Interest

None declared.

Multimedia Appendix 1
An example of generalizing the dashboard tutorial playbook to maternal mortality.
[DOCX File, 25 KB - publichealth_v12i1e83157_app1.docx]

References

1. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis 2020 May;20(5):533-534. [doi: 10.1016/S1473-3099(20)30120-1] [Medline: 32087114]

XSL•FO
RenderX

2. Dong E, Ratcliff J, Goyea TD, et al. The Johns Hopkins University Center for Systems Science and Engineering COVID-19 dashboard: data collection process, challenges faced, and lessons learned. Lancet Infect Dis 2022 Dec;22(12):e370-e376. [doi: 10.1016/S1473-3099(22)00434-0] [Medline: 36057267]
3. Wu E, Villani J, Davis A, et al. Community dashboards to support data-informed decision-making in the HEALing communities study. Drug Alcohol Depend 2020 Dec 1;217:108331. [doi: 10.1016/j.drugalcdep.2020.108331] [Medline: 33070058]
4. Arcia A, Benda NC, Wu DTY. Advancing the science of visualization of health data for lay audiences. J Am Med Inform Assoc 2024 Jan 18;31(2):283-288. [doi: 10.1093/jamia/ocad255] [Medline: 38238784]
5. Ancker JS, Benda NC, Zikmund-Fisher BJ. Do you want to promote recall, perceptions, or behavior? The best data visualization depends on the communication goal. J Am Med Inform Assoc 2024 Jan 18;31(2):525-530. [doi: 10.1093/jamia/ocad137] [Medline: 37468448]
6. Valdez RS, Rogers CC, Claypool H, et al. Ensuring full participation of people with disabilities in an era of telehealth. J Am Med Inform Assoc 2021 Feb 15;28(2):389-392. [doi: 10.1093/jamia/ocaa297] [Medline: 33325524]
7. Almasi S, Bahaadinbeigy K, Ahmadi H, Sohrabei S, Rabiei R. Usability evaluation of dashboards: a systematic literature review of tools. Biomed Res Int 2023;2023:9990933. [doi: 10.1155/2023/9990933] [Medline: 36874923]
8. HEALing Communities Study Consortium. The HEALing (Helping to End Addiction Long-term SM) communities study: protocol for a cluster randomized trial at the community level to reduce opioid overdose deaths through implementation of an integrated set of evidence-based practices. Drug Alcohol Depend 2020 Dec 1;217:108335. [doi: 10.1016/j.drugalcdep.2020.108335] [Medline: 33248391]
9. Sprague Martinez L, Rapkin BD, Young A, et al. Community engagement to implement evidence-based practices in the HEALing communities study. Drug Alcohol Depend 2020 Dec 1;217:108326. [doi: 10.1016/j.drugalcdep.2020.108326] [Medline: 33059200]
10. Balvanz P, Harris D, Olvera R, et al. Lessons from the community-engaged, data-driven selection of evidence-based practice strategies in the HEALing communities study. J Public Health Manag Pract 2026;32(1):25-33. [doi: 10.1097/PHH.0000000000002252] [Medline: 41260589]
11. Brooke J. SUS-a quick and dirty usability scale. In: Usability Evaluation in Industry: CRC Press; 1996:189-194. [doi: 10.1201/9781498710411]
12. Fareed N, Olvera RG, Wang Y, et al. Lessons learned from developing dashboards to support decision-making for community opioid response by community stakeholders: mixed methods and multisite study. JMIR Hum Factors 2024 Sep 9;11:e51525. [doi: 10.2196/51525] [Medline: 39250216]
13. Kortum PT, Bangor A. Usability ratings for everyday products measured with the system usability scale. Int J Hum Comput Interact 2013 Jan;29(2):67-76. [doi: 10.1080/10447318.2012.681221]
14. Slavova S, LaRochelle MR, Root ED, et al. Operationalizing and selecting outcome measures for the HEALing communities study. Drug Alcohol Depend 2020 Dec 1;217:108328. [doi: 10.1016/j.drugalcdep.2020.108328] [Medline: 33091844]
15. Olvera RG, Plagens C, Ellison S, Klingler K, Kuntz AK, Chase RP. Community-Engaged Data Science (CEDS): a case study of working with communities to use data to inform change. J Community Health 2024 Dec;49(6):1062-1072. [doi: 10.1007/s10900-024-01377-y] [Medline: 38958892]
16. Book reviews. USENIX Association. 2008. URL: https://www.usenix.org/system/files/login/articles/267-bookreviews1208.pdf [accessed 2025-01-02]
17. Few S. Information Dashboard Design: Displaying Data for at-a-Glance Monitoring, th edition: Burlingame, California: Analytics Press; 2013.
18. Cleveland WS. Robust locally weighted regression and smoothing scatterplots. J Am Stat Assoc 1979 Dec;74(368):829-836. [doi: 10.1080/01621459.1979.10481038]
19. Cleveland WS. LOWESS: a program for smoothing scatterplots by robust locally weighted regression. Am Stat 1981 Feb;35(1):54. [doi: 10.2307/2683591]
20. Cleveland WS, Devlin SJ. Locally weighted regression: an approach to regression analysis by local fitting. J Am Stat Assoc 1988 Sep;83(403):596-610. [doi: 10.1080/01621459.1988.10478639]
21. Diggle P. A kernel method for smoothing point process data. Appl Stat 1985;34(2):138. [doi: 10.2307/2347366]
22. Armstrong MP, Rushton G, Zimmerman DL. Geographically masking health data to preserve confidentiality. Statist Med 1999 Mar 15;18(5):497-525. [doi: 10.1002/(SICI)1097-0258(19990315)18:5<497::AID-SIM45>3.0.CO;2-#]
23. Wang J, Kim J, Kwan MP. An exploratory assessment of the effectiveness of geomasking methods on privacy protection and analytical accuracy for individual-level geospatial data. Cartogr Geogr Inf Sci 2022 Sep 3;49(5):385-406. [doi: 10.1080/15230406.2022.2056510]
24. The functional art: an introduction to information graphics and visualization. Google Books. 2012. URL: https://books.google.com/books?hl=en&lr=&id=xwjH6WuVUC8C&oi=fnd&pg=PT21&dq=The+functional+art+an+introduction+to+information+graphics+and+visualization&ots=nOxPLlGQ48&sig=kKoY3PYF9SSD5pnOOPOM [accessed 2026-01-30]

XSL•FO
RenderX

25. The truthful art: data, charts, and maps for communication. Google Books. 2016. URL: https://books.google.com/books?hl=en&lr=&id=8dKKCwAAQBAJ&oi=fnd&pg=PT15&dq=The+truthful+art+data+charts+and+maps+for+communication&ots=8U0zVJKJN&sig=dCgNwVr1UmhL55IhyuRIxGNY [accessed 2026-01-30]
26. Cleveland WS. Visualizing Data: Hobart Press; 1993:1-300.
27. Cleveland WS. The Elements of Graphing Data, th edition: Hobart Press; 1994:1-297.
28. Robbins NB. Creating More Effective Graphs: A Succinct and Highly Readable Guide to Creating Effective Graphs: Wiley; 2012:1-128.
29. Ware C. Information Visualization: Perception for Design, th edition: Morgan Kaufmann, Inc; 2021:1-560.
30. Tufte ER. The Visual Display of Quantitative Information, th edition: Graphics Press; 2018:1-197.
31. Joshi A, Amadi C, Katz B, Kulkarni S, Nash D. A human-centered platform for HIV infection reduction in New York: development and usage analysis of the Ending the Epidemic (ETE) dashboard. JMIR Public Health Surveill 2017 Dec 11;3(4):e95. [doi: 10.2196/publichealth.8312] [Medline: 29229592]
32. Burgan K, McCollum CG, Guzman A, et al. A mixed methods evaluation assessing the feasibility of implementing a PrEP data dashboard in the Southeastern United States. BMC Health Serv Res 2024 Jan 18;24(1):101. [doi: 10.1186/s12913-023-10451-5] [Medline: 38238697]
33. Sauro J. SUPR-Q: a comprehensive measure of the quality of the website user experience. J Usability Stud 2015;10(2):68-86 [FREE Full text]
34. van Elten HJ, Sülz S, van Raaij EM, Wehrens R. Big data health care innovations: performance dashboarding as a process of collective sensemaking. J Med Internet Res 2022 Feb 22;24(2):e30201. [doi: 10.2196/30201] [Medline: 35191847]
35. How to build opioid data dashboards for community decision-making: best practices from the HCS. YouTube. 2024 May 10. Video URL: https://www.youtube.com/watch?v=qgmOm-z0zVk [accessed 2026-03-31]
36. Data dashboard learning collaborative portal. 2024 Oct 11. URL: https://www.addictionpolicy.org/data-dashboard-learning-collaborative-portal [accessed 2026-03-16]
37. AMIA-healing. 2024 Mar 11. URL: https://github.com/AMIA-healing [accessed 2026-03-16]
38. Pinto AD, Birdi S, Durant S, et al. Machine learning applications in population and public health: guidelines for development, testing, and implementation. JMIR Public Health Surveill 2025 Oct 24;11:e68952. [doi: 10.2196/68952] [Medline: 41134979]
39. Sultana R, Master of Science in Management Information Systems, College of Business, Lamar University, Texas, USA. Artificial intelligence in data visualization: reviewing dashboard design and interactive analytics for enterprise decision-making. Int J Bus Econ Insights 2025 Sep 15;5(3):01-29 [FREE Full text] [doi: 10.63125/cp51y494]

Abbreviations

ADA: Americans with Disabilities Act
AI: artificial intelligence
CDC: Centers for Disease Control and Prevention
DUA: data use agreement
HCS: HEALing Communities Study
HIPAA: Health Insurance Portability and Accountability Act
ML: machine learning
NIDA: National Institute on Drug Abuse
NIH: National Institutes of Health
SAMSHA: Substance Abuse and Mental Health Services Administration
SUPR-Q: Standard User Experience Percentile Rank Questionnaire
SUS: System Usability Survey
WCAG: Web Content Accessibility Guidelines

Edited by A Mavragani, T Sanchez; submitted 28.Aug.2025; peer-reviewed by F Adjei, K Triep, O Rosales, P Usuemerai; accepted 24.Feb.2026; published 09.Apr.2026.

Please cite as:
Wu E, Balise R, Katz B, Harris D, Bullard M, Fareed N, Larochelle M, Villani J
Building Public Health Data Dashboards: Tutorial Playbook
JMIR Public Health Surveill 2026;12:e83157
URL: https://publichealth.jmir.org/2026/1/e83157
doi:10.2196/83157

XSL•FO
RenderX

© Elwin Wu, Raymond Balise, Benjamin Katz, Daniel Harris, Matthew Bullard, Naleef Fareed, Marc Larochelle, Jennifer Villani. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 9.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Bayesian Models to Generate Small Area Estimates of Population Health: Tutorial for Using Rate Stabilizing Tools and Their Output

David DeLara[1], MS; Ryan Zomorrodi[2], MS; Harrison Quick[3], PhD; Joshua Tootoo[2], MS; Ruiyang Li[2], MS; Justan Baker[4], MPH; Jihyeon Kwon[5], MS; Michele Casper[1], PhD; Adam Vaughan[1], PhD

[1]Division for Heart Disease and Stroke Prevention, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, 4770 Buford Highway, Atlanta, GA, United States

[2]Children's Environmental Health Initiative, University of Illinois Chicago, Chicago, IL, United States

[3]Division of Biostatistics and Health Data Science, University of Minnesota, Minneapolis, MN, United States

[4]Diabetes and Cardiovascular Health Program, Rhode Island Department of Health, Providence, RI, United States

[5]Department of Epidemiology and Biostatistics, Drexel University, Philadelphia, PA, United States

Corresponding Author:
David DeLara, MS
Division for Heart Disease and Stroke Prevention, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, 4770 Buford Highway, Atlanta, GA, United States

Abstract

The demand for high-quality population health data at the local level calls for expanded tools for those working to enhance the health of communities across the country to easily calculate small area estimates. Statistical models that generate small area estimates often use Bayesian estimation techniques, which are computationally complex and not readily accessible to most public health professionals. We developed 2 tools to facilitate small area estimation. For ArcGIS Pro users, we developed the Rate Stabilizing Toolbox ArcGIS plugin (RSTbx), and for R users, we developed the Rate Stabilizing Tool R package (RSTr). In this tutorial, we demonstrate how to use these tools to calculate small area estimates and evaluate their reliability. We also demonstrate 3 key benefits from using either of these tools: (1) decreased number of geographic units with suppressed estimates, (2) flexibility to set the threshold for statistical reliability, and (3) credible intervals that can be used to identify statistically significant differences between geographic units. Additionally, both tools offer built-in age-standardization capabilities. We created census tract–level maps from North Carolina mortality data and Rhode Island hospitalization data to showcase the benefits of generating small area estimates with these tools. Rate Stabilizing Toolbox and Rate Stabilizing Tool for R are powerful tools that can be used to meet the demand for high-quality local-level data to inform public health programs and tailor health promotion activities to the needs of communities across the country.

(JMIR Public Health Surveill 2026;12:e83498) doi:10.2196/83498

KEYWORDS

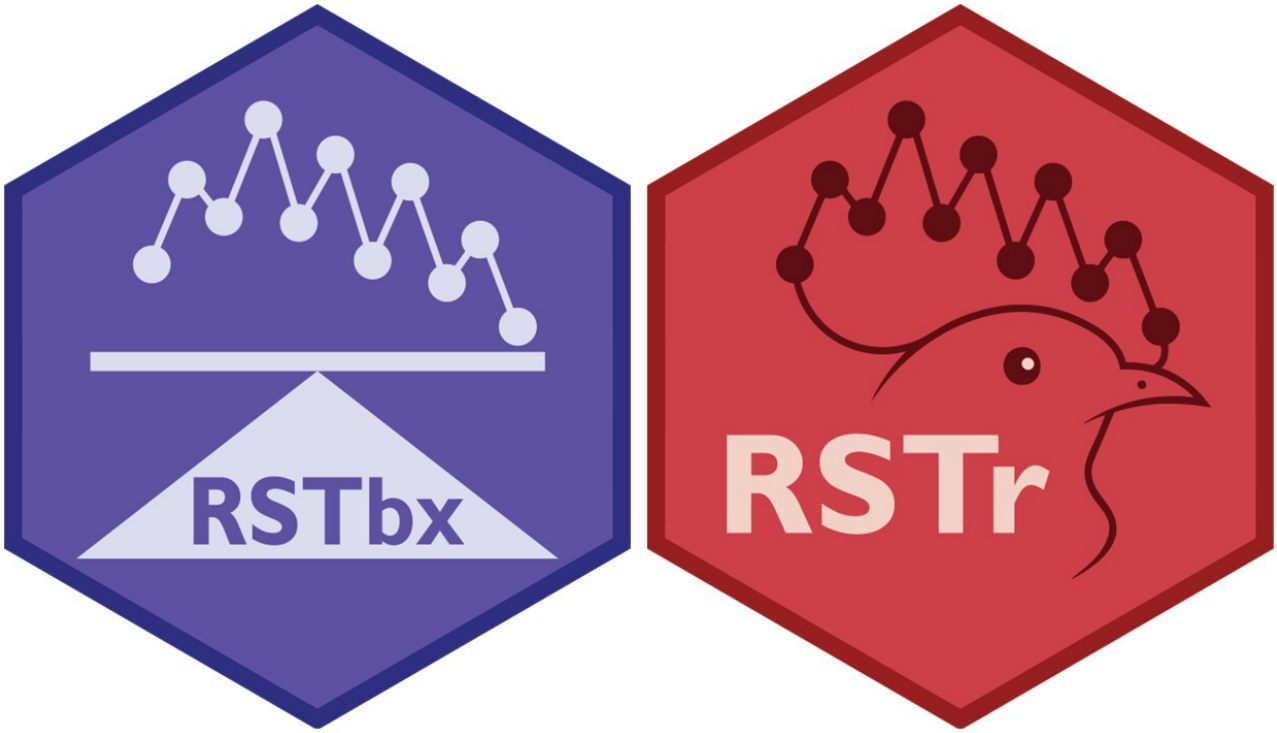
small area estimates; spatial analysis; software; geographic information system; GIS; spatiotemporal models; R; Bayesian statistics

Introduction

The demand for high-quality population health data at the local level calls for expanded tools that public health professionals and others can use to easily calculate robust small area estimates [1-3]. A key challenge for calculating robust small area estimates occurs in regions with small population sizes or few numbers of events; these small numbers introduce statistical uncertainty, leading to challenges in the meaningful interpretation of results [4]. Statistical methods that appropriately address these challenges are often computationally complex and may not be readily accessible to those working to enhance the health of communities across the country. In response to this need, we implemented established statistical models into 2 distinct tools: an ArcGIS toolbox (the Rate Stabilizing Toolbox [RSTbx]) and an R package (the Rate Stabilizing Tool for R [RSTr]) (Figure 1). These tools enable users to input local-level data and calculate robust small area estimates. This tutorial provides step-by-step instructions for how to use each tool along with a demonstration of the benefits derived from using these tools.

XSL•FO
RenderX

Figure 1. Icons for Rate Stabilizing Toolbox (RSTbx; left) and Rate Stabilizing Tool for R (RSTr; right).



Based on Tobler's First Law of Geography [5], which asserts that closer places are more alike than further places, spatial statistical models improve estimate precision by leveraging the data's underlying spatial structure (ie, spatial smoothing). Among these existing models, RSTbx and RSTr are based on the conditional autoregressive (CAR) model developed by Besag, York, and Mollié (BYM) [6], which is used extensively in spatial epidemiology and disease mapping [7]. Outcomes as varied as excess cardiovascular disease death rates in the United States during the COVID-19 pandemic [8], tuberculosis relative risk in Indonesia [9], and under-nutrition among under-five children in Ethiopia [10] have recently been investigated with methodology based on the BYM model. The popularity of the BYM model is further enhanced due to its extensions into spatiotemporal and multivariate settings through the MCAR of Gelfand and Vounatsou [11] and the MSTCAR of Quick et al [12]. Recent developments in BYM models involve strategies to quantify model informativeness and avoid oversmoothing [13,14] and outlining of standards for estimate reliability [15].

The RSTbx and RSTr both offer key benefits for calculating and mapping small area estimates of population health. First, by leveraging spatial and other dependencies in the data, the estimates produced by the RSTbx and RSTr are more precise than those based solely on the observed data. As such, the estimates produced by these tools will be more reliable, resulting in fewer estimates being suppressed, thereby permitting the documentation of geographic patterns with a more comprehensive geographic coverage. Second, both tools allow users to relax the threshold for reliability based on measures of uncertainty (eg, basing thresholds for reliability on the 80% credible intervals [CIs] rather than the standard 95% CIs). Third, these measures of uncertainty can be used to assess statistically significant differences between estimates for geographic units and other domains.

We have divided this tutorial into 3 sections. We begin with an overview of each tool, including how to use the tool, the input and output datasets, and their modeling capabilities. Then, we demonstrate the benefits of using the tools by mapping small area estimates of mortality using RSTbx and hospitalization rates using RSTr. Finally, we review the tools' strengths and limitations and provide a table comparing their major features.

RSTbx: An ArcGIS Toolbox

Overview

RSTbx is a Python-based set of tools designed for Environmental Systems Research Institute's ArcGIS Pro software [16]. Users can input their own local-level data and calculate local-level estimates using Bayesian spatial smoothing methods. RSTbx uses a BYM model framework that smooths across space using data from adjacent geographic units [6]. RSTbx also includes options for data processing, age-standardization, and the generation of CIs. RSTbx, with detailed instructions, is available for download on GitHub [16].

RSTbx is an upgrade to the original Rate Stabilizing Tool (RST) that was created in 2019 [17]. Enhancements include the following:

- A CAR model implementation based on the BYM model to replace the Poisson-gamma empirical Bayes method used in the original RST [7,8].
- The user interface has been revamped for greater ease of use.
- Users can import custom features; they are no longer limited to 2010 US Census geographies.
- Users can import their own population tables, whereas the original RST allowed only population data from the 2010 Census of the Population.

XSL•FO
RenderX

- Age-standardization to 10-year age groups is now possible using either the 2000 US Standard Population or the 2010 US Standard population [18].
- Population tables from the decennial US Census and American Community Survey (ACS; by census tract or county) can now be downloaded directly into an ArcGIS Pro project along with US Census TIGER and cartographic geographic boundaries [19].

RSTbx includes 3 tools: the Census Data Retriever (CDR), the Individual Data Processor (IDP), and the RST. The CDR generates population tables at the county or census tract level using data from the US Census' Decennial Census or the ACS. It can also be used to download TIGER or cartographic boundaries. The IDP aggregates individual-level event data by calculating the number of events within each geographic unit and joining the aggregate counts to the provided population table. The IDP can aggregate individual-level event data by age group to produce aggregate data for each geographic unit. The IDP also performs several data validation checks (eg, the IDP will check for null data, incorrect datatypes, or duplicate geographic units within population data). Finally, the RST runs a BYM model and generates small area estimates.

Installing RSTbx

To install RSTbx, users first download the latest release from the GitHub repository as a zip file [16]. After unzipping this file, users should open an ArcGIS Project, navigate to the Catalog Pane, right-click on Toolboxes, click on Add Toolbox, and find the rate_stabilizing_toolbox.pyt file within the unzipped RSTbx folder. After completing these steps, RSTbx is ready to use.

Input Data Requirements

RSTbx requires a minimum of 3 sets of data: event data, population data, and a boundary file. First, users enter event data for the health outcome of interest. RSTbx is designed to accommodate event data at either the individual or group level. For individual-level event data, RSTbx requires a unique identifier for each geographic unit (eg, GEOID). The age of each individual is also required to calculate age-standardized rates. For group-level event data, RSTbx requires a unique identifier for each geographic unit, the number of events in each geographic unit (ie, the numerator for the health outcome of interest), and the age group. The age group is optional and only required if users want to generate age-standardized rates. If the age group is included, the groups are restricted to: "0-4," "5-14," "15-24," "25-34," "35-44," "45-54," "55-64," "65-74," "75-84," "85up." For both the event and population data, the typical use case is single-year data, but data can be analyzed along any temporal aggregation.

Second, users must provide a table with the population data for each geographic unit or geographic unit–age group combination. Users can either supply their own population-level data or access the CDR within RSTbx to retrieve population data for census geographies. Users with noncensus geographies must provide their own population-level data. Importantly, RSTbx assumes that data for each geographic unit are a census (ie, data represent the entire population).

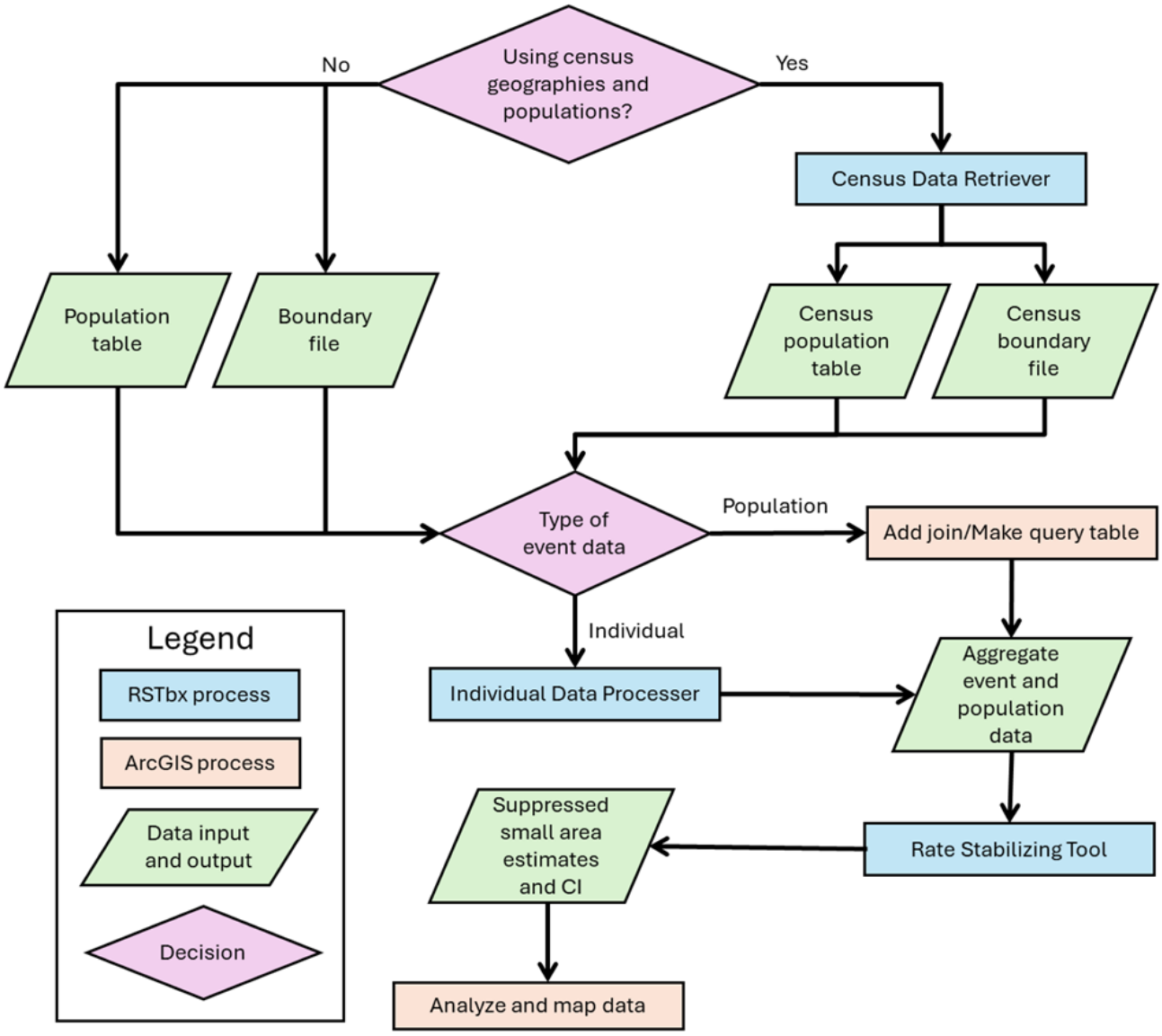
Finally, users enter a boundary file for the geographic area of interest. RSTbx supports most major file types through ArcGIS Pro, including but not limited to Geodatabase features, GeoPackage features, GeoParquets, and shapefiles. Each geographic unit must have at least 1 neighbor. If users are not using their own geographic boundary files, the CDR can be used to download TIGER or cartographic boundaries from the US Census Bureau.

Generating Small Area Estimates With RSTbx

Figure 2 outlines the steps involved in using RSTbx to calculate local-level measures of population health. The first steps include determining the geographic unit of analysis and obtaining population data and a boundary file for the chosen geographic unit. If the event data (eg, health data) are at the individual level, they must be aggregated to the geographic unit and joined to the population table with the IDP; population-level event data may be directly joined to the population table. The next step entails choosing the desired threshold for reliability. Finally, if conducting age-standardization, it is necessary to choose the age groups and a standard population year to be used in the age-standardization process. Detailed directions are provided within RSTbx and within the GitHub repository [16].

XSL•FO
RenderX

Figure 2. Flowchart for generating small area estimates using the Rate Stabilizing Toolbox (RSTbx).



Output

After running the models, RSTbx generates samples from the posterior distribution for each geographic unit, from which an output table is created. The table includes spatially smoothed estimates (ie, the posterior medians), lower and upper bounds of the user-defined CIs, and the level of reliability (eg, 95%, 90%, 75%) for each geographic unit. If age-standardization is requested, results will be produced for the age-standardized age range and all composite age groups within the data.

RSTr R Package

Overview

RSTr is an R package using Rcpp and RcppArmadillo [20-23] that generates small area estimates of health outcomes using CAR models [24]. Users may choose from one of many BYM-based models to generate small area estimates, including the univariate CAR, multivariate CAR (MCAR), and multivariate spatiotemporal CAR (MSTCAR). RSTr's CAR model smooths across space using data from neighboring geographic units [6,13]. The MCAR expands upon the CAR model by smoothing across both geographic units and domains (eg, sociodemographic groups) [11]. The MSTCAR further expands the MCAR model by smoothing across geographic units, domains, and time [12]. Statistical and technical details of these models are available in the package documentation [24].

Package Setup

RSTr uses the R statistical software [25] and can be installed from the Comprehensive R Archive Network (CRAN) [24].

Input Data Requirements

RSTr requires 3 pieces of input data: event counts, population data, and the adjacency structure of the geographic units. RSTr uses aggregate, rather than individual-level, data. Requirements for the structure of these inputs are described in detail in the package documentation [24]. Briefly, datasets for the event and population data for each geographic unit should be combined into an R list object. Requirements for the data structure differ based on the selected model. For example, the CAR model uses vectors, the MCAR model uses matrices, and the MSTCAR model uses 3-dimensional arrays. Importantly, RSTr assumes

XSL•FO
RenderX

that data for each geographic unit represent a census (ie, data represent the entire population). For both event and population data, the typical use case is single-year data, but data can be analyzed along any temporal aggregation. For the MSTCAR model, any temporal aggregation can be used, but all time periods must be an equal distance apart. Users must supply their own population count data, but can easily acquire data through the use of the tidycensus package for downloading 1-year and 5-year ACS estimates, along with 10-year decennial census data [26]. Finally, RSTr's models will impute data for estimates censored due to privacy reasons. For additional information about the data setup process, refer to the vignette titled "01: Understanding and Preparing Your Event Data."

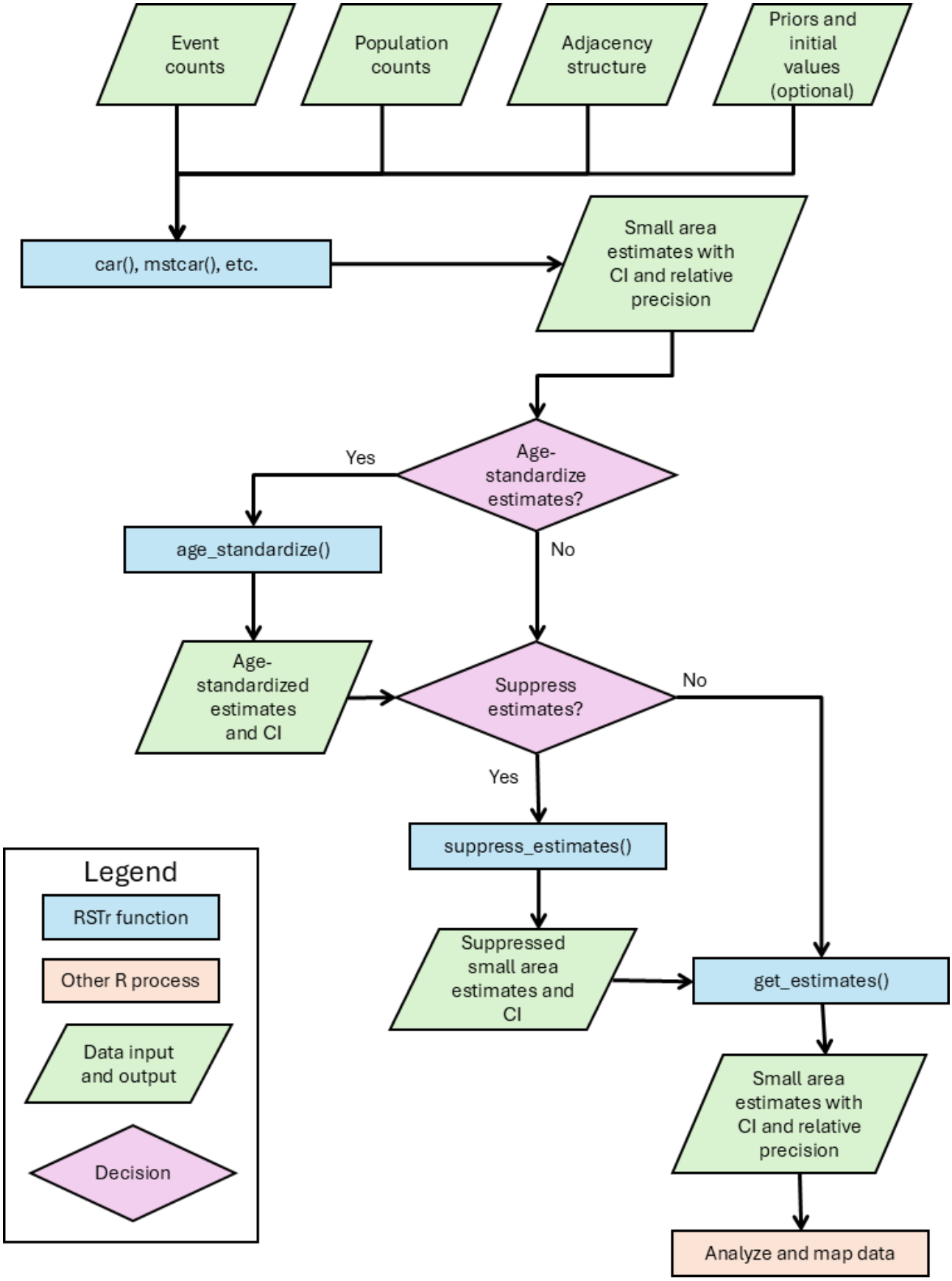
Adjacency information tells RSTr the neighbors of each geographic unit. Each geographic unit must have at least 1 neighbor. These adjacency data should be formatted in a structure the same as that generated by the poly2nb() function from the spdep package [27,28]. RSTr supports any filetype supported by GDAL through the use of the sf package [29,30], which has drivers for nearly every geospatial format.

Generating Small Area Estimates With RSTr

Figure 3 provides an outline of how to run models with RSTr. When running the model, users specify, at a minimum, the event data, population data, and adjacency structure. However, users may also specify the distribution of the event data (either Poisson or binomial as appropriate), initial values, priors, and other parameters to tune the model. Historically, birth and death data have been assumed to arise from Poisson processes [31], thus prompting the use of Poisson distributions to model birth [14,15] and death data [32,33]. However, because the Poisson distribution can be used to approximate the binomial distribution—and because a binomial might be more appropriate for other types of data—RSTr defaults to using a binomial distribution.

XSL•FO
RenderX

Figure 3. Flowchart for generating small area estimates using the Rate Stabilizing Tool for R (RSTr).



The decision of which model to select depends on the nature of the underlying data. In particular, the UCAR model is appropriate for a single year of data and a single demographic group, while the MCAR model is appropriate for a single year and multiple demographic groups (ie, multiple age groups to generate age-standardized estimates) and the MSTCAR model

XSL•FO
RenderX

is appropriate for data with both multiple years of data and multiple groups.

Output

Typical of Bayesian models, the RSTr models generate output in the form of samples. These samples are then used to generate small area estimates and CIs. Samples are included as a standalone output; estimates based on those samples, along with their corresponding CIs and relative precisions, are included as part of a larger model object from which these elements can be extracted.

Estimates are defined as the median of these samples and are extracted using the get_estimates() function on the RSTr model object. The output of get_estimates() is by default a long table containing medians (ie, estimates) for each region, group, and time period included in the model. The get_estimates() function also returns the credible interval, event and population counts, and relative precision, defined by Quick et al [13] as the ratio of the posterior median and the width of the user-specified CI, for each estimate. Relative precisions greater than 1 correspond to "reliable" estimates, per the framework of Quick et al [13].

Users may specify thresholds for reliability from 0 to 1 to define the CI used for relative precision calculations. At the default level of 0.95, the CI is defined as the 2.5 and 97.5 percentiles of the samples. This function may be especially beneficial when data are limited and few, if any, estimates would be deemed reliable at the traditional 0.95 level. For example, presenting maps of the estimates that are reliable at the 0.80 level will allow more estimates to be displayed while also acknowledging that a relaxed standard of reliability is being used. This functionality is demonstrated below.

Benefits of Mapping Small Area Estimates Generated by RSTbx and RSTr

Overview

There are many benefits to mapping small area estimates of population health generated by the RSTbx and RSTr. In this section, we demonstrate 3 key benefits: (1) decreased number of geographic units with suppressed estimates, (2) flexibility to set a threshold for reliability, and (3) CIs that can be used to identify statistically significant differences between geographic units.

Datasets Used to Demonstrate the Benefits of RSTbx and RSTr

RSTbx: North Carolina Mortality Data

To demonstrate the benefits of the RSTbx, we used North Carolina mortality data at the census tract level. Specifically, we examined heart disease deaths among adults aged 35 to 64 years in North Carolina census tracts for the years 2017 to 2019. Heart disease deaths were defined as deaths with International Classification of Disease, 10th revision (*ICD-10*) I00–I09, I11, I13, I20–I51 listed as the underlying cause. Results were age-standardized in 10-year age bands to the 2010 US Standard Population. Underlying data were made available through an agreement with the North Carolina Department of Health and Human Services, Division of Public Health.

RSTr: Rhode Island Hospitalization Data

To demonstrate the RSTr, we use Rhode Island hospitalization data from the Rhode Island Department of Health Hospital Discharge Data Program [34]. Specifically, we examine myocardial infarction and stroke-related inpatient hospitalizations in acute care hospitals among adults aged 20 to 69 years for Rhode Island census tracts in 2021 to 2023. Myocardial infarction and stroke hospitalizations were defined as hospitalizations with *ICD-10* I21-22, I60-63, I65-66 as the primary diagnosis. We ran an MSTCAR model for the years 2021 to 2023 by sex and 10-year age group and used the 2000 US Standard Population in 10-year age bands for age-standardization [18]. We then aggregated across sex and age-standardized across age groups. Underlying data were made available through an agreement with the Rhode Island Department of Health Hospital Discharge Data Program and are not publicly available.

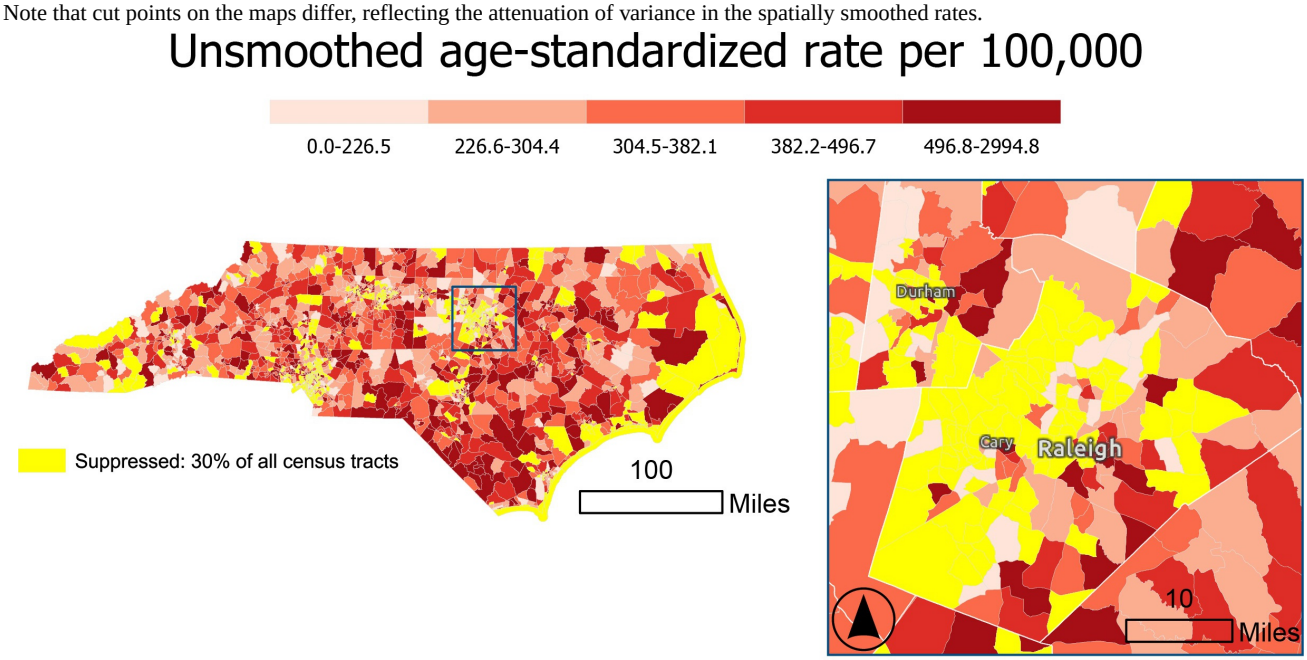
Benefit 1: Decreased Number of Geographic Units With Suppressed Data

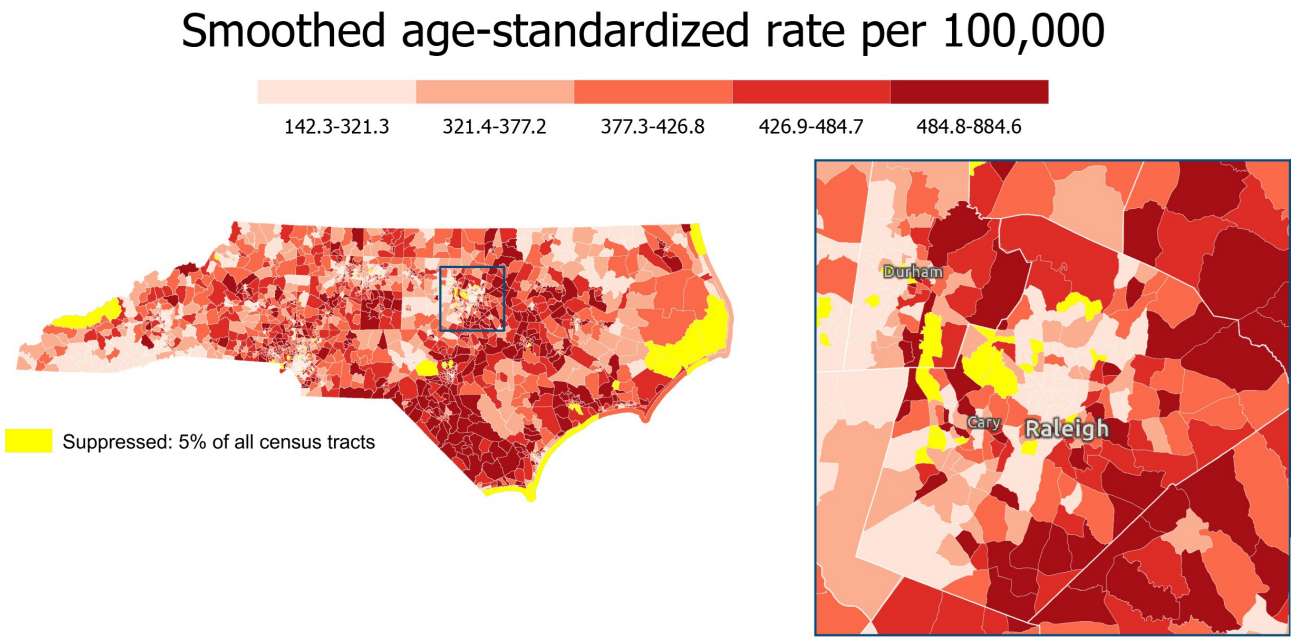
Overview

An important advantage of the underlying Bayesian statistical models in RSTbx and RSTr is their ability to increase the precision of estimates, thereby offering the potential to produce reliable estimates even when the event counts and/or population sizes are small. Compared to crude, unmodeled estimates, RSTbx and RSTr generate reliable estimates for a greater number of geographic units and therefore display fewer suppressed geographic units on maps (Figures 4 and 5). In these figures, estimates are deemed reliable based on their 95% CI. Additionally, the models used by these tools attenuate outliers and therefore narrow the range of estimates after spatial smoothing. This change reflects the improved precision and reliability of estimates that are otherwise sporadically high or low.

XSL•FO
RenderX

Figure 4. Rate Stabilizing Toolbox (RSTbx)-generated heart disease death rates by North Carolina census tract, adults aged 35 years and older, 2017 - 2019. The rates displayed in the top map are unsmoothed and suppressed according to United States Cancer Statistics suppression criteria; the rates displayed in the bottom map are spatially smoothed and use suppression criteria based on relative precision using a 95% CI. Comparison of the smoothed and unsmoothed rates shows the decreased percentage of census tracts that have suppressed rates when using RSTbx, demonstrating Benefit 1. Note that cut points on the maps differ, reflecting the attenuation of variance in the spatially smoothed rates.





XSL•FO
RenderX

Figure 5. Rate Stabilizing Tool for R (RSTr)-generated myocardial infarction and stroke inpatient hospitalization rates by Rhode Island census tract, adults aged 20‑69 years, 2021‑2023. The top map displays unsmoothed hospitalization rates suppressed according to the Rhode Island Department of Health Small Numbers Policy suppression criteria; the bottom map shows spatially smoothed hospitalization rates and use suppression criteria based on relative precision using 95% CIs and an overall census tract population threshold of 100. Comparison of the smoothed and unsmoothed rates shows the decreased percentage of census tracts that have suppressed rates when using RSTr (50% and 7%, respectively), demonstrating Benefit 1. Note that cut points on the maps differ, reflecting the attenuation of variance in the spatially smoothed estimates.

XSL•FO
RenderX

Benefit 1: RSTbx Demonstration

Figure 4 displays age-standardized heart disease death rates for adults aged 35 - 64 years by census tract in North Carolina for the years 2017 - 2019. In the top map of unsmoothed death rates, census tracts with fewer than 16 deaths are suppressed according to United States Cancer Statistics guidelines [35]. Using these guidelines, 30% (n=656) of census tracts were suppressed. The suppressed census tracts were distributed across the state, particularly in coastal regions and urban areas. The bottom map of Figure 4 displays death rates that were spatially smoothed using RSTbx. Here, applying the default threshold for reliability (which is based on the 95% CI) resulted in 5% (n=105) of census tracts being suppressed.

Benefit 1: RSTr Demonstration

Figure 5 displays age-standardized myocardial infarction and stroke inpatient hospitalization rates for adults aged 20 to 69 years by census tract in Rhode Island for the years 2021 to 2023. In the top map, estimates are suppressed according to the Rhode Island Department of Health Small Numbers Reporting Policy [36]; the bottom map is spatially smoothed and suppressed based on precision and population. The Rhode Island Department of Health Small Numbers Policy features data suppression recommendations for a wide variety of data types and data reporting scenarios. In the case of the hospitalization data, several layers of consideration about the data are needed to determine whether or not to suppress rates. In brief, the rates suppressed in the unsmoothed map take into consideration small numerators (number of hospitalizations) paired with large denominators, leading to unreliable or unstable rates. The suppression criteria result in 50% (n=122) of census tracts being suppressed. The suppressed census tracts were primarily located in small Rhode Island towns and villages, especially small coastal towns and villages.

For the spatially smoothed map, relative precision was calculated using a 95% CI and resulted in 7% (n=17) of census tracts being suppressed with an overall census tract population threshold of 100. Additionally, the smoothed map demonstrates an extension of the spatial patterns in the unsmoothed map. The inner census tracts of Providence have more tracts in the higher categories than in the unsmoothed data, and many less dense areas have low hospitalization rates. In this example, the increase in the number of reliable estimates using the smoothed estimates tells a clearer version of the story presented by the unsmoothed data.

Benefit 2: Flexibility to Set the Threshold for Reliability for Small Area Estimates

Overview

The default threshold for reliability for both RSTbx and RSTr is set at the commonly used 95%. However, RSTbx and RSTr both allow users to relax that threshold to less than 95%. Both tools use the definition in Quick and Song [13] to deem an estimate “reliable” if its relative precision (ie, the ratio of the posterior median and the width of the CI) is greater than 1. RSTbx output includes the maximum CI for which each estimate is deemed reliable, known as its level of reliability. RSTr allows users to relax the threshold for reliability by calculating the relative precision using narrower CIs (eg, 90% or 80%).

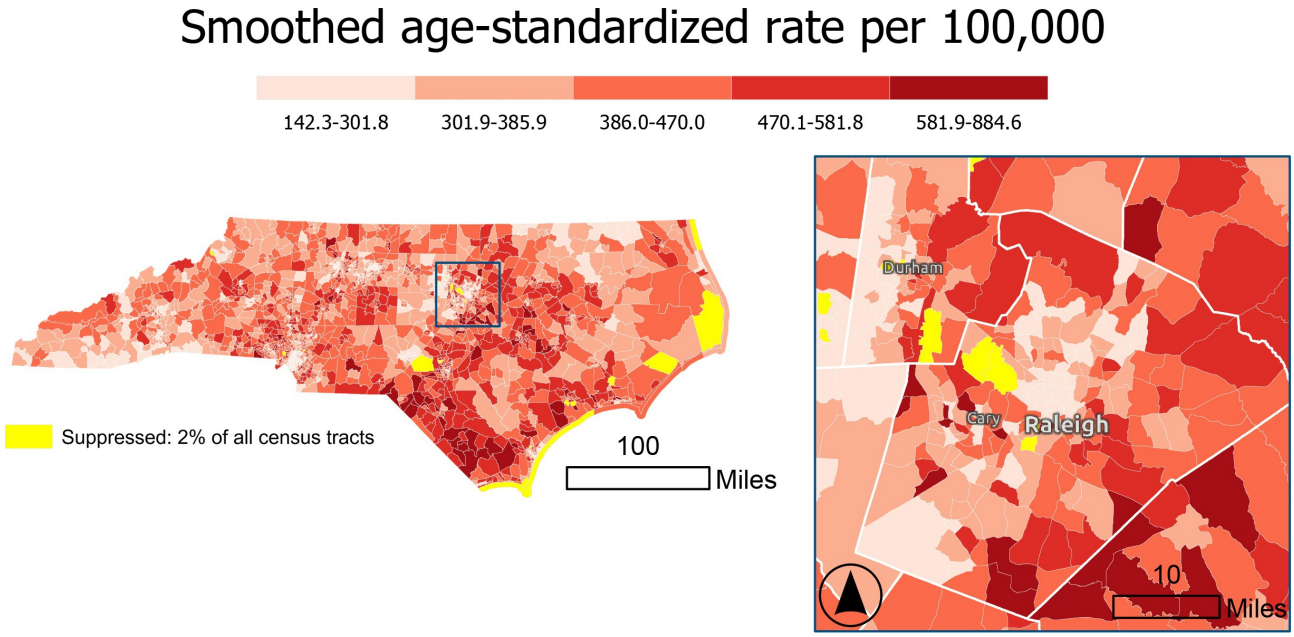
This feature allows users to more completely visualize geographic patterns due to less suppression at lower levels of reliability. This advantage, however, comes with a caveat: at some point, with a low enough threshold for reliability (eg, 70%), all geographic units would be considered to have “reliable” estimates. However, estimates with low levels of reliability may be based on very small numbers, leading to potential privacy concerns. Therefore, relaxing the reliability level may need to be accompanied by additional suppression criteria based upon population size.

Benefit 2: RSTbx Demonstration

Figure 6 displays North Carolina heart disease mortality rates for census tracts that meet the threshold for reliability using the 90% CIs. Relaxing the threshold for reliability from 95% in Figure 4 to 90% decreases the percent of census tracts that are suppressed from 5% to 2%.

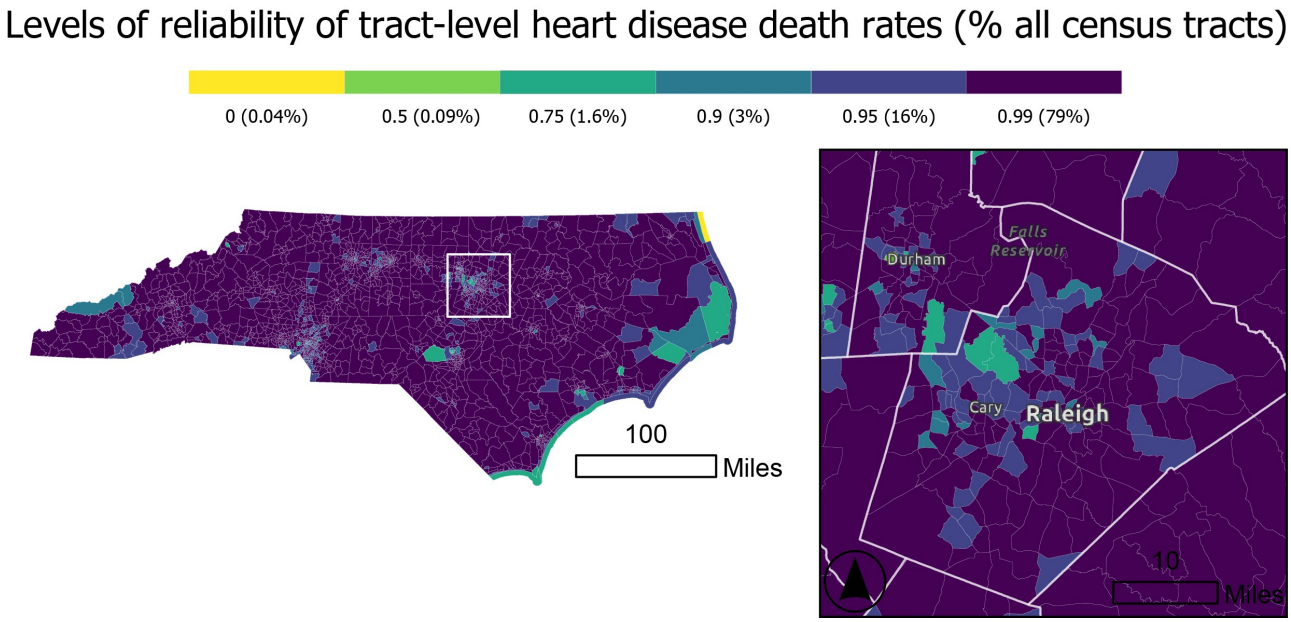
XSL•FO
RenderX

Figure 6. Rate Stabilizing Toolbox (RSTbx)-generated heart disease death rates by North Carolina census tract, adults aged 35 years and older, 2017 - 2019. The map displays spatially smoothed rates, and reliability is defined by relative precision using 90% CIs. Comparison of this figure with Figure 4 shows the decrease in the percentage of census tracts with suppressed rates when using a relaxed precision level, demonstrating Benefit 2.



RSTbx also permits users to map geographic units by their level of reliability since those are included in the RSTbx output. Figure 7 displays the level at which the estimate for each census tract is reliable. Estimates for much of the state are reliable at or above 95%. From this map, users can easily see geographic patterns in the reliability of the estimates and the range of reliability across estimates. As demonstrated in Quick and Song, these maps can be viewed as mapping a measure of uncertainty, as they highlight the geographic variation in the amount of information contributed by each geographic unit [15].

Figure 7. Levels of reliability of Rate Stabilizing Toolbox (RSTbx)-generated age-standardized heart disease death rates by North Carolina census tract, adults aged 35 years and older, 2017 - 2019. This map shows the maximum credible interval level at which stable and reliable smoothed estimates can be produced, illustrating the spatial variation in the quality of RSTbx-derived estimates and demonstrating Benefit 2. Wider credible intervals reflect greater uncertainty, which tends to occur in tracts with smaller total population estimates and lower case counts.

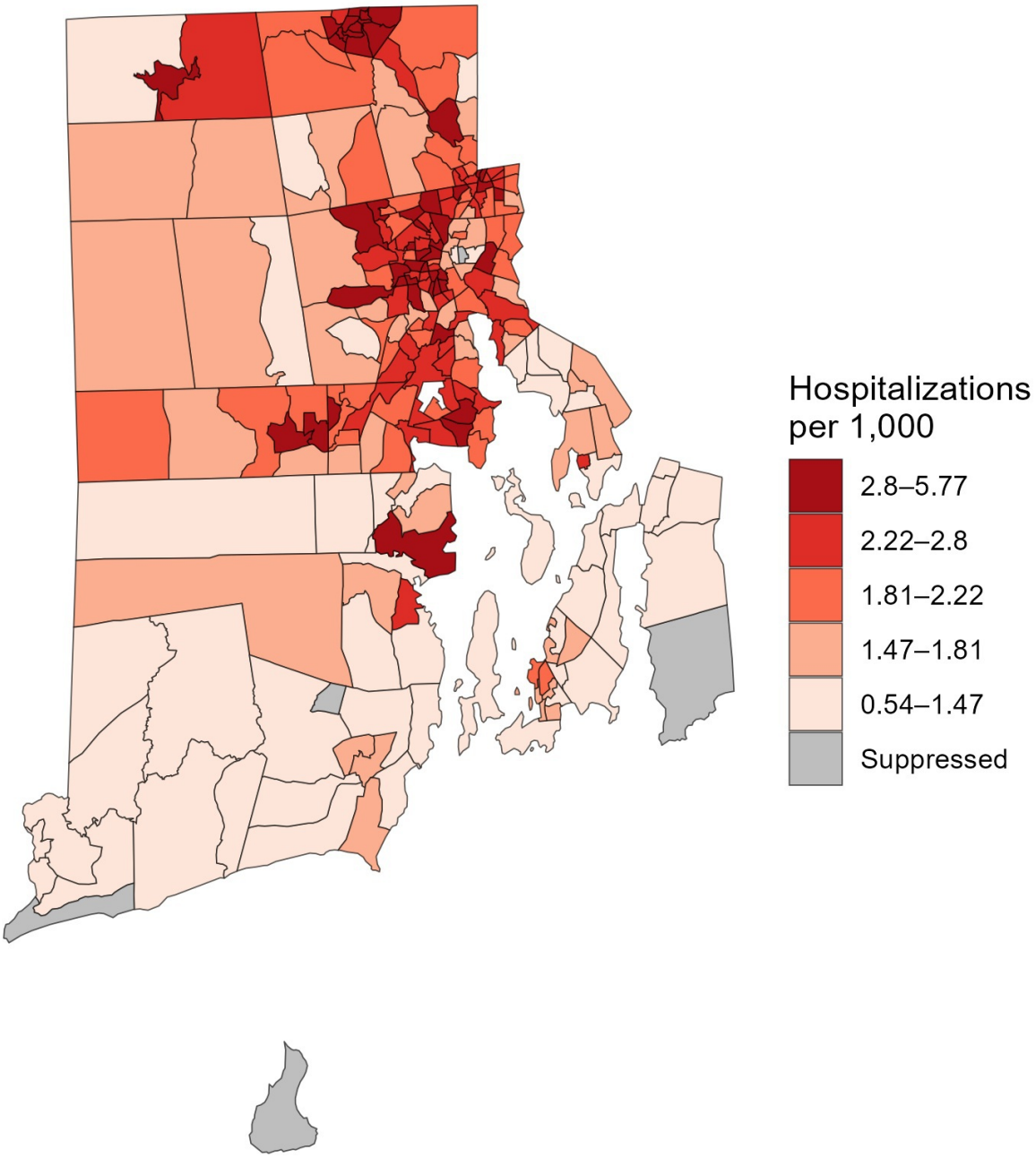


Benefit 2: RSTr Demonstration

Figure 8 displays Rhode Island myocardial infarction and stroke in-patient hospitalization rates for census tracts that meet the reliability criteria using a 90% CI. Relaxing the CI for the reliability criteria from 95% in Figure 5 to 90% further decreases the percent of census tracts that are suppressed (from 7% to 2%), but includes less precise estimates.

XSL•FO
RenderX

Figure 8. Rate Stabilizing Tool for R (RSTr)-generated myocardial infarction and stroke hospitalization rates by Rhode Island census tract, adults aged 20 - 69 years, 2021 - 2023. The map displays spatially smoothed rates. Reliability is defined by relative precision using a 90% CI. Comparison of this figure with Figure 5 shows the decrease in the number of census tracts with suppressed rates using a relaxed precision level, demonstrating Benefit 2. A total of 5 (2%) of census tracts are suppressed.



Geographic patterns in the levels of reliability (ie, the maximum CI whose width is less than the posterior median) will closely align with the population sizes of the geographic units. However, unlike maps of population sizes—where a given population size may be sufficient to produce reliable estimates for a common outcome and be insufficient to produce reliable estimates for a rare outcome—maps of the reliability levels will be standardized across outcomes. That is, estimates from analyses of different outcomes in different datasets with the same level of reliability can be viewed as being equally precise on a relative basis.

Benefit 3: Using Credible Intervals to Identify Statistically Significant Differences Between Places

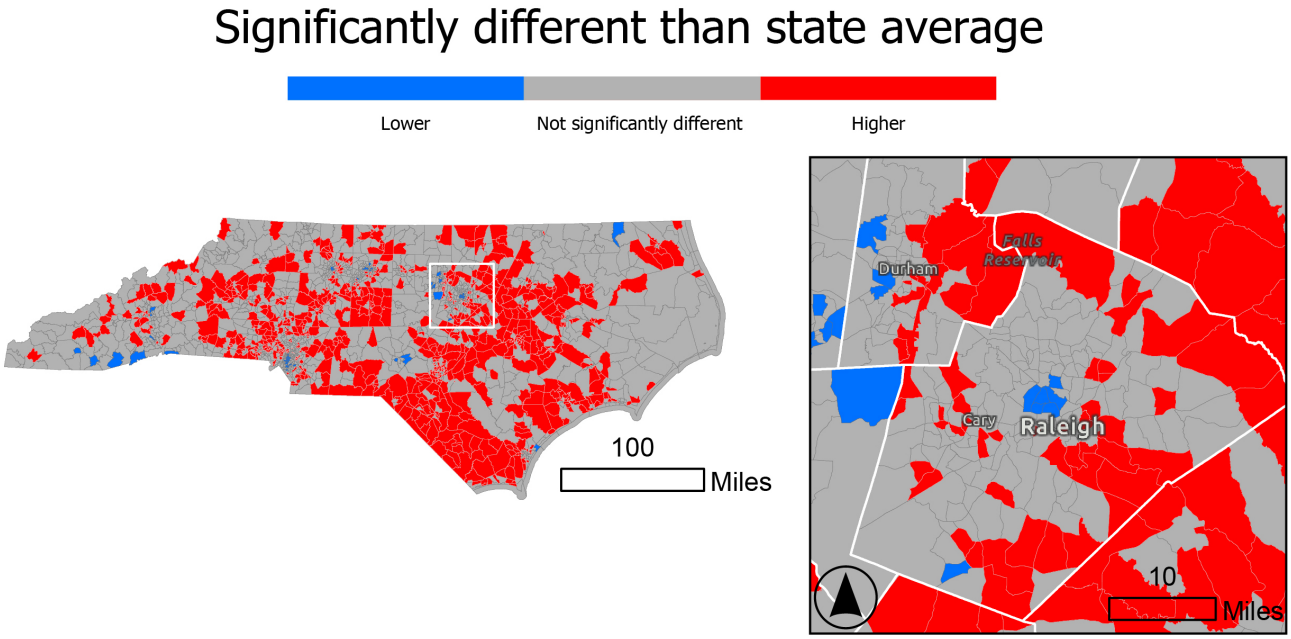
Overview

Another benefit of RSTbx and RSTr is the ability to compare estimates for geographic units to another value. This other value may be a single value (eg, a state rate) or another estimate (eg, the census tract with the lowest estimate). This comparison is made by using the CI at a specified level (eg, 95%). Estimates with a CI that excludes a value or that does not overlap with another CI are determined to be statistically different. Importantly, estimates that are unreliable can still be statistically significantly different from a comparison rate. As demonstrated in Figures 9 and 10, this approach can be used to highlight

XSL•FO
RenderX

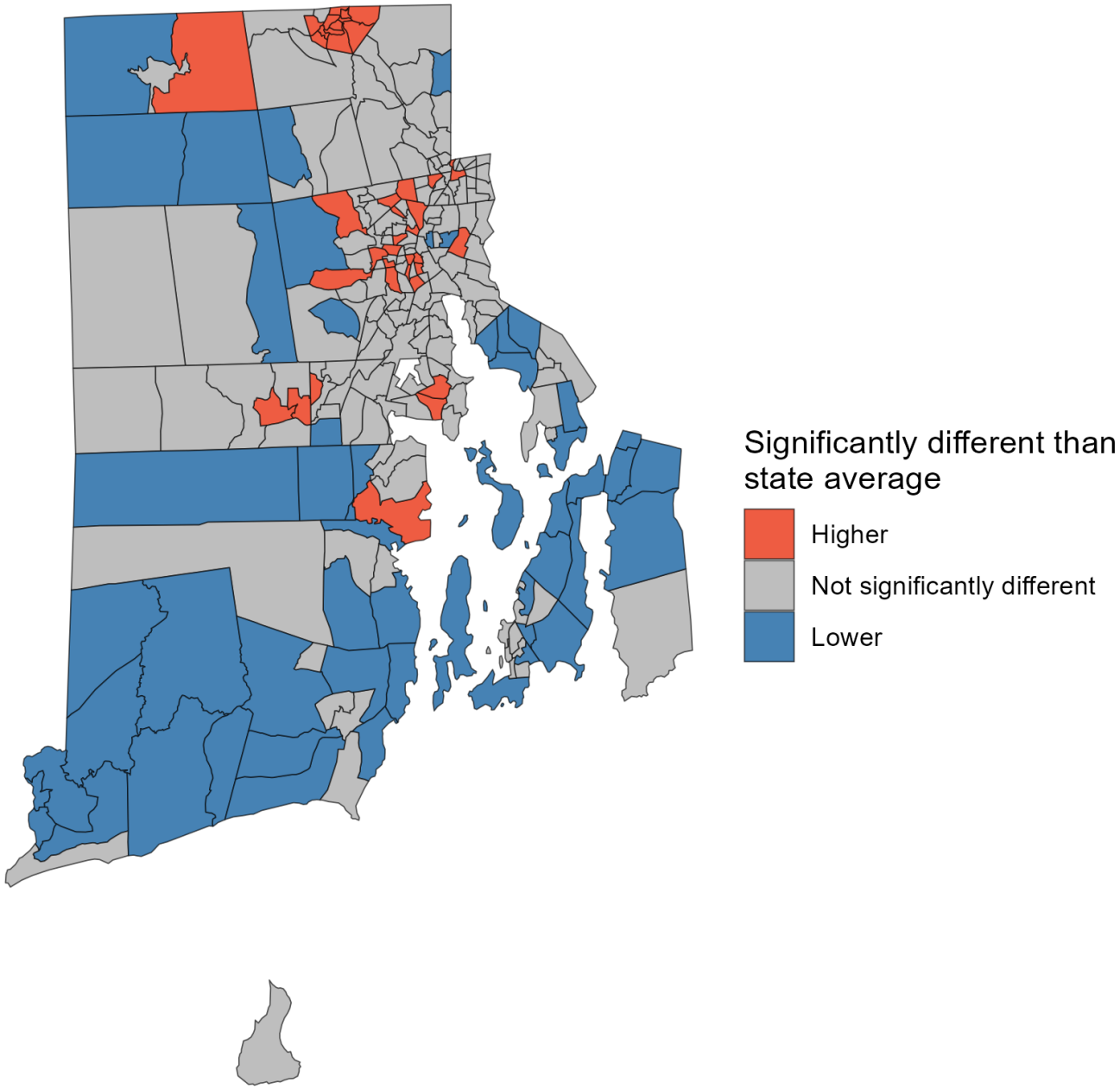
geographic units (eg, census tracts) that are statistically higher or lower than the state rate.

Figure 9. Comparison of Rate Stabilizing Toolbox (RSTbx)-generated census tract level heart disease death rates to the state rates, ages ≥35 years, 2017 - 2019, North Carolina. By using 95% CIs that are generated for each census tract, this map shows census tracts with heart disease death rates that are statistically significantly higher or lower than the North Carolina heart disease death rate, demonstrating Benefit 3.



XSL•FO
RenderX

Figure 10. Comparison of Rate Stabilizing Tool for R (RSTr)-generated census tract level myocardial infarction and stroke hospitalization rates to the state rates aged 20 - 69 years, 2021 - 2023, Rhode Island. By using 95% CIs, this map shows census tracts with hospitalization rates that are statistically significantly higher or lower than the Rhode Island hospitalization rate, demonstrating Benefit 3.



Benefit 3: RSTbx Demonstration

Figure 9 displays census tracts in North Carolina for which the age-standardized, spatially smoothed heart disease death rates are statistically higher (red) or lower (blue) than the state level. Statistical significance is determined when the 95% CI for the tract level death rates do not include the state level value. Census tracts with rates that are statistically different from the state rate are found across the state. Notably, statistically higher rates are concentrated in the state's urban areas.

Benefit 3: RSTr Demonstration

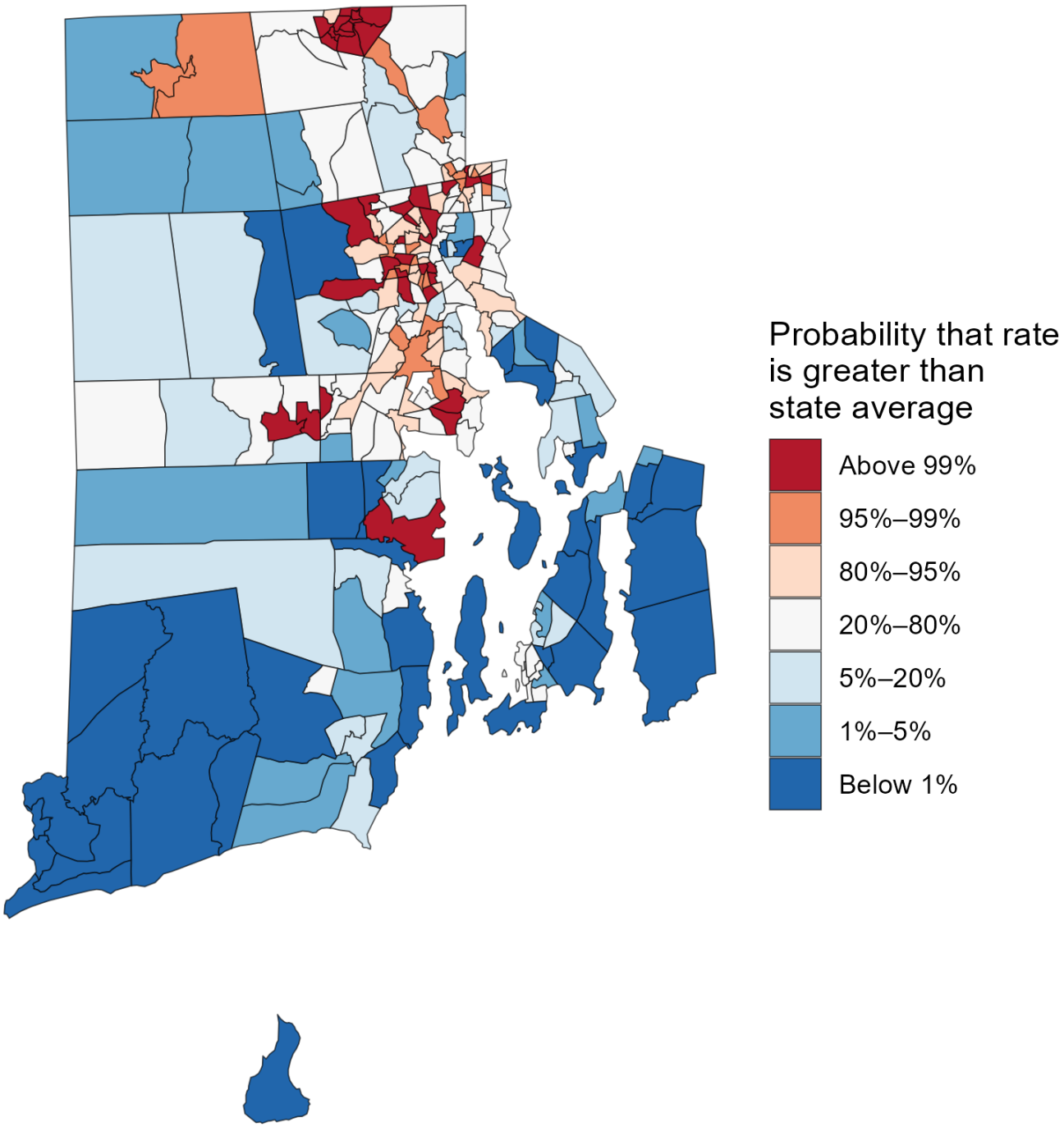
Figure 10 displays census tracts in Rhode Island for which the age-standardized, spatially smoothed myocardial infarction and stroke in-patient hospitalization rates are statistically higher (red) or lower (blue) than the state rate. Statistical significance is determined when the 95% CI for the tract level death rates do not include the state rate. Census tracts shown in gray have estimates that are not significantly different from the state hospitalization rate.

Additionally, the output from RSTr allows users to calculate the probability that one value is greater than another, frequently referred to as an exceedance probability. In Figure 11, we demonstrate the probability that the tract level hospitalization rates in Rhode Island exceed that of the state level. Unlike the approach in Figure 10, we are not limited to binary differences in probability, revealing a nuanced look at the statistical strength of the estimates. With RSTr's samples, users can calculate the percent of samples for a given region, group, or time that exceed a given value. In this approach, the percent estimate implies its opposite as well: an 80% chance for an estimate to be greater than the state hospitalization rate implies a 20% chance for the estimate to be less than that rate. In Figure 11, most of the

XSL•FO
RenderX

strongest probabilities of an estimate being greater than the rate for Rhode Island are clustered in urban areas, whereas rural areas tend to have the lowest probabilities of being greater than the state's hospitalization rate, indicating that their rates may be significantly lower than the state rate.

Figure 11. Probability that Rate Stabilizing Tool for R (RSTr)-generated myocardial infarction and stroke hospitalization rates by census tract are greater than the state value, adults aged 20 - 69 years, 2021 - 2023, Rhode Island. This is an extension of Benefit 3.



Discussion

Strengths and Limitations

RSTbx and RSTr have many strengths. First, as we have demonstrated, the outputs of both RSTbx and RSTr can be used to map small area estimates or statistical comparisons, with varying thresholds for reliability. Second, these tools allow users to quickly run complex Bayesian models. Third, users can easily generate age-standardized rates. RSTbx includes integrated standardization to either the 2000 US Standard Population or the 2010 US Standard Population in 10-year age groups. RSTr users may age-standardize to any population and are able to aggregate by domains other than age, such as race, sex, year, or region.

Though software such as JAGS, Stan, and BUGS [37-39] have addressed many of the same issues as RSTr and RSTbx and provide many of the statistical models used by our tools in a more generic framework, our tools come with distinct advantages. First, RSTr features a wider family of CAR models to choose from with more complex specifications available;

XSL•FO
RenderX

Stan and JAGS, for example, provide at best an MCAR model for estimation, and BUGS only provides a BYM CAR model. Additionally, since RSTbx and RSTr are designed specifically to run CAR models, their parameters and initial values are pre-specified based on existing work in the literature. More general software like Stan and BUGS requires a more hands-on approach and can lead to nonideal CAR model specification. Finally, RSTbx and RSTr utilize recent advances in CAR model methodology not seen in other statistical packages. The UCAR model implemented in both tools features enhancements that restrict the strength of the spatial smoothing using models which aim to establish a minimum on the amount of data required to yield a reliable estimate [13,15].

Both RSTbx and RSTr have limitations. First, the input data are assumed to be a census or representative sample of the population being studied. Neither tool is designed to accommodate survey weights or potential disparities in reporting quality (eg, undercounts). Any data quality issues that would cause the crude values (ie, the event counts divided by the population sizes) to be biased would also affect the quality of the estimates. Spatial data in general are prone to many limitations; nuances such as the modifiable areal unit problem, the ecological fallacy, and edge effects should all be considered when interpreting results generated by RSTbx and RSTr. Additionally, RSTbx includes only the UCAR model, meaning that smoothing occurs only over space and not over time or other domains.

Finally, our statistical models have some limitations. In contrast to the UCAR models used in both tools, RSTr's MCAR and MSTCAR models do not feature enhancements to restrict the strength of the spatial smoothing and thus have the potential to yield estimates with relative precisions greater than one when zero events have been observed. As a result, these models may produce estimates that are overly smooth (ie, reduced geographic disparities between adjacent geographic units) and overly precise (ie, inflated levels of reliability). To address this limitation, users can flag estimates generated by RSTr as unreliable when population sizes are not sufficiently large in addition to the use of relative precision. While Quick and Song do not provide guidance regarding requirements for population sizes, others have suggested requiring population sizes be at least 30 or 100 in order to display estimates [12]. Development of restricted MCAR and MSTCAR models is an active area of research, and future updates to RSTr and RSTbx aim to include such models. Additionally, there are future plans to expand RSTbx to use open-source GIS software, such as QGIS. Future versions of RSTbx will also include imputation of censored data and further streamline the data setup process.

Conclusion

RSTbx and RSTr facilitate the calculation of small area estimates of population health for ArcGIS and R users, respectively. Table 1 summarizes the main features for RSTbx and RSTr. Overall, both tools simplify the implementation of complicated Bayesian spatial and spatiotemporal models. Both tools also take advantage of recent methodological developments for the CAR model and easily age-standardize estimates. Using census tract-level data from North Carolina and Rhode Island, we demonstrated the benefits of using RSTbx and RSTr to generate and map small area estimates, especially when there are small numbers of events or population sizes. We focused on three key benefits of using these tools: (1) decreased number of geographic units with estimates being suppressed based on reliability criteria, (2) flexibility to set the threshold for reliability, and (3) using credible intervals to identify statistically significant differences between geographic units. In summary, RSTbx and RSTr are powerful tools that can be used to meet the demand for high-quality, local-level data to inform public health programs and tailor health promotion activities to the needs of communities across the country.

XSL•FO
RenderX

Table . Comparison of Rate Stabilizing Toolbox and Rate Stabilizing Tool for R functionality.

	RSTbx[a]	RSTr[b]
Platform	Esri[j] ArcGIS Geoprocessing Tool(s)	R package
Geographic units	Any user-specified geographic unit[d]	Any user-specified geographic unit[d]
Census interface	The CDR[e] is a component of RSTbx that can be used to generate age-stratified or unstratified tables using US Census decennial or American Community Survey data at the county or census tract level	No census interface; users gather any census data of interest on their own. There are R packages to assist with this, eg, tidycensus, censusapi, etc
Age-standardization	Yes	Yes
CAR[f] models[g]	CAR	CAR, MCAR[h], MSTCAR[i]

Input data format	Event/population data:
	• Geodatabase table
	• dBASE (.dbf)
	• Delimited files
	  • Comma-delimited files (.csv, .txt, .asc)
	  • Tab-delimited files (.tsv and .tab)
	  • Pipe-delimited files (.psv)
	
	Boundary file:
	• Geodatabase feature
	• GeoPackage feature
	• GeoParquet
	• Shapefile
		Event/population data: A list of arrays whose dimensions depend on the input data, including vector (UCAR), matrix (MCAR), and three-dimensional array (MSTCAR)
		Adjacency data: A list of vectors which represent the index of their neighboring regions

Additional tools/software needed	ArcGIS Pro	R

Output data format	The following table formats are supported as outputs:
	• Geodatabase table
	• dBASE (.dbf)
	• Delimited files
	  • Comma-delimited files (.csv, .txt, and .asc)
	  • Tab-delimited files (.tsv and .tab)
	  • Pipe-delimited files (.psv)
	
	By default, tables are written to the current ArcGIS project default geodatabase
		Outputs an RSTr model object from which a long table containing estimates, events, populations, CIs, and relative precisions can be extracted. The model object requires additional processing for age-standardization and suppressed estimates

Group aggregation	None	Geographic units, race/ethnicity, sex, year
Reliability calculations	Generated automatically in the output data	Generated automatically in the output data
Probability calculations	Not calculable with the output data	Not directly available, but can be generated by user

[a]RSTbx: Rate Stabilizing Toolbox.

[b]RSTr: Rate Stabilizing Tool for R.

[c]Esri: Environmental Systems Research Institute.

[d]Assumes that the user has numerator and denominator data for the same geographic unit (eg, census tract, county)

[e]CDR: Census Data Retriever.

[f]CAR: conditional autoregressive.

[g]For a description of the CAR models used in RSTbx and RSTr, please refer to the overview section for that specific software.

[h]MCAR: multivariate CAR.

[i]MSTCAR: multivariate spatiotemporal CAR.

Acknowledgments

This research uses data from the North Carolina Center for Health Statistics (SCHS) to illustrate the use of the RSTbx. The authors wish to acknowledge SCHS but are responsible for the analysis and interpretation of these data.

XSL•FO
RenderX

The authors attest that there was no use of generative artificial intelligence (AI) technology in the generation of text, figures, or other informational content of this manuscript. The findings and conclusions in this article are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention. Use of trade names and commercial sources is for identification only and does not imply endorsement by the U.S. Department of Health and Human Services

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this manuscript.

Conflicts of Interest

None declared.

References

1. Wang F. Why public health needs GIS: a methodological overview. Ann GIS 2020;26(1):1-12. [doi: 10.1080/19475683.2019.1702099] [Medline: 32547679]
2. Brissette I, Casper M, Huston SL, et al. Application of geographic information systems to address chronic disease priorities: experiences in state and local health departments. Prev Chronic Dis 2019 May 23;16:E65. [doi: 10.5888/pcd16.180674] [Medline: 31124434]
3. Cromley EK. Using GIS to address epidemiologic research questions. Curr Epidemiol Rep 2019 Jun;6(2):162-173. [doi: 10.1007/s40471-019-00193-6]
4. Greenland S, Mansournia MA, Altman DG. Sparse data bias: a problem hiding in plain sight. BMJ 2016 Apr 27;352:i1981. [doi: 10.1136/bmj.i1981] [Medline: 27121591]
5. Tobler WR. A computer movie simulating urban growth in the Detroit region. Econ Geogr 1970 Jun;46:234-240. [doi: 10.2307/143141]
6. Besag J, York J, Mollié A. Bayesian image restoration, with two applications in spatial statistics. Ann Inst Stat Math 1991 Mar;43(1):1-20. [doi: 10.1007/BF00116466]
7. Louzada F, Ribeiro Diniz CA, Dawodu OA, Egbon OA. Bayesian spatio-temporal modelling: a systematic review. Res Stat 2025 Dec 31;3(1):2568891. [doi: 10.1080/27684520.2025.2568891]
8. Vaughan AS, Quick H, Beck KB, Woodruff RC, DeLara D, Casper M. Applying a Bayesian spatiotemporal model to examine excess county-level cardiovascular disease death rates during the COVID-19 pandemic. Am J Epidemiol 2025 Jun 3;194(6):1556-1565. [doi: 10.1093/aje/kwae330] [Medline: 39218426]
9. Sasmita NR, Arifin M, Kesuma ZM, Mardalena S, Rahayu L, Kruba R. Spatial estimation for tuberculosis relative risk in Aceh Province, Indonesia: a Bayesian conditional autoregressive approach with the Besag-York-Mollie (BYM) model. J Appl Data Sci 2024 May 15;5(2):342-356 [FREE Full text] [doi: 10.47738/jads.v5i2.185]
10. Habtewold FG, Arero BG. Modeling and mapping under-nutrition among under-five children in Ethiopia: a Bayesian spatial analysis. Front Public Health 2025;13:1553908. [doi: 10.3389/fpubh.2025.1553908] [Medline: 40520306]
11. Gelfand AE, Vounatsou P. Proper multivariate conditional autoregressive models for spatial data analysis. Biostatistics 2003 Jan;4(1):11-25. [doi: 10.1093/biostatistics/4.1.11] [Medline: 12925327]
12. Quick H, Waller L, Casper M. Multivariate spatiotemporal modeling of age-specific stroke mortality. Ann Appl Stat 2017 Dec 1;11(4):2165-2177. [doi: 10.1214/17-AOAS1068]
13. Quick H, Song G, Tabb LP. Evaluating the informativeness of the Besag-York-Mollié CAR model. Spat Spatiotemporal Epidemiol 2021 Jun;37:100420. [doi: 10.1016/j.sste.2021.100420] [Medline: 33980402]
14. Song G, Tabb LP, Quick H. Restricted spatial models for the analysis of geographic and racial disparities in the incidence of low birthweight in Pennsylvania. Spat Spatiotemporal Epidemiol 2024 Jun;49:100649. [doi: 10.1016/j.sste.2024.100649] [Medline: 38876562]
15. Quick H, Song G. Reliable event rates for disease mapping. J Off Stat 2024 Jun;40(2):333-347. [doi: 10.1177/0282423x241244917] [Medline: 39364189]
16. CEHI-code-repos/ratestabilizingtoolbox. GitHub Repository. 2025. URL: https://github.com/CEHI-code-repos/RateStabilizingToolbox [accessed 2026-01-19]
17. Quick H, Tootoo J, Li R, et al. The rate stabilizing tool: generating stable local-level measures of chronic disease. Prev Chronic Dis 2019 Mar 28;16:E38. [doi: 10.5888/pcd16.180442] [Medline: 30925140]
18. Klein RJ, Schoenborn CA. Age adjustment using the 2000 projected U.S. population. : U.S. Department of Health and Human Services; 2001 583772012-001 URL: https://www.cdc.gov/nchs/data/statnt/statnt20.pdf [accessed 2026-01-19]
19. Census mapping files. United States Census Bureau. URL: https://www.census.gov/geographies/mapping-files.html [accessed 2026-01-19]
20. Eddelbuettel D. Seamless R and C++ Integration with RCPP: Springer New York; 2013. [doi: 10.1007/978-1-4614-6868-4]

XSL•FO
RenderX

21. Eddelbuettel D, Balamuta JJ. Extending R with C++: a brief introduction to RCPP. Am Stat 2018 Jan 2;72(1):28-36. [doi: 10.1080/00031305.2017.1375990]
22. Eddelbuettel D, François R. RCPP: seamless R and C++ integration. J Stat Soft 2011;40(8). [doi: 10.18637/jss.v040.i08]
23. Sanderson C, Curtin R. Armadillo: a template-based C++ library for linear algebra. J Open Source Softw 2016 Jun 10;1(2):26. [doi: 10.21105/joss.00026]
24. RSTr: Gibbs samplers for bayesian spatiotemporal CAR models. GitHub, Inc. 2025. URL: https://github.com/CEHI-code-repos/RSTr [accessed 2026-01-09]
25. R: The R project for statistical computing. R Core Team. 2024. URL: https://www.R-project.org/ [accessed 2026-01-19]
26. Walker K, Herman M, Eberwein K. Load US census boundary and attribute data as “tidyverse” and ’sf’-ready data frames. The Comprehensive R Archive Network. 2017.
27. Bivand R. R packages for analyzing spatial data: a comparative case study with areal data. Geogr Anal 2022 Jul;54(3):488-518 [FREE Full text] [doi: 10.1111/gean.12319]
28. Bivand RS, Wong DWS. Comparing implementations of global and local indicators of spatial association. TEST 2018 Sep;27(3):716-748. [doi: 10.1007/s11749-018-0599-x]
29. Pebesma E, Bivand R. Spatial Data Science: With Applications in R, 1st edition: Chapman and Hall/CRC; 2023. [doi: 10.1201/9780429459016]
30. Pebesma E. Simple features for R: standardized support for spatial vector data. R J 2018;10(1):439. [doi: 10.32614/RJ-2018-009]
31. Brillinger DR. The natural variability of vital rates and associated statistics. Biometrics 1986 Dec;42(4):693-734. [Medline: 3814721]
32. Vaughan AS, Schieb L, Quick H, Kramer MR, Casper M. Before the here and now: what we can learn from variation in spatiotemporal patterns of changing heart disease mortality by age group, time period, and birth cohort. Soc Sci Med 2018 Nov;217:97-105. [doi: 10.1016/j.socscimed.2018.09.045] [Medline: 30300762]
33. Vaughan AS, Quick H, Schieb L, Kramer MR, Taylor HA, Casper M. Changing rate orders of race-gender heart disease death rates: an exploration of county-level race-gender disparities. SSM Popul Health 2019 Apr;7:100334. [doi: 10.1016/j.ssmph.2018.100334] [Medline: 30581967]
34. Hospital discharge data. State of Rhode Island Department of Health. URL: https://health.ri.gov/data/hospital-discharge-data [accessed 2026-01-19]
35. Suppression of rates and counts. US Centers for Disease Control and Prevention. 2025. URL: https://www.cdc.gov/united-states-cancer-statistics/technical-notes/suppression.html [accessed 2025-07-07]
36. Rhode island small numbers reporting policy. : Rhode Island Department of Health; 2025 URL: https://health.ri.gov/sites/g/files/xkgbur1006/files/2025-02/SmallNumbersReporting.pdf [accessed 2026-01-19]
37. The BUGS project. MRC Biostatistics Unit. 2007. URL: https://www.mrc-bsu.cam.ac.uk/software/bugs-project [accessed 2026-01-19]
38. Stan-dev / stan. GitHub, Inc. 2024. URL: https://github.com/stan-dev/stan [accessed 2026-01-19]
39. Plummer M. Just another Gibbs sampler. SourceForge. 2023. URL: https://mcmc-jags.sourceforge.io/

Abbreviations

ACS: American Community Survey
BYM: Besag, York, and Mollié
CAR: conditional autoregressive
CDR: Census Data Retriever
ICD-10: International Classification of Disease, 10th revision
IDP: Individual Data Processor
MCAR: multivariate CAR
MSTCAR: multivariate spatiotemporal CAR
RST: Rate Stabilizing Tool
RSTbx: Rate Stabilizing Toolbox
RSTr: Rate Stabilizing Tool for R
UCAR: univariate CAR

XSL•FO
RenderX

Edited by A Mavragani; submitted 03.Sep.2025; peer-reviewed by NR Geyer, R Janicki; revised version received 22.Dec.2025; accepted 23.Dec.2025; published 30.Jan.2026.

Please cite as:
DeLara D, Zomorrodi R, Quick H, Tootoo J, Li R, Baker J, Kwon J, Casper M, Vaughan A
Bayesian Models to Generate Small Area Estimates of Population Health: Tutorial for Using Rate Stabilizing Tools and Their Output
JMIR Public Health Surveill 2026;12:e83498
URL: https://publichealth.jmir.org/2026/1/e83498
doi:10.2196/83498

© David DeLara, Ryan Zomorrodi, Harrison Quick, Joshua Tootoo, Ruiyang Li, Justan Baker, Jihyeon Kwon, Michele Casper, Adam Vaughan. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 30.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Early Detection Intervals for Evaluating Event-Based Surveillance System: Reference Dataset Development Study

Yannan Shen[1], PhD; Philip AbdelMalik[2], PhD; Russell J Steele[3], PhD; David L Buckeridge[1], MD, PhD

[1]Department of Epidemiology, Biostatistics and Occupational Health, School of Population and Global Health, Faculty of Medicine and Health Sciences, McGill University, Suite 1200, 2001 McGill College Avenue, Montreal, QC, Canada

[2]Department of Pandemic and Epidemic Intelligence Systems, World Health Organization, Berlin, Germany

[3]Department of Mathematics and Statistics, Faculty of Science, McGill University, Montreal, QC, Canada

Corresponding Author:
David L Buckeridge, MD, PhD
Department of Epidemiology, Biostatistics and Occupational Health, School of Population and Global Health, Faculty of Medicine and Health Sciences, McGill University, Suite 1200, 2001 McGill College Avenue, Montreal, QC, Canada

Abstract

Background: Early detection of health threats is an objective of public health surveillance, and event-based surveillance (EBS) using unstructured information from diverse sources has played an increasingly important role in achieving this objective. However, the evaluation of EBS systems has been hindered by the lack of reference data on outbreak onsets.

Objective: We introduce the concept of an "early detection interval" and create a dataset of these intervals across multiple countries for the epidemic caused by the Omicron variant of SARS-CoV-2.

Methods: We defined the early detection interval as the time between the date of introduction of an infectious agent to a country and the date at which an increase is detectable in traditional public health surveillance data. To determine the date of the introduction of the Omicron variant, we analyzed phylogenetic studies and genome databases. We estimated the end of the interval by applying Bayesian online change point detection to reported COVID-19 case counts. In addition to the early detection intervals, this dataset also contains variables indicating data quality. To further understand the variation in the lengths of the early detection intervals, stratified analysis and univariate Cox proportional hazards were implemented.

Results: This dataset contains early detection intervals for the Omicron variant in 117 countries. The intervals have a median length of 28 (IQR 18-44) days, with a median beginning date of November 27, 2021 (IQR November 17, 2021, to December 12, 2021), and a median ending date of January 2, 2022 (IQR December 19, 2021, to January 9, 2022). Countries with high sequencing availability tend to have earlier start dates with a maximum difference across data sources of only 15 (IQR 7 - 39) days and consequently a prolonged interval length with a median length of 29 (20-47) days. Countries with low incomes were underrepresented in this dataset, with only 12 (29.27%) out of 41 included, and they tended to have shorter intervals with a median duration of 16 (IQR 12 - 23) days. The univariate Cox proportional hazards ratio regression analysis confirmed prolonged interval length in countries with high sequencing availability (hazard ratio 0.59, 95% CI 0.38 - 0.92) and shortened interval length in low-income countries (hazard ratio 2.37, 95% CI 1.29 - 4.36).

Conclusions: The dataset of early detection intervals created in this study can serve as reference data and facilitate the evaluation of the timeliness of alerts generated by EBS systems. Systematic and comprehensive evaluation of EBS is important to guide the development of EBS and motivate the integration of EBS into public health practice. Our study also highlights cross-country disparities in data quality, particularly for genomic evidence, and the need for data collection and sharing focused on low-resource settings.

(JMIR Public Health Surveill 2026;12:e87030) doi:10.2196/87030

KEYWORDS

event-based surveillance; public health surveillance; surveillance; evaluation; infectious disease outbreak; dataset; data quality; detection; COVID-19; SARS-CoV-2

Introduction

Infectious diseases continue to pose a major threat to global health [1,2]. As a recent example, the COVID-19 pandemic resulted in over 777 million cases and 7 million deaths reported as of May 2025 [3]. Early detection of infectious disease outbreaks is important for initiating timely prevention and mitigation measures to minimize health and social impacts [1,2,4]. To improve the timeliness and sensitivity of outbreak detection, public health agencies conduct surveillance using various information sources, such as online news and social

XSL•FO
RenderX

media (known as event-based surveillance; EBS), as a complement to traditional case- and laboratory-based surveillance data (known as indicator-based surveillance; IBS) [5-7].

Evaluating the performance of EBS is crucial to ensure the objective of early detection is achieved, to understand the value of different surveillance methods, and to identify opportunities for performance improvement, but limited studies have comprehensively assessed EBS performance [4,8-12]. In particular, timeliness is one of the most important attributes for evaluating EBS systems aiming at the early detection of outbreaks [4,6,9,13]. Intrinsic timeliness measures the time from the start of an outbreak to an EBS alert [6,13]. Although the start of an outbreak is usually defined as "the earliest epidemiologically linked symptom onset or death," the intrinsic timeliness of EBS alerts was found in several studies to range from 19 days to 3 months after the first retrospectively identified cases [14,15]. However, retrospective analyses of genomic data have revealed that the transmission of viruses among a population can occur weeks before the index case is retrospectively identified, which may provide a better approximation of the ground truth compared to the index case approach, for evaluating EBS timeliness [6,16-20]. For example, phylogenetic estimation has shown that the earliest transmission of the Omicron variant in South Africa occurred over a month before the reporting of the index case [18,21].

Extrinsic timeliness measures the time interval from EBS alerts to the notification or report of the same event from official sources [6,8,11,12,22-24]. In previous evaluation studies, the timeliness of EBS detection compared to official reporting ranged from −1 (ie, 1 day lagged behind) to 20 days, with results varying by disease, EBS system, and geographic region [8,11,23,24]. Official reports are generally well documented and relatively easy to access but are subject to reporting delays, which may lead to overestimating extrinsic timeliness. Meanwhile, the interpretation of extrinsic timeliness alone can be ambiguous, as alerts occurring much earlier than the official report may not be related to the corresponding outbreak. Additionally, a few studies use a fixed time window to define timely EBS detection. For example, Ganser et al [10] categorized EBS detections as timely if they occurred within a fixed, 4-week time window and found that only 9.2% of seasonal influenza outbreaks were detected in a timely manner by EBS systems. However, such a fixed time is arbitrary, even with some justification based on disease-specific knowledge, and, more importantly, ignores possible different contexts across countries, which may consequently lead to the misclassification of alerts.

In this study, we develop a dataset of intervals based on genomic evidence and IBS data that can serve as a reference framework for evaluating the timeliness of EBS, allowing the assessment of the potential contribution of EBS to early outbreak detection, prior to detection through IBS. We introduce the concept of an early detection interval, which is the time between the earliest evidence of pathogen introduction in a population and the earliest signals from routine analysis of traditional IBS data. Specifically, using the example of the Omicron variant, we build a dataset of early detection intervals that begin with the earliest introduction of the Omicron variant based on genomic evidence and end with the detected increase in reported COVID-19 case counts. We present a descriptive analysis of this dataset to summarize the temporal distribution of the start and end of the early detection intervals and explore the country-level factors associated with the length of the early detection interval using Cox proportional hazards regression models.

Methods

Study Design

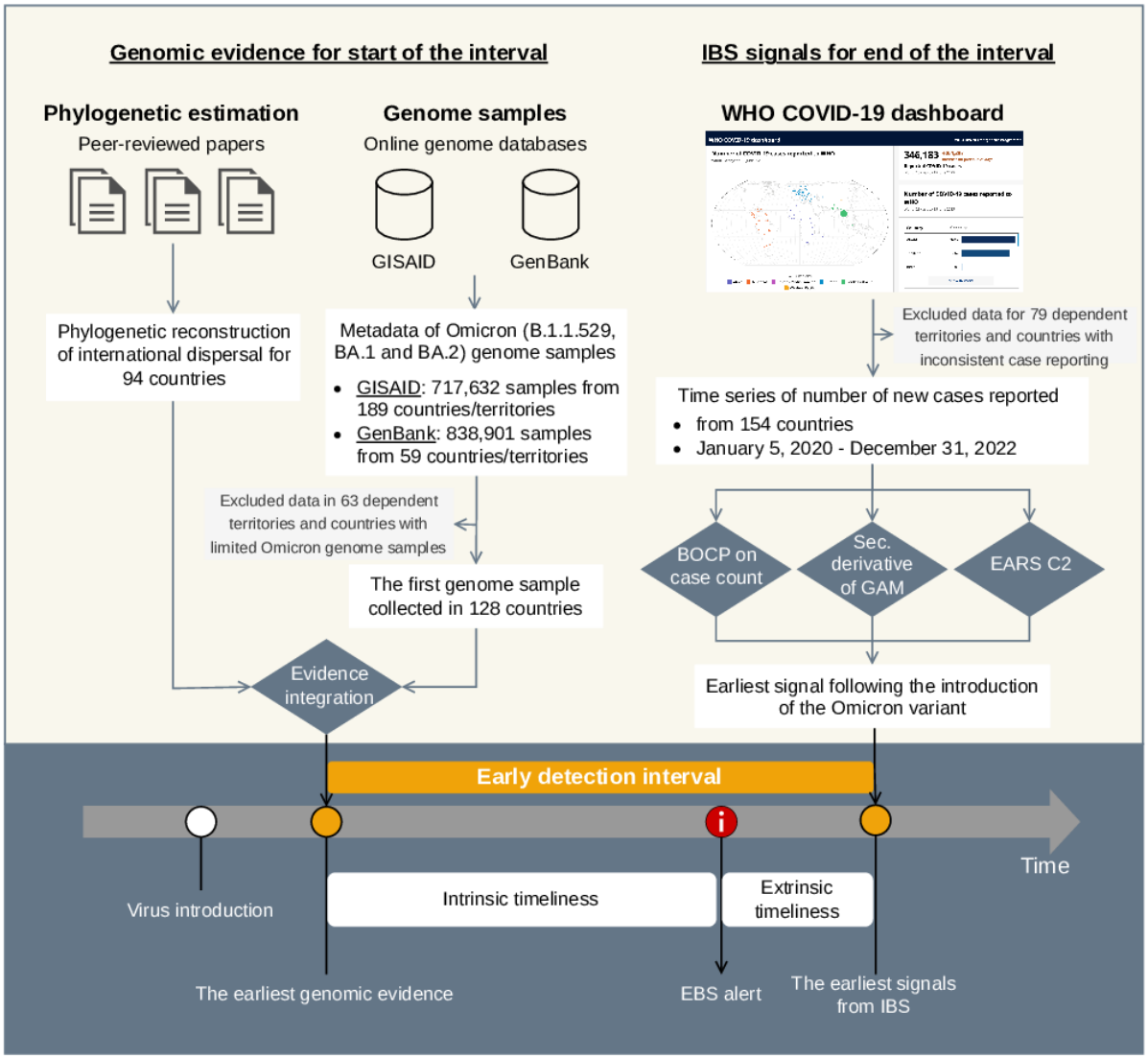
In this study, we created a dataset of early detection intervals for the earliest subvariants of the Omicron variant of SARS-CoV-2 (ie, BA.1 and BA.2) in countries worldwide from the end of 2021 to the beginning of 2022. The early detection interval was defined as the time from the earliest genetic evidence of the Omicron variant's introduction among the population in a country to the detectable increase in reported COVID-19 incidence cases.

To identify the beginning and end of the early detection interval in different countries, we conducted a retrospective analysis by compiling and linking genomic evidence from published papers and 2 online genome databases and detecting change points from COVID-19 incidence cases reported in different countries, as shown in Figure 1. In addition to the early detection interval, we also created supplementary variables to support users in assessing the data quality and potential biases of the early detection intervals.

We included countries with at least 1 data source for identifying the beginning of the interval and available case data. We excluded dependent territories from our study as our analysis was at the country level. The dependent territories tended to have different transmission patterns and capacities for public health surveillance and reporting compared to their sovereign states, while the available data sources had limited granularity to reflect within-country variation. To ensure the reliability of the interval, we also excluded countries with limited genome surveillance, data accessibility, and unreliable case reporting, as described in the following subsections.

XSL•FO
RenderX

Figure 1. Illustration of the study design for creating the early detection intervals for the Omicron variant in countries worldwide, including the data sources, data inclusion and exclusion, and analytical methods used for generating the beginning and end of early detection intervals. BOCP: Bayesian online change point; EBS: event-based surveillance; EARS: Early Aberration Reporting Systems; GAM: generalized additive model; GISAID: Global Initiative on Sharing All Influenza Data; IBS: indicator-based surveillance.



The Start of the Early Detection Intervals

The early detection interval in a country should begin when a pathogen (ie, the Omicron variant in this study) is first introduced to and begins circulating among the population. Among the multiple subvariants of the Omicron variant, we considered BA.1 and BA.2 only, as they are the earliest subvariants emerging at a similar time. Despite their genetic and epidemiological differences, both BA.1 and BA.2 have high transmissibility contributing to a new wave of COVID-19 cases in different countries. However, this ground truth of virus introduction and the start of transmission among the population is generally not possible to capture exactly, given the limited sample collection and sequencing capacity in most countries. Therefore, our definition of the start of an early detection interval was the earliest date of genomic evidence of the circulation of either the BA.1 or BA.2 Omicron variant in a country.

The primary data source for such genomic evidence is published studies estimating the initial introduction of the Omicron variant in different countries using phylogenetic reconstruction methods (Table S1 in Multimedia Appendix 1) [18,25-30]. One of the main studies we used was conducted by Tegally et al [25], which provides the reconstruction of the international dispersal patterns of variants of concern from Alpha to Omicron inferred from a representative sample of genome sequences using ancestral state reconstruction of discrete spatial locations for each variant of concern. From this study, we obtained global dispersal data, including the origin, destination, and date of international dispersal, estimated by 10 repeated phylogenetic analyses. We identified the earliest introduction of the Omicron variant in a destination country as the median of the 10 earliest international dispersals estimated. There are two countries, Bulgaria and Romania, where the median of the earliest exportation occurred a few days earlier than that of the earliest introduction. In this case, we modified the earliest introduction to the lower quartile

XSL•FO
RenderX

of its estimation to ensure that the introduction occurred before exportation and to avoid confusion.

As many countries are not covered in the published studies identified, we also used the first collection date of Omicron genome samples to define the beginning of the interval. We collected the metadata of genome samples from 2 of the most popular genome databases that host large numbers of genome data on SARS-CoV-2 and have been widely used by researchers worldwide, the Global Initiative on Sharing All Influenza Data (GISAID) and GenBank hosted and maintained by the Federal Ministry of Food and Agriculture in Germany and the National Institutes of Health in the United States, respectively [31-33]. Both databases rely on the voluntary submission of genome data from laboratories worldwide with unknown sampling processes. Consequently, the two databases have different geographic coverage and focus despite a large degree of overlap. We tried to mitigate these biases to some extent by using these databases together so that we could identify the earliest possible collection of Omicron genome samples. Using the filter function of each platform, we selected Omicron genome samples collected after October 1, 2021, and obtained their metadata. Based on the date of sample collection in the metadata, we identified the first Omicron genome sample collected from human or environmental samples (eg, wastewater) in different countries.

For each country, we matched phylogenetic estimates from published studies with the earliest Omicron genome collection dates from GISAID and GenBank and then used the earliest of these as the beginning of early detection intervals. To indicate the quality of the data sources and the reliability of the genome evidence, we documented the number of data sources available for defining the beginning of an interval and computed the minimal and maximum temporal gap of the earliest evidence among these data sources. Meanwhile, for each genome database, countries were marked as having questionable data quality if a limited number of Omicron genome samples (ie, lower than 10% percentile in a database) or all samples were submitted on the same date, which might indicate a lack of sufficient and consistent sequencing and reporting efforts in these countries to provide reliable genome evidence. We excluded countries if they had evidence from only 1 genome database, and their data quality in this genome database was marked as questionable.

The End of the Early Detection Intervals

The end of the early detection interval was defined as the earliest alert of a potential outbreak captured by traditional IBS. We assumed that alerts from other surveillance systems would add little value to early detection after this date. The number of new cases indexed by time and geolocation is one of the most commonly used indicators for early outbreak detection in IBS, as case counts are a timelier indicator than other metrics, such as the frequency of hospitalizations and deaths. We collected the weekly number of new COVID-19 cases reported from January 5, 2020, to December 31, 2022, from the World Health Organization (WHO) COVID-19 dashboard, which uses the date of reporting (rather than the date of symptom onset) for the WHO definition of a confirmed case of SARS-CoV-2 infection: “(a) a person with a positive Nucleic Acid Amplification Test, regardless of clinical criteria OR epidemiological criteria; (b) a person meeting clinical criteria AND/OR epidemiological criteria with a positive professional-use or self-test SARS-CoV-2 Antigen rapid diagnostic test” [3]. We excluded countries when the first case was reported later than May 31, 2020 (ie, no COVID-19 cases have been reported for nearly 3 months after the announcement of the global pandemic, which is unlikely to reflect transmission dynamics but reporting errors) or when no incident cases were reported in over one-third of the interval, both of which suggest inconsistent or delayed case identification and public health surveillance reporting.

We applied the Bayesian online change point (BOCP) detection model, the second derivative of the generalized additive model (GAM), and the Early Aberration Reporting Systems (EARS) C2 method to the entire time series of the number of new COVID-19 cases reported [34-40]. The change point detection methods have been applied to detect the beginning and end of outbreaks, and both models selected are reported to have good performance in some previous studies [10,35,36,41]. The BOCP algorithm assumed that the differenced case count could be divided into nonoverlapping partitions, and within each partition, it followed a Gaussian distribution. For countries with low case counts reported (ie, a median weekly count lower than 100), a BOCP model assuming a Poisson distribution was used. The differenced case count was used to remove the autocorrelation and stabilize the data. The BOCP analysis was implemented using the R package ocp (version 0.1.1). The change points were selected based on the global maximum of the posterior distribution of all change points. We also detected the change points of the growth rate of the case count using a GAM, assuming a negative binomial distribution with a log link. In the GAM, a single explanatory variable of time in weeks was included and modeled using thin plate splines. We also implemented a simple aberration detection method, that is, the EARS C2 method, as a baseline method for the BOCP and GAM models, using the R package surveillance (version 1.24.1) [38-40]. The EARS C2 method compares the observed case count with the most recent baseline, calculated as the moving average of the case count reported 3 to 5 weeks before the observed case count.

We identified the end of the early detection interval as the earliest change point detected by any of the models that occurred between the 7 days after the introduction of the Omicron variant (ie, the beginning of the interval) and the date of the maximum number of case counts reported within the subsequent 120 days. By doing so, we enforced a minimal interval length of 7 days, which is approximately 2 incubation periods of Omicron. We also recorded (1) the number of change point detection models that agreed on the selected point and (2) the temporal gap from the selected point to the following peak of case counts (Table 1). These measures provide future users with information about the reliability of the estimated interval end.

XSL•FO
RenderX

Table . The definitions of variables in the dataset of early detection intervals, including the beginning and end of the early intervals and data quality variables.

Variable	Working definition
Define the interval	
Country	Country where the transmission activity of the Omicron variant was taking place.
Date when the interval starts	The date when the earliest genomic evidence that indicates the dissemination of the Omicron variant among the population was presented, considering data from phylogenetic estimation and the earliest genome samples collected and submitted to GISAID[a] and GenBank.
Date interval end	The date when the earliest increasing change point was detected from the case count time series by the highest number of models during the time between the 7 days after the beginning of the interval and the following peak of incidence case count reported within 120 days.
Length of interval	The number of days from the beginning to the end of an early detection interval.
Data quality for interval start	
Data source for interval start	In a country, the data source, among phylogenetic estimation, GISAID, and GenBank, that provides the earliest genomic evidence.
Number of data sources available for identifying the start of interval	In a country, the number of data sources, among the phylogenetic estimation, GISAID, or GenBank, that are available.
The minimal temporal gap across different data sources	In a country, if more than 2 data sources are available, the minimum difference in the time when the earliest evidence of transmission of the Omicron variant across data sources.
The maximum temporal gap across different data sources	In a country, if more than 2 data sources are available, the maximum difference in the time when the earliest evidence of transmission of the Omicron variant across data sources.
Data quality for interval end	
Completeness score	In a country, the proportion of weeks with new cases reported in the case count dataset.
Number of change point detection models that agree on the selected end point	In a country, the number of the change point detection models, including BOCP[b] on differenced count, second derivative of GAM[c] and EARS[d] C2, detected the selected end point.
Temporal gap from the interval end to the following peak of case count	For a country, the number of days from the end of the early detection interval to the following peak of incidence case count reported within 120 days.

[a]GISAID: Global Initiative on Sharing All Influenza Data.

[b]BOCP: Bayesian online change point.

[c]GAM: generalized additive model.

[d]EARS: Early Aberration Reporting Systems.

Statistical Analysis

We computed the median and IQR of the timing of the start and end of the interval and length of the early detection intervals for all included countries and by strata of countries based on the WHO regions and income levels (ie, low-income countries [LICs], lower-middle-income countries, upper-middle-income countries, and high-income countries [HICs]), based on the World Bank's classification of the economies and public availability of genomic sequencing of countries (ie, low, moderate, and high availability based on a previous study) [42]. For levels of sequencing availability, we merged countries with low and moderate availability into 1 group to ensure sufficient observations for stable and reliable estimates. We compared the median of the length of intervals in different strata using the Wilcoxon rank-sum test (for 2-level strata) and the Kruskal-Wallis test (for multilevel strata). We also used univariate Cox proportional hazards regression models to assess factors associated with the hazard of the end of the interval. With the interval length as the outcome, the factors explored included income level (LICs vs non-LICs), sequencing availability, and indicators of data quality for genomic evidence, such as the number of data sources available and the maximum difference in the timing for the earliest genomic evidence across data sources. We verified if the proportional hazards assumption held by testing the correlation of scaled Schoenfeld residuals with time.

XSL•FO
RenderX

Ethical Considerations

The original data in this study were aggregated and nonidentified, which are publicly accessible from the WHO COVID-19 dashboard [3], GenBank [33], and GISAID [31,32]. Thus, ethical approval and participant consent were not required for this study.

Results

Interval Characteristics

In this study, we collected phylogenetic estimates of viral introduction for 94 countries; metadata of 717,632 and 838,901 Omicron variant genome samples collected in 189 and 59 countries and territories from GISAID and GenBank, respectively; and case count data for 233 countries and territories from the WHO COVID-19 dashboard. After applying the inclusion and exclusion criteria, 117 countries were included in our study, with 46 out of 70 (65.71%) HICs and 12 out of 41 (29.27%) LICs included. The early detection intervals began on a median date of November 27, 2021 (IQR November 17, 2021, to December 12, 2021) and ended on a median date of January 2, 2022 (IQR December 19, 2021, to January 9, 2022), with a median length of 28 (IQR 18-44) days (Table 2).

XSL•FO
RenderX

Table . Distribution of the intervals for the early detection of the Omicron variant and data quality variables for 117 countries included and different groups of countries.

	Number of countries	Beginning of interval, median (IQR)	End of interval, median (IQR)	Interval length (days), median (IQR)	Max difference across data sources (days), median (IQR)	Gap to the peak of case count (days), median (IQR)
Total	117	November 27, 2021 (November 17, 2021, to December 12, 2021)	January 2, 2022 (December 19, 2021, to January 9, 2022)	28 (18-44)	16 (8-41)	28 (14-42)
Sequencing availability						
High	83	November 22, 2021 (November 16, 2021, to December 6, 2021)	December 26, 2021 (December 15, 2021, to January 9, 2022)	29 (20-47)	15 (7-39)	21 (14-28)
Low to moderate	27	December 12, 2021 (November 26, 2021, to December 16, 2021)	January 2, 2022 (December 12, 2021, to January 9, 2022)	23 (13-34)	27 (16-50)	28 (21-49)
Country income group						
LICs[a]	12	November 29, 2021 (November 23, 2021, to December 14, 2021)	December 19, 2021 (December 10, 2021, to December 27, 2021)	16 (12-23)	23 (17-43)	14 (12-16)
LMICs[b]	25	November 30, 2021 (November 13, 2021, to December 17, 2021)	January 2, 2022 (December 12, 2021, to January 16, 2022)	30 (23-43)	22 (13-54)	21 (14-28)
UMICs[c]	34	November 27, 2021 (November 19, 2021, to December 12, 2021)	January 2, 2022 (December 26, 2021, to January 9, 2022)	33 (21-47)	29 (13-58)	28 (21-35)
HICs[d]	46	November 22, 2021 (November 16, 2021, to December 7, 2021)	December 26, 2021 (December 5, 2021, to January 2, 2022)	27 (18-41)	8 (4-26)	35 (21-54)
WHO regions						
African Region	21	November 21, 2021 (October 14, 2021, to November 30, 2021)	December 12, 2021 (December 5, 2021, to December 19, 2021)	21 (12-47)	33 (16-59)	14 (14-28)
Region of the Americas	19	December 2, 2021 (November 20, 2021, to December 16, 2021)	January 2, 2022 (December 22, 2021, to January 5, 2022)	27 (19-37)	16 (11-30)	28 (18-32)
Eastern Mediterranean Region	16	December 1, 2021 (November 25, 2021, to December 9, 2021)	December 26, 2021 (December 19, 2021, to January 10, 2022)	19 (11-31)	26 (5-46)	25 (19-42)
European Region	42	November 22, 2021 (November 15, 2021, to December 12, 2021)	January 2, 2022 (December 19, 2021, to January 9, 2022)	28 (18-45)	11 (5-26)	28 (21-47)
South-East Asia Region	7	November 27, 2021 (November 24, 2021, to November 28, 2021)	January 2, 2022 (January 2, 2022, to January 12, 2022)	44 (36-46)	26 (8-42)	28 (25-46)

XSL•FO
RenderX

	Number of countries	Beginning of interval, median (IQR)	End of interval, median (IQR)	Interval length (days), median (IQR)	Max difference across data sources (days), median (IQR)	Gap to the peak of case count (days), median (IQR)
Western Pacific Region	12	December 4, 2021 (November 19, 2021, to December 17, 2021)	January 16, 2022 (January 2, 2022, to January 23, 2022)	37 (29-52)	32 (12-57)	35 (19-53)

[a]LIC: low-income country.

[b]LMIC: lower-middle-income country.

[c]UMIC: upper-middle-income country.

[d]HIC: high-income country.

When defining the beginning of an interval, 33 out of 117 (28.21%) countries were based on the phylogenetic estimation, whereas 84 out of 117 (71.79%) countries were based on the date of the first collection of the Omicron genome sample, mainly from GISAID. The phylogenetic estimations were, on average, 7 (IQR 4 - 11) days earlier than the first collection in GISAID for 35 countries and 24 (IQR 14 - 39) days earlier than the first collection in GenBank for 20 countries. Among 100 out of 117 (85.47%) countries with more than 1 data source to define the beginning of an interval, the maximum difference across different data sources was a median of 16 (IQR 8 - 41) days, and the minimum value for this difference was a median of 13 (IQR 4 - 32) days. Meanwhile, when defining the end of an interval, the completeness score of the COVID-19 case count data for the included countries had a median of 94.23% (IQR 92.31% - 94.87%). Most countries (101/117, 86.32%) had the end of their interval detected by either the BOCP model or the GAM model, and in only 16 out of 117 (13.68%) countries, the EARS C2 method detected an increase in case counts prior to detection by the BOCP or GAM models. The number of new COVID-19 cases reported peaked after a median of 18 (IQR 14 - 42) days after the end of the interval.

Factors Associated with the Length of the Early Detection Interval

In a stratified analysis, countries with high sequencing availability had the smallest range in earliest transmission evidence across 3 data sources and a lead of 12 (95% CI 5 - 21) days at the beginning of the interval, compared with countries with moderate or low sequencing availability (Table S2 and Figure S1 in Multimedia Appendices 2 and 3, respectively). Conversely, the intervals began on similar dates across countries with various income levels and ended 7 (95% CI −0.00004 to 21) days earlier in LICs compared to non-LICs, resulting in intervals that were a median of 11 (95% CI 3 - 20) days shorter. However, the case count peaked about 2 weeks after the end of intervals in LICs, which was 14 (95% CI 7 - 21) days shorter than in non-LICs. Meanwhile, in HICs, the case count peaked at a median of 35 (IQR 21 - 54) days after the end of intervals. HICs also had consistent genomic evidence for the introduction of the Omicron variant, with the maximum gap across different sources of 8 (IQR 4 - 26) days, which was 12 (95% CI 5 - 20) days shorter than that of other countries. Across different WHO regions, the earliest signals from case counts occurred first in the African Region on December 19, 2021 (IQR December 10, 2021 to December 27, 2021), which was about 2 weeks before the peak of new COVID-19 cases occurred, while this signal was the latest in the Western Pacific Region, which was 1 month after the same signal in the African Region and 5 weeks before the peak of COVID-19 cases in the Western Pacific Region.

The univariate Cox proportional hazards ratio regression analysis confirmed prolonged interval length in countries with high sequencing availability, with a hazard ratio of 0.59 (95% CI 0.38 - 0.92), and shortened interval length in LICs, with a hazard ratio of 2.37 (95% CI 1.29 - 4.36; Table 3). The other 2 variables, the number of data sources available and the maximum difference across different data sources for defining the beginning of intervals, were significantly associated with prolonged intervals with an estimated hazard ratio below 1, both of which indicated the reliability of the data quality of the beginning of intervals.

XSL•FO
RenderX

Table . The estimated hazard ratio and its 95% CI from univariate Cox proportional hazards regression analysis of the lengths of the early detection interval in 117 countries included.

Predictor	Number of events	Hazard ratio (95% CI)	*P* value
Maximum difference across data sources for interval start	100	0.98 (0.97-0.99)	.004
Number of data sources available for interval start	117	0.65 (0.48, 0.88)	.005
Country income group			
Non-LICs[a]	105	Reference	
LICs	12	2.37 (1.29-4.36)	.006
Sequencing availability			
Low to moderate	27	Reference	
High	83	0.59 (0.38-0.92)	.02

[a]LIC: low-income country.

Discussion

Principal Findings

In this study, we defined an early detection interval as the time between the date of the earliest genomic evidence of the presence of a pathogen in a country and the date of the first alert from an IBS system, which is conceptually tailored to support evaluating early outbreak alerts with well-defined boundaries. We created a dataset of such intervals for the early detection of the Omicron variant in 117 countries worldwide with a median start date of November 27, 2021 (IQR November 17, 2021, to December 12, 2021) and an end date of January 2, 2022 (IQR December 19, 2021, to January 9, 2022) [43]. We also observed interval lengths ranging from 7 to 66 days across countries, which may reflect the influence of factors such as the capacity for sequencing and data sharing in genomic surveillance, the capacity for case identification and reporting of IBS systems, and the transmission dynamics of the Omicron variant. This dataset can serve as a reference dataset in future evaluation studies to evaluate both intrinsic and extrinsic timeliness of EBS.

The early detection interval begins, in theory, with the initial introduction of the Omicron variant into a population, as only then can the variant cause excess infections and be detected in the population. Yet, the timing of introduction is difficult to measure consistently across countries with existing data, so we approximated this date using the earliest genomic evidence of the presence of the variant, retrospectively identified. Retrospective analysis of genomic data has been widely adopted to understand transmission dynamics and the impact of virus variants, and it can generate evidence uncovering the transmission of specific substrains of viruses such as SARS-CoV-2 among a population, in some cases weeks or even months before other data sources [16-20]. High-quality genomic evidence that closely approximates the introduction of the Omicron variant allows capturing the contribution of early EBS alerts. When used as a reference dataset in evaluation studies, this early detection interval can avoid the underestimation of the intrinsic timeliness of EBS alerts compared to previous studies that have used the first retrospectively identified case to approximate the beginning of an outbreak [14,15].

We found that countries with high sequencing availability and income levels tended to have concordant genomic evidence across data sources (ie, an indicator of the data quality of genomic evidence we compiled) and therefore a relatively earlier start of the intervals than those of other countries. Genomic data may have uneven quality across countries due to differences in sequencing capacities, sampling strategies, and data reporting and sharing, which consequently translate into biases in the measurement of the start of intervals [16,42]. We integrated several data sources to identify the earliest possible genomic evidence for the introduction and transmission of the Omicron variant. We found that integrating multiple sources for genomic evidence may help reduce disparities in early detection, compared to using a single database. For example, the phylogenetic estimations were earlier than the first collection in GISAID in 35 countries, with a median of 7 (IQR 4 - 11) days of lead time. Specifically, South Africa had its retrospectively identified first collection of the Omicron genome sample 28 days before the index case reported [21], yet 11 days after the phylogenetic estimation of the earliest introduction.

The early detection interval ended with the detection of the earliest alerts from IBS data. Using this end of early detection intervals as a reference, we can assess the extrinsic timeliness that highlights the added value of EBS in comparison with the IBS alerts and the potential for accelerated response. Unlike several studies that compared EBS alerts with the first official reporting [6,8,11,44], this measurement also avoids overestimating the contribution of EBS alerts that occurred after the earliest IBS alerts. Using 3 different methods to detect signals from IBS data, we found that the EARS C2 method was likely to generate detections later than BOCP and GAM and missed the gradual and subtle increases, which aligns with findings in previous studies showing EARS methods to be inferior to regression-based methods for aberration detection [45,46]. We also found that the temporal distribution of the end of intervals mainly varied across countries in different geographic regions, aligning with the global dispersal pattern of the Omicron variant [25]. The reported case increase,

XSL•FO
RenderX

however, may lag behind the increase in actual infections due to asymptomatic cases. Such a lag may be more evident in countries that were first impacted by the Omicron variant due to a lack of awareness before official identification [47,48]. A delayed end may cause the overestimation of the interval length and, when combined with a delayed start, shift the early detection interval later in time, both contributing to an overestimation of EBS's extrinsic timeliness.

As for the length of early detection intervals, while it is not a measurement of the timeliness of IBS, the median length of approximately 1 month indicates the potential for improvement in early outbreak detection of a newly emerged variant of SARS-CoV-2. Our findings also demonstrate the substantial variation in the lengths of early detection intervals across countries. Our analysis mainly revealed the association between varying lengths and genomic data quality. Given this variation in data quality, the interval lengths should be interpreted with caution, considering potential biases. For example, shorter interval lengths observed in LICs compared to those in non-LICs in our study may indicate extensive underestimation of the introduction of the Omicron variant due to limited genomic surveillance, rather than a lower value of adopting EBS.

Limitations

This study is subject to several limitations. First, the countries included in this dataset are not a representative sample. This dataset includes 117 countries, yet is overly represented by HICs, with 46 out of 70 (65.7%) HICs included. Conversely, less than half of the LICs (12/31, 38.7% countries) and countries in the African Region (21/51, 41.2% countries), the Region of the United States (19/54, 35.2% countries), and the Western Pacific Region (12/41, 29.3% countries) were included due to limited data availability. Therefore, although EBS generated timely signals in LIC included in our study, caution is warranted in generalizing our findings to all LICs. There is also a need for additional research to generate the data that would allow this reference dataset to be expanded to a larger and more representative sample of LICs.

Second, as we discussed earlier in this section, our measurements of both the start and end of the intervals are likely to be biased toward later in time, which can further bias the lengths of early detection intervals and the Cox analysis of the interval lengths. To minimize such biases, we designed our inclusion and exclusion criteria and compiled genomic evidence from several data sources. We also included several supporting variables alongside the early detection interval to support potential users in assessing the data quality and corresponding biases. However, the data quality variables may not fully reveal the potential biases, and we could not include additional quality metrics due to the limited metadata available for some sources. For example, the data collection and sampling processes and representativeness of both genomic data from GISAID and GenBank and reported incidence cases are unknown but very likely to be different across countries and be uneven within a country [16,49]. Moreover, although the models selected in our study are reported to have good performance in some previous studies [10,35,36,41], further evaluation using simulation data could help us to understand the performance of these models and identify important factors that could impact the accuracy of the measurement of early detection intervals. Therefore, the timeliness metrics measured using these early detection intervals should be interpreted in the context of each country and carefully considering measured and unmeasured biases.

Third, in terms of generalizability, the concept of the early detection interval can be applied to other infectious diseases for evaluating the timeliness of detection. However, it requires corresponding data of relatively high quality to be available. As the early detection intervals are designed to evaluate the timeliness of EBS, a substantial delay in case identification that is hard to resolve by improving practice within IBS will likely increase the value of EBS for early outbreak detection. Therefore, this early detection interval may also apply to some vector-borne diseases, such as dengue virus disease and Zika virus disease, as well as avian influenza. Early indication of outbreaks of these diseases can be observed from environmental, zoonotic, entomologic, and genomic data, while asymptomatic infections make early detection through IBS challenging. When generalizing this early detection interval to other contexts, the definition and measurement of the interval should be adapted to the study context, which in turn would alter the interpretation of the evaluation results. Conversely, this approach might not work well for diseases such as Ebola because of the small number of cases and the relatively low delay in case identification due to severe and abrupt symptoms.

Finally, while the dataset of early detection intervals can enable the evaluation of the added value of EBS relative to the timeliness of outbreak detection, it would also be informative if timeliness metrics could consider other outbreak milestones, for example, verification of a signal and the initiation of interventions [9,50,51]. In our study, we were unable to include these milestones in our dataset due to limited data available across countries. To allow the extension of evaluation to these milestones, public health agencies should document, store, and publish the data on the verification of EBS and IBS signals and on the initiation of interventions. Routine and systematic capture of these data would provide reference data to expand evaluation to allow a more comprehensive and accurate quantification of the impacts of EBS in accelerating response [9,13,51,52]. It should also be noted that this definition of early detection interval may exclude legitimate early EBS signals (eg, reports of unusual clusters prior to sequencing confirmation), which could underestimate the full potential contribution of EBS. In addition to timeliness, other attributes, such as sensitivity, precision, usefulness, representativeness, and flexibility, are also important for comprehensively evaluating the early outbreak detection capacity of EBS [4,6,9]. For instance, the early detection interval focuses on the temporal adjacency of EBS alerts to the target event yet may not reflect their geographic coverage and representativeness. Further analysis is needed to measure these attributes in parallel using quantitative and qualitative analyses [6,9,11].

Conclusion

Our study created a dataset of early detection intervals for the Omicron variant in countries worldwide by retrospectively compiling and analyzing data from multiple sources. Using this

XSL•FO
RenderX

dataset of early detection intervals as the reference data enables future evaluation studies to assess both intrinsic and extrinsic timeliness of EBS. Our study also highlights cross-country disparities in data quality, particularly for genomic evidence and the need for data collection and sharing, focusing on low-resource settings. While many EBS systems have been developed and improved over the past few decades, the dearth of systematic and structured evaluation has limited the integration of EBS systems into public health surveillance practice. The early detection interval dataset presented in this study addresses important barriers of the lack of reference data and enables evaluation studies to generate valuable evidence on the added value of EBS compared to traditional IBS and the contribution of EBS systems with different designs. This evidence should guide the development of EBS and motivate acceptance and integration of EBS into public health practice, which is key to achieving the goal of timely outbreak detection and response and building collaborative public health surveillance.

Acknowledgments

We gratefully acknowledge all data contributors, that is, the authors and their originating laboratories responsible for obtaining the specimens and their submitting laboratories for generating the genetic sequence and metadata and sharing via the Global Initiative on Sharing All Influenza Data Initiative, on which this research is based. The authors confirm that no generative artificial intelligence tools were used in the writing or preparation of this manuscript.

Funding

This work was supported by Canadian Institutes of Health Research and UK Research and Innovation under project EPI-AI: Automated Understanding and Alerting of Disease Outbreaks from Global News Media (Canadian Institutes of Health Research project code: UAI-169305 and UK Research and Innovation grant number ES/T012277/1). YS is grateful for financial support from the Fonds de recherche du Québec–Santé (FRQS, Doctoral training award 313602). DLB is supported by a Canada Research Chair in Population Health Informatics (950-232679).

Data Availability

To ensure accessibility and transparency, we have adhered to the Findable, Accessible, Interoperable, and Reusable data principles throughout the creation and curation of the dataset. The data and associated metadata are available via Borealis Dataverse, ensuring that researchers from diverse disciplines can easily access and use the resource [43].

Authors' Contributions

Conceptualization: DLB, PA, YS
Data curation: YS
Formal analysis: YS
Investigation: DLB, YS
Methodology: PA, RJS, YS
Project administration: DLB, YS
Resources: DLB
Supervision: DLB, PA, RJS
Validation: DLB, PA
Visualization: DB, YS
Writing – original draft: DLB, YS
Writing – review & editing: DLB, PA, RJS, YS

Conflicts of Interest

None declared.

Multimedia Appendix 1
List of literature reporting the phylogenetic estimations of the introduction of the Omicron variant in countries.
[DOCX File, 15 KB - publichealth_v12i1e87030_app1.docx]

Multimedia Appendix 2
List of countries included in the dataset of early detection intervals for the Omicron variant.
[XLS File, 37 KB - publichealth_v12i1e87030_app2.xls]

Multimedia Appendix 3

XSL•FO
RenderX

Temporal distribution of the beginning and end of early detection intervals.
[DOCX File, 37 KB - publichealth_v12i1e87030_app3.docx]

References

1. Sachs JD, Karim SSA, Aknin L, et al. The Lancet Commission on lessons for the future from the COVID-19 pandemic. Lancet 2022 Oct 8;400(10359):1224-1280. [doi: 10.1016/S0140-6736(22)01585-9] [Medline: 36115368]
2. Baker RE, Mahmud AS, Miller IF, et al. Infectious disease in an era of global change. Nat Rev Microbiol 2022 Apr;20(4):193-205. [doi: 10.1038/s41579-021-00639-z] [Medline: 34646006]
3. WHO COVID-19 dashboard. World Health Organization. URL: https://data.who.int/dashboards/covid19/summary [accessed 2025-06-16]
4. Groseclose SL, Buckeridge DL. Public health surveillance systems: recent advances in their use and evaluation. Annu Rev Public Health 2017 Mar 20;38(1):57-79. [doi: 10.1146/annurev-publhealth-031816-044348] [Medline: 27992726]
5. Balajee SA, Salyer SJ, Greene-Cramer B, Sadek M, Mounts AW. The practice of event-based surveillance: concept and methods. Glob Secur Health Sci Policy 2021 Jan 1;6(1):1-9. [doi: 10.1080/23779497.2020.1848444]
6. Early detection, assessment and response to acute public health events: implementation of early warning and response with a focus on event-based surveillance: interim version. : World Health Organization; 2014 URL: https://iris.who.int/server/api/core/bitstreams/e68010ea-66d9-428a-ac4f-f0a309de81e7/content [accessed 2026-04-07]
7. International health regulations. : World Health Organization; 2005 URL: https://iris.who.int/server/api/core/bitstreams/1632f577-e897-4da2-8d4b-d0d37751b65c/content [accessed 2026-04-07]
8. Williams GS, Koua EL, Abdelmalik P, et al. Evaluation of the epidemic intelligence from open sources (EIOS) system for the early detection of outbreaks and health emergencies in the African region. BMC Public Health 2025 Mar 4;25(1):857. [doi: 10.1186/s12889-025-21998-9] [Medline: 40038598]
9. Crawley AW, Mercy K, Shivji S, et al. An indicator framework for the monitoring and evaluation of event-based surveillance systems. Lancet Glob Health 2024 Apr;12(4):e707-e711. [doi: 10.1016/S2214-109X(24)00034-2] [Medline: 38364834]
10. Ganser I, Thiébaut R, Buckeridge DL. Global variations in event-based surveillance for disease outbreak detection: time series analysis. JMIR Public Health Surveill 2022 Oct 31;8(10):e36211. [doi: 10.2196/36211] [Medline: 36315218]
11. Barboza P, Vaillant L, Mawudeku A, et al. Evaluation of epidemic intelligence systems integrated in the early alerting and reporting project for the detection of A/H5N1 influenza events. PLOS ONE 2013;8(3):e57252. [doi: 10.1371/journal.pone.0057252] [Medline: 23472077]
12. Barboza P, Vaillant L, Le Strat Y, et al. Factors influencing performance of internet-based biosurveillance systems used in epidemic intelligence for early detection of infectious diseases outbreaks. PLOS ONE 2014;9(3):e90536. [doi: 10.1371/journal.pone.0090536] [Medline: 24599062]
13. Fieldhouse JK, Randhawa N, Fair E, Bird B, Smith W, Mazet JAK. One Health timeliness metrics to track and evaluate outbreak response reporting: a scoping review. EClinicalMedicine 2022 Nov;53:101620. [doi: 10.1016/j.eclinm.2022.101620] [Medline: 36097540]
14. Rotureau B, Barboza P, Tarantola A, Paquet C. International epidemic intelligence at the Institut de Veille Sanitaire, France. Emerg Infect Dis 2007 Oct;13(10):1590-1592. [doi: 10.3201/eid1310.070522] [Medline: 18258016]
15. Hasan M, Hossain MA, Chowdhury S, et al. Human monkeypox and preparedness of Bangladesh: a knowledge and attitude assessment study among medical doctors. J Infect Public Health 2023 Jan;16(1):90-95. [doi: 10.1016/j.jiph.2022.11.032] [Medline: 36508945]
16. Brito AF, Semenova E, Dudas G, et al. Global disparities in SARS-CoV-2 genomic surveillance. Nat Commun 2022 Nov 16;13(1):7003. [doi: 10.1038/s41467-022-33713-y] [Medline: 36385137]
17. Han AX, Toporowski A, Sacks JA, et al. SARS-CoV-2 diagnostic testing rates determine the sensitivity of genomic surveillance programs. Nat Genet 2023 Jan;55(1):26-33. [doi: 10.1038/s41588-022-01267-w] [Medline: 36624344]
18. Viana R, Moyo S, Amoako DG, et al. Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa. Nature New Biol 2022 Mar;603(7902):679-686. [doi: 10.1038/s41586-022-04411-y] [Medline: 35042229]
19. Cavallaro M, Dyson L, Tildesley MJ, Todkill D, Keeling MJ. Spatio-temporal surveillance and early detection of SARS-CoV-2 variants of concern: a retrospective analysis. J R Soc Interface 2023 Nov;20(208):20230410. [doi: 10.1098/rsif.2023.0410] [Medline: 37963560]
20. Kumar A, Asghar A, Dwivedi P, et al. A bioinformatics tool for predicting future COVID-19 waves based on a retrospective analysis of the second wave in India: model development study. JMIR Bioinform Biotechnol 2022;3(1):e36860. [doi: 10.2196/36860] [Medline: 36193192]
21. Classification of Omicron (B11529): SARS-CoV-2 variant of concern. World Health Organization. URL: https://www.who.int/news/item/26-11-2021-classification-of-omicron-(b.1.1.529)-sars-cov-2-variant-of-concern [accessed 2023-07-29]
22. Brownstein JS, Freifeld CC, Chan EH, et al. Information technology and global surveillance of cases of 2009 H1N1 influenza. N Engl J Med 2010 May 6;362(18):1731-1735. [doi: 10.1056/NEJMsr1002707] [Medline: 20445186]
23. Chunara R, Andrews JR, Brownstein JS. Social and news media enable estimation of epidemiological patterns early in the 2010 Haitian cholera outbreak. Am J Trop Med Hyg 2012 Jan;86(1):39-45. [doi: 10.4269/ajtmh.2012.11-0597] [Medline: 22232449]

XSL•FO
RenderX

24. Kim S, Lee J, Oh J, et al. Event-based surveillance in Republic of Korea: assessment of the effectiveness of Epidemic Intelligence from Open Sources. Western Pac Surveill Response J 2025;16(3):1-7. [doi: 10.5365/wpsar.2025.16.1151] [Medline: 40778189]
25. Tegally H, Wilkinson E, Tsui JLH, et al. Dispersal patterns and influence of air travel during the global expansion of SARS-CoV-2 variants of concern. Cell 2023 Jul 20;186(15):3277-3290. [doi: 10.1016/j.cell.2023.06.001] [Medline: 37413988]
26. Lee DW, Kim JM, Park AK, et al. Genomic epidemiology of SARS- CoV-2 Omicron variants in the Republic of Korea. Sci Rep 2022 Dec 27;12(1):22414. [doi: 10.1038/s41598-022-26803-w] [Medline: 36575217]
27. Tsui JLH, McCrone JT, Lambert B, et al. Genomic assessment of invasion dynamics of SARS-CoV-2 Omicron BA.1. Science 2023 Jul 21;381(6655):336-343. [doi: 10.1126/science.adg6605] [Medline: 37471538]
28. Castelán-Sánchez HG, Martínez-Castilla LP, Sganzerla-Martínez G, Torres-Flores J, López-Leal G. Genome evolution and early introductions of the SARS-CoV-2 Omicron variant in Mexico. Virus Evol 2022;8(2):veac109. [doi: 10.1093/ve/veac109] [Medline: 36582501]
29. Li YT, Polotan FGM, Sotelo GIS, et al. Lineage BA.2 dominated the Omicron SARS-CoV-2 epidemic wave in the Philippines. Virus Evol 2022;8(2):veac078. [doi: 10.1093/ve/veac078] [Medline: 36090771]
30. Bergna A, Lai A, Sagradi F, et al. Genomic epidemiology of the main SARS-CoV-2 variants circulating in Italy during the Omicron era. J Med Virol 2025 Feb;97(2):e70215. [doi: 10.1002/jmv.70215] [Medline: 39936851]
31. Shu Y, McCauley J. GISAID: global initiative on sharing all influenza data—from vision to reality. Euro Surveill 2017 Mar 30;22(13):30494. [doi: 10.2807/1560-7917.ES.2017.22.13.30494] [Medline: 28382917]
32. GISAID. URL: https://gisaid.org/ [accessed 2025-06-30]
33. GenBank overview. National Center for Biotechnology Information (NCBI). URL: https://www.ncbi.nlm.nih.gov/genbank [accessed 2025-06-30]
34. Adams RP, MacKay DJC. Bayesian online changepoint detection. arXiv. Preprint posted online on Oct 19, 2007. [doi: 10.48550/arXiv.0710.3742]
35. van den Burg GJJ, Williams CKI. An evaluation of change point detection algorithms. arXiv. Preprint posted online on Mar 13, 2020. [doi: 10.48550/arXiv.2003.06222]
36. Pritchard E, Vihta KD, Eyre DW, et al. Detecting changes in population trends in infection surveillance using community SARS-CoV-2 prevalence as an exemplar. Am J Epidemiol 2024 Dec 2;193(12):1848-1860. [doi: 10.1093/aje/kwae091] [Medline: 38808625]
37. Fewster RM, Buckland ST, Siriwardena GM, Baillie SR, Wilson JD. Analysis of population trends for farmland birds using generalized additive models. Ecology 2000 Jul;81(7):1970-1984. [doi: 10.1890/0012-9658(2000)081[1970:AOPTFF]2.0.CO;2]
38. Hutwagner L, Browne T, Seeman GM, Fleischauer AT. Comparing aberration detection methods with simulated data. Emerg Infect Dis 2005 Feb;11(2):314-316. [doi: 10.3201/eid1102.040587] [Medline: 15752454]
39. Hutwagner L, Thompson W, Seeman GM, Treadwell T. The bioterrorism preparedness and response early aberration reporting system (EARS). J Urban Health 2003 Jun;80(2 Suppl 1):i89-i96. [doi: 10.1007/pl00022319] [Medline: 12791783]
40. Salmon M, Schumacher D, Höhle M. Monitoring count time series in R: aberration detection in public health surveillance. J Stat Soft 2016;70(10):1-35. [doi: 10.18637/jss.v070.i10]
41. Sharpe JD, Hopkins RS, Cook RL, Striley CW. Evaluating Google, Twitter, and Wikipedia as tools for influenza surveillance using Bayesian change point analysis: a comparative analysis. JMIR Public Health Surveill 2016 Oct 20;2(2):e161. [doi: 10.2196/publichealth.5901] [Medline: 27765731]
42. Chen Z, Azman AS, Chen X, et al. Global landscape of SARS-CoV-2 genomic surveillance and data sharing. Nat Genet 2022 Apr;54(4):499-507. [doi: 10.1038/s41588-022-01033-y] [Medline: 35347305]
43. Shen Y, Steele R, AbdelMalik P, Buckeridge D. Early detection intervals for epidemics caused by SARS-cov-2 omicron variant in multiple countries. Borealis. URL: https://borealisdata.ca/dataset.xhtml?persistentId=doi:10.5683/SP3/JKP6CO [accessed 2025-10-31]
44. Thomas MC, Kite-Powell A, Atrubin D, Hamilton JJ. Evaluating the utility of HealthMap as a supplementary surveillance tool. Online J Public Health Inform 2015;7(1):e61658. [doi: 10.5210/ojphi.v7i1.5833]
45. Noufaily A, Morbey RA, Colón-González FJ, et al. Comparison of statistical algorithms for daily syndromic surveillance aberration detection. Bioinformatics 2019 Sep 1;35(17):3110-3118. [doi: 10.1093/bioinformatics/bty997] [Medline: 30689731]
46. Fricker Jr RD, Hegler BL, Dunfee DA. Comparing syndromic surveillance detection methods: EARS' versus a CUSUM-based methodology. Stat Med 2008 Jul 30;27(17):3407-3429. [doi: 10.1002/sim.3197] [Medline: 18240128]
47. Mustapha M, Togo K, Volkman H, et al. Under-reporting of SARS-cov-2 infections in 27 countries, 2020–2022. SSRN. Preprint posted online on Nov 9, 2024. [doi: 10.2139/ssrn.5014145]
48. Wu SL, Mertens AN, Crider YS, et al. Substantial underestimation of SARS-CoV-2 infection in the United States. Nat Commun 2020 Sep 9;11(1):4507. [doi: 10.1038/s41467-020-18272-4] [Medline: 32908126]
49. Tegally H, San JE, Cotten M, et al. The evolving SARS-CoV-2 epidemic in Africa: insights from rapidly expanding genomic surveillance. Science 2022 Oct 7;378(6615):eabq5358. [doi: 10.1126/science.abq5358] [Medline: 36108049]

XSL•FO
RenderX

50. Fieldhouse J, Nakiire L, Kayiwa J, et al. How feasible or useful are timeliness metrics as a tool to optimise One Health outbreak responses? BMJ Glob Health 2024 Jul 10;9(7):e013615. [doi: 10.1136/bmjgh-2023-013615] [Medline: 38991578]
51. Crawley AW, Divi N, Smolinski MS. Using timeliness metrics to track progress and identify gaps in disease surveillance. Health Secur 2021;19(3):309-317. [doi: 10.1089/hs.2020.0139] [Medline: 33891487]
52. Shen Y, Powell G, Ganser I, et al. Monitoring non-pharmaceutical public health interventions during the COVID-19 pandemic. Sci Data 2021 Aug 24;8(1):225. [doi: 10.1038/s41597-021-01001-x] [Medline: 34429423]

Abbreviations

BOCP: Bayesian online change point
EARS: Early Aberration Reporting Systems
EBS: event-based surveillance
GAM: generalized additive model
GISAID: Global Initiative on Sharing All Influenza Data
HIC: high-income country
IBS: indicator-based surveillance
LIC: low-income country
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 03.Nov.2025; peer-reviewed by P Florentino, X Shi; revised version received 11.Feb.2026; accepted 23.Mar.2026; published 05.May.2026.

Please cite as:
Shen Y, AbdelMalik P, Steele RJ, Buckeridge DL
Early Detection Intervals for Evaluating Event-Based Surveillance System: Reference Dataset Development Study
JMIR Public Health Surveill 2026;12:e87030
URL: https://publichealth.jmir.org/2026/1/e87030
doi:10.2196/87030

© Yannan Shen, Philip AbdelMalik, Russell J Steele, David L Buckeridge. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 5.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Endemic Channel Parametrization in Dengue Surveillance: Methodological Assessment of Retrospective Windows, Outbreak Trends, and Zero-Case Periods in Colombia

Juan D Umaña[1*], BSc, MSc; Juan Montenegro-Torres[1*], BSc, MSc; Julian Otero[1,2], BSc, MSc; Maria Camila Tavera-Cifuentes[1], BSc, MSc; Natalia Niño-Machado[3], BA, MA, MSc, PhD; Catalina González-Uribe[3], BA, BSc, MA, MSc, PhD; Juan Manuel Cordovez[1], BSc, MSc, PhD; Mauricio Santos-Vega[1,4], BSc, MSc, PhD

[1]Grupo de Biología Matemática y Computacional (BIOMAC), Universidad de Los Andes, Carrera 1 # 18A - 12 Edificio A, Bogotá, Colombia

[2]Department of Community Medicine, Faculty of Health Sciences, UiT The Arctic University of Norway, Tromsø, Norway

[3]Centro de los Objetivos de Desarrollo Sostenible para América Latina y el Caribe (CODS), Universidad de Los Andes, Bogotá, Colombia

[4]Departamento Ciencias Biológicas, Universidad de Los Andes, Bogotá, Colombia

[*]these authors contributed equally

Corresponding Author:
Mauricio Santos-Vega, BSc, MSc, PhD
Grupo de Biología Matemática y Computacional (BIOMAC), Universidad de Los Andes, Carrera 1 # 18A - 12 Edificio A, Bogotá, Colombia

Abstract

Background: The endemic channel is a surveillance method that presents statistical indicators and visual representations of a disease's historical dynamics. Its epidemic curve defines the central tendency of cases and their expected variation, providing 3 levels (ie, "safety," "warning," and "epidemic") to assess the epidemiological status of a region. Parameters include the central tendency used as the epidemiological warning threshold (EWT), the size of the retrospective window, and the handling of previous outbreaks and zero values in data. The absence of clear guidelines for the selection of these parameters may compromise reproducibility and hinder outbreak definitions and responses for endemic diseases such as dengue.

Objective: This study aimed to review the parameters of the endemic channel used for the definition and monitoring of dengue outbreaks in Colombia while quantitatively assessing the performance of the method.

Methods: We reviewed institutional epidemiological bulletins in Colombia and quantitatively assessed the endemic channel in two main aspects: (1) the impact on the EWT of parameter selection regarding the retrospective data window, previous epidemic years handling, and zero-value handling, using a statistical framework; and (2) the endemic channel's performance based on the windows of opportunity, outbreak detection capacity, and the ratio of warnings that correspond to actual outbreaks.

Results: The endemic channel's performance is higher as transmission increases due to more robust data that facilitate a timely detection of outliers, while lower-transmission areas show a sharper rise in cases when outbreaks are missed, indicating limited detection capacity. Reducing the retrospective data window improved metrics across all transmission profiles by 6.34% on average, while extending it decreased performance due to changes in detection capacity. There was no significant difference (P value >.01) in performance when data from epidemic years were included or excluded for municipalities with high or very high transmission levels. Instead of adding an entire unit, shifting the data by 0.001 prevents the estimation of null values for the EWT and thresholds and significantly improves performance across all transmission levels (P value <.01) by 23.07% on average.

Conclusions: The endemic channel's performance varies with the outbreak definition and the municipality's transmission level. Encouraging an optimal retrospective window is challenging, as data are computed over the years. Nevertheless, the improved performance with shorter retrospective windows is likely due to reduced overlap in seasonal outbreaks. Shifting data by a limiting-to-zero value, instead of adding a complete unit, improves performance and can be easily integrated into existing surveillance templates. Windows of opportunity should be considered when selecting the parameter combination. Finally, reassessing outbreak definitions and method parameters underpinning surveillance tools is essential to ensure their validity and effectiveness, especially when used to inform early warning systems and public policies.

(JMIR Public Health Surveill 2026;12:e79914) doi:10.2196/79914

KEYWORDS

endemic channel; epidemiological monitoring; dengue; disease surveillance; disease outbreaks; early warning systems; data accuracy

XSL•FO
RenderX

Introduction

States and policymakers often use various methods to set dengue epidemic warning thresholds in endemic regions, resulting in different outcomes. One prevalent method is the endemic channel approach, which establishes thresholds based on recent data (often excluding epidemic years) by analyzing variability measures, such as the SD, and central tendency measures, such as the mean to, define an expected upper limit for cases (eg, mean plus 2 SDs) [1]. The endemic channel, introduced in 1999 by Bortman [2], exemplifies a surveillance tool that presents statistical metrics and a visual schematic of a disease's historical dynamics in a specific region. It uses case data to construct an epidemic curve, showing cases over time and establishing the central tendency along with upper and lower limits where cases are expected to fluctuate. This method creates a graph with 3 thresholds, defining "safety," "warning," and "epidemic" zones, which are used to assess the region's epidemiological situation based on current case counts.

The World Health Organization (WHO) reports an 8-fold increase in dengue cases over the last two decades, from approximately 505,000 cases in 2000 to more than 2.4 million in 2010 and 5.2 million in 2019 [3]. More recently, in 2024, the Americas region reported an unprecedented rise in dengue cases, exceeding 13 million notifications [4]. Colombia has been under epidemiological warning due to an increase in cases since the second quarter of 2022, prompting the intensification and strengthening of prevention, comprehensive care, surveillance, and control measures [5]. To support these activities, the endemic channel has been widely adopted in the Americas, including Colombia, the Dominican Republic, and Peru, to assess dengue surveillance and response [6]. The method, along with other outbreak definitions, such as deviation of bar charts and moving averages, has been evaluated by experts from the WHO's Special Program for Research and Training in Tropical Diseases [6]. Their assessment highlights several key advantages of this approach: first, it provides a straightforward workflow for monitoring epidemic thresholds, allowing for the quick detection of outbreaks; second, its standardized outbreak definition promotes more transparent and consistent communication across diverse audiences, making messages easier to understand and compare; third, the endemic channel enables public health authorities to evaluate outbreak severity by examining factors, such as duration and total case numbers, thereby simplifying comparisons across different countries; and ultimately, it serves as a crucial tool for evaluating response effectiveness and informing decisions on whether to intensify or reduce intervention efforts, thereby enhancing outbreak management and control.

It is important to recall that the endemic channels focus on intraannual trends, calculating monthly or seasonal means to create a baseline, which simplifies operations but can lead to significant uncertainty [1]. As a result, the endemic channel can be quite uncertain, with outbreak thresholds being occasionally exceeded. These challenges have led to discussions and efforts to develop evidence-based models for dengue surveillance and outbreak management [6]. Methodologically, the endemic channel encompasses several parameters to be considered in its implementation. First, the central tendency measure used is the epidemiological warning threshold (EWT). Then, the size of the retrospective window (in years) for historical data is considered. Finally, the handling of previous outbreaks and zero values in the data is addressed. Owing to the direct influence of these parameters on algorithmic outcomes, an absence of clear institutional pipelines regarding their selection and underlying assumptions could hinder reproducibility and complicate the interpretation of results. This ambiguity can also foster experience-dependent usage and limit the comparability of sensitivity and specificity analyses across regions (eg, Bowman et al [7] and Schlesinger et al [8]) with differing transmission profiles, such as hypoendemic, endemic, and hyperendemic areas. Analytical pipelines implemented in open-access tools, such as the epiCo R package (R software developed by JDU, JM-T, and JO at Epiverse-TRACE) [9], could provide reproducible endemic channel workflows for other diseases or regions and promote the review of the mathematical assumptions embedded in current surveillance systems.

This paper examined how institutions define dengue outbreaks and evaluate the guidelines used to estimate the endemic channel at both national and departmental levels in Colombia. Specifically, we analyze the endemic channel's outcomes by two main aspects: (1) testing how parameter choices—such as the retrospective data window, handling of previous epidemic years, and zero values affect results; and (2) proposing epidemiological metrics to assess the method's performance based on the resulting windows of opportunity, outbreak detection capability, and epidemiological outcomes following a warning.

Methods

Overview

This study was conducted in 2 main stages. First, we performed a documentary analysis to characterize how endemic channels for dengue are constructed and used across Colombia. This involved a review of publicly available epidemiological reports at national and departmental levels. The second stage used previous results to establish an institutional baseline, understood as the set of parameters to use in the endemic channel, to assess experiments on the method using official linelist data.

Ethical Considerations

The TRACE-LAC (Tools for Response, Analytics, and Control Enhancing of Epidemics in Latin America and the Caribbean) research project has a no risk classification according to the Ethics Committee of Universidad de los Andes, expresser in the Acta 1632 of 2022 and based on the Colombian resolutions 008430 of 1993 and 2378 of 2008. This work does not involve the manipulation of sensitive data or interaction with vulnerable communities or individuals, nor the performance of any experiment on subjects.

Exploration of Institutional Reports and Outbreak Definitions

We first developed an analytical baseline to characterize how Colombia constructed and used endemic channels. To support

XSL•FO
RenderX

this analysis, we retrieved the annual dengue reports from 2008 to 2022 from the National Institute of Health (Instituto Nacional de Salud [INS]) [10] and Health Secretariats' websites, covering both the national and department (administrative division one) levels. From these reports, we aimed to identify the type of data and parameters used in constructing endemic channels and assess the consistency and availability of these information methods for constructing endemic channels nationwide. The compiled reports allowed us to establish parameters, such as the method, the retrospective window, zero-value handling, and the previous outbreak handling, which were subsequently used as a reference for the experiments applied to the data described below.

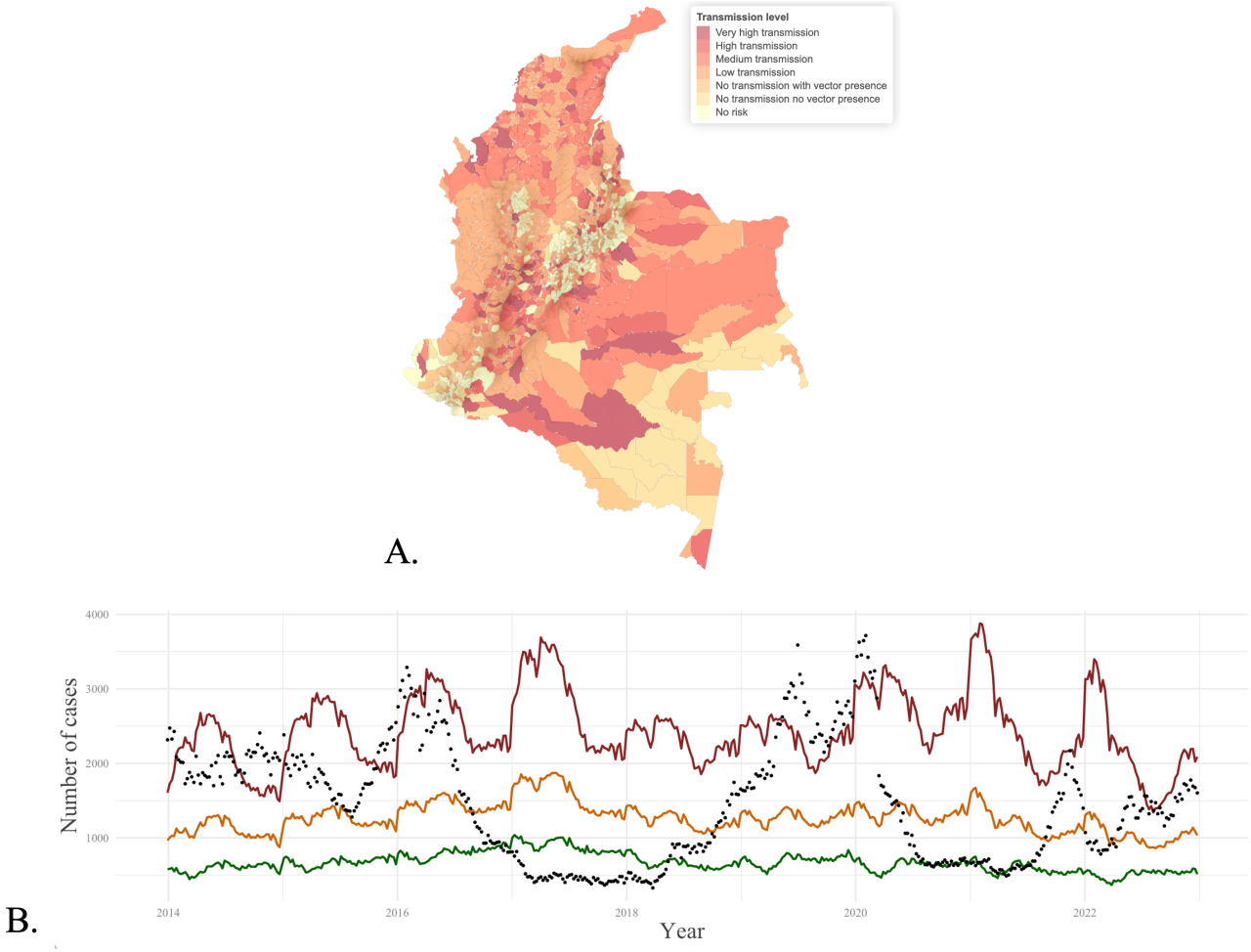
While analyses were outlined for Colombian municipalities with dengue transmission, we replicated the following protocol at departmental resolution to explore how endemic channel outcomes vary when aggregating data from settings with heterogeneous transmission levels. The stratification of transmission at the departmental level was assigned as the level to which most inhabitants within each department are exposed.

Data Description

Epidemiological Data

Dengue linelist data between 2007 and 2022 in Colombia were downloaded from the National Surveillance System (SIVIGILA) website using the *sivirep* R package [11]. The dataset comprises anonymized individual case notifications, which were cleaned and harmonized by location codes. Cases were aggregated solely by municipalities or departments of occurrence and grouped by epidemiological weeks according to the symptom onset. Municipalities in Colombia are typically classified into 7 transmission levels by the Ministry of Health and the National Institute of Health [12]; the spatial distribution of the transmission levels and the national weekly incidence are shown in Figure 1A and B, respectively.

Figure 1. Epidemiological overview of dengue notification in Colombia. (**A**) Spatial distribution of dengue transmission levels for 2023, (**B**) Continuous endemic channels (from 2014 to 2022) at the national level. Black dots report weekly incidence; green, yellow, and red lines account for "safety," "warning," and "epidemic" levels, respectively.



Estimation of the Endemic Channels

The endemic channels were estimated using the *epiCo* R package [9], which provides statistical analyses and visualizations of vector-borne disease outbreaks in Colombia. *epiCo* is a product of the TRACE-LAC Initiative and implements functions to estimate the expected number of cases as the EWT and the associated CIs used as safety and outbreak thresholds, as proposed by Bortman and explained in detail in Multimedia Appendix 1. We used the case data to calculate the endemic channel in the municipalities of Colombia. Figure 1B

XSL•FO
RenderX

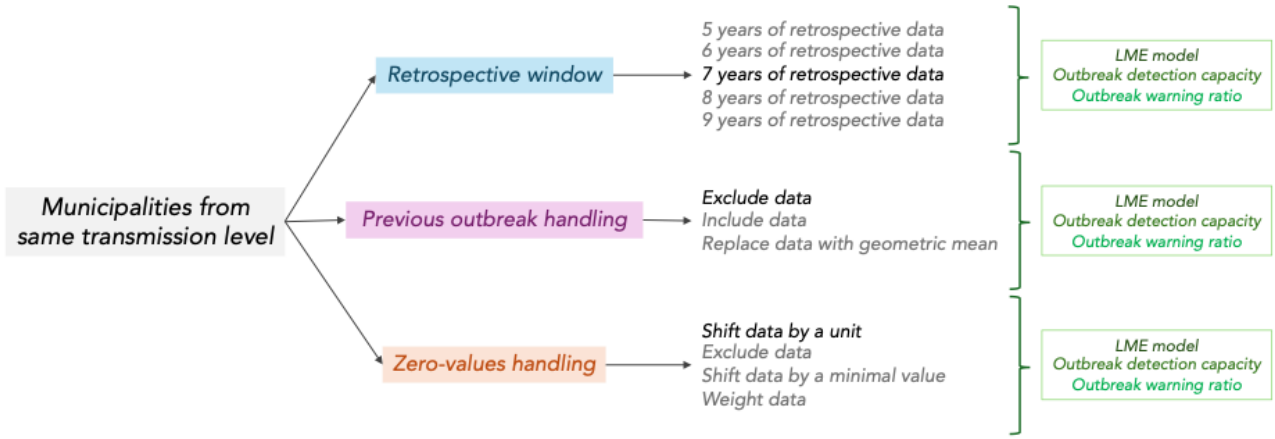
shows the endemic channels from 2014 to 2022 in a continuous graph.

Endemic Channel Assessment Protocol

Overview

We quantitatively assessed the impact of the endemic channel parameters, focusing on 2 key aspects: their effect on the EWT and their effect on the endemic channel's performance. Figure 2 outlines the implemented protocol. We tested parameter configurations using municipalities grouped by similar transmission levels. Three parameters of the endemic channel method were systematically assessed: the retrospective window, the handling of previous outbreaks, and the handling of zero values. The process was conducted separately for the sake of interpretation, assuming that it is not possible for a region to exhibit exceptional outbreaks and systematic zero cases within its historical records.

Figure 2. The quantitative assessment of the endemic channel parameters for dengue surveillance consisted of 3 experiments using the endemic channels of Colombian municipalities with the same transmission level, from 2014 to 2022. The changes in the epidemiological warning threshold of each experiment were evaluated with a linear mixed-effects (LME) model. The outbreak detection capacity and the warning ratio metric assess the effectiveness of the outbreak warning; both metrics contribute to the overall performance evaluation of the endemic channels. In each experiment, parameters shown in bold text represent the resulting baselines used for comparison.



The retrospective window, defined as the number of years of historical data used for threshold estimation, was tested, ranging from 5 to 9 years. To handle previous outbreaks within the retrospective window, three strategies were considered: (1) excluding outbreak periods, (2) including them as-is, or (3) replacing them with the geometric mean of the remaining dataset. To evaluate the effect of zero-value handling strategies, we (1) shifted the number of cases by a unit to avoid zero values as stated by Bortman [1], prior to the EWT calculation, to later be subtracted; (2) excluded data with zero values from the computation; (3) shifted data by a minimal value (0.001 for the purpose of this study) prior to EWT calculation to later be subtracted, as an alternative approach to Bortman; and (4) weighted the estimations by the fraction of zero values present in the data; these last 2 approaches are inspired by de la Cruz and Kreft's [13] and Habib's [14] approaches to estimating the geometric mean of a dataset with zeros. Each parameter combination was then evaluated using a linear mixed-effects (LME) model and a composite performance metric for outbreak detection capacity and warning ratio.

LME Model on the EWT

The following analysis evaluates whether the experiments shown in Figure 2 produce a significant increase or decrease in the EWT. The epidemiological interpretation of the assessment is that changes in the EWT ultimately define the "warning" threshold level, which greatly influences the triggering of the method and subsequent performance evaluations. To ensure a valid evaluation, the analysis recognizes that observations within the same municipality are not independent but influenced by underlying trends over time.

Therefore, we applied an LME model with the EWT as the response variable and each experiment depicted in Figure 2 as a fixed effect. We also incorporated a random effect structure to control for within-group correlations in the observations of the same municipality and date (Equations 3 to 5).

Equation 1EWTij=β0+β1Experimenti+uj(i)+ ij

Equation 2uj(i)~N(0,σu2)

Equation 3 ij~N(0,σ2)

where $u_{j(i)}$ is the random intercept for the group $j(i)$, that is, municipality j with observation at time i, σ_u^2 represents the variance of this intercept, and t describe the variability across the municipalities The model coefficients' analysis focused on whether the EWT was significantly impacted (P value <.01) by each experiment, captured by β_1, and the percentage of change that this coefficient induces.

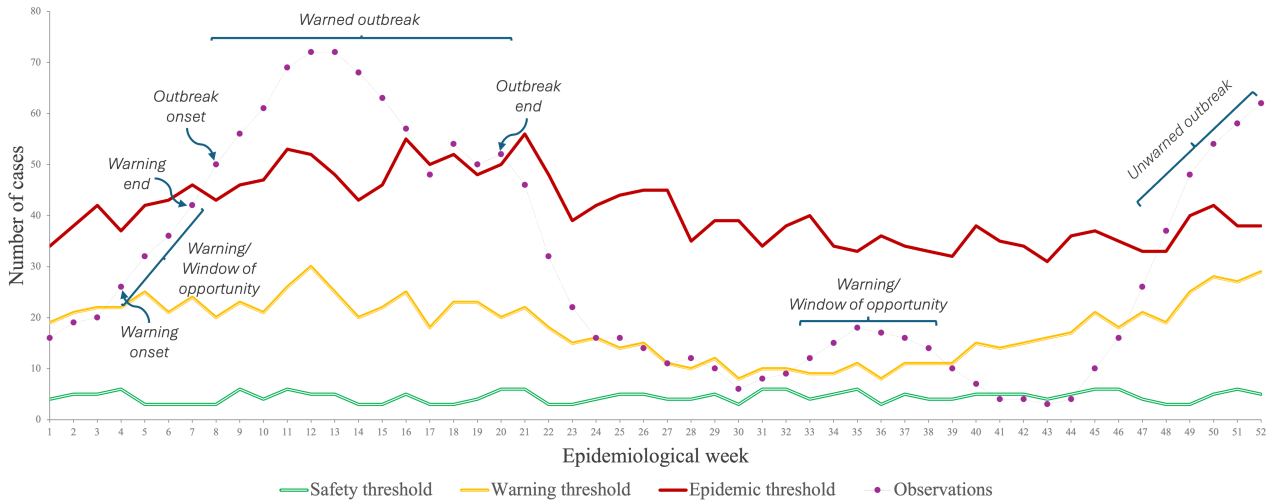
Epidemic Features and the Endemic Channel Performance

We developed an algorithm that calculates several epidemic features, such as the outbreak's onset (the first of 3 or more consecutive weeks when cases are above the upper limit after being at least 2 weeks below it), the total number of warnings (the first of 2 or more consecutive weeks when cases are between the EWT and the upper limit), the number of warnings that resulted in an outbreak, the number of outbreaks that occurred without triggering a warning, and the windows of

XSL•FO
RenderX

opportunity. A more detailed definition of these features can be found in Table S1 in Multimedia Appendix 1 and is exemplified in Figure 3.

Figure 3. Graphical representation of the epidemic features based on the Joint External Communication [15] from the Ministry of Health and the National Institute of Health. Two warnings at weeks 4 and 33 are shown: the first one resulted in an outbreak with a window of opportunity of 4 weeks, while during the second warning, the increase in cases was controlled in a window of 6 weeks. Cases from weeks 18, 19, and 20 were not classified as a new outbreak because the previous warning did not last 2 or more weeks. Similarly, there was no triggered warning at weeks 25, 28, and 47 because cases were over the epidemiological warning threshold for only 1 week, the last scenario leading to an unwarned outbreak.



On the basis of these epidemic features, we propose 3 metrics to assess the endemic channel methodology: the outbreak detection capacity, the outbreak warning ratio, and the endemic channel performance. Metrics are described in equations 6, 7, and 8. To weigh the impact of outbreak detection capacity and the outbreak warning ratio equally, the parameter is set to 0.5.

Equation 4 Outbreak detection capacity = Number of warned outbreaks / Total number of outbreaks

Equation 5 Outbreak warning ratio = Number of warned outbreaks / Number of triggered warnings

Equation 6 Endemic channel performance = α × Outbreak detection capacity + (1 − α) × Outbreak warning ratio

Although we did not assess the endemic channel as a forecasting tool, we aimed to gain insight into the window of opportunity that the method offers in the studied context. The window, then, was defined for all experiments as the delay in weeks between a warning onset (first of 2 or more consecutive weeks when cases are between the EWT and the upper limit) and an outbreak onset (first of 3 or more successive weeks when cases are above the upper limit after being at least 2 weeks below it). Significance in epidemic features was assessed with a mixed-effects ANOVA for a P value <.01.

Results

Overview

Our exploration of institutional reports and outbreak definitions did not show a standardized set of parameters; however, the review of national surveillance guidelines and Bortman's original reference informed the selection of the following baseline parameters for evaluating endemic channel outcomes: (1) method: geometric mean; (2) zero-value handling: addition of one unit to the data; (3) outlier year: 2019; (4) outlier handling: exclusion; and (5) retrospective window: 7 years. Table S2 in Multimedia Appendix 1 displays the findings on the method used to estimate the endemic channel for each year, the number of years to consider a retrospective window, and the handling of epidemic years. None of the reports displayed the numerical values of the EWT or the epidemic thresholds.

The same previous baseline was used in the assessment at the departmental level due to the limited number of epidemiological bulletins at this administrative level that reported the used methodology and the heterogeneity in the selection of its parameters when documented. For example, in 2022, from the 28 departments with dengue transmission, only 12 displayed an endemic channel graph, 8 of these documented the used central tendency measure (12.5% used the mean, 12.5% used the geometric mean, 37.5% used the median, and 37.5% reported using more than one method during the year); and, in the same year, the window size was reported only in 16.7% of the departments, varying between 5 and 9 years of data. Tables S3 and S4 in Multimedia Appendix 1 contain the resulting classification of dengue transmission at the departmental level and the corresponding parameters, respectively.

The following subsections detail the results by parameter and by transmission level and describe the behaviors of the windows of opportunity and spatial aggregation when analyzed at the departmental level. As noted in the methodology, only one parameter was varied at a time relative to the baseline, and therefore, parameter combinations were not assessed. Consequently, it cannot be determined whether combining the best-performing settings for each parameter would yield superior overall results.

Assessment of the Retrospective Window

On the basis of the 7-year retrospective data window, the EWT decreased relative to the baseline in all cases. The reduction was particularly pronounced when the extremes of the data window were assessed—specifically, 5 and 9 years of

XSL•FO
RenderX

retrospective data—as illustrated in Table S5 in Multimedia Appendix 1 and Figure 4A. The only exception was the 6-year retrospective data window for municipalities with low and medium transmission levels, which did not significantly change the EWT. Figure 5A depicts the influence of retrospective window size on the performance of the endemic channel. Across all transmission levels, performance declined as the retrospective window increased, primarily due to a reduction in outbreak detection capacity. Notably, municipalities with higher transmission levels demonstrated better endemic channel performance overall. Furthermore, municipalities with low and medium dengue transmission were the most substantially affected by variations in the retrospective data window, highlighting the sensitivity of these settings to the parameters studied.

Figure 4. Endemic channel's parameters effect on the epidemiological warning threshold: a linear mixed-effects model analysis. (**A**) Retrospective data windows experiment, (**B**) previous outbreaks handling experiment, (**C**) zero-value handling experiment. Percentages displayed as the logarithm of the value. Detailed data can be accessed in Tables S5, S7, and S9 in Multimedia Appendix 1.

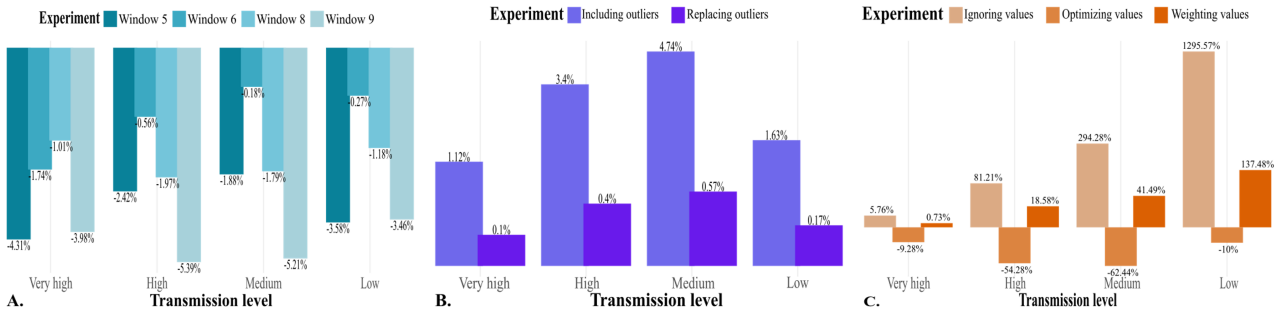
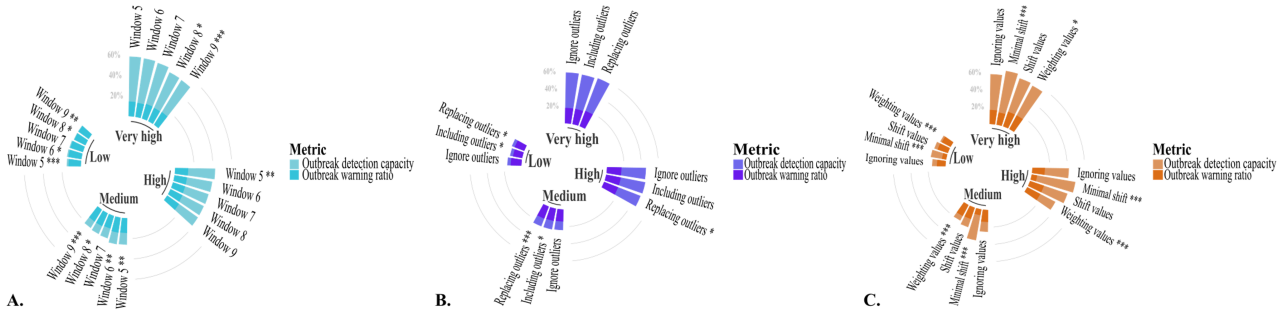


Figure 5. Endemic channel's parameters effect on the endemic channel performance: a composite metric including the number of warned outbreaks over the total number of outbreaks (outbreak detection capacity) and the number of warned outbreaks over the total number of triggered warnings (outbreak warning ratio). (**A**) Retrospective windows experiment, (**B**) previous outbreaks handling experiment, (**C**) zero-value handling experiment. Performance is proposed as the contribution of the detection capacity (light colors) and the outbreak warning ratio (dark colors). Detailed data can be accessed in Tables S6, S8, and S10 in Multimedia Appendix 1. Significance levels: **P*<.05, ***P*<.01, ****P*<.001.



Assessment of How Previous Outbreaks Are Handled

For both treatments, the EWT increased compared to the baseline strategy of ignoring previous outbreaks, with a greater increase observed when data from prior outbreaks were incorporated, as shown in Table S7 in Multimedia Appendix 1 and Figure 4B. The impact of handling previous outbreaks on endemic channel performance is illustrated in Figure 5B. In low-transmission municipalities, the highest performance was achieved when ignoring data from previous outbreak years. In contrast, replacing outbreak data with a geometric mean resulted in better performance for medium and very-high-transmission municipalities. For high transmission municipalities, the best performance was observed when data from previous outbreaks were included. Overall, as transmission levels increase, the performance of the endemic channel improved across different handling methods.

Effect of Zero-Value Handling

The minimal shift was the only strategy that decreased the EWT compared to the baseline (shifting the data by a unit), as seen in Table S9 in Multimedia Appendix 1 and Figure 4C. The percentage of change was noticeably greater in low- and medium-transmission municipalities, as expected given their higher frequency of zero values.

The impact of zero-value handling strategies on endemic channel performance is depicted in Figure 5C. Metrics for the ignored-value and weighted-value strategies are presented for municipalities where we could estimate these metrics; however, the endemic channel consistently failed when these strategies were applied in cases where retrospective data lacked sufficient nonzero values. Across all transmission levels, performance significantly improved when using the minimal shift strategy. Notably, higher transmission levels corresponded with better endemic channel performance. Moreover, on average, all municipalities, regardless of their dengue transmission levels, were significantly influenced by applying the minimal shift and weighting strategies.

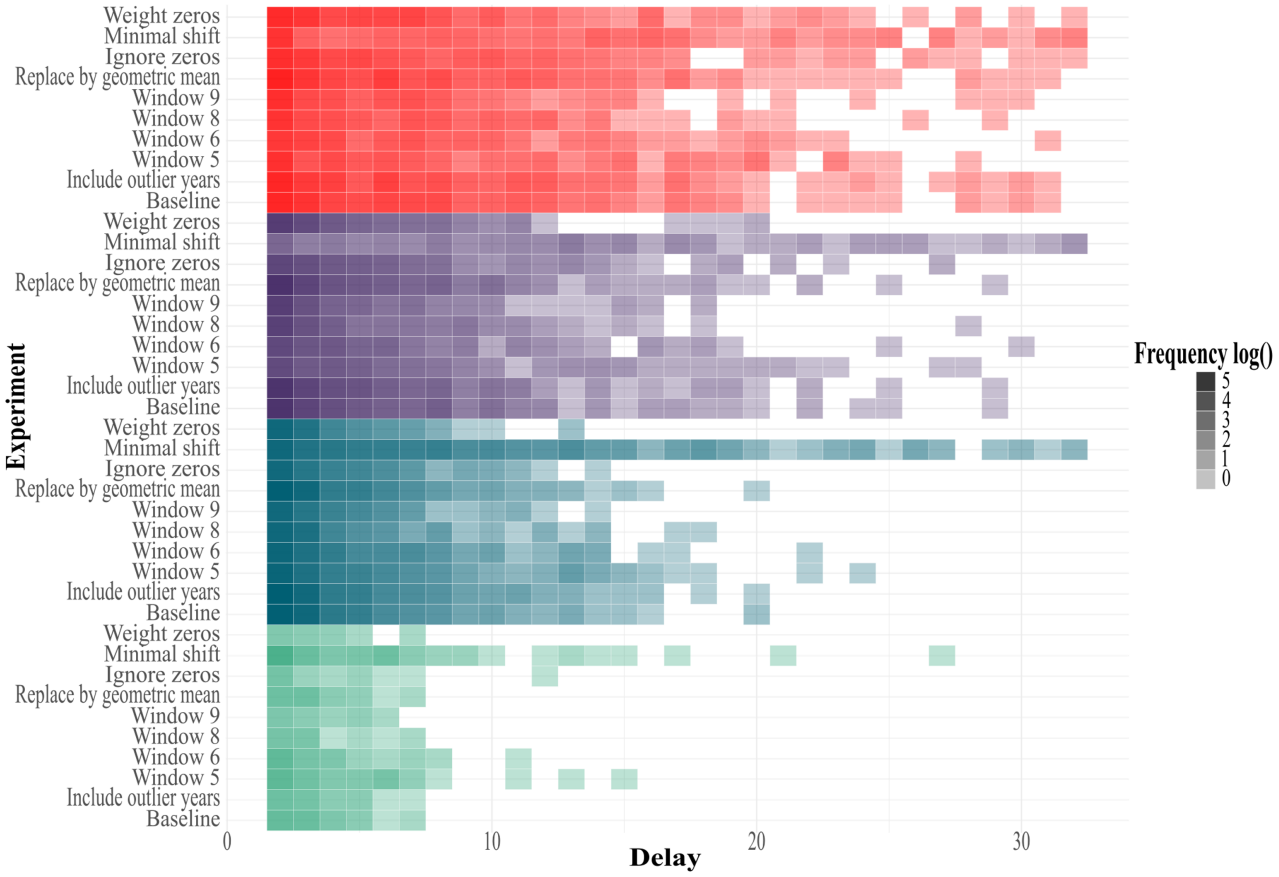
Windows of Opportunity Behavior

We identified that the 2-week delay is the shortest and most predominant window of opportunity in all transmission levels and experiments, as expected from the definition of the outbreak

XSL•FO
RenderX

onset used in this study. Figure 6 displays the distribution of the windows of opportunity for all experiments across the different transmission levels. Neither the endemic channel nor our predictive assessment considered the magnitude or duration of the outbreak following the warning or the intraannual trend in case numbers.

Figure 6. Windows of opportunity distributions across experiments on very high (red), high- (violet), medium- (turquoise), and low-transmission (mint) municipalities. Color intensity is the logarithm of the frequency of observed delays, baselines refer to the institutional protocol, and the figure is shown up to a 32-week delay. Over-a-year outliers were found in municipalities with very high, high, and medium transmission.



In general, at the municipality level, the experiments allowed us to identify 2 settings that consistently improved the performance of endemic channels across all transmission levels: reducing the retrospective window relative to the baseline and applying a minimal shift to data (0.001) instead of adding a complete unit to handle zero values. In contrast, the management of previous outbreaks produced heterogeneous effects across transmission levels. When evaluated at departmental scales, the endemic channel significantly outperformed municipal results in all experiments and all transmission profiles, as seen in Tables S11, S12, and S13 in Multimedia Appendix 1. For instance, the baseline protocol showed a performance approximately 4 times greater for departments with low transmission, 2.3 times greater for those with medium transmission, 1.4 times greater in high transmission settings, and 1.2 times greater for departments with very-high transmission, compared to the performance at the municipality level. In contrast to the experiments at the municipal level, the effect of the retrospective window size, previous outbreak handling, and zero-value handling did not exhibit a statistically significant difference compared to the baseline, leading to similar conclusive insights and recommendations provided in this study.

Discussion

Principal Findings

Our assessment of the endemic channel parameters shows that reducing the retrospective data window improved metrics across all transmission profiles by 6.34% on average, while extending it decreased performance due to changes in detection capacity. It also highlights that the additional step of excluding outbreak years, such as 2019, in the estimation of the EWT does not significantly affect the method's performance in municipalities with very high and high transmission but produces significant differences in medium- and low-transmission units, where performance changed by approximately +1.48% and –14.82%, respectively. Experiments on handling zero values demonstrated that adding a small number to observed cases (such as 0.001) not only prevents null estimates of epidemic thresholds but also improves overall performance by an average of 23.07% across all transmission levels. An analysis of the window of opportunity at the baseline showed that its duration varies with transmission intensity; with an average of 8 weeks for very high, 5 weeks for high, 4 weeks for medium, and 3 weeks for low-transmission municipalities, reflecting that the sharpness of windows as transmission decays may make the anticipation of outbreaks difficult.

XSL•FO
RenderX

The use of metrics such as the outbreak detection capacity or the endemic channel performance proposed in this study faces challenges comparable to the definition of an outbreak. For example, Brady et al [16] modeled 5 candidate definitions of an outbreak (recent mean, monthly mean, moving mean, cumulative mean, and fixed incidence threshold) and found that resulting outbreaks varied significantly in terms of frequency, duration, and impact. Additionally, outbreak definitions vary in timing, affecting their suitability for early intervention. Likewise, our results show that the behavior of the endemic channel performance differs according to the selected outbreak definition (tuned by the parameters used in the method and the definitions of the epidemic features) and the transmission context. In addition, implementing performance metrics on outbreak detection methods that exhibit different definitions hinders the comparison of epidemiological scenarios within regions or periods. For example, previous assessments on epidemic thresholds report a limited sensitivity of 40% (accounting for warnings that lead to a significant increase in cases) and a specificity of 90% in Puerto Rico when using 2 SDs as CIs [17]. In contrast, a sensitivity of 66% in Brazil was reported when using 1 SD [18]. These assessments differ from the endemic performance proposed in this study because they do not account for diverse definitions of outbreaks based on other parameters of the surveillance method besides the SD as a threshold size.

To the best of our knowledge, no prior study has examined the effect of retrospective window size, previous outbreak handling, and alternative strategies to handle zero values on the outcomes generated by the endemic channel method. It is challenging to encourage a shorter window, as the selection of retrospective data is defined in year units and hinders the ability to identify an optimal window size. Excluding specific years from the retrospective data could lead to inconsistent usage of the endemic channel, primarily because the selection uses the same criteria once the potential epidemic year has already occurred. In addition, our analyses support the decision to include previous outbreak data with minimal impact on the number of triggered warnings. When evaluating the different experiments to handle the occurrence of zero values in the retrospective data, we identified that the selection is restricted to the shifting strategies in the scenario where all data are zero. Therefore, shifting data with a minimal value is equally feasible to implement in existing surveillance templates in official institutions. It is important to recall that the minimal shift strategy does lead to a significant reduction in the EWT, especially in municipalities with high and medium transmission. This will generate more warnings, and local authorities may review their response protocols to avoid exhaustion of surveillance and control programs. This behavior is better understood when the delay between the warning signal and the outbreak onset is considered, as the following sections in the study do. Given the nature of surveillance data, it is challenging to evaluate if the persistence of the warning is caused by a slow-paced transmission or a mitigated outbreak, for example. Ultimately, for municipalities with very high, high, and medium transmission, the window of opportunity provided more information regarding the amount of time that the system is in a warning status than regarding the timeliness and specificity of the endemic channel to trigger a warning for nonseasonal and significant outbreaks. On the other hand, the low-transmission municipalities display short periods between warnings and outbreaks, reflecting their outbreaks' sharp, intermittent, and difficult-to-anticipate behavior. The temporal criteria in the onset definition, while needed to address that the threshold lines are often surpassed briefly and intermittently [6], condition the analysis of the timeliness of the endemic channel warning, as the minimum delay is policy rather than data driven.

Study Limitations

The results presented in this study rely on the assumption of a constant guideline for constructing endemic channels, which we subsequently used as baselines regardless of spatial or temporal setting. To enable analyses across all municipalities and years, we did not validate the warnings triggered by these baselines or the experiments against public or private records of health responses. If outbreak definitions were influenced by information beyond reported case data and its expected epidemiological limits, our approach does not capture such adjustments, as it does not account for external variables.

Similarly, the performance metrics used to assess the outcomes of the endemic channel were not compared with other methodologies or with external variables that might better reflect the true epidemiological situation of the regions. Consequently, there is limited certainty that governmental institutions would classify as outbreaks the same events we define as such. Furthermore, our interpretation of the metrics assumes an ideal performance of maximizing the detection of actual outbreaks while minimizing the number of triggered warnings, which may also be biased or constrained from the perspective of a public health officer. This absence of a ground truth dataset, combined with the statistical limitations inherent to the endemic channel, led us to avoid forecasting evaluations or other assessments that may exceed the practical uses of the surveillance method.

Finally, the statistical aggregation of our experiments relies exclusively on the Ministry of Health classification of dengue transmission levels. This stratification was not contrasted with additional approaches and may hinder population and disease dynamics that shape the expected variation of cases and the occurrence of outbreaks. The heterogeneity of our results and their dependency on the transmission level prevent us from recommending a single guideline; instead, they support the need to evaluate the appropriateness of the endemic channel in settings with low disease incidence.

Conclusions

Our results show that the endemic channel's performance significantly varies according to the outbreak definition and transmission level. The metrics exhibit better performance as the transmission level increases, likely due to (1) more robust retrospective data for high transmission settings, which supports the identification of transmission outliers; and (2) a sharper increase in cases in municipalities with lower transmission when outbreaks occur that are not detected by the endemic channel on time (ie, low outbreak detection capacity in these settings). Therefore, the timeliness of the warnings needs to be considered, especially in departments with a considerable proportion of

XSL•FO
RenderX

low- and medium-transmission municipalities because notification aggregation can obscure the occurrence of incipient outbreaks in these areas.

Concerns about the definition and usage of outbreak thresholds have been previously expressed by the WHO's Special Program for Research and Training in Tropical Diseases [6,19]. In their study, the committee gathered evidence of methods' weaknesses, such as health officers may mislead the outbreak threshold as a "warning sign" rather than an already triggered epidemic; a satisfactory algorithm for recognizing past aberrations has yet to be developed [20]; previous outbreaks may lead to an exaggerated increase in thresholds; seasonal outbreaks may differ in frequency from the fixed retrospective window; and, when geographical units are small, the variation in cases tends to increase, potentially resulting in significant oscillations. This study provides insights into some of these concerns and aims to promote the frequent assessment of methodologies and definitions used in institutional outbreak surveillance, especially in the increasing interest in automated surveillance systems.

Given the relevance of the disease in the region, it is crucial to develop and implement tools to predict possible outbreaks at different geographical scales. In this context, the WHO Global Arbovirus Initiative, for example, has emphasized the importance of innovative tools to promote integration of social and mathematical sciences with public health and include climate and environmental context in early warning and forecasting systems [21]. The initiative also highlights methodological gaps when integrating surveillance systems, field and/or laboratory confirmation, and epidemiological investigations. For example, whereas a predictive system is a deterministic or probabilistic model that uses historical data and identified patterns to infer future values [22], tools such as the endemic channel do not meet this definition, as it is a descriptive statistical method that does not make any inference of future cases. Therefore, to use the endemic channel as a forecasting system, it is necessary to incorporate complementary evidence from other tools, such as stochastic methods [23] or probabilistic models [24]. Additionally, as noted in the various methodologies, it is essential to consider the multiple factors that can affect the presence and biology of the vector, such as climate variability, water storage, population density, and socioeconomic conditions that have been shown to impact the disease behavior [25]. An early literature review of dengue epidemiological surveillance found that, although 46 of the 56 major dengue-endemic countries had dengue surveillance systems in place, only 4 had systems robust enough for epidemic prediction [26,27], and that no plan provided for the use of climate data to predict an outbreak [28]. For this reason, it is advisable not to use the endemic channel as the only source of information for decision-making but to consider other tools that incorporate these factors.

While the endemic channel is a methodology that overlooks the pathogen's transmission dynamics and local contexts, the assessment of parameters for other epidemiological events should be carefully considered. For example, we encountered challenges when defining a consensus of the institutional baselines after reviewing the guidelines and definitions published by official institutions in Colombia. Therefore, caution should be exercised when extrapolating our insights from this study to other contexts and diseases. With this study, we extend an invitation to frequently question and review the mathematical assumptions embedded in current surveillance methods to ensure their correct use and interpretation, especially when their results are used as warning systems in public health response policies.

Policy Recommendations

Governmental institutions may benefit from public communication of their methodology and parameters when constructing an endemic channel. The usage of the method, as analyzed through the epidemiological bulletins, is heterogeneous across regions and years. Notably, it is often difficult to identify the underlying assumptions and data workflows used in its development, which hinders the ability to replicate results, to establish a baseline, and to compare the epidemiological situation across municipalities or departments.

Including longitudinal (intraannual) analysis in a surveillance methodology, such as the endemic channel, is necessary to strengthen the tool's capacity to account for dengue transmission dynamics and weight notification data according to its recency. Meanwhile, we observed that the superior performance of shorter windows could be explained by the lower overlap of seasonal outbreaks, which occur approximately every 3 years in Colombia. Encouraging an optimal retrospective window is challenging, as data are computed by years, but the endemic channel for dengue surveillance in other regions may improve its performance if its retrospective data window is synchronized with the regional seasonality of the disease.

Alternative analyses to account for seasonal outbreaks should be incorporated when evaluating the epidemiological situation of a region. While removing previous epidemic years from the retrospective data helps mitigate outlier effects on the endemic channel, this practice did not yield a systematic improvement in performance metrics. Despite statistically significant differences in medium- and low-transmission municipalities, we found that this approach compromises the method's reproducibility in higher transmission settings, where its consistency may be more critical.

The addition of a small value to observed cases before the EWT estimation, rather than a complete unit, can be easily implemented for the benefit of the endemic channel's performance metrics. Although this change can increase the number of warnings, additional surveillance activities and tools can enlighten the correct outbreak response in these settings.

XSL•FO
RenderX

Acknowledgments

According to the GAIDeT (Generative Artificial Intelligence Delegation Taxonomy) taxonomy (2025), the generative artificial intelligence–assisted technology ChatGPT was solely used for proofreading and editing and translation in the manuscript. Authors wrote, reviewed, and edited all the content as needed and took responsibility for the originality of the article.

Funding

This work is part of the TRACE-LAC research project funded by the International Research Center in Ottawa, Canada (109848-001-). The funders had no role in study design, data collection and analysis, or the decision to publish or prepare manuscripts, reports, or working papers. The views expressed herein do not represent those of IDRC or its Board of Governors.

Data Availability

Source code can be found in a GitHub repository [29] for reproducibility and implementation in other epidemiological events.

Authors' Contributions

Conceptualization: JDU, JM-T
Exploration of departmental guidelines: MCT-C
Exploration of national guidelines: JDU
Formal analysis: JDU, JM-T
Stratification of transmission at the departmental level: JO
Supervision: NN-M, CG-U, JMC, MS-V
Writing – original draft: JDU, JM-T
Writing – review & editing: NN-M, CG-U, JMC, MS-V

Conflicts of Interest

None declared.

Multimedia Appendix 1
Endemic channel algorithm, the exploration of institutional reports and outbreak definitions, and assessment results.
[DOCX File, 66 KB - publichealth_v12i1e79914_app1.docx]

References

1. Wagner MM, Tsui FC, Espino JU, et al. The emerging science of very early detection of disease outbreaks. J Public Health Manag Pract 2001 Nov;7(6):51-59. [doi: 10.1097/00124784-200107060-00006] [Medline: 11710168]
2. Bortman M. Elaboración de corredores o canales endemicos mediante planillas de calculos. PAHO. 1999. URL: https://iris.paho.org/handle/10665.2/8562 [accessed 2025-02-11]
3. Dengue and severe dengue. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue [accessed 2025-01-11]
4. Epidemiological alert risk of dengue outbreaks due to increased circulation of DENV-3 in the Americas region. : Pan American Health Organization; 2025 URL: https://www.paho.org/sites/default/files/2025-05/2025-feb-7-phe-epi-alert-dengue-final2.pdf [accessed 2025-11-24]
5. Reyes AJR, Melo HEP, Herrera LCP, Rangel SYG, Morales RMC. Informe de evento 2023 dengue. : Instituto Nacional de Salud; 2022 Jan. [doi: 10.33610/infoeventos.4.1]
6. Badurdeen S, Valladares DB, Farrar J, et al. Sharing experiences: towards an evidence based model of dengue surveillance and outbreak response in Latin America and Asia. BMC Public Health 2013 Jun 24;13:607. [doi: 10.1186/1471-2458-13-607] [Medline: 23800243]
7. Bowman LR, Tejeda GS, Coelho GE, et al. Alarm variables for dengue outbreaks: a multi-centre study in Asia and Latin America. PLoS One 2016;11(6):e0157971. [doi: 10.1371/journal.pone.0157971] [Medline: 27348752]
8. Schlesinger M, Prieto Alvarado FE, Borbón Ramos ME, et al. Enabling countries to manage outbreaks: statistical, operational, and contextual analysis of the early warning and response system (EWARS-csd) for dengue outbreaks. Front Public Health 2024;12:1323618. [doi: 10.3389/fpubh.2024.1323618] [Medline: 38314090]
9. Umaña JD, Montenegro-Torres J, Otero J. EpiCo: statistical and viz tools for vector-borne diseases in Colombia. R package version 100 R. URL: https://epiverse-trace.github.io/epiCo/ [accessed 2025-02-11]
10. Informes de evento. Instituto Nacional de Salud. URL: https://www.ins.gov.co/buscador-eventos/Informesdeevento/Forms/AllItems.aspx [accessed 2025-02-12]
11. Gómez-Millán G, Cucunubá ZM, Mendez-Romero A, Huguett-Aragón C, International Development Research Center (IDRC). Sivirep: data wrangling and automated reports from “SIVIGILA” source. URL: https://epiverse-trace.github.io/sivirep/ [accessed 2025-02-11]

XSL•FO
RenderX

12. Lineamiento metodológico para la estratificación y estimación de la población en riesgo para arbovirosis en Colombia 2020-2023. Ministerio Nacional de Salud y Protección Social and Instituto Nacional de Salud. 2022. URL: https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/ED/VSP/lineamiento-metodologico-estimacion-poblacion-arbovirosis-colombia-2020-2023.pdf [accessed 2025-06-24]
13. de la Cruz R, Kreft JU. Geometric mean extension for data sets with zeros. arXiv. Preprint posted online on Apr 4, 2019. [doi: 10.48550/arXiv.1806.06403]
14. Habib EA. Geometric mean for negative and zero values. Semantic Scholar. 2012. URL: https://www.semanticscholar.org/paper/GEOMETRIC-MEAN-FOR-NEGATIVE-AND-ZERO-VALUES-Habib/4d0946f3051945509175d7d9f2cfe9d8c290a219 [accessed 2025-02-11]
15. Ministerio Nacional de Salud y Protección Social. Instrucciones para la organización y respuesta para el control del dengue en Colombia. 2023 May. URL: https://www.minsalud.gov.co/Normatividad_Nuevo/Circular%20Conjunta%20Externa%20No.13%20de%202023.pdf [accessed 2025-06-24]
16. Brady OJ, Smith DL, Scott TW, Hay SI. Dengue disease outbreak definitions are implicitly variable. Epidemics 2015 Jun;11:92-102. [doi: 10.1016/j.epidem.2015.03.002] [Medline: 25979287]
17. Rigau-Pérez JG, Millard PS, Walker DR, Deseda CC, Casta-Vélez A. A deviation bar chart for detecting dengue outbreaks in Puerto Rico. Am J Public Health 1999 Mar;89(3):374-378. [doi: 10.2105/ajph.89.3.374] [Medline: 10076488]
18. Hussain-Alkhateeb L, Kroeger A, Olliaro P, et al. Early warning and response system (EWARS) for dengue outbreaks: recent advancements towards widespread applications in critical settings. PLOS ONE 2018;13(5):e0196811. [doi: 10.1371/journal.pone.0196811] [Medline: 29727447]
19. Technical handbook for dengue surveillance, outbreak prediction/detection and outbreak response. World Health Organization. 2016. URL: https://www.who.int/publications/i/item/9789241549738 [accessed 2025-02-11]
20. Farrington P, Andrews N. Outbreak detection: application to infectious disease surveillance. In: Brookmeyer R, Stroup DF, editors. Monitoring the Health of Populations: Statistical Principles and Methods for Public Health Surveillance: Oxford University Press; 2003:203-231. [doi: 10.1093/acprof:oso/9780195146493.003.0008]
21. Global arbovirus initiative: preparing for the next pandemic by tackling mosquito-borne viruses with epidemic and pandemic potential. World Health Organization. URL: https://iris.who.int/bitstream/handle/10665/376630/9789240088948-eng.pdf?sequence=1 [accessed 2025-06-24]
22. Montgomery DC, Jennings CL, Kulahci M. Introduction to Time Series Analysis and Forecasting, 2nd edition: Wiley-Interscience; 2015.
23. Cascante-Vega J, Cordovez JM, Santos-Vega M. Estimating and forecasting the burden and spread of Colombia's SARS-CoV2 first wave. Sci Rep 2022 Aug 9;12(1):13568. [doi: 10.1038/s41598-022-15514-x] [Medline: 35945249]
24. Colón-González FJ, Soares Bastos L, Hofmann B, et al. Probabilistic seasonal dengue forecasting in Vietnam: a modelling study using superensembles. PLOS Med 2021 Mar;18(3):e1003542. [doi: 10.1371/journal.pmed.1003542] [Medline: 33661904]
25. Kache P, Ruiz-Carrascal D, Lowe R, et al. Climate extremes increase dengue risk along elevation and socio-economic gradients in Colombia. medRxiv. Preprint posted online on Apr 5, 2024. [doi: 10.1101/2024.04.02.24304484]
26. Gubler DJ. Epidemic dengue/dengue hemorrhagic fever as a public health, social and economic problem in the 21st century. Trends Microbiol 2002 Feb;10(2):100-103. [doi: 10.1016/s0966-842x(01)02288-0] [Medline: 11827812]
27. Gubler DJ. How effectively is epidemiological surveillance used for dengue programme planning and epidemic response. 2002 URL: https://iris.who.int/server/api/core/bitstreams/1f0456e3-6f06-484e-b994-ae8118ecc617/content [accessed 2026-04-28]
28. Harrington J, Kroeger A, Runge-Ranzinger S, O'Dempsey T. Detecting and responding to a dengue outbreak: evaluation of existing strategies in country outbreak response planning. J Trop Med 2013;2013:756832. [doi: 10.1155/2013/756832] [Medline: 24222774]
29. Umaña JD, Montenegro-Torres J, Otero J. Supporting code for the paper 'Endemic channel parametrization in dengue surveillance: a methodological assessment of retrospective windows, outbreak trends, and zero-case periods in Colombia. GitHub. URL: https://github.com/biomac-lab/endemic-channel [accessed 2025-11-24]

Abbreviations

EWT: epidemiological warning threshold
GAIDeT: Generative Artificial Intelligence Delegation Taxonomy
INS: Instituto Nacional de Salud
LME: linear mixed-effects
TRACE-LAC: Tools for Response, Analytics, and Control Enhancing of Epidemics in Latin America and the Caribbean
WHO: World Health Organization

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 02.Jul.2025; peer-reviewed by I marcilio, J Nejati; revised version received 02.Jan.2026; accepted 06.Jan.2026; published 08.May.2026.

Please cite as:
Umaña JD, Montenegro-Torres J, Otero J, Tavera-Cifuentes MC, Niño-Machado N, González-Uribe C, Cordovez JM, Santos-Vega M
Endemic Channel Parametrization in Dengue Surveillance: Methodological Assessment of Retrospective Windows, Outbreak Trends, and Zero-Case Periods in Colombia
JMIR Public Health Surveill 2026;12:e79914
URL: https://publichealth.jmir.org/2026/1/e79914
doi:10.2196/79914

© Juan D Umaña, Juan Montenegro-Torres, Julian Otero, Maria Camila Tavera-Cifuentes, Natalia Niño-Machado, Catalina González-Uribe, Juan Manuel Cordovez, Mauricio Santos-Vega. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 8.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Active Surveillance for COVID-19 Vaccine Safety Using Sequential Analysis in Korea: Population-Based Retrospective Observational Study

Na-Young Jeong[1], PhD; Haerin Cho[2], MS; Heehyun Won[2], MS; Suvin Park[2], BS; Joongyub Lee[3], MD, PhD; Hyesook Park[4], MD, PhD; Nam-Kyong Choi[2,5], PhD

[1]Health Science Convergence Research Institute, College of Science & Industry Convergence, Ewha Womans University, Seoul, Republic of Korea

[2]Department of Health Convergence, College of Science & Industry Convergence, Ewha Womans University, 52 Ewhayeodae-gil, Seodaemun-gu, Seoul, Republic of Korea

[3]Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea

[4]Department of Preventive Medicine, College of Medicine, Ewha Womans University, Seoul, Republic of Korea

[5]Graduate School of Industrial Pharmaceutical Science, College of Pharmacy, Ewha Womans University, Seoul, Republic of Korea

Corresponding Author:
Nam-Kyong Choi, PhD
Department of Health Convergence, College of Science & Industry Convergence, Ewha Womans University, 52 Ewhayeodae-gil, Seodaemun-gu, Seoul, Republic of Korea

Abstract

Background: With the advent of new vaccines, including the COVID-19 vaccines introduced during the recent pandemic, the need for near real-time active surveillance has increased to support timely regulatory decision-making.

Objective: This study aimed to assess the feasibility of sequential monitoring for potential adverse events following immunization in Korea, focusing on COVID-19 vaccines.

Methods: This population-based study used a linked database that combined the COVID-19 registry with national health insurance claims data. Participants included individuals older than 12 years who received either monovalent or bivalent COVID-19 vaccines in Korea between February 2021 and March 2023. Monthly retrospective sequential testing was performed for 3 prespecified outcomes (acute myocardial infarction, myocarditis, and anaphylaxis) as well as a negative control event (colonic diverticulitis). A Poisson-based maximized sequential probability ratio test was applied to compare postvaccination incidence rates with historical background rates, accounting for multiple testing and claims processing delays. Analyses were stratified by age group, vaccine platform, and dose.

Results: This study included over 43 million monovalent and 6.3 million bivalent vaccine recipients. Sequential analyses identified statistical signals for myocarditis following mRNA vaccines in individuals aged 12 to 64 years and protein subunit vaccines in those aged 40 to 64 years. Signals for anaphylaxis were observed following mRNA and nonreplicating viral vector vaccines in individuals older than 18 years. No safety signals were identified for acute myocardial infarction or colonic diverticulitis. Sequential monitoring detected signals for myocarditis and anaphylaxis before regulatory authorities took safety actions, with the earliest signals observed on September 30, 2021, and April 30, 2021, respectively.

Conclusions: Near real-time sequential testing detected statistical safety signals for myocarditis and anaphylaxis following COVID-19 vaccination. These signals were recognized by the regulatory authority as being associated with the vaccines, demonstrating the potential of this approach to detect signals requiring further causality assessments, particularly for newly introduced vaccines at an early stage.

(JMIR Public Health Surveill 2026;12:e75094) doi:10.2196/75094

KEYWORDS

vaccines; safety; active surveillance; near real-time monitoring; COVID-19 vaccine

Introduction

Active surveillance for vaccine safety is an ongoing and systematic process designed to comprehensively identify adverse events (AEs) in a population [1,2]. This approach plays a critical role in maintaining trust in vaccines by promptly providing data to evaluate potential associations between vaccinations and AEs, thereby supporting regulatory decision-making [3]. Active surveillance can be broadly classified into two types: (1) near real-time (NRT) sequential monitoring using electronic health

XSL•FO
RenderX

records (EHRs) or claims databases generated during routine health care services, and (2) cohort event monitoring, which collects primary data through individual surveys to assess postvaccination experiences [4]. NRT sequential monitoring is particularly effective for the early detection of vaccine safety signals when a new vaccine is introduced, a vaccine brand is changed, or the recommended target population is expanded [5]. Sequential monitoring has been widely conducted in various countries, including the United States and European nations, to evaluate the safety of various vaccines.

In the United States, NRT sequential monitoring is primarily performed by the Vaccine Safety Datalink (VSD) [1,6] under the Centers for Disease Control and Prevention (CDC) and by the Biologics Effectiveness and Safety (BEST) Initiative [7] under the Food and Drug Administration (FDA). Since the mid-2000s, the VSD has conducted rapid cycle analysis using weekly updated data files [8], and the results have been published for several vaccines, including seasonal influenza [9-11], H1N1 influenza [12], rotavirus [13], human papillomavirus [14,15], pneumococcal conjugate [16], and meningococcal conjugate vaccines [8]. More recently, sequential analyses have been applied to studies on COVID-19 vaccines to rapidly generate safety evidence following vaccine rollout [17-19]. Similarly, the post-licensure rapid immunization safety monitoring system [20], part of the FDA Sentinel Initiative, initially conducted sequential monitoring for influenza vaccines [21]. Subsequently, the BEST Initiative has employed sequential testing methods for active COVID-19 vaccine safety surveillance [22-24].

In Europe, the Accelerated Development of Vaccine Benefit-Risk Collaboration in Europe, a public-private partnership, assessed the feasibility of NRT monitoring for vaccine coverage, benefits, and risks [5]. Building on the methodologies and tools developed by the Accelerated Development of Vaccine Benefit-Risk Collaboration in Europe, the Vaccine Monitoring Collaboration for Europe leveraged EHRs from various European countries to actively monitor the postmarketing safety and effectiveness of COVID-19 vaccines [25]. Additionally, studies have evaluated the feasibility of sequential monitoring using national datasets, such as the Clinical Practice Research Datalink in the United Kingdom [26,27]. These systems facilitate safety monitoring during the early stages of new vaccine introductions and support ongoing monitoring of annual vaccination programs, thereby reinforcing public trust in vaccination.

Early detection of potential safety signals through NRT sequential monitoring enables timely decisions regarding the need for further comprehensive analyses. By using longitudinal observational data, such as EHRs or health insurance claims data, sequential monitoring comprehensively identifies safety signals in real-world populations, including those excluded from clinical trials [28]. In South Korea, studies using nationwide health insurance claims data have demonstrated the potential to detect postmarketing safety signals [29,30]. However, the feasibility of NRT monitoring of vaccine safety using these data has not yet been assessed. Furthermore, prior studies have not used real-world data to investigate whether sequential analysis can accurately identify potential AEs requiring further rigorous epidemiological assessment, or how quickly such events can be detected.

During the COVID-19 pandemic, the importance of early detection of safety signals became even more apparent due to limited preapproval safety evidence on vaccines. NRT sequential monitoring is particularly valuable for identifying risks when new vaccines are introduced, vaccine types or brands are updated, or vaccination eligibility expands. The results of such monitoring can help prioritize new epidemiological studies or inform regulatory actions. Given the ongoing development of various new vaccines, the need for NRT safety monitoring continues to grow. Therefore, this study aims to assess the feasibility of applying sequential testing methods using a nationwide linked database in South Korea. We conducted an exploratory analysis using recently introduced COVID-19 vaccination data. Additionally, the results of this sequential analysis were compared with known safety evidence and the timing of regulatory actions related to COVID-19 vaccine safety in Korea.

Methods

Settings and Study Population

This study used a nationwide database combining the COVID-19 registry from the Korea Disease Control and Prevention Agency (KDCA) and the health insurance claims data from the National Health Insurance Service (NHIS). Korea has a universal health insurance system that provides coverage for the entire population. Claims data are generated when health care providers deliver services to patients and subsequently file claims with the NHIS. The NHIS compiles a comprehensive database that includes information on insurance eligibility, premiums, medical services, and health care institutions [31]. This database encompasses records of inpatient, outpatient, and emergency department visits. Since the launch of the national COVID-19 vaccination program on February 26, 2021, the KDCA has managed vaccination records for the entire population [32]. The linked database includes information on COVID-19 vaccinations, laboratory-confirmed SARS-CoV-2 infections, diagnoses, prescribed medications (eg, drug codes, daily dosage, and duration), and health care procedures. It also contains timestamps for data recording, enabling us to track the dates when the data were accrued.

For this study, we used the KDCA COVID-19 registry data from February 26, 2021, to June 30, 2023, and NHIS health insurance claims data from January 1, 2017, to June 30, 2023. The study population consisted of individuals aged 12 years and older who received either monovalent or bivalent COVID-19 vaccines in Korea. Only individuals with Korean nationality who were enrolled in the national health insurance during the study period were included. Individuals participating in clinical trials for COVID-19 vaccines, those vaccinated abroad, or those with erroneous or duplicate vaccination records were excluded.

Exposure

The exposures of interest included all COVID-19 monovalent vaccines used for primary doses (ie, first and second doses) and

XSL•FO
RenderX

booster doses, as well as bivalent COVID-19 vaccines administered in Korea. Korea introduced 6 types of COVID-19 monovalent vaccines, including those based on mRNA, nonreplicating viral vector, and recombinant protein platforms (Table S1 in Multimedia Appendix 1). Our analysis focused on all primary and booster doses of monovalent COVID-19 vaccines administered from February 26, 2021, to September 30, 2022. During the 2022-2023 winter season, 4 types of mRNA bivalent vaccines were introduced in Korea (Table S2 in Multimedia Appendix 1). As part of the COVID-19 winter booster vaccination program, 1 dose of a bivalent vaccine was recommended regardless of the specific vaccine type. This study included all bivalent COVID-19 vaccines administered from October 11, 2022, to March 31, 2023.

To account for vaccine platforms and recommended dose schedules, we categorized the exposure groups based on the type of vaccines and dose number: mRNA monovalent vaccine (doses 1, 2, and 3), non-replicating viral vector monovalent vaccine (doses 1 and 2), recombinant protein–based monovalent vaccine (doses 1, 2, and 3), and mRNA bivalent vaccine doses. All analyses were stratified by platform and dose number.

Outcomes

This study selected outcomes of interest by considering both the associations between COVID-19 vaccines and AEs, and the timing of regulatory actions. The selection process involved reviewing the priority lists of adverse events of special interest (AESIs) for COVID-19 vaccines published by the Safety Platform for Emergency Vaccines [33], the VSD [17], and the BEST Initiative [34]. We also considered a list of AEs deemed causally or suspiciously related to COVID-19 vaccines by the KDCA (Table S3 in Multimedia Appendix 1). Through this process, 3 conditions were chosen as the outcomes of interest based on different causal association scenarios: anaphylaxis, myocarditis, and acute myocardial infarction (AMI). Each of these conditions was commonly listed as AESIs by the Safety Platform for Emergency Vaccines, VSD, and BEST Initiative. Anaphylaxis was considered an AESI that was causally associated with COVID-19 vaccines by the KDCA even before vaccine introduction, given that it can potentially be induced by all vaccine types [35]. Subsequent literature further supported this association [36-38]. Myocarditis was included in the AESI lists and, following the vaccine rollout, causal associations were recognized for mRNA and NVX-CoV2373 vaccines [39]. AMI, initially listed as an AESI, has not shown a definitive causal relationship with vaccination in postvaccination epidemiological studies.

In addition, a negative control event was selected—an event for which there is no evidence of a causal association with vaccination and that has not been flagged as a signal [40]. Colonic diverticulitis was chosen as the negative control event, as it was not included in COVID-19 AESI lists and is generally considered unrelated to vaccination. Moreover, this acute condition typically requires immediate treatment, making it unlikely to show significant variations in health care use between the prepandemic and pandemic periods [41].

Table S4 in Multimedia Appendix 1 presents the operational definitions for the selected AESIs and the negative control event. To maintain sensitivity and enable timely analysis for NRT sequential monitoring, operational definitions were based solely on diagnostic codes. The diagnostic codes and the lengths of risk periods were determined by referencing definitions employed by VSD [17] and BEST [34], as well as expert opinions from clinicians in the relevant medical fields.

Study Design

This study conducted monthly sequential testing to compare the background incidence rates of outcomes of interest with the incidence rates observed during the risk period following COVID-19 vaccination. The study design consisted of three steps: (1) estimation of background incidence rates for the outcomes of interest, (2) analysis of their incidence rates during the postvaccination risk periods, and (3) application of the sequential testing method. Sensitivity analyses were performed when a signal met the statistical threshold.

Sequential testing, which adjusts for multiple comparisons through repeated analyses over time, can rapidly identify potential safety concerns [1,42]. It generates a statistical signal as soon as there is sufficient evidence to reject the null hypothesis of no excess risk, thereby providing timely safety information that can enhance public confidence in vaccines [42]. Sequential testing methods are categorized into continuous sequential methods, which analyze data frequently (eg, weekly or monthly), and group sequential methods, which analyze data at discrete intervals after a specific amount of data has been accumulated [43]. This study adopted a continuous sequential method, known for its superior performance in using data aggregated over short intervals [44]. Among the continuous sequential methods, the maximized sequential probability ratio test (MaxSPRT), developed by VSD for sequential monitoring, is particularly suited for ongoing surveillance as it continues monitoring until the statistical threshold is reached or the maximum sample size is attained [44]. This study used Poisson MaxSPRT (PMaxSPRT) to compare the observed number of cases of AEs with the expected number derived from historical controls, especially for rare events during a large-scale vaccination campaign.

Statistical Analysis

Calculation of Background Rates for AESIs

Background incidence rates for AESIs, serving as historical controls, were estimated from the 2018 and 2019 incidence data. The prepandemic period was selected to avoid potential pandemic-related changes in health care use during the early phase of COVID-19.

The cohort for calculating background rates was defined as individuals with national health insurance eligibility as of January of each respective year. The cohort entry date was set as the later of either January 1 of the respective year or the date when the individual met the clean period criteria. The clean period was defined as the interval preceding cohort entry during which the AESI must not have occurred, serving as a requirement for inclusion in the study cohort. The specific duration varied depending on each AESI and negative control event. Person-time at risk for each individual was calculated from the cohort entry date to the earliest of the following: the

XSL•FO
RenderX

date of occurrence of the event, December 31 of the same year, or the date of death. If follow-up ended due to an event, individuals could re-enter the cohort if they subsequently met the clean period criteria (Figure S1 in Multimedia Appendix 1).

For each AESI, the annual incidence rates per 100,000 person-years with exact Poisson 95% CIs were calculated by age group. To enhance the sensitivity of the analysis, the lower annual incidence rate between 2018 and 2019 was used as the background rate for each AESI and age group. Monthly incidence rates were also calculated for each year to assess potential seasonality. This approach allowed for sensitivity analysis by using the average monthly background rate for the corresponding months in each sequential test.

Analysis of Incidence Rates During the Risk Period

Although this study retrospectively analyzed data, it was designed under the assumption of conducting sequential analyses starting from the initiation of COVID-19 vaccination. Considering that health care claims from medical care institutions are typically submitted on a weekly or monthly basis, the analysis frequency was set to monthly intervals. Analyses were conducted using data accrued to the database by the end of each month, starting from March 31, 2021. For eligible individuals exposed at each sequential analysis point, the incidence rates of AESIs during the risk period were calculated by age group, vaccine platform, and dose. Follow-up ended upon death or the analysis date, if it occurred before the risk period concluded (Figure S2 in Multimedia Appendix 1).

The follow-up period accounted for delays between the date of medical care and data accrual, which may vary by condition, clinical setting, and institution type. Delays in data accrual during the study period were evaluated using patterns from the prepandemic (January 2019-December 2019) and early pandemic (February 2020-June 2020) periods. For each AESI, monthly accrual rates of claims associated with diagnoses in a given month were calculated, and average delays were derived by analyzing their distribution across subsequent months. The period with higher accrual rates was used to minimize false signal detection. Claims were considered fully accrued once 95% of claims were collected following the month of diagnosis, and the entire follow-up period was used thereafter.

Although vaccination records may also be subject to data delays, they were not considered in this study, as individuals were followed from the recorded exposure date, minimizing the impact on the results. Figure S3 in Multimedia Appendix 1 illustrates the diagram for calculating the follow-up period using historical data on accrual delays.

Sequential analyses incorporated updated data at each testing point. Cases excluded in earlier analyses due to delays in data accrual could be included in subsequent analyses. Similarly, cases initially classified as incident AESIs could be reclassified as having a prior history based on the updated data.

Implementation of the Sequential Testing Method

Sequential analyses were conducted monthly following the initiation of the COVID-19 vaccination program, using the PMaxSPRT method. At each analysis point, the number of observed events during the post-vaccination risk period was compared to the expected number of events calculated based on background rates. The expected number of events for individuals vaccinated between month t and s was calculated by multiplying the number of person-days (A) during the risk period (from w d to T_i d for each individual i) by the background rate (θ) and adjusting for data delay by incorporating $P(s,t)$. To account for reporting delays arising from partially accrued data, we introduced an adjustment factor $P(s,t)$, defined as the proportion of AESIs that occurred in study month and were observed by month s [34]. The process used to adjust for data accrual delays is illustrated in Figure S4 in Multimedia Appendix 1.

$$(1)\sum t=1s\sum i=1nt\sum w=0TiAstiw\times\theta\times P(s,t)$$

Sequential analyses began on March 31, 2021, or when at least 1 cumulative case was observed within the follow-up period. Data accumulated by the last day of each month were analyzed. The relative risk (RR) and log-likelihood ratio were calculated to compare the likelihood of the observed RR to the null hypothesis. A statistical signal was declared if the log-likelihood ratio exceeded a prespecified critical value. Surveillance ended when one of the following occurred: (1) the statistical signal threshold was met, (2) observed events equaled expected events under the null hypothesis, or (3) the study period ended [42]. Expected event counts under the null hypothesis were estimated using the expected number of vaccinated individuals in the recommended age group within the resident population.

The null hypothesis (H_0: RR≤m) was tested using a one-tailed design with a significance level of 0.05. At least 2 cases were required to declare a signal. Threshold m values were determined based on severity and background incidence of each condition, proposed by Lloyd et al [22] (Table S5 in Multimedia Appendix 1). For colonic diverticulitis, m was set equal to that of anaphylaxis, a severe condition with a comparable background incidence that generally requires hospitalization or emergency department visits. Sensitivity analyses were conducted for all age groups, vaccine platforms, or doses that met the statistical signal threshold, adjusting for time-varying confounders such as seasonality. For outcomes that met the statistical threshold in sequential testing, the timing of the signal was compared with the timing and rationale behind safety measures implemented by the regulatory authority in Korea.

Analyses were performed using SAS Enterprise Guide version 7.1 (SAS Institute, Inc), R version 4.3.1 (R Foundation for Statistical Computing), and the R package *Sequential* version 4.3.3.

Ethical Considerations

This study was approved by the institutional review board of Ewha Womans University (ewha-202404-0014-01). Informed consent was not required due to the utilization of anonymized patient data.

XSL•FO
RenderX

Results

Baseline Characteristics

Table 1 summarizes the sociodemographic and vaccination characteristics of the Korean population who received COVID-19 vaccines during the study period, stratified by vaccine platform and dose number. Over 43 million individuals received monovalent vaccines, with mRNA vaccines being the most frequently administered across all dose levels. Among individuals aged 65 years and older, ChAdOx1 was the vaccine most commonly administered. Since NVV was recommended only for those aged 18 years or older, most individuals aged 12 to 17 years received other vaccine platforms. Recombinant protein–based vaccines, introduced later than other platforms, had fewer recipients. Nonreplicating viral vector vaccines for booster doses were excluded from sequential monitoring as they were not recommended for booster doses.

XSL•FO
RenderX

Table . Demographic and vaccination characteristics of COVID-19 vaccine recipients.

Characteristics	Type of vaccine platform		
	mRNA vaccine	Nonreplicating viral vector vaccine	Recombinant protein–based vaccine
Monovalent vaccine (dose 1)			
Total, n (%)	30,772,595 (100)	12,116,724 (100)	126,269 (100)
Sex, n (%)			
Male	15,173,824 (49.3)	6,220,454 (51.3)	55,755 (44.2)
Female	15,598,771 (50.7)	5,896,270 (48.7)	70,514 (55.8)
Age (y)			
Mean (SD)	42.8 (18.8)	59.0 (12.9)	46.9 (17.5)
12 - 17, n (%)	2,282,889 (7.4)	1 (0)	320 (0.3)
18 - 39, n (%)	11,365,058 (36.9)	1,544,309 (12.7)	51,105 (40.5)
40 - 64, n (%)	13,588,828 (44.2)	5,803,270 (47.9)	53,390 (42.3)
≥65, n (%)	3,535,820 (11.5)	4,769,144 (39.4)	21,454 (16.7)
Vaccine brand, n (%)			
Ad26.COV2.S	0 (0)	1,255,754 (10.4)	0 (0)
BNT162b2	24,290,565 (78.9)	0 (0)	0 (0)
ChAdOx1	0 (0)	10,860,970 (89.6)	0 (0)
GBP510/AS03	0 (0)	0 (0)	110 (0.1)
mRNA-1273	6,482,030 (21.1)	0 (0)	0 (0)
NVX-CoV2373	0 (0)	0 (0)	126,159 (99.9)
Monovalent vaccine (dose 2)			
Total, n (%)	32,172,217 (100)	9,034,713 (100)	103,456 (100)
Sex, n (%)			
Male	15,828,203 (49.2)	4,302,525 (47.6)	44,921 (43.4)
Female	16,344,014 (50.8)	4,732,188 (52.4)	58,535 (56.6)
Age (y)			
Mean (SD)	43.11 (8.4)	64.1 (9.0)	48.0 (17.5)
12 - 17, n (%)	2,157,466 (6.7)	0 (0)	42 (0)
18 - 39, n (%)	11,698,042 (36.4)	269,704 (3)	39,033 (37.7)
40 - 64, n (%)	14,698,914 (45.7)	4,177,738 (46.2)	45,336 (43.8)
≥65, n (%)	3,617,795 (11.2)	4,587,271 (50.8)	19,045 (18.4)
Vaccine brand, n (%)			
Ad26.COV2.S	0 (0)	0 (0)	0 (0)
BNT162b2	25,849,287 (80.3)	0 (0)	0 (0)
ChAdOx1	0 (0)	9,034,713 (100)	0 (0)
GBP510/AS03	0 (0)	0 (0)	0 (0)
mRNA-1273	6,322,930 (19.7)	0 (0)	0 (0)
NVX-CoV2373	0 (0)	0 (0)	103,456 (100)
Monovalent vaccine (dose 3)			
Total, n (%)	31,806,466 (100)	14,917 (100)	97,145 (100)
Sex, n (%)			
Male	15,753,046 (49.5)	11,189 (75)	44,237 (45.5)
Female	16,053,420 (50.5)	3728 (25)	52,908 (54.5)

XSL•FO
RenderX

Characteristics	Type of vaccine platform		
	mRNA vaccine	Nonreplicating viral vector vaccine	Recombinant protein–based vaccine
Age (y)			
Mean (SD)	51.9 (17.7)	45.8 (12.2)	48.6 (16.0)
12 - 17, n (%)	302,753 (1)	0 (0)	107 (0.1)
18 - 39, n (%)	7,762,384 (24.4)	6281 (42.1)	29,347 (30.2)
40 - 64, n (%)	15,847,317 (49.8)	7252 (48.6)	51,861 (53.4)
≥65, n (%)	7,894,012 (24.8)	1384 (9.3)	15,830 (16.3)
Vaccine brand, n (%)			
Ad26.COV2.S	0 (0)	14,820 (99.3)	0 (0)
BNT162b2	21,504,295 (67.6)	0 (0)	0 (0)
ChAdOx1	0 (0)	97 (0.7)	0 (0)
GBP510/AS03	0 (0)	0 (0)	56 (0.1)
mRNA-1273	10,302,171 (32.4)	0 (0)	0 (0)
NVX-CoV2373	0 (0)	0 (0)	97,089 (99.9)
Bivalent vaccine			
Total	6,393,809 (100)	—[a]	—
Sex, n (%)			
Male	3,236,617 (50.6)	—	—
Female	3,157,192 (49.4)	—	—
Age (y)			
Mean (SD)	63.3 (17.0)	—	—
12-17, n (%)	33,629 (0.5)	—	—
18-39, n (%)	697,503 (10.9)	—	—
40 - 64, n (%)	2,062,883 (32.3)	—	—
≥65, n (%)	3,599,794 (56.3)	—	—
Vaccine brand, n (%)			
BNT162b2 BA.1	878,184 (13.7)	—	—
BNT162b2 BA.4/5	3,393,011 (53.1)	—	—
mRNA-1273 BA.1	1,910,876 (29.9)	—	—
mRNA-1273 BA.4/5	211,738 (3.3)	—	—

[a]Not applicable.

Approximately 6.3 million individuals received bivalent vaccines, with the majority being aged 65 years or older, reflecting the prioritization of older adults, immunocompromised individuals, and high-risk groups. BNT162b2 BA.4/5 was the most frequently administered bivalent vaccine. Figures S5 and S6 in Multimedia Appendix 1 illustrate the weekly distribution of vaccinations by vaccine type and dose number for both monovalent and bivalent vaccines.

Calculation of Background Rates for AESIs

Table 2 presents the annual background rates of AESI derived from data for 2018, 2019, and 2020. For most AESI across all age groups, the background rates in 2020 were either statistically significantly lower or similar to those in the pre-COVID-19 period. This study used the lower incidence rate between 2018 and 2019 as the background rate for each AESI and age group.

XSL•FO
RenderX

Table . Background incidence rates of prespecified adverse events of special interest by age group.

Adverse events of special interest and age group	Pre-COVID-19 pandemic, incidence rates (95% CI)[a]		COVID-19 pandemic (2020[b]), incidence rates (95% CI)
	2018	2019	
Acute myocardial infarction			
Total	108.36 (107.43-109.29)	113.75 (112.81-114.70)	99.29 (98.36-100.23)
12–17 years	4.88 (4.14-5.76)	5.66 (4.85-6.61)	3.86 (3.17-4.70)
18–39 years	13.00 (12.45-13.58)	15.88 (15.26-16.52)	13.04 (12.46-13.66)
40–64 years	96.34 (95.01-97.68)	101.68 (100.32-103.05)	89.67 (88.34-91.02)
≥65 years	386.04 (381.59-390.54)	381.71 (377.44-386.04)	317.96 (313.95-322.03)
Myocarditis			
Total	3.00 (2.84-3.16)	3.50 (3.34-3.67)	2.57 (2.43-2.73)
12–17 years	8.74 (7.73-9.88)	8.44 (7.43-9.58)	3.94 (3.24-4.78)
18–39 years	1.93 (1.73-2.16)	2.75 (2.50-3.03)	1.66 (1.46-1.89)
40–64 years	2.26 (2.07-2.48)	2.68 (2.47-2.91)	1.93 (1.75-2.14)
≥65 years	4.97 (4.50-5.50)	5.32 (4.84-5.85)	5.32 (4.84-5.86)
Anaphylaxis			
Total	37.25 (36.71-37.81)	39.86 (39.30-40.44)	32.28 (31.75-32.81)
12–17 years	23.63 (21.93-25.46)	25.13 (23.36-27.05)	18.06 (16.50-19.78)
18–39 years	24.72 (23.95-25.51)	26.42 (25.62-27.24)	23.94 (23.14-24.77)
40–64 years	45.99 (45.08-46.92)	48.46 (47.53-49.41)	38.15 (37.30-39.03)
≥65 years	44.46 (42.99-45.97)	48.49 (47.00-50.03)	37.28 (35.95-38.66)
Colonic diverticulitis			
Total	44.34 (43.74-44.94)	47.45 (46.83-48.07)	46.02 (45.40-46.66)
12–17 years	6.47 (5.61-7.46)	7.73 (6.78-8.83)	6.48 (5.57-7.54)
18–39 years	31.32 (30.46-32.21)	34.86 (33.94-35.80)	34.05 (33.10-35.03)
40–64 years	51.29 (50.32-52.27)	54.46 (53.47-55.46)	52.55 (51.54-53.58)
≥65 years	66.60 (64.80-68.46)	67.43 (65.66-69.25)	64.34 (62.59-66.14)

[a]Incidence rate per 100,000 person-years.

[b]The background incidence rate for 2020 was calculated for individuals with eligibility information as of February of that year, considering the observation period from February 2020 to December 2020.

Incidence Rates During the Risk Period

Figures S6-S9 in Multimedia Appendix 1 present the distribution of delays in data accrual for claims made before the introduction of COVID-19 vaccination. For all AESIs, comparisons revealed that diagnoses from inpatient settings had faster data accrual than those from combined inpatient and outpatient settings (Figures S7-S10 in Multimedia Appendix 1). Consequently, the study used data delay distributions based on inpatient diagnostic claims. During the COVID-19 pandemic, over 95% of claims data were accrued faster compared to the prepandemic period. Thus, this study used data accrual delay distributions from inpatient diagnoses recorded between February 2020 and June 2020 to calculate the contribution of the risk period to sequential testing (Table S6 in Multimedia Appendix 1).

The person-time at risk and the number of events at each sequential testing point were calculated, and the expected number of events within the risk period was estimated using background rates. These calculations were performed separately for each age group, vaccine platform, and dose at each testing.

Implementation of the Sequential Monitoring Method

For AMI and colonic diverticulitis, none of the analyses stratified by age group, vaccine platform, or dose met the statistical signal threshold (Tables S7 and S8 in Multimedia Appendix 1).

For myocarditis, the statistical signal threshold was met for mRNA vaccines in the 12 - 17 and 18 - 39 age groups across all doses and for the first dose in the 40 - 64 age group. The threshold was also met for recombinant protein vaccines in the 40 - 64 age group for both the first and second doses. Significant results from sequential testing for myocarditis are summarized in Table 3, while the complete results for all age and dose groups are provided in Table S9 in Multimedia Appendix 1.

XSL•FO
RenderX

Table . Summary of significant sequential testing results for myocarditis.

Vaccine platform and dose	At data cutoff point				At statistical signal detection point				
	Person-days	Events	RR[a]	LLR[b,c]	Statistical signal	Date	Events	RR	LLR
12 - 17 years									
mRNA									
Dose 1	69,097,626	54	3.38	13.83	Yes	January 31, 2022	17	3.79	5.48
Dose 2	70,515,241	54	3.31	13.22	Yes	January 31, 2022	12	6.70	8.65
Dose 3	7,720,439	12	6.70	8.65	Yes	June 30, 2022	8	11.27	9.20
18 - 39 years									
mRNA									
Dose 1	349,949,622	146	7.89	124.16	Yes	September 30, 2021	11	6.92	8.20
Dose 2	375,637,912	115	5.79	70.13	Yes	September 30, 2021	7	9.33	6.92
Dose 3	233,868,158	32	2.59	3.99	Yes	February 28, 2022	11	5.21	5.87
40 - 64 years									
mRNA									
Dose 1	408,331,078	61	2.41	5.92	Yes	December 31, 2021	47	2.10	2.40
Recombinant protein–based									
Dose 1	1,676,909	3	30.00	6.14	Yes	June 30, 2022	2	22.22	3.53
Dose 2	1,386,915	3	33.33	6.44	Yes	June 30, 2022	2	40.00	6.64

[a]RR: relative risk.

[b]LLR: log-likelihood ratio.

[c]The log-likelihood ratio is artificially set to 0 when the relative risk is less than 1.

For anaphylaxis, the statistical signal threshold was exceeded for mRNA vaccines in the 18 - 39 age group across all doses and for the first and second doses in individuals aged 40 years and older. For nonreplicating viral vector vaccines, the threshold was met following the first and second doses in the 18 - 39 and 40 - 64 age groups, as well as after the first dose in individuals aged 65 years and older. A summary of significant sequential testing results for anaphylaxis is presented in Table 4, and detailed findings across all age and dose groups are available in Table S10 in Multimedia Appendix 1.

XSL•FO
RenderX

Table . Summary of significant sequential testing results for anaphylaxis.

Vaccine platform and dose	At data cutoff point				At statistical signal detection point				
	Person-days	Events	RR[a]	LLR[b,c]	Statistical signal	Date	Events	RR	LLR
18 - 39 years									
mRNA									
Dose 1	11,355,193	133	17.30	203.72	Yes	June 30, 2021	2	66.67	5.63
Dose 2	11,687,130	44	5.56	25.49	Yes	August 31, 2021	3	25.00	5.62
Dose 3	7,713,760	13	2.49	1.42	Yes	January 31, 2022	4	9.30	3.94
Nonreplicating viral vector									
Dose 1	1,543,045	43	40.95	100.77	Yes	April 30, 2021	9	900.00	48.59
Dose 2	269,520	4	27.32	9.79	Yes	July 31, 2021	2	20.00	3.33
40 - 64 years									
mRNA									
Dose 1	13,578,335	105	6.14	68.59	Yes	June 30, 2021	3	42.86	7.16
Dose 2	14,686,963	48	2.59	6.04	Yes	July 31, 2021	2	28.57	4.00
Nonreplicating viral vector									
Dose 1	5,799,204	64	8.76	59.87	Yes	April 30, 2021	3	42.86	7.16
Dose 2	4,173,320	11	2.09	0.55	Yes	July 31, 2021	2	28.57	4.00
≥65 years									
mRNA									
Dose 1	3,147,253	23	5.35	12.69	Yes	July 31, 2021	14	5.69	8.36
Dose 2	3,613,725	15	3.41	3.91	Yes	August 31, 2021	10	3.25	2.34
Nonreplicating viral vector									
Dose 1	4,759,853	16	2.76	2.45	Yes	July 31, 2021	7	3.65	2.10

[a]RR: relative risk.

[b]LLR: log-likelihood ratio.

[c]The log-likelihood ratio is artificially set to 0 when the relative risk is less than 1.

Sensitivity analyses were conducted for myocarditis and anaphylaxis to account for temporal trends in monthly incidence rates. For myocarditis, no differences were found between the sensitivity and primary analyses (Table S11 in Multimedia Appendix 1). For anaphylaxis, however, the sensitivity analysis showed that the signal threshold was no longer exceeded for the second dose of non-replicating viral vector vaccines in the 40 - 64 age group or the first dose in individuals aged 65 years and older (Table S12 in Multimedia Appendix 1).

Comparison of the Timing of Signal Detection and Implementation of Domestic Safety Measures

The timing of statistical signal detection using the sequential monitoring method was compared to the timing of safety measures implemented by the KDCA (Figure S11 in Multimedia Appendix 1).

For myocarditis, the KDCA initially did not recognize an association between COVID-19 vaccines and myocarditis during the early vaccination phase. However, international reports of myocarditis following mRNA COVID-19 vaccination and

XSL•FO
RenderX

reviews by global regulatory agencies prompted further investigation [45,46]. Based on domestic causality assessments, the KDCA announced on March 14, 2022, that myocarditis was causally associated with mRNA COVID-19 vaccination. Subsequently, on December 29, 2022, myocarditis was identified as a potential AE related to recombinant protein–based vaccines [39].

In this study, sequential testing detected statistical signals for myocarditis in different age groups after receiving mRNA vaccines. For the 12 - 17 and 18 - 39 age groups, the signals were detected in January 2022 and September 2021, respectively. In the 40 - 64 age group, signals were detected for mRNA vaccines in December 2021 and for recombinant protein–based vaccines in June 2022. The earliest signal was observed on September 30, 2021, for mRNA vaccines in the 18 - 39 age group.

For anaphylaxis, the KDCA recognized a causal relationship with COVID-19 vaccination from the beginning of the vaccination program. Sequential testing detected signals for nonreplicating viral vector vaccines in the 18 - 39 age group in April 2021 and for mRNA vaccines in June 2021. In the 40 - 64 age group, signals for nonreplicating viral vector vaccines were detected in April 2021, while mRNA vaccines reached the threshold in June 2021, after approximately 80,000 doses. Among individuals aged 65 years and older, signals for both nonreplicating viral vector and mRNA vaccines were detected in July 2021. The earliest signal was observed on April 30, 2021, for the first dose of nonreplicating viral vector vaccines in the 18 - 39 and 40 - 64 age groups.

Discussion

Principal Findings

This study conducted an exploratory analysis to evaluate the applicability of NRT surveillance for recently developed COVID-19 vaccines using a sequential testing method. Monthly analyses were performed using PMaxSPRT to compare background rates with postvaccination incidence rates, accounting for data accrual delays. All COVID-19 vaccines introduced in Korea were analyzed by vaccine platform, age group, and dose for 4 outcomes of interest.

The results showed that the statistical signal threshold for myocarditis was met in individuals younger than 64 years after mRNA vaccination and in those aged 40 to 64 years after recombinant protein–based vaccines. For anaphylaxis, the statistical signal threshold was met following mRNA and nonreplicating viral vector vaccines in individuals aged 18 years and older. No statistical signals were detected for AMI or colonic diverticulitis by the end of surveillance.

Comparison with prior studies revealed consistent findings for myocarditis and anaphylaxis. Lloyd et al [22] identified statistical signals for myocarditis, pericarditis, and anaphylaxis following COVID-19 vaccination in the US population aged 12 - 64 years. Similarly, studies analyzing BNT162b2 in adolescents aged 12 years and older also detected signals for myocarditis and pericarditis [23,24]. Wong et al [47], studying US adults aged 65 years and older, reported signals for AMI and other outcomes, but sensitivity analyses accounting for time trends and pandemic background rates showed no signals. These findings align with our study, which detected statistical signals for myocarditis and anaphylaxis but not for AMI.

This study compared the timing of statistical signals with the regulatory actions taken by the KDCA. In Korea, the KDCA delegated COVID-19 vaccine safety monitoring to CoVaSC, which made causality assessments prior to regulatory decisions [32]. Anaphylaxis was recognized as causally associated with COVID-19 vaccination early in the campaign. Sequential testing revealed statistical signals for nonreplicating viral vector vaccines in April 2021 and for mRNA vaccines in June 2021. Subsequent epidemiological studies confirmed the association between COVID-19 vaccines and incident anaphylaxis risk [36,39]. For myocarditis, causality was recognized for mRNA vaccines in March 2022 and for recombinant protein–based vaccines in December 2022 [39]. Sequential testing showed that statistical signals were first detected in September 2021 for mRNA vaccines and in June 2022 for recombinant protein–based vaccines. These findings suggest that NRT sequential testing could have facilitated earlier epidemiological evaluations of high-priority conditions using domestic data.

Sequential monitoring conducted early during the COVID-19 vaccination campaign in countries such as the United States demonstrated the rapid detection of safety signals when applying NRT methods. For instance, sequential testing in the US FDA surveillance program identified statistical signals for AMI, pulmonary embolism, disseminated intravascular coagulation, immune thrombocytopenia, myocarditis, and pericarditis after BNT162b2 vaccination. It also detected a signal for anaphylaxis after mRNA-1273 vaccination [22,47]. In a US pediatric population aged 5 - 17 years, sequential testing detected signals for myocarditis and pericarditis [24]. Initial VSD analyses in the United States did not detect signals for AESIs [17], but extended studies later identified signals for Guillain-Barré syndrome with Ad26.COV2.S and myocarditis or pericarditis in younger adults after mRNA vaccination [48,49].

Countries with established NRT sequential monitoring systems identified safety concerns relatively early in their vaccination campaigns. In contrast, Korea's vaccination campaign began later without a sequential monitoring system in place. As a result, regulatory actions largely relied on international findings, such as safety concerns about myocarditis raised abroad before vaccines were administered to Korean adolescents and young adults. This reliance may have influenced diagnosis coding practices. Cases resembling myocarditis might have been coded as myocarditis, which could have impacted studies based on claims data. Our study, which used a health insurance claims database for sequential analysis, may also have been influenced by these coding practices in analyzing myocarditis, one of our outcomes of interest. Such diagnostic awareness and heightened attention following international safety reports may have increased the likelihood of coding myocarditis, potentially inflating the observed signal strength. If an NRT monitoring system had been established before these issues were raised internationally, or if sequential analysis had been conducted globally at the onset of vaccinations, the time required to detect a myocarditis signal in our study might have been longer. In

XSL•FO
RenderX

that case, a lower relative risk might have been estimated. However, had sequential monitoring been implemented domestically, safety signals might have been detected earlier, as demonstrated by this study.

To implement NRT monitoring effectively within the country, further exploration of data sources capable of supporting NRT monitoring is needed. In this study, we used health insurance claims data and addressed data latency by calculating delay distributions based on recent historical data. One advantage of using domestic health insurance claims data is that they offer medical information for the entire population, allowing for the surveillance of even rare AEs. However, given the nature of claims data, inherent delays may occur during several stages of data processing. These include the provision of medical services, submission of claims by medical institutions, review procedures, and subsequent reimbursement by the NHIS. This process can take approximately 3 months or, in some cases, more than 6 months before the data becomes available. Therefore, to enhance NRT monitoring in the future, collaboration with sample hospitals might be considered to collect data based on EHRs, instead of claims data, which may accelerate the data accrual process. Integrating multiple data sources, such as NRT hospital data and EHRs, would allow timely validation of diagnostic codes and help mitigate potential diagnostic biases and misclassification. Furthermore, to prevent spurious signal detection driven by heightened diagnostic awareness, such as increased hospital visits and changes in coding practices observed for myocarditis, future surveillance systems should incorporate clinical validation processes. These should include case verification using detailed medical record review or laboratory findings. Linking diagnostic data with laboratory results or chart reviews could help distinguish genuine cases from those arising due to increased diagnostic activity.

Strengths and Limitations

To our knowledge, this is the first study in Korea to apply NRT sequential monitoring based on a nationwide vaccination registry and health care data. We used a data accrual delay adjustment method tailored to claims data. This approach enabled early safety surveillance shortly after vaccine rollout. However, this study has several limitations. First, while analyses were stratified by age group, vaccine platform, and dose—key factors influencing safety and timing of COVID-19 vaccination—the study did not account for sex differences. This was because vaccine recommendations and rollout timing differed across age groups. These factors were expected to have a greater impact on safety signal detection during the signal detection phase. Considering potential biological and behavioral differences between sexes, future studies should incorporate subgroup analyses by sex to identify potential variations. Second, as this study was designed as a feasibility assessment, the scope was intentionally limited to 3 AESIs and 1 negative control outcome. The selected outcomes were mostly diagnosed in Korea after being reported as AEs in countries that began vaccination earlier. Such international reports may have influenced health care–seeking behavior, potentially leading to overdiagnosis of conditions with similar clinical presentations. For instance, myocarditis cases in Korea were diagnosed predominantly after international reports, where heightened awareness may have contributed to an increase in the diagnosis of related conditions. Future research should expand sequential monitoring to a broader range of AESIs, such as Guillain-Barré syndrome or thrombosis with thrombocytopenia, to evaluate its applicability across diverse clinical presentations. Third, the operational definitions of AESIs relied solely on diagnostic codes for timely signal detection, which may limit diagnostic validity. For AESIs that met the statistical signal threshold in sequential testing, follow-up epidemiological studies with refined operational definitions are necessary to improve diagnostic accuracy. Fourth, the average data accrual rate, calculated from historical data, may not fully reflect actual delays or unexpected changes during the study period. However, this study minimized potential inaccuracies by using recent historical data. Further efforts should be made to continuously update and incorporate monthly accrual rates from the most recent data to enhance accuracy. Finally, although the sequential analysis controlled for repeated testing within each analysis stratum, no adjustment was made for multiple testing across all strata. As a result, the possibility of an inflated type I error rate at the overall study level cannot be entirely ruled out, and the findings should be interpreted as hypothesis-generating rather than confirmatory.

Conclusions

In conclusion, this study demonstrated that sequential monitoring can be effectively applied using Korean data sources in situations that require active surveillance, such as monitoring newly developed vaccines. Comparisons between signal detection timelines and KDCA regulatory actions indicated that NRT sequential monitoring could have enabled earlier recognition of safety concerns, such as anaphylaxis and myocarditis, thereby supporting more timely regulatory decisions. Future studies should explore alternative data sources, such as EHRs, to incorporate a broader range of clinical information and improve monitoring timeliness.

Acknowledgments

This study was conducted as part of the COVID-19 Vaccine Safety Research Center in the Republic of Korea. The authors sincerely thank the Korea Disease Control and Prevention Agency and the National Health Insurance Service for providing the database used in their study. Special gratitude also goes to Song Vogue Ahn and Minah Park for their insightful comments, which greatly contributed to the development of this manuscript. The authors declare the use of generative artificial intelligence in a limited capacity for language editing during manuscript preparation, under full human supervision.

XSL•FO
RenderX

Funding

This research was supported by a grant from the Korea Disease Control and Prevention Agency and by a grant from the Ministry of Food and Drug Safety (RS-2024-00332011). The funders played no role in study design, data collection, analysis and interpretation of data, or the writing of this manuscript.

Data Availability

The study used pseudonymized patient-level data from the Korea Disease Control and Prevention Agency and the National Health Insurance Service, which is protected by strict confidentiality under the Personal Information Protection Act. Access to personal data is granted solely for research purposes by the responsible authority upon submission of a formal application.

Authors' Contributions

NYJ, HC, and NKC were involved in the conceptualization and design of the study. NYJ conducted the statistical analyses with full access to the data and takes responsibility for ensuring the accuracy of the data analysis. NYJ, HW, SP, JL, HP, and NKC interpreted the data. NYJ and NKC wrote the first draft. All authors critically reviewed the manuscript and approved the final version of the manuscript. All authors had final responsibility for the decision to submit for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Exposure and AESI definitions, study design diagram, sequential testing results, and supporting tables and figures.
[DOCX File, 2605 KB - publichealth_v12i1e75094_app1.docx]

References

1. Yih WK, Kulldorff M, Fireman BH, et al. Active surveillance for adverse events: the experience of the Vaccine Safety Datalink project. Pediatrics 2011 May;127:S54-S64. [doi: 10.1542/peds.2010-1722I] [Medline: 21502252]
2. COVID-19 vaccines: safety surveillance manual. : World Health Organization; 2020 URL: https://www.who.int/docs/default-source/covid-19-vaccines-safety-surveillance-manual/covid19vaccines_manual_aefi_20210104.pdf [accessed 2026-03-21]
3. Bellitto C, Luxi N, Ciccimarra F, et al. What is the safety of COVID-19 vaccines in immunocompromised patients? Results from the European “COVID Vaccine Monitor” active surveillance study. Drug Saf 2024 Oct;47(10):1011-1023. [doi: 10.1007/s40264-024-01449-x] [Medline: 38907947]
4. Gee J, Shimabukuro TT, Su JR, et al. Overview of U.S. COVID-19 vaccine safety surveillance systems. Vaccine (Auckl) 2024 Sep 17;42:125748. [doi: 10.1016/j.vaccine.2024.02.065] [Medline: 38631952]
5. Bollaerts K, de Smedt T, McGee C, et al. ADVANCE: towards near real-time monitoring of vaccination coverage, benefits and risks using European electronic health record databases. Vaccine (Auckl) 2020 Dec 22;38 Suppl 2:B76-B83. [doi: 10.1016/j.vaccine.2019.08.012] [Medline: 31677951]
6. McNeil MM, Gee J, Weintraub ES, et al. The Vaccine Safety Datalink: Successes and challenges monitoring vaccine safety. Vaccine (Auckl) 2014 Sep 22;32(42):5390-5398. [doi: 10.1016/j.vaccine.2014.07.073] [Medline: 25108215]
7. Vaccines & allergenics. Biologics Effectiveness and Safety (BEST) Initiative. 2023. URL: https://bestinitiative.org/vaccines-and-allergenics [accessed 2025-02-15]
8. Lieu TA, Kulldorff M, Davis RL, et al. Real-time vaccine safety surveillance for the early detection of adverse events. Med Care 2007 Oct;45(10 Supl 2):S89-S95. [doi: 10.1097/MLR.0b013e3180616c0a] [Medline: 17909389]
9. Greene SK, Kulldorff M, Lewis EM, et al. Near real-time surveillance for influenza vaccine safety: proof-of-concept in the Vaccine Safety Datalink Project. Am J Epidemiol 2010 Jan 15;171(2):177-188. [doi: 10.1093/aje/kwp345] [Medline: 19965887]
10. Tse A, Tseng HF, Greene SK, Vellozzi C, Lee GM, VSD Rapid Cycle Analysis Influenza Working Group. Signal identification and evaluation for risk of febrile seizures in children following trivalent inactivated influenza vaccine in the Vaccine Safety Datalink Project, 2010-2011. Vaccine (Auckl) 2012 Mar 2;30(11):2024-2031. [doi: 10.1016/j.vaccine.2012.01.027] [Medline: 22361304]
11. Perez-Vilar S, Hu M, Weintraub E, et al. Guillain-Barré syndrome after high-dose influenza vaccine administration in the United States, 2018-2019 season. J Infect Dis 2021 Feb 13;223(3):416-425. [doi: 10.1093/infdis/jiaa543] [Medline: 33137184]
12. Lee GM, Greene SK, Weintraub ES, et al. H1N1 and seasonal influenza vaccine safety in the Vaccine Safety Datalink project. Am J Prev Med 2011 Aug;41(2):121-128. [doi: 10.1016/j.amepre.2011.04.004] [Medline: 21767718]

XSL•FO
RenderX

13. Belongia EA, Irving SA, Shui IM, et al. Real-time surveillance to assess risk of intussusception and other adverse events after pentavalent, bovine-derived rotavirus vaccine. Pediatr Infect Dis J 2010 Jan;29(1):1-5. [doi: 10.1097/INF.0b013e3181af8605] [Medline: 19907356]
14. Gee J, Naleway A, Shui I, et al. Monitoring the safety of quadrivalent human papillomavirus vaccine: findings from the Vaccine Safety Datalink. Vaccine (Auckl) 2011 Oct 26;29(46):8279-8284. [doi: 10.1016/j.vaccine.2011.08.106] [Medline: 21907257]
15. Donahue JG, Kieke BA, Lewis EM, et al. Near real-time surveillance to assess the safety of the 9-valent human papillomavirus vaccine. Pediatrics 2019 Dec;144(6):e20191808. [doi: 10.1542/peds.2019-1808] [Medline: 31740498]
16. Tseng HF, Sy LS, Liu ILA, et al. Postlicensure surveillance for pre-specified adverse events following the 13-valent pneumococcal conjugate vaccine in children. Vaccine (Auckl) 2013 May 24;31(22):2578-2583. [doi: 10.1016/j.vaccine.2013.03.040] [Medline: 23579258]
17. Klein NP, Lewis N, Goddard K, et al. Surveillance for adverse events after COVID-19 mRNA vaccination. JAMA 2021 Oct 12;326(14):1390-1399. [doi: 10.1001/jama.2021.15072] [Medline: 34477808]
18. Hause AM, Shay DK, Klein NP, et al. Safety of COVID-19 vaccination in United States children ages 5 to 11 years. Pediatrics 2022 Aug 1;150(2):e2022057313. [doi: 10.1542/peds.2022-057313] [Medline: 35581698]
19. Goddard K, Donahue JG, Lewis N, et al. Safety of COVID-19 mRNA vaccination among young children in the vaccine safety datalink. Pediatrics 2023 Jul 1;152(1):e2023061894. [doi: 10.1542/peds.2023-061894] [Medline: 37278199]
20. Baker MA, Nguyen M, Cole DV, Lee GM, Lieu TA. Post-licensure rapid immunization safety monitoring program (PRISM) data characterization. Vaccine (Auckl) 2013 Dec 30;31:K98-112. [doi: 10.1016/j.vaccine.2013.04.088] [Medline: 24331080]
21. Yih WK, Kulldorff M, Sandhu SK, et al. Prospective influenza vaccine safety surveillance using fresh data in the Sentinel System. Pharmacoepidemiol Drug Saf 2016 May;25(5):481-492. [doi: 10.1002/pds.3908] [Medline: 26572776]
22. Lloyd PC, Hu M, Wong HL, et al. Near real-time surveillance of safety outcomes in US COVID-19 vaccine recipients aged 12 to 64 years. Vaccine (Auckl) 2022 Oct 26;40(45):6481-6488. [doi: 10.1016/j.vaccine.2022.09.060] [Medline: 36195472]
23. Hu M, Shoaibi A, Feng Y, et al. Safety of ancestral monovalent BNT162b2, mRNA-1273, and NVX-CoV2373 COVID-19 vaccines in US children aged 6 months to 17 years. JAMA Netw Open 2024 Apr 1;7(4):e248192. [doi: 10.1001/jamanetworkopen.2024.8192] [Medline: 38656578]
24. Hu M, Wong HL, Feng Y, et al. Safety of the BNT162b2 COVID-19 vaccine in children aged 5 to 17 years. JAMA Pediatr 2023 Jul 1;177(7):710-717. [doi: 10.1001/jamapediatrics.2023.1440] [Medline: 37213095]
25. VAC4EU studies. Vaccine Monitoring Collaboration for Europe. 2023. URL: https://vac4eu.org/studies/ [accessed 2025-02-15]
26. Leite A, Andrews NJ, Thomas SL. Assessing recording delays in general practice records to inform near real-time vaccine safety surveillance using the Clinical Practice Research Datalink (CPRD). Pharmacoepidemiol Drug Saf 2017 Apr;26(4):437-445. [doi: 10.1002/pds.4173] [Medline: 28156036]
27. Leite A, Thomas SL, Andrews NJ. Implementing near real-time vaccine safety surveillance using the Clinical Practice Research Datalink (CPRD). Vaccine (Auckl) 2017 Dec 14;35(49 Pt B):6885-6892. [doi: 10.1016/j.vaccine.2017.09.022] [Medline: 29056424]
28. Suling M, Pigeot I. Signal detection and monitoring based on longitudinal healthcare data. Pharmaceutics 2012 Dec 13;4(4):607-640. [doi: 10.3390/pharmaceutics4040607] [Medline: 24300373]
29. Choi NK, Chang Y, Kim JY, Choi YK, Park BJ. Comparison and validation of data-mining indices for signal detection: using the Korean national health insurance claims database. Pharmacoepidemiol Drug Saf 2011 Dec;20(12):1278-1286. [doi: 10.1002/pds.2237] [Medline: 22114005]
30. Choi NK, Chang Y, Choi YK, Hahn S, Park BJ. Signal detection of rosuvastatin compared to other statins: data-mining study using national health insurance claims database. Pharmacoepidemiol Drug Saf 2010 Mar;19(3):238-246. [doi: 10.1002/pds.1902] [Medline: 20049851]
31. Cheol Seong S, Kim YY, Khang YH, et al. Data resource profile: the national health information database of the National Health Insurance Service in South Korea. Int J Epidemiol 2017 Jun 1;46(3):799-800. [doi: 10.1093/ije/dyw253] [Medline: 27794523]
32. Jeong NY, Park H, Oh S, et al. A framework for nationwide COVID-19 Vaccine safety research in the Republic of Korea: the COVID-19 Vaccine Safety Research Committee. Osong Public Health Res Perspect 2023 Feb;14(1):5-14. [doi: 10.24171/j.phrp.2023.0026] [Medline: 36944340]
33. Law B, Pim C. Third update—priority list of adverse events of special interest: COVID-19. Zenodo. 2022. URL: https://zenodo.org/records/6669790 [accessed 2026-03-21]
34. COVID-19 vaccine safety surveillance: active monitoring master protocol. : Center for Biologics Evaluation and Research (CBER), Office of Biostatistics and Epidemiology; 2021 URL: https://bestinitiative.org/wp-content/uploads/2021/02/C19-Vaccine-Safety-Protocol-2021.pdf [accessed 2026-03-21]
35. McNeil MM, DeStefano F. Vaccine-associated hypersensitivity. J Allergy Clin Immunol 2018 Feb;141(2):463-472. [doi: 10.1016/j.jaci.2017.12.971] [Medline: 29413255]
36. Yoon D, Kim JH, Jeong HE, et al. Risk of anaphylaxis after COVID-19 vaccination in South Korea: a nationwide self-controlled case series analysis. Allergy 2024 Mar;79(3):754-757. [doi: 10.1111/all.15975] [Medline: 38084802]

XSL•FO
RenderX

37. Boufidou F, Hatziantoniou S, Theodoridou K, et al. Anaphylactic reactions to COVID-19 vaccines: an updated assessment based on pharmacovigilance data. Vaccines (Basel) 2023 Mar 8;11(3):613. [doi: 10.3390/vaccines11030613] [Medline: 36992197]
38. Sobczak M, Pawliczak R. The risk of anaphylaxis behind authorized COVID-19 vaccines: a meta-analysis. Clin Mol Allergy 2022 Jan 17;20(1):1. [doi: 10.1186/s12948-022-00167-y] [Medline: 35039051]
39. Jeong NY, Park H, Oh S, et al. The COVID-19 Vaccine Safety Research Center: a cornerstone for strengthening safety evidence for COVID-19 vaccination in the Republic of Korea. Osong Public Health Res Perspect 2024 Apr;15(2):97-106. [doi: 10.24171/j.phrp.2023.0343] [Medline: 38621762]
40. Schuemie MJ, Arshad F, Pratt N, et al. Vaccine safety surveillance using routinely collected healthcare data-an empirical evaluation of epidemiological designs. Front Pharmacol 2022;13:893484. [doi: 10.3389/fphar.2022.893484] [Medline: 35873596]
41. Moll K, Lufkin B, Fingar KR, et al. Background rates of adverse events of special interest for COVID-19 vaccine safety monitoring in the United States, 2019-2020. Vaccine (Auckl) 2023 Jan 9;41(2):333-353. [doi: 10.1016/j.vaccine.2022.11.003] [Medline: 36404170]
42. Kulldorff M. Sequential statistical methods for prospective postmarketing safety surveillance. In: Strom BL, Kimmel SE, Hennessy S, editors. Pharmacoepidemiology, 5th edition: Wiley; 2012:852-867. [doi: 10.1002/9781119959946]
43. Silva IR, Kulldorff M. Continuous versus group sequential analysis for post-market drug and vaccine safety surveillance. Biometrics 2015 Sep;71(3):851-858. [doi: 10.1111/biom.12324] [Medline: 26011024]
44. Kulldorff M, Davis RL, Kolczak† M, Lewis E, Lieu T, Platt R. A maximized sequential probability ratio test for drug and vaccine safety surveillance. Seq Anal 2011 Jan 21;30(1):58-78. [doi: 10.1080/07474946.2011.539924]
45. Salvo F, Pariente A, Valnet-Rabier MB, et al. Role of spontaneous reporting in investigating the relationship between mRNA COVID-19 vaccines and myocarditis: the French perspective. Therapie 2023;78(5):509-515. [doi: 10.1016/j.therap.2023.03.003] [Medline: 37012151]
46. Meeting highlights from the Pharmacovigilance Risk Assessment Committee (PRAC) 3-6 May 2021. European Medicines Agency. URL: https://www.ema.europa.eu/en/news/meeting-highlights-pharmacovigilance-risk-assessment-committee-prac-3-6-may-2021 [accessed 2025-02-15]
47. Wong HL, Tworkoski E, Ke Zhou C, et al. Surveillance of COVID-19 vaccine safety among elderly persons aged 65 years and older. Vaccine (Auckl) 2023 Jan 9;41(2):532-539. [doi: 10.1016/j.vaccine.2022.11.069] [Medline: 36496287]
48. Hanson KE, Goddard K, Lewis N, et al. Incidence of Guillain-Barré syndrome after COVID-19 vaccination in the vaccine safety datalink. JAMA Netw Open 2022 Apr 1;5(4):e228879. [doi: 10.1001/jamanetworkopen.2022.8879] [Medline: 35471572]
49. Goddard K, Lewis N, Fireman B, et al. Risk of myocarditis and pericarditis following BNT162b2 and mRNA-1273 COVID-19 vaccination. Vaccine (Auckl) 2022 Aug;40(35):5153-5159. [doi: 10.1016/j.vaccine.2022.07.007]

Abbreviations

AE: adverse event
AESI: adverse events of special interest
AMI: acute myocardial infarction
BEST: Biologics Effectiveness and Safety
CDC: Centers for Disease Control and Prevention
EHR: electronic health record
FDA: Food and Drug Administration
KDCA: Korea Disease Control and Prevention Agency
MaxSPRT: maximized sequential probability ratio test
NHIS: National Health Insurance Service
NRT: near real-time
PMaxSPRT: Poisson maximized sequential probability ratio test
RR: relative risk
VSD: Vaccine Safety Datalink

XSL•FO
RenderX

Edited by TH Kwan; submitted 28.Mar.2025; peer-reviewed by B Han, C Asuzu, M Halim; revised version received 02.Dec.2025; accepted 04.Dec.2025; published 23.Apr.2026.

Please cite as:
Jeong NY, Cho H, Won H, Park S, Lee J, Park H, Choi NK
Active Surveillance for COVID-19 Vaccine Safety Using Sequential Analysis in Korea: Population-Based Retrospective Observational Study
JMIR Public Health Surveill 2026;12:e75094
URL: https://publichealth.jmir.org/2026/1/e75094
doi:10.2196/75094

© Na-Young Jeong, Haerin Cho, Heehyun Won, Suvin Park, Joongyub Lee, Hyesook Park, Nam-Kyong Choi. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 23.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Multimodal Data Approaches for Examining the 2024-2025 Highly Pathogenic Avian Influenza Outbreak in the United States: Descriptive Study

Juliana Sopko[1], MPH; Aimee R Han[1], MS; Jacqueline Powers[1], MPH; Jacquelin A Sauer[1], MPH; Mansi Avunoori[1], MPH; Stanislaw Zakrzewski[2], MS; Allison Krugman[3], MPH; Abhishek Dasgupta[4,5], DPhil; Kara Sewalk[1], MPH; Autumn Gertz[1], MS; Benjamin Rader[1,6], PhD; James Sheldon[7], MS; Brennan Klein[8,9], PhD; Jessica Malaty Rivera[1], MS; Moritz U G Kraemer[4,5], PhD; Samuel V Scarpino[7,8,10,11,12], PhD; John S Brownstein[1,13], PhD

[1]Computational Epidemiology Lab, Boston Children's Hospital, 401 Park Drive, 7th Floor West, Boston, MA, United States
[10]Department of Public Health and Health Sciences, Northeastern University, Boston, MA, United States
[11]Vermont Complex Systems Institute, University of Vermont, Burlington, VT, United States
[12]Santa Fe Institute, Santa Fe, NM, United States
[13]Harvard Medical School, University of Ngaoundéré, Boston, Cameroon
[2]Technical University of Lodz, Lodz, Poland
[3]Council on Foreign Relations, New York, NY, United States
[4]Department of Biology, University of Oxford, Oxford, United Kingdom
[5]Pandemic Sciences Institute, University of Oxford, Oxford, United Kingdom
[6]Harvard Medical School, Harvard University, Boston, MA, United States
[7]Institute for Experiential AI, Northeastern University, Boston, MA, United States
[8]Network Science Institute, Northeastern University, Boston, MA, United States
[9]Department of Physics, Northeastern University, Boston, MA, United States

Corresponding Author:
Juliana Sopko, MPH
Computational Epidemiology Lab, Boston Children's Hospital, 401 Park Drive, 7th Floor West, Boston, MA, United States

Abstract

Background: Highly pathogenic avian influenza (HPAI) A(H5N1) clade 2.3.4.4b, a globally predominant strain, was introduced into poultry in the United States in 2022 via spillover from wild birds, and has since been regularly reported, posing ongoing risks to animal and human health. In 2024, the United States reported the first known HPAI A(H5N1) clade 2.3.4.4b infection in dairy cattle, rapidly evolving into a multispecies outbreak among cattle and poultry, with spillover into humans. Publicly available data remained siloed and fragmented, hindering timely response. Innovative multimodal surveillance methods can enhance situational awareness through comprehensive, standardized data collection, integration, and visualization.

Objective: This study aimed to describe observations from the application of enhanced surveillance methods that collect, integrate, and visualize multimodal data for real-time tracking of the 2024‑2025 HPAI A(H5) outbreak in the United States as an innovative, transparent, repeatable, and scalable approach for open-source public health surveillance.

Methods: Global.health conducted real-time, multimodal surveillance of the United States 2024‑2025 HPAI A(H5) outbreak using publicly available data for human cases (Centers for Disease Control and Prevention), animal outbreaks (United States Department of Agriculture), wastewater monitoring (WastewaterSCAN), genomic data (public genomic databases), research updates (scholarly communication), and policy changes and response measures (media and government) for the study period from February 1, 2024 through February 28, 2025. This digital data stream was used to create outbreak resources—an epidemiological linelist, event timeline, and interactive map—using a One Health framework to track emerging hotspots.

Results: Global.health curated 70 confirmed human HPAI A(H5) cases across 13 states in a linelist, with exposure for nearly all (n = 65, 92.9%) cases associated with commercial agriculture and related operations. We curated 682 timeline entries across 6 distinct categories: human, cattle, response (eg, research, policy changes, and public health guidance), birds, genome, wastewater, and mammals. The map integrated human cases (n=70) and animal outbreaks (commercial cattle: n=977 and commercial poultry: n=325) into a single view. California was identified as the outbreak epicenter with high numbers of human cases (n=38, 54.3%), commercial cattle outbreaks (n=748, 76.6%), and commercial poultry outbreaks (n=66, 20.3%) during the study period. Wastewater surveillance detected the virus in California, with an unknown source at least 81 days before the first confirmed commercial dairy cattle case.

XSL•FO
RenderX

Conclusions: Global.health's approach for integrating traditional and nontraditional public health surveillance data within a One Health framework enhanced early situational awareness during the United States 2024 - 2025 HPAI A(H5) outbreak, creating open access to resources that improve contextual understanding of the scope and evolution of this emerging zoonotic event. Further research should seek to understand the full potential of multimodal data in outbreak surveillance.

(JMIR Public Health Surveill 2026;12:e86209) doi:10.2196/86209

KEYWORDS

one health; digital health; public health surveillance; zoonotic infectious diseases; zoonotic spillover; avian influenza; wastewater surveillance; data curation; data analysis; data visualization

Introduction

Background

The global spread of the highly pathogenic avian influenza (HPAI) virus (A/goose/Guangdong/96-lineage virus of the A(H5) subtype) has resulted in many different clades and genotypes due to genetic mutations, drift, and reassortment [1,2]. The H5Nx goose/Guangdong lineage in clade 2.3.4.4, circulating since 2013, has evolved into 8 different clade variants (2.3.4.4a-h), with clade 2.3.4.4b emerging as a notably fit virus with expanding geographic and host range and becoming widespread globally [3-5]. In early 2022, HPAI A(H5) clade 2.3.4.4b virus was detected in the United States poultry via spillover from wild birds and has since been regularly reported on US poultry farms, continuing to pose a threat to human and animal health [1,6-8].

Early in 2024, farmers in the United States began to report the emergence of a mysterious illness among their cattle herds. This cluster of symptoms was later linked to HPAI A(H5N1) clade 2.3.4.4b virus infection [9]. HPAI is better known for infecting wild and domesticated bird populations, with occasional spillovers into mammals and humans [10]. This appears to be the first time that HPAI (H5N1) has been identified in dairy cattle, aside from controlled infection experiments conducted in 2006 by the Centers for Disease Control and Prevention (CDC) [11-15]. On April 1, 2024, the Texas Department of State Health Services announced the first human case of HPAI A(H5N1). This was the first human case identified since the country's first and only previous confirmed human case of HPAI A(H5N1) in 2022 [16-18]. The individual was a dairy farm worker who had come in contact with infected cattle while at work. Genomic testing of the virus provided strong evidence that the farm worker had been infected via a spillover from dairy cattle and is believed to be the first mammal-to-human transmission of the HPAI H5N1 virus [14,17]. This spillover into humans marked the start of a multistate zoonotic outbreak affecting humans, cattle, poultry, and other mammals [19].

The emergence of a public health threat, such as a zoonotic outbreak, initiates a series of mitigation efforts. Within the United States, multiple local, state, and federal agencies are responsible for protecting and monitoring human, environmental, and animal health [20-22]. This often results in a complex and fragmented response and surveillance system for monitoring multispecies diseases such as HPAI. A key challenge arising from this surveillance landscape is that data collected by responding agencies are often siloed, which in turn can hamper timely analysis and response [23]. Barriers to implementing One Health surveillance systems—including data privacy, quality, and resource capacity—mirror longstanding challenges in the evolution of public health surveillance, highlighting the need for innovative methods to improve monitoring of zoonotic diseases [24,25].

Global.health was formed during the early stages of the SARS-CoV-2 pandemic in 2020 to address the challenge of integrating fragmented data during multiagency, multijurisdictional outbreak responses [23,26-28]. The technology powering Global.health is an open-source, scalable, and accessible platform for the rapid curation, analysis, and visualization of early outbreak data using a common data language for standardized database entry at a global scale [29]. The platform has been deployed in response to mpox, Ebola, and Marburg outbreaks, evolving to address the specific needs presented by each event [30-32]. Metadata associated with individual epidemiological cases can be of great importance to understand early disease dynamics, including origins, transmission dynamics, and key epidemiological parameters.

Study Objectives

To increase access to data and support One Health analysis of the 2024 - 2025 HPAI A(H5) outbreak in the United States, Global.health launched a data collection effort integrating human cases, animal outbreaks, wastewater surveillance, genomic data, research, response efforts, and policy actions. Accompanying this effort was the creation of enhanced data visualizations to improve outbreak insights. In this study, we examine Global.health's multimodal data collection, integration, and visualization efforts. Subsequently, we discuss observations regarding the spread of HPAI A(H5) in the United States during 2024 and 2025. Using information collected between February 2024 and February 2025, this paper highlights the value of integrating disparate datasets into infectious disease surveillance systems to improve situational awareness.

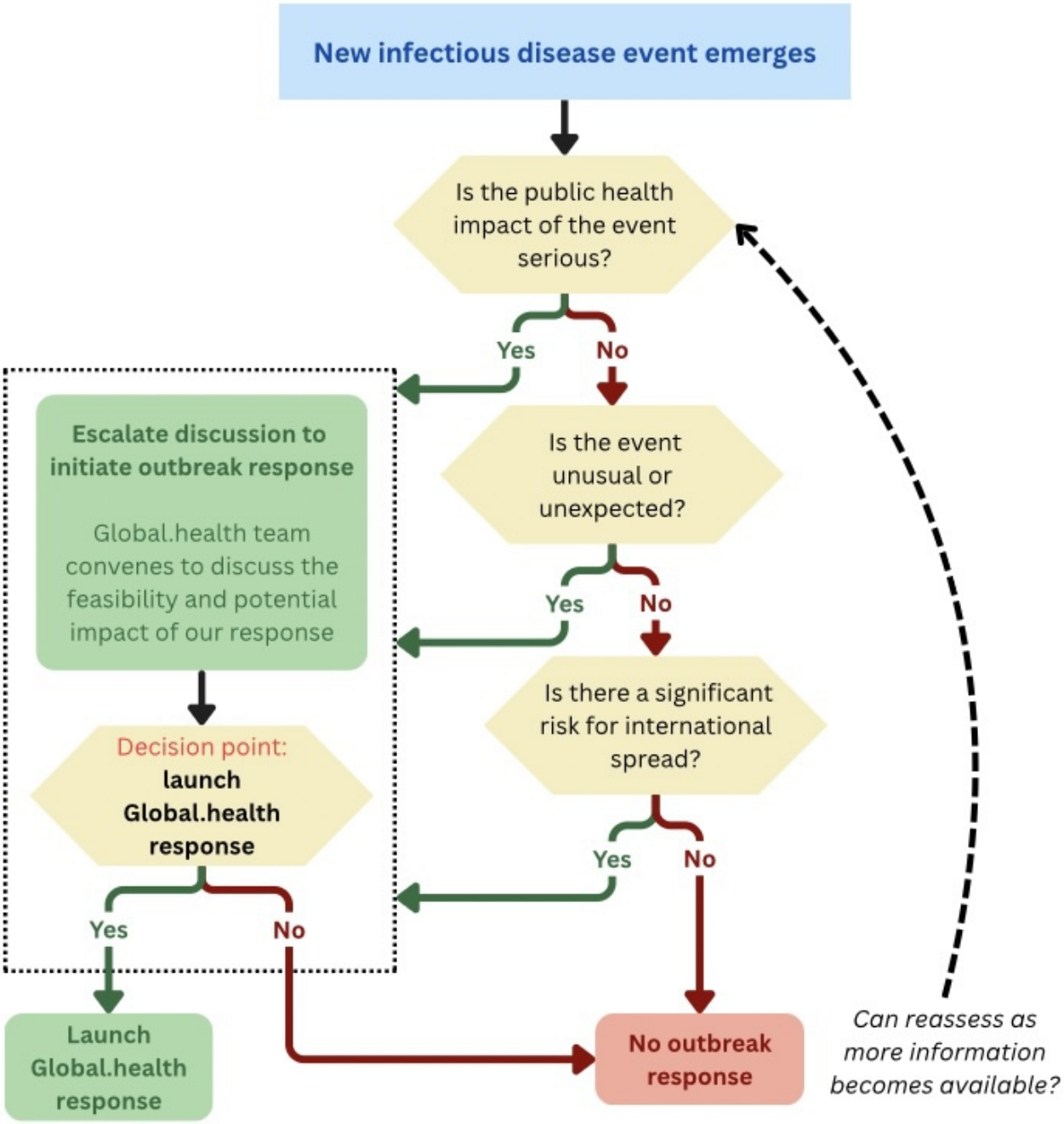
Methods

Outbreak Initiation

A set of established protocols are used to analyze emerging public health events and determine if they require the launch of Global.health outbreak response [29]. Early in 2024, the Global.health team initiated an outbreak response to the 2024 - 2025 HPAI A(H5) outbreak in the United States. The launch of an outbreak response triggers a cascade of planning and preparatory measures outlined in Figures 1 and 2.

XSL•FO
RenderX

Figure 1. Global.health decision instrument for the assessment of an emerging infectious disease event.



XSL•FO
RenderX

Figure 2. Outbreak response workflow. AWS: Amazon Web Services; CSV: comma-separated values.

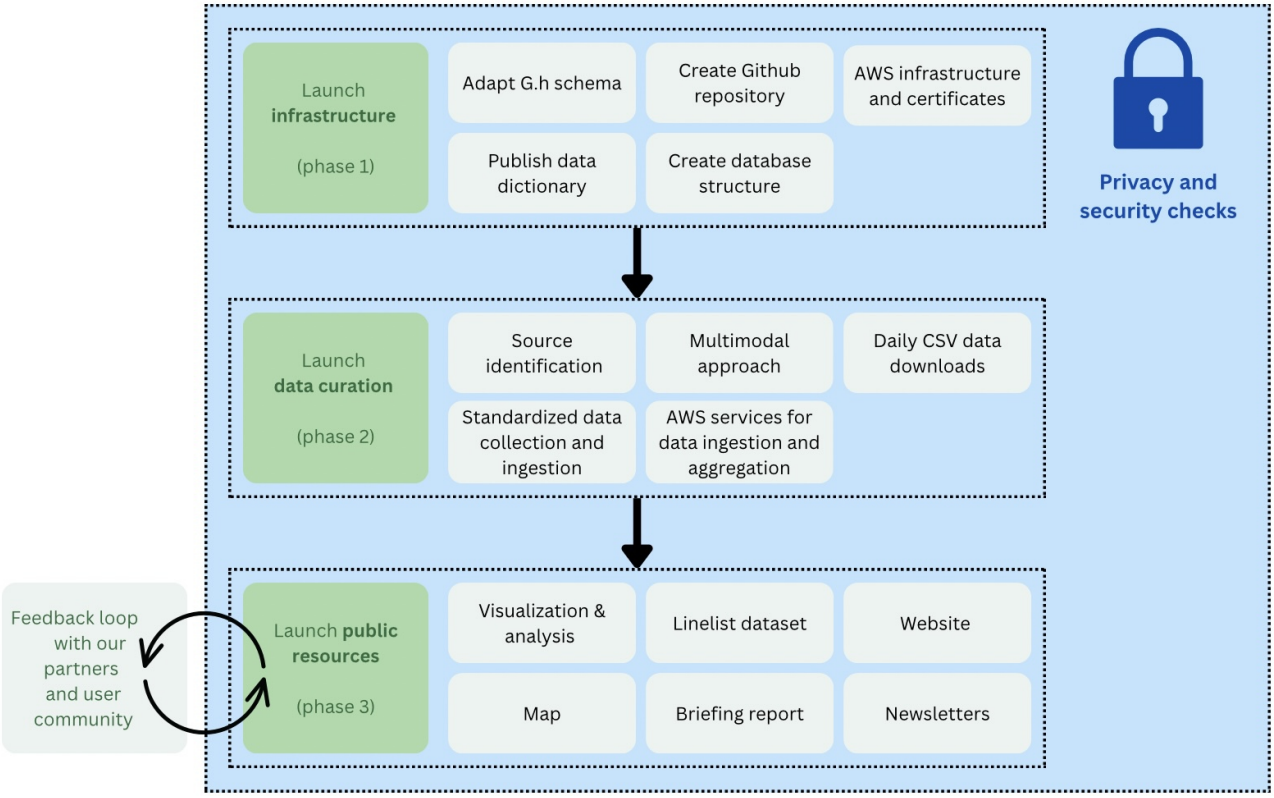


Figure 1 modeled after the World Health Organization's International Health Regulations (2005) Third Edition Annex 2, a cascade of yes or no binary questions are used to support informed decision-making [33]. A "yes" response to these questions will escalate internal discussion around the feasibility and potential impact of our work, considering timing, team resources, publicly available information, and current event coverage. This assessment is used to determine whether or not a new infectious disease event warrants the activation of a Global.health outbreak response. The outcome of this discussion reaches a decision point (yes or no) to launch the Global.health outbreak response. A "yes" decision will move forward with the launch of a Global.health outbreak response outlined in Figure 2. A "no" decision can be reassessed at any time as more information becomes available. Serious public health events include those with high case-fatality rates, limited access to effective prophylaxis or treatment, challenges in detecting the pathogen, spread in community and health care settings, and the need for an enhanced response at multiple levels. A subset of pathogens that Global.health considers as likely triggers for serious public health events are listed by the United Kingdom Health Security Agency as high-consequence infectious diseases [34]. Global.health may respond to other high-consequence infectious diseases such as mpox and measles, as well as vector-borne diseases such as dengue, chikungunya, and yellow fever.

Decision to launch a Global.health response triggers a triphased series of actions that first sets up the infrastructure (phase 1), then launches data curation efforts (phase 2), and later provides public resources to our global user community (phase 3). Examples of the types of activities or products created to support each phase are listed in the relevant section. A feedback loop is built into the model to support a transparent, crowdsourced approach to sharing outbreak information and to enable meaningful collaboration and fact checks through our platform. We consider relevant data privacy laws and regulations and use robust data protection measures and mitigation strategies, including deidentification, data validation, data quality checks, and access control.

Data Collection: Outbreak Linelist, Human Data

Global.health developed and maintains a common data schema used to structure infectious disease data. During a response, the Global.health outbreak schema is applied and may be adapted for novel data streams or pathogens to create the outbreak linelist that records granular epidemiological data pertaining to human cases in a standardized manner, building on lessons learned from COVID-19 data curation [35,36]. The outbreak schema is used for every outbreak (Multimedia Appendix 1) and can be modified to add new fields of interest. The schema format standardizes key epidemiological outbreak data that are deidentified. The availability of each data element varies by pathogen, location, and source, among other factors. For the 2024 - 2025 HPAI A(H5) outbreak, the Pathogen_name field was expanded to capture subtype detail for HPAI; Pathogen_subtype1 and Pathogen_subtype2 capture data for hemagglutinin and neuraminidase proteins, respectively. The genomics field was expanded to capture distinct information for sequence, clade, and genotype when that information was publicly available. The study period captures and reviews outbreak data in the United States between February 1, 2024, and February 28, 2025. Systematic data collection started in April 2024 and retrospectively added earlier events of significance. Research and data collection were conducted using

XSL•FO
RenderX

secondary data from publicly available, deidentified sources and did not involve human participants. This study followed the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) reporting guideline. Data collection is conducted by a team of curators who undergo extensive training for data collection and standardization. Curators accessed the same publicly available data directly from source websites. No restricted databases were used. Each entry is reviewed and validated by at least one additional curator before posting. Data cleaning included standardization of variable names, removal of duplicate entries, and reconciliation of conflicting reports across sources. No imputation of missing data was performed.

The outbreak linelist includes both confirmed and probable human cases of HPAI in the United States, but only confirmed cases are analyzed in this study. Confirmed cases are defined by the Council of State and Territorial Epidemiologists as those that have tested positive for H5 at a public health laboratory and have been confirmed by testing at the CDC [19,37]. Global.health's primary source for human case data was the CDC in addition to reports from local or state health departments.

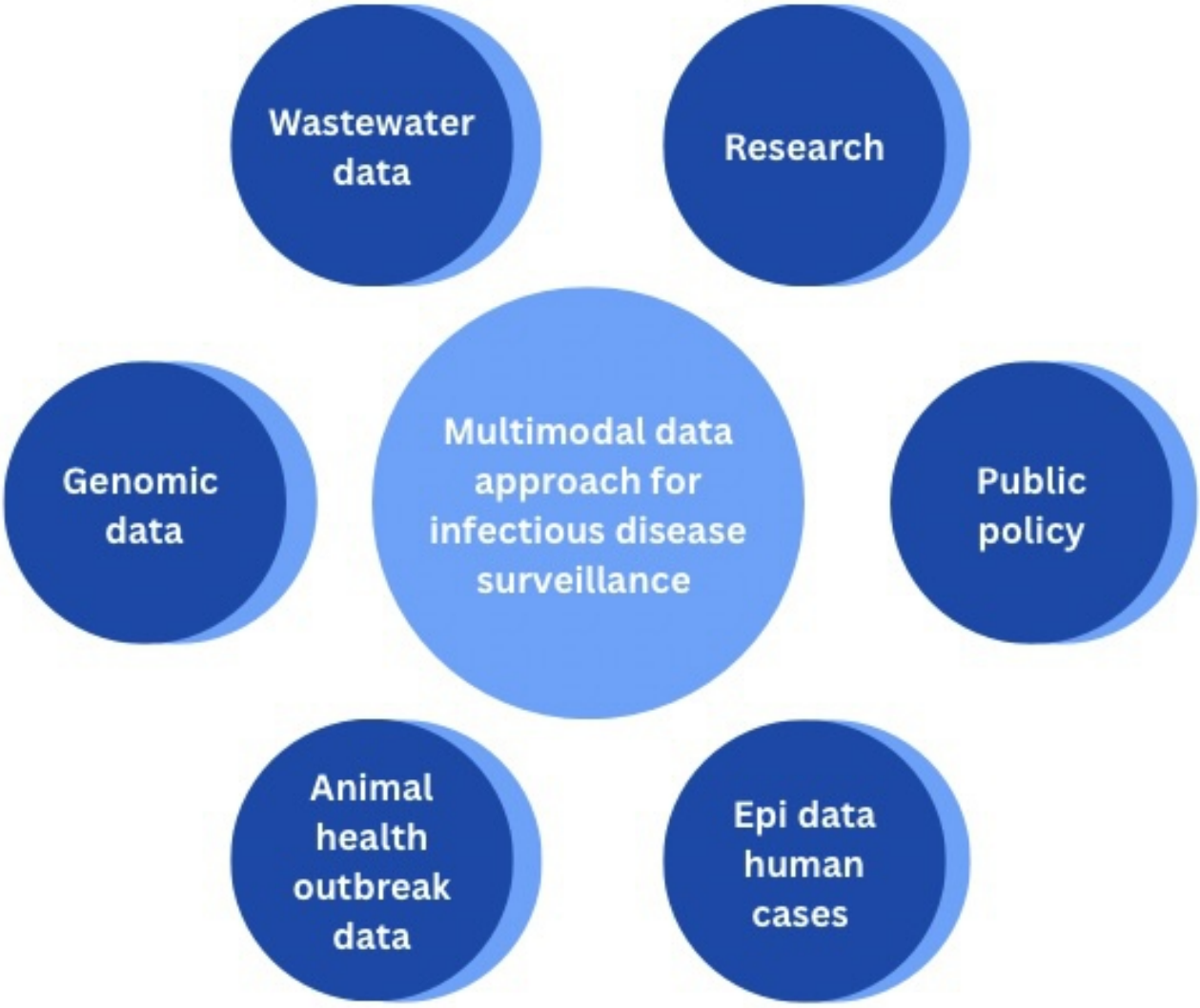
Exposure source information reflects what was publicly reported by the CDC and was not defined by the researchers. The exposure source can be one of four categories: dairy herds, poultry farms and culling operations, other animal exposure, and exposure source unknown [19]. For dairy herds and poultry farms, the exposure was associated with commercial agriculture and related operations [19]. For other animal exposure, the exposure was related to other animals such as backyard flocks, wild birds, or other mammals [19]. When the exposure source is unknown, it means the exposure source was unable to be identified following public health investigations [19].

Data Collection: Outbreak Timeline, Multimodal Data

Complementing the outbreak linelist and to help place the events in context, an outbreak timeline was created [38]. Because the outbreak linelist is exclusive to human cases, an outbreak timeline was developed that used multimodal data to enhance outbreak surveillance methods and weaved together publicly available, diverse data sources (outlined in Figure 3) into a single, integrated chronology that adopted a One Health approach, highlighting the interconnectedness of humans, animals, and our shared environment for this zoonotic outbreak [39]. The outbreak timeline includes entries related to the 2024 - 2025 HPAI A(H5) outbreak during the study period, February 1, 2024 through February 28, 2025. Systematic data collection started in April 2024 and retrospectively added earlier events of significance. Data collection for this event also extended beyond the study period but is not evaluated in this study.

XSL•FO
RenderX

Figure 3. A multimodal data approach for integrated infectious disease surveillance.



This figure outlines the types of sources collected to track the emerging H5 HPAI outbreak in the United States. Disparate data sources included human epidemiologic data, commercial cattle and poultry outbreak reports, wastewater detections, genomic sequences, research updates, response measures, and policy actions. The integrated view supports early situational awareness for zoonotic events.

Outbreak timeline entries were organized under four variables: date, event, read more (source link), and category. Each entry was tagged under one of the following categories for easy filtering and organization: human, cattle, response, birds, wastewater, genome, and mammals.

The “human” category included information related to HPAI infections in humans, including case counts and research updates. Our primary sources for human data were press releases and reports from the CDC and local and state health departments.

The primary source used to identify animal outbreaks was the United States Department of Agriculture (USDA) [40]. The USDA website reports HPAI in commercial and backyard flocks, wild birds, mammals, and livestock. For poultry outbreaks included in the “birds” category, using the USDA website, curators were able to ascertain the confirmation date, state and county location, production type, control area release, and the number of birds affected [8]. Only poultry or birds classified under commercial production, including the World Organization for Animal Health (WOAH) poultry category, were included in the descriptive analyses for this study. For livestock outbreaks in the “cattle” category, curators used the USDA website to ascertain confirmation date, state location, production type, and species [41]. Outbreaks among “cattle” species and “dairy milking cows” as the production type were included in these analyses. The “mammals” category provides updates on HPAI in other mammals besides cattle and humans.

For the collection of wastewater data, curators manually reviewed publicly available WastewaterSCAN reports from 192 wastewater testing sites across the United States [42]. WastewaterSCAN is an academic partner of the CDC’s National Wastewater Surveillance System (NWSS), and system data are collected and reported to the CDC [43]. WastewaterSCAN provided early, open-access to routinely tested wastewater surveillance data for the H5 subtype with results displayed on a public dashboard, updated daily, starting in May 2024 [42,44]. Filters were selected for pathogen (influenza), view by location, subtype (H5), locations (national, all wastewater sites, n=192).

XSL•FO
RenderX

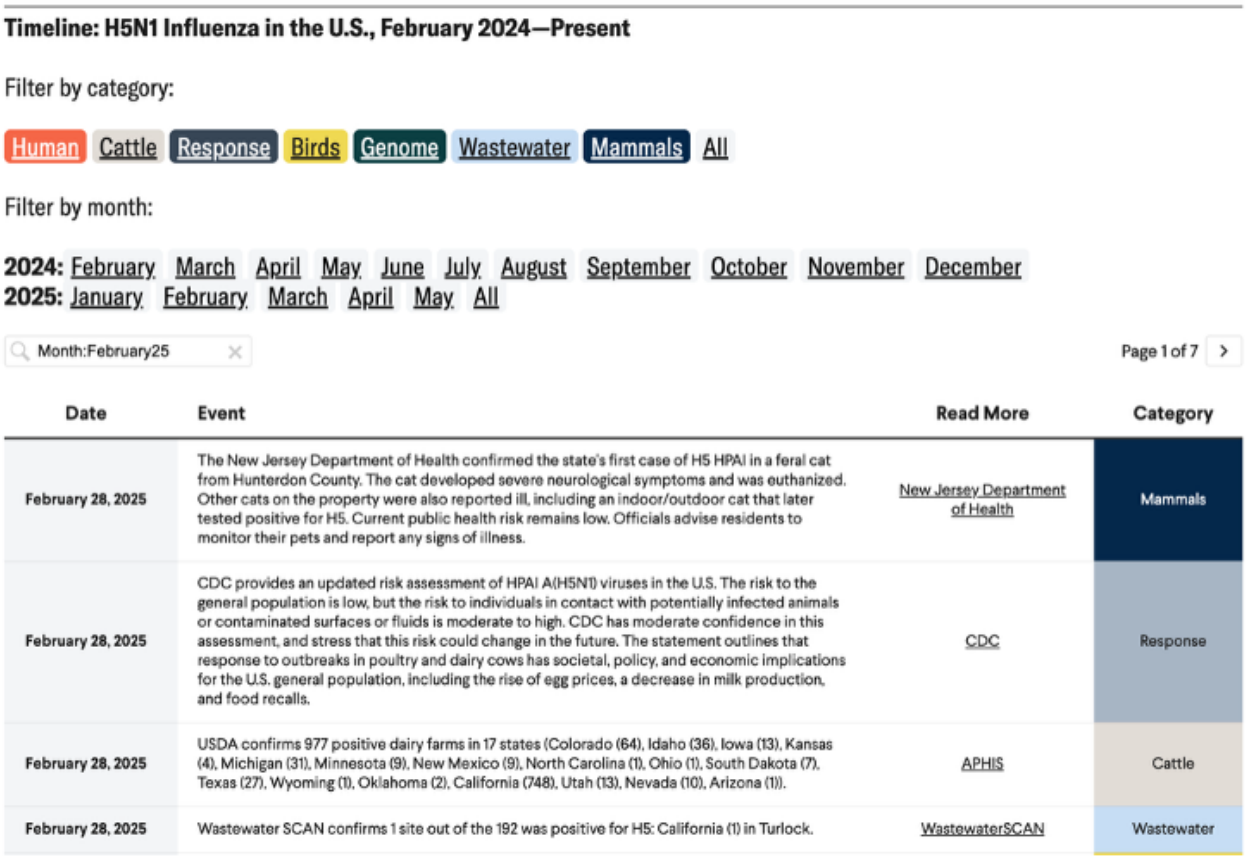
Data are available from WastewaterSCAN as of May 7, 2024, and are available for the remainder of the study period. Positive detections should be interpreted with consideration of the complexities of H5 in wastewater, primarily noting the inability to distinguish between human and animal sources (both domesticated and wild), as discussed in this research.

For the collection of genomic data, including clade and genotype information, sequence name, and accession number, the curation team gathered details from research and technical updates and information submitted to the Global Initiative on Sharing All Influenza Data virus genomic database or to GenBank, the genetic sequence database run by the US National Institutes of Health [45,46]. Data collection and reconciliation efforts for HPAI are ongoing.

The "response" category highlights response efforts, including policy mandates and changes to recommendations and guidelines as the outbreak evolved. This category also encompasses new research related to the response to the outbreak, such as the development of diagnostics and animal vaccines. Sources for "response" entries include news outlets, government agencies, preprint servers, and research journals.

Curators performed daily source checks, and all timeline entries were reviewed by at least one additional curator before posting to help ensure accuracy and consistency in reporting. Global.health collaborated with Think Global Health to publish a weekly version of the outbreak timeline (Figure 4) [38].

Figure 4. Outbreak timeline.



Visualizing Multimodal Data

The outbreak visualization map was created to accompany the outbreak linelist and multimodal outbreak timeline as publicly available resources [47]. Based on our development experience with maps from COVID-19, mpox (2022 and 2024), Ebola (Uganda 2022), and Marburg (Equatorial Guinea 2023), we used the Mapbox Boundaries and Mapbox GL JS library to create a state-level human HPAI map and later added new data overlays for county-level animal outbreaks in commercial cattle and poultry. Further visualization methods are described in Multimedia Appendix 2 [48].

Ethical Considerations

No institutional review board approval was applied for. Research was conducted using secondary data from publicly available sources that are deidentified. The research monitored public health signals without involving human participants and was designed to provide timely situational awareness and contribute to generalizable knowledge.

Results

Outbreak Linelist, Human Data

As of February 28, 2025, there were a total of 70 confirmed human cases of HPAI A(H5N1) reported in 13 states across the United States [5,19]. Only one confirmed human case of HPAI A(H5N1) had ever been documented in the United States before this outbreak, underscoring a sharp rise in incidence [10,17,49]. Included cases were publicly reported (eg, confirmed human HPAI cases meeting source-defined criteria for our setting and

XSL•FO
RenderX

study period). Data availability varied across sources. due to public reporting practices; no data linkage was performed. Nearly all confirmed human cases (92.9%) were from exposure associated with commercial agriculture and related operations. Among the confirmed human cases, 58.6% (n=41) occurred after exposure to infected commercial dairy herds. Additionally, 34.3% (n=24) resulted from exposure to infected commercial poultry farms and culling operations. Two human cases were exposed to other animals (such as backyard flocks, wild birds, or other mammals). For 3 cases, 1 Missouri resident and 2 Californian children, the source of exposure was unknown [19].

H5N1 was the only pathogen subtype that had been identified among these human cases. The sequencing of specimens showed that clade 2.3.4.4b of H5N1 was responsible for these infections, which was the same clade reportedly causing the outbreaks in cattle, poultry, and wild birds [5,50-52]. During the study period, of these sequenced human specimens where data were publicly available, 3 genotypes of clade 2.3.4.4b had been identified: B3.13, D1.1, and D1.3 [10,52,53]. There was no evidence of human-to-human transmission of the virus as of February 28, 2025 [19]. All confirmed human cases during the study period with a known exposure source had resulted from zoonotic, animal-to-human transmission. As of February 28, 2025, there had been only one death reported within the United States [19]. The case fatality rate for the US outbreak was 1.4%. Sixty-one confirmed human cases reported some kind of symptom associated with HPAI as of February 28, 2025. When specific symptoms were noted (Table 1), the 3 most common symptoms reported were fever, conjunctivitis or eye discomfort or watery discharge, and cough. Global.health findings align with those reported in publications in the *New England Journal of Medicine* and *Nature Medicine*, whose researchers analyzed the symptomatology of individuals with confirmed HPAI infections [54,55].

Table . Symptom variables and their respective definitions used in our clinical analysis of confirmed human HPAI cases. Definitions include alternate terms for variables, which were grouped to avoid duplicates. These variables were extracted from the human case line list dataset. These symptoms were reported among the 61 individuals with confirmed HPAI infections in the United States between February 1, 2024, and February 28, 2025.

Variable	Definition
pink_eye_conjunctivitis	“pink eye” “conjunctivitis” “eye discomfort” “eye redness”
watery_eyes	“watery discharge” “eye tearing” “drainage” “discharge”
subconjunctival_hemorrhage	“subconjunctival hemorrhage”
upper_respiratory_tract	“upper respiratory tract symptoms” “mild upper respiratory symptoms”
respiratory_symptoms	“common respiratory infection symptoms” “severe respiratory illness” “respiratory symptoms”
sore_throat	“sore throat”
runny_nose	“runny nose”
chest_pain	“chest pain”
nausea	“nausea”
weakness	“weakness”
vomiting	“vomiting”
diarrhea	“diarrhea”
nasal_congestion	“nasal congestion”
fatigue	“fatigue”
fever	“fever”
cough	“cough”

Table 1 outlines the symptom variables and their respective definitions used in our clinical analysis of confirmed human HPAI cases. Definitions include alternate terms for variables, which were grouped to avoid duplicates. These variables were extracted from the human case linelist dataset. These symptoms were reported among the 61 individuals with confirmed HPAI infections in the United States between February 1, 2024, and February 28, 2025.

Outbreak Timeline, Multimodal Data

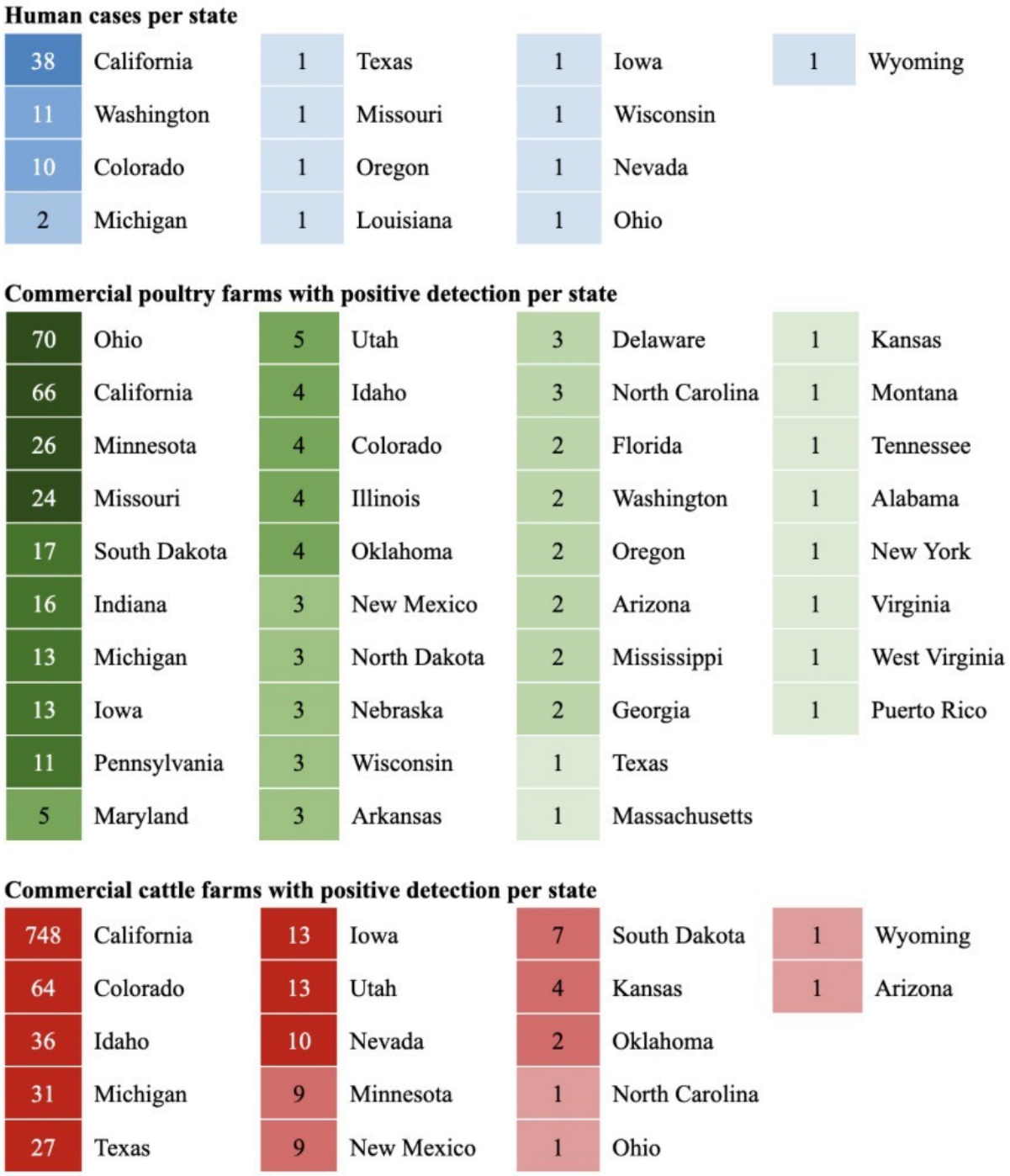
Early in the outbreak, the Global.health team partnered with Think Global Health, a digital publication by the Council on Foreign Relations supported by Bloomberg Philanthropies. Think Global Health is an online bulletin that examines the ways in which changes in health are reshaping economies, societies, and the everyday lives of people around the globe. Think Global Health published data curated by the Global.health team, creating the outbreak timeline (Figure 4) [38]. During the study period, the outbreak timeline was the 16th most viewed story on the Think Global Health website and recorded over 5800 page views from its publication through February 28, 2025. It also experienced periodic surges in traffic such as when it was first published in June 2024, with additional spikes in October 2024 and January 2025. The outbreak Timeline is accessible directly through the Global.health website, GitHub, and ThinkGlobalHealth.org [26,38,56].

XSL•FO
RenderX

Figure 4 outlines the outbreak timeline published on the Think Global Health website, which presents events related to HPAI in the United States between February 1, 2024, and February 28, 2024 [38]. The timeline allows readers to sort entries by category or month. Each entry provides the date, a brief description of the event, a source link for more details, and the assigned category.

As of February 28, 2025, the outbreak timeline data consisted of a total of 682 entries across all categories. Sixty entries in the "human" category included information related to HPAI infections in humans (eg, case counts and research updates). As case counts rapidly increased in the fall of 2024, some timeline entries reported multiple newly confirmed human infections on the same day, which is why the number of human infections exceeds the number of entries tagged in the "human" category. However, summing the case reports included in our 60 "human" entries provides the total number of confirmed cases reported in the United States. For animal outbreak data, there were 116 entries in the "birds" category, with the latest timeline entry capturing a total of 325 positive detections at commercial poultry farms across 37 states and 1 US territory (Figure 5). There were 191 entries under the "cattle" category, with the latest timeline entry capturing a total of 977 positive detections at commercial dairy farms across 17 states (Figure 5).

Figure 5. Highly pathogenic avian influenza (HPAI) detections at commercial poultry and cattle farms per state, 2024 - 2025.



XSL•FO
RenderX

Figure 5 outlines HPAI detections among humans, commercial poultry farms, and commercial cattle farms by state in the United States between February 1, 2024, and February 28, 2025. These charts depict the number of human cases (blue), commercial poultry farms with positive detections (green), and commercial cattle farms with positive detections (red) per state collected through a multimodal data approach. The state sums correspond to the total case reports collected in the timeline, human cases (n=70), commercial poultry farm outbreaks (n=325), and commercial cattle farm outbreaks (n=977). Shading indicates the number of outbreaks per state, where the shading gets darker for states with more detections.

A total of 189 entries in the "wastewater" category displayed wastewater detections across the country. The "wastewater" category displayed results from our observation of positive H5 detections using WastewaterSCAN's public dashboard in Michigan, Idaho, Texas, Minnesota, Iowa, California, South Dakota, Colorado, Arkansas, Florida, Nevada, Tennessee, Utah, Maine, Delaware, New Jersey, Illinois, Kansas, Connecticut, and Massachusetts as of February 28, 2025 [42].

The timeline also incorporated 31 entries in the "genome" category detailing genomic information about the outbreak and research related to the avian influenza genome. There were 29 entries in the "mammals" category that provided updates on HPAI detections in other mammals besides cattle and humans, such as domestic and feral cats, black rats, pigs, zoo animals (eg, harbor seals), house mice, deer mice, prairie voles, and desert cottontails. Sixty-six entries in the "response" category highlighted response efforts, including policy mandates, and changes to recommendations and guidelines as the outbreak evolved.

Visualizing Multimodal Data

The outbreak visualization map allows users to explore human cases and animal outbreaks in a single view, supporting a One Health approach to outbreak surveillance [47].

Figure 6 shows the spatial distribution of reported cases of HPAI in humans, commercial poultry, and commercial cattle outbreaks as of February 28, 2025 (confirmation dates were used to pinpoint cases in time). During the study period, there were 13 states with confirmed human cases, 17 states with infected commercial cattle, and 37 states plus 1 US territory with infected commercial poultry. Of the confirmed human cases, 41 were the result of exposure to commercial dairy herds (cattle), 24 were the result of exposure to commercial poultry farms and culling operations, 2 resulted from exposure to other animals (ie, other noncommercial birds, including backyard poultry), and 3 had an unknown exposure source. California emerged as a hotspot during the study period, reporting the highest numbers of human cases and commercial cattle outbreaks in the United States. Overall, 54.3% (n=38) of all human cases in the United States were reported in California, including 36 human cases associated with cattle exposure and 2 cases with an unknown exposure source [19].

Figure 6. Outbreak visualization map.

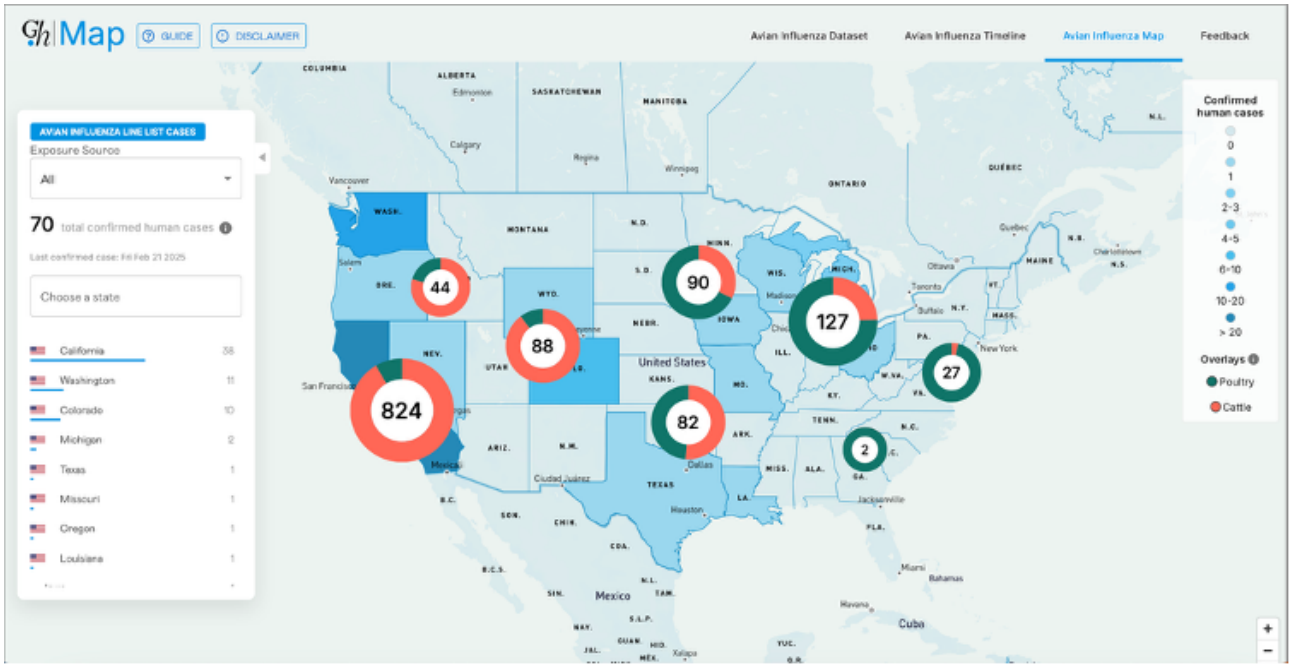


Figure 6 displays the outbreak visualization map with human and animal surveillance data compiled between February 1, 2024, and February 28, 2025. The map shows the geographic distribution of confirmed human cases and overlays corresponding poultry and cattle outbreak data in a single view to show emerging hotspots. Circle markers indicate regional- or state-level outbreak totals for each species, with the ability to filter and zoom for spatial granularity. A further description of the methods is described in Multimedia Appendix 2.

Case Study: California

This case study focuses on the temporal progression of HPAI A(H5) detections in public health surveillance data. We examined human health, animal outbreak (commercial cattle and commercial poultry), and wastewater surveillance data

XSL•FO
RenderX

reported during the study period from February 1, 2024 through February 28, 2025, filtered to events occurring in California. During the study period, we also recorded response actions, policy changes, and research updates in our outbreak timeline; however, these activities are not the focus of this case study and are acknowledged only for contextual completeness.

In April 2024, WastewaterSCAN developed and evaluated a test targeting the H5 hemagglutinin gene of influenza A for use in wastewater surveillance [44]. Routine H5 testing was implemented at sites nationwide in May 2024, with results publicly reported beginning in June [42,44]. On June 10, 2024, WastewaterSCAN reported the first positive H5 detection at a wastewater treatment site in California [42]. Additional positive wastewater detections were observed in July and September, representing early signals of viral presence in the state before confirmed infections in animals or humans [42], but notably not before the detection of HPAI in wild birds. Although the interpretation of wastewater detections is complex and does not allow attribution to a specific host, wastewater surveillance is recognized as a useful early-warning component of outbreak monitoring [44,57,58].

Eighty-one days after the first positive wastewater detection, on August 30, 2024, the California Department of Food and Agriculture confirmed the state's first HPAI detection in commercial dairy cattle in Sacramento County within the Central Valley [59].

Before the 2024‑2025 HPAI A(H5) outbreak, the USDA reported HPAI infections among commercial poultry in Fresno County, California, in August 2022 (outbreak reference: Fresno 01). From 2022 to 2024, periods of active detection were interspersed with intervals without reported cases, alongside ongoing efforts by poultry producers to strengthen biosecurity measures to protect flocks and prevent disease spread [8]. The state was subsequently self-declared free of HPAI in poultry on June 28, 2024, following completion of required response actions and surveillance [60]. On September 18, 2024, 100 days after the initial wastewater detection, the USDA confirmed a new HPAI outbreak in commercial poultry in Merced County (outbreak reference: Merced 12), marking the first detection since the state's HPAI-free declaration and resulting in the loss of that status [60].

On October 3, 2024, the California Department of Public Health announced the state's first confirmed human cases of HPAI A(H5) in 2 adult dairy farm workers with occupational exposure to infected cattle, occurring 115 days after the first wastewater detection [61]. No epidemiological link between the 2 individuals was identified, and there was no evidence of human-to-human transmission [61].

Positive H5 detections in wastewater continued throughout the remainder of the study period, with an increasing trend observed beginning in October 2024 and peaking between November 2024 and January 2025 [42]. Human infections among individuals with occupational exposure followed a similar temporal pattern, with confirmed cases peaking in October (n=16), November (n=12), and December (n=8) of 2024 [19]. Viral genome sequencing findings supported spillover from dairy cattle as the source of infection among affected farm workers, consistent with observations in other states [62]. However, in most cases, the definitive source of H5 detections in wastewater has not been determined.

By the end of the study period, California had reported 748 positive detections among commercial dairy farms and 66 positive detections among commercial poultry farms, illustrating rapid expansion of the outbreak within animal populations [8,41]. These trends are displayed in Figure 7.

XSL•FO
RenderX

Figure 7. Graph of key data surveillance events in the California outbreak of highly pathogenic avian influenza (HPAI) A(H5) (February 1, 2024, to February 28, 2025).

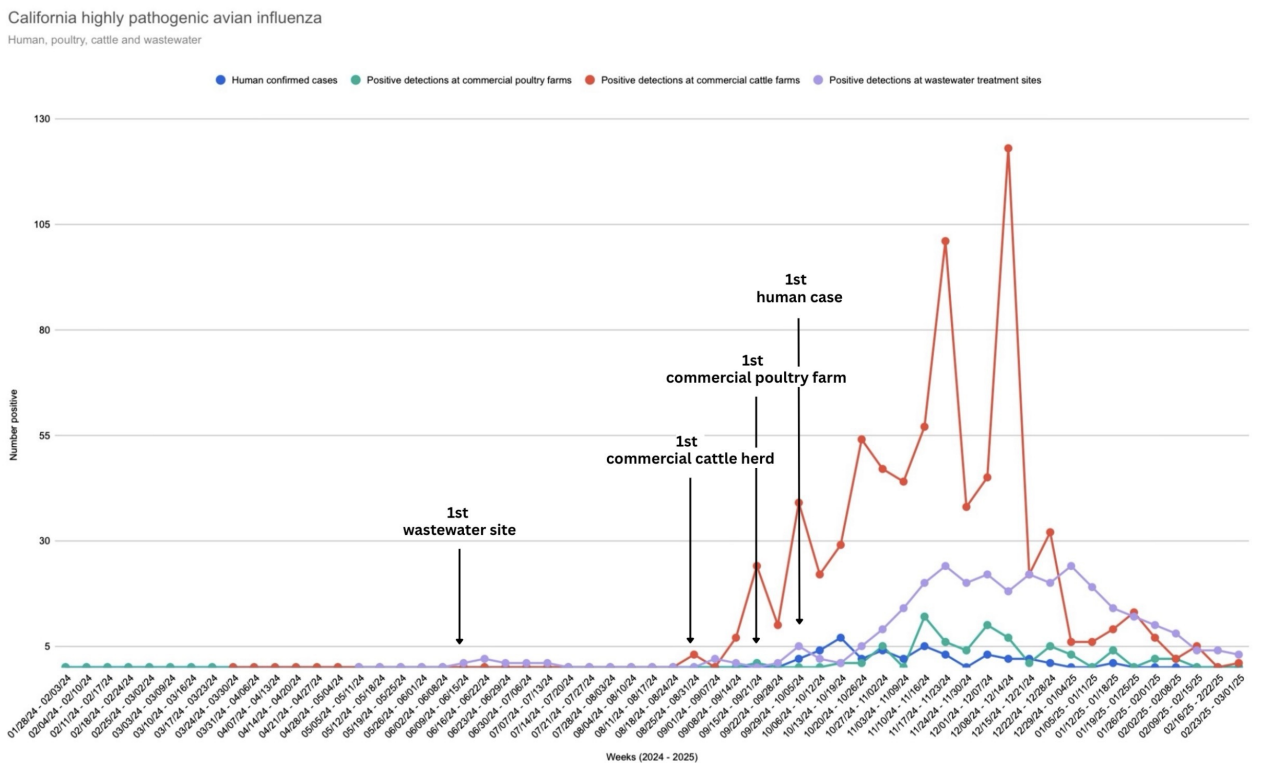


Figure 7 displays the time-series analysis used to track HPAI A(H5) activity across multiple data modalities in California. The dataset includes the number of positive human cases, detections at commercial poultry farms, detections at commercial dairy cattle farms, and wastewater treatment site detections. The study period is from February 1, 2024, through February 28, 2025, and dates were matched to CDC epidemiological weeks, ranging from January 28, 2024 (week 5, 2024) through March 1, 2025 (week 9, 2025). Data from WastewaterSCAN were not available before May 7, 2024. Cattle data from the USDA were not available before March 25, 2024. The USDA reports historical poultry outbreak data dating back to the introduction of HPAI in California in 2022. The first positive detection of H5 at a wastewater site in California occurred on June 10, 2024, followed by the first positive detection at a commercial cattle farm on August 30, 2024. There was also a positive detection at a commercial poultry farm on September 18, 2024, the first detection of HPAI in poultry since California was declared HPAI-free in June 2024. The first human cases of HPAI in California were announced on October 3, 2024, 115 days after the first positive detection at a wastewater site.

Discussion

Principal Findings

In this study, we examined Global.health's multimodal data collection, integration, and visualization efforts in response to the 2024‑2025 HPAI A(H5) outbreak using data collected in the United States between February 2024 and February 2025. The data collection efforts resulted in the development of publicly accessible resources to improve contextual understanding and track the outbreak, including the outbreak linelist dataset (human) and outbreak timeline (multimodal data collection: human, animal, wastewater, genomic, research, response, and policy). The data collected were also used to produce an outbreak visualization map (human and animal), which overlaid instances among humans, cattle, and poultry for enhanced data visualization and insight capabilities.

When disparate data points were examined holistically, in real-time—tracking human, animal, and wastewater data together—this approach helped identify emerging hotspots, highlighted outbreak trends, and improved situational awareness. California is examined in this study as an outbreak hotspot that quickly became an epicenter. Wastewater was an early indication of the virus in the state, with positive A(H5) detection of unknown source origin in June 2024 [42].

In the context of the 2024‑2025 HPAI A(H5) outbreak, the time between initial detection in wastewater and cases in animals and humans highlights an important window of opportunity for early response. Monitoring of wastewater data coupled with trend analysis, taking into account events such as annual bird migration, can provide insight into changes in trends that can point to emerging outbreaks. Monitoring of these trends within the context of contributing sources is important for the identification of zoonotic diseases with pandemic potential, such as HPAI [63]. Further, a coordinated and collaborative One Health approach to better understand the source could guide public health actions to hone methods and protocols for future respiratory illness seasons or outbreaks (eg, a positive H5 detection in wastewater may prompt further monitoring or testing in animals, milk, or humans; or identify the need for

XSL•FO
RenderX

deployment of wastewater testing to high-priority locations to focus on areas with a higher risk of disease transmission) [58].

This outbreak demonstrated key challenges and complexities with wastewater surveillance of HPAI noted in prior literature: an inability to distinguish between human and animal sources [64]; nonhuman and noncattle animal inputs, including wild and migratory birds [64,65]; industrial and agricultural inputs into wastewater [63]; stormwater contributions [64-66]; timely data integration to address knowledge gaps [58]; unknown thresholds for sensitivity in testing and positive detection [58]; current testing is not specific for H5N1 and may also detect low-pathogenic H5 influenza viruses [67]; and differences in laboratory approaches and reporting practices may complicate interpretation [58].

Although waste processing from industrial dairy farms in the United States is highly regulated and, as a result, most farms have internal systems such that waste is not designed to be dumped into municipal sewage systems, research has demonstrated that animal inputs have been responsible for some wastewater detections [44,63,68]. Dairy processing discharge into sewage systems has been studied with evidence to support the hypothesis of dairy industry contributions and the likelihood of multiple sources feeding into wastewater treatment plants [63]. Retrospective testing in other affected areas (Texas and North Carolina) has also shown that H5 was detectable in wastewater before the state identified H5-infected cattle, providing early warning for outbreaks that may contribute to the sewershed [63].

Further, the A(H5) virus is enzootic in wild birds, and animal sources should be considered when interpreting influenza subtype detections in wastewater [64]. Wild birds can be asymptomatic carriers of HPAI [6,69], and the movement and congregation of birds during migration, among all 4 US migratory flyways [70], and on a broader intercontinental scale, where birds may live or stop in different habitats [71], have been shown to be important factors in the geographic spread and ongoing circulation and phylogenetic changes of avian influenza viruses [52,72-77]. California is in the Pacific Flyway, a major migratory route for more than a billion visiting birds each year [78], creating risk for spillover events and point-source transmissions [72], where previous studies have aimed to model high-risk areas in the state [79,80]. The USDA conducts wild bird surveillance for avian influenza viruses of concern and reports positive detections of HPAI in wild birds and poultry in California before and during our study period [8,69]. However, wild bird surveillance data were not collected as part of this study. Wild bird detections are discussed in this context as a potential source contributing to the sewershed and subsequent positive detection in wastewater surveillance data [65]. Further research is needed to assess the geographic representativeness of site detections in California, including associations with migratory bird activity, seasonality of positive detections [81], historical avian influenza detections, and proximity to commercial agriculture.

While the interpretation of the virus in wastewater is complex, it is recognized as an important tool in the surveillance of emerging pathogens and has the potential to improve our understanding of outbreak dynamics and strengthen ongoing avian influenza surveillance efforts, with implications for animal and human health [44,64]. Wastewater data can complement traditional epidemiological data and existing health-monitoring systems [58,82], as demonstrated by Global.health's multimodal data approach to comprehensive disease surveillance.

Future work should focus on determining the origin of H5 in US municipal wastewater systems and include efforts to sequence viral genome fragments present in wastewater samples. Further, there is a critical need for clear communication to explain the meaning and interpretation of detection for influenza A virus and subtypes in wastewater.

The study demonstrates how the compilation of open-source multimodal data can provide early situational awareness and a more complete view of emerging zoonotic outbreaks. Open-source tools that are able to adapt and scale in real time as an outbreak evolves and presents new challenges, integrating new data sources for enhanced surveillance, are in line with broader calls to modernize pandemic intelligence by leveraging cross-sectoral and nontraditional data streams [83]. Global.health was able to quickly adapt the outbreak schema for the standardized collection of key information pertinent to this pathogen. In this study, we highlighted observations gained from the integration of data from official and nonofficial media sources. Our open-source format supports a transparent approach to information sharing, especially in the early days when there may be delays in official reporting, siloed and fragmented data, and data vacuums. This integration and methodology presents an innovative, valuable, and repeatable approach for open-source public health surveillance for zoonotic or other emerging or reemerging pathogens.

Limitations

Global.health strives to ensure that the data collated and shared on our platform are timely and accurate. In addition to the limitations associated with interpretation discussed in the Discussion section, (eg, the source of H5 signals in wastewater), there are important limitations associated with the data collected by Global.health during the 2024 - 2025 HPAI A(H5) outbreak. As part of quality assurance, data requires continuous reconciliation and validation across local, state, national, and international datasets to ensure alignment and to track changes. All Global.health data collection relies on publicly available data that are deidentified. Further research is needed to realize the full potential of open-source, multimodal data integration to support outbreak response and the impact it can have on public health as a whole.

Curators faced many challenges while building an emerging infectious disease dataset in real time, including delays in reporting for human cases, animal outbreaks, and wastewater detections; delays in confirmatory testing; limited data standardization within and across government organizations; lack of precision of virus nomenclature and terminology (eg, avian influenza, bird flu, AI, HPAI, HPAI A(H5), HPAI A(H5N1), HPAIV, H5, H5Nx, H5N1, and H5-type being used inconsistently and interchangeably in media sources and scientific literature); presentation of aggregated case data; limited case metadata; and retroactive removal and addition of

XSL•FO
RenderX

impacted herds, flocks, and wastewater sites. USDA reports confirmed outbreak events by date and location using outbreak reference numbers (also referred to as special IDs). Global.health conducted regular data reconciliation efforts to mitigate retroactive, unannounced, and unexplained changes and data discrepancies between Global.health and our primary sources (WastewaterSCAN and USDA). The average look-back period was 1 week for Global.health to check and update data and was limited by time and available resources. USDA published the total number of cattle outbreaks on their website, and we were able to align counts on a daily basis. However, cumulative numbers were not available from USDA (Poultry) or WastewaterSCAN. Therefore, publicly available data from these sources may differ from the results of this study. These challenges reflect longstanding gaps in US surveillance systems; however, advanced tools and approaches create opportunities to integrate data not traditionally accessible to public health [84]. Innovative methods, such as artificial intelligence and automated quality control tools, may also help enhance data quality in real time [85,86]. To ensure data transparency, all outbreak resources are publicly accessible and point back to the original sources of information.

Primary sources for multimodal data collection are listed for this research. Additional relevant sources may be available, but were not included in this study.

Conclusions

As outbreaks and the public health landscape evolve, so too must surveillance methods; how data are gathered, analyzed, and visualized matters. Further, the 2024 - 2025 HPAI A(H5) outbreak in the United States again highlighted the importance of a One Health approach as the pandemic potential of zoonotic outbreaks grows [87]. Focusing on the interplay of humans, animals, and the shared environment will assist researchers, public health teams, and communities in better understanding the origins, evolution, and scope of emerging zoonotic events. Our data collection efforts and research showed both the value and the challenges of integrating disparate sources to modernize surveillance methods and create public resources that enhance early situational awareness. The potential for wastewater surveillance, and other emerging data modalities, to provide an early indication of disease in a community calls for increased research into how to interpret positive detections in the context of multimodal datasets. Accurate, accessible, real-time data are the cornerstone of prudent outbreak response. Early, open access to data and visualization tools, such as those put forward by Global.health, can demonstrate value in a public health response approach that is adaptable and resilient in the face of emerging threats.

Funding

This work was supported by the Rockefeller Foundation (grant PC-2022-POP-005) and the Wellcome Trust (grant 228186/Z/23/Z). The funding sources had no involvement in the study design, data collection, data analysis, data interpretation, writing of the report, or submission decisions.

Data Availability

Data collected for this study will be made available upon publication. Our data products include the linelist dataset (human cases), data dictionary, interactive map, and event timeline. No additional related documents will be available. Data can be accessed through the Global.health website. Global.health commits to releasing code and software tools under the MIT License and data, where possible, under the CC BY 4.0 license. Both the University of Oxford and Boston Children's Hospital serve as data controllers. If you choose to access, download, or use Global.health data, we ask that you adhere to our terms of use, privacy policy, and citation guidelines provided on the Global.health website.

Authors' Contributions

JS, AH, JP, SZ, AK, MUGK, and SVS contributed to conceptualization. JS, AH, JP, MA, SZ, AK, and SVS contributed to data curation. JS, AH, JP, MA, and SZ contributed to formal analysis. JS, AH, JP, and SZ contributed to investigation. JS, AH, JP, MA, SZ, AK, AD, KS, AG, BR, BK, MUGK, SVS, and JSB contributed to methodology. JS, AH, JP, MA, and SZ contributed to visualization. JS, AH, JP, and SZ contributed to writing – original draft. JS, AH, JP, MA, SZ, KS, AG, BR, JMR, MUGK, SVS, and JSB contributed to writing – review and editing. JS, AH, KS, and JSB contributed to project administration. JSB contributed to resources. JS and JSB contributed to supervision. MUGK, SVS, and JSB contributed to funding acquisition.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Schema
[PDF File, 220 KB - publichealth_v12i1e86209_app1.pdf]

Multimedia Appendix 2
Methods for visualizing multimodal data.

XSL•FO
RenderX

[DOCX File, 15 KB - publichealth_v12i1e86209_app2.docx]

Checklist 1
RECORD checklist.
[DOCX File, 1825 KB - publichealth_v12i1e86209_app3.docx]

References

1. Hatta Y, De La Cruz JA, Murray T, et al. Highly pathogenic avian influenza A(H5N1) Clade 2.3.4.4b virus infection in poultry farm workers, Washington, USA, 2024. Emerg Infect Dis 2025 Dec;31(12):2297-2301. [doi: 10.3201/eid3112.251118] [Medline: 41490523]
2. Abdelwhab EM, Beer M. Panzootic HPAIV H5 and risks to novel mammalian hosts. Npj Viruses 2024 May 29;2(1):22. [doi: 10.1038/s44298-024-00039-z] [Medline: 40295707]
3. Caserta LC, Frye EA, Butt SL, et al. Spillover of highly pathogenic avian influenza H5N1 virus to dairy cattle. Nature New Biol 2024 Oct;634(8034):669-676. [doi: 10.1038/s41586-024-07849-4] [Medline: 39053575]
4. Evolution of the influenza A(H5) haemagglutinin: WHO/OIE/FAO H5 working group reports a new clade designated 2.3.4.4. World Health Organization. URL: https://www.who.int/publications/m/item/evolution-of-the-influenza-a(h5)-haemagglutinin-who-oie-fao-h5-working-group-reports-a-new-clade-designated-2.3.4.4 [accessed 2026-04-11]
5. Mostafa A, Nogales A, Martinez-Sobrido L. Highly pathogenic avian influenza H5N1 in the United States: recent incursions and spillover to cattle. Npj Viruses 2025 Jul 5;3(1):54. [doi: 10.1038/s44298-025-00138-5] [Medline: 40617962]
6. Xie Z, Yang J, Jiao W, et al. Clade 2.3.4.4b highly pathogenic avian influenza H5N1 viruses: knowns, unknowns, and challenges. J Virol 2025 Jun 17;99(6):e0042425. [doi: 10.1128/jvi.00424-25] [Medline: 40340397]
7. Caliendo V, Lewis NS, Pohlmann A, et al. Transatlantic spread of highly pathogenic avian influenza H5N1 by wild birds from Europe to North America in 2021. Sci Rep 2022 Jul 11;12(1):11729. [doi: 10.1038/s41598-022-13447-z] [Medline: 35821511]
8. Confirmations of highly pathogenic avian influenza in commercial and backyard flocks. US Department of Agriculture. URL: https://www.aphis.usda.gov/livestock-poultry-disease/avian/avian-influenza/hpai-detections/commercial-backyard-flocks [accessed 2025-05-19]
9. Commissioner Miller says mystery dairy cow disease has been identified (3/25/2024). Texas Department of Agriculture. URL: https://texasagriculture.gov/News-Events/Article/9950/COMMISSIONER-MILLER-SAYS-MYSTERY-DAIRY-COW-DISEASE-HAS-BEEN-IDENTIFIED [accessed 2025-05-19]
10. Technical Report: June 2024 highly pathogenic avian influenza A(H5N1) viruses. : Centers for Disease Control and Prevention; 2024 URL: https://www.cdc.gov/bird-flu/php/technical-report/h5n1-06052024.html [accessed 2025-05-19]
11. Highlights in the history of avian influenza (bird flu). Centers for Disease Control and Prevention. 2024. URL: https://www.cdc.gov/bird-flu/avian-timeline/index.html [accessed 2025-05-19]
12. Sonnberg S, Webby RJ, Webster RG. Natural history of highly pathogenic avian influenza H5N1. Virus Res 2013 Dec 5;178(1):63-77. [doi: 10.1016/j.virusres.2013.05.009] [Medline: 23735535]
13. Current situation: bird flu in dairy cows. Centers for Disease Control and Prevention. 2025. URL: https://www.cdc.gov/bird-flu/situation-summary/mammals.html [accessed 2025-08-21]
14. Mostafa A, Naguib MM, Nogales A, et al. Avian influenza A (H5N1) virus in dairy cattle: origin, evolution, and cross-species transmission. MBio 2024 Dec 11;15(12):e0254224. [doi: 10.1128/mbio.02542-24] [Medline: 39535188]
15. Kalthoff D, Hoffmann B, Harder T, Durban M, Beer M. Experimental infection of cattle with highly pathogenic avian influenza virus (H5N1). Emerg Infect Dis 2008 Jul;14(7):1132-1134. [doi: 10.3201/eid1407.071468] [Medline: 18598640]
16. DSHS reports first human case of avian influenza in Texas. Texas Health and Human Services. 2024. URL: https://www.dshs.texas.gov/news-alerts/dshs-reports-first-human-case-avian-influenza-texas [accessed 2025-05-19]
17. Avian influenza A(H5N1) – United States of America. World Health Organization. 2022. URL: https://www.who.int/emergencies/disease-outbreak-news/item/2022-DON379#:~:text=Description%20of%20the%20situation,(CDC)%20for%20further%20confirmation [accessed 2025-09-16]
18. Highly pathogenic avian influenza A (H5N1) virus infection reported in a person in the US. Centers for Disease Control and Prevention. 2024. URL: https://www.cdc.gov/media/releases/2024/p0401-avian-flu.html [accessed 2026-04-11]
19. A(H5) bird flu: current situation. Centers for Disease Control and Prevention. 2025. URL: https://www.cdc.gov/bird-flu/situation-summary/index.html [accessed 2025-05-19]
20. About outbreak investigations in healthcare settings. Centers for Disease Control and Prevention. 2024. URL: https://www.cdc.gov/healthcare-associated-infections/about/outbreak-investigations-in-healthcare.html [accessed 2025-05-19]
21. About APHIS. US Department of Agriculture. 2024. URL: https://www.aphis.usda.gov/about [accessed 2025-05-19]
22. What does FDA regulate? US Food and Drug Administration. 2024. URL: https://www.fda.gov/about-fda/what-we-do/what-does-fda-regulate [accessed 2025-05-19]

XSL•FO
RenderX

23. US data federation. Resources.data.gov. URL: https://resources.data.gov/ [accessed 2025-05-19]
24. Choi BCK. The past, present, and future of public health surveillance. Scientifica (Cairo) 2012;2012:875253. [doi: 10.6064/2012/875253] [Medline: 24278752]
25. Koopmans MPG, Barton Behravesh C, Cunningham AA, et al. The panzootic spread of highly pathogenic avian influenza H5N1 sublineage 2.3.4.4b: a critical appraisal of one health preparedness and prevention. Lancet Infect Dis 2024 Dec;24(12):e774-e781. [doi: 10.1016/S1473-3099(24)00438-9] [Medline: 39134084]
26. Global.health: a data science initiative. Globalhealth. URL: https://global.health/ [accessed 2025-04-04]
27. Xu B, Kraemer MUG, Xu B. Open access epidemiological data from the COVID-19 outbreak. Lancet Infect Dis 2020 May;20(5):534. [doi: 10.1016/S1473-3099(20)30119-5] [Medline: 32087115]
28. Xu B, Gutierrez B, Mekaru S, et al. Epidemiological data from the COVID-19 outbreak, real-time case information. Sci Data 2020 Mar 24;7(1):106. [doi: 10.1038/s41597-020-0448-0] [Medline: 32210236]
29. Benjamin A, Brownstein J, Cansler A, et al. Global.health: a scalable plaform for pandemic data integration, analytics, and preparedness [in review]. Research Square. Preprint posted online on Jul 19, 2022. [doi: 10.21203/rs.3.rs-1528783/v1]
30. Kraemer MUG, Tegally H, Pigott DM, et al. Tracking the 2022 monkeypox outbreak with epidemiological data in real-time. Lancet Infect Dis 2022 Jul;22(7):941-942. [doi: 10.1016/S1473-3099(22)00359-0] [Medline: 35690074]
31. Ebola Uganda 2022 repository. GitHub. 2022. URL: https://github.com/globaldothealth/ebola [accessed 2025-07-07]
32. Marburg 2023 Equatorial Guinea Outbreak 2023-05-08. Global.health: A Data Science Initiative. 2023. URL: https://www.marburg.global.health/2023-05-08.html [accessed 2025-07-07]
33. International health regulations (2005): as amended in 2014 and 2022. World Health Organization. 2024. URL: https://apps.who.int/gb/bd/pdf_files/IHR_2022-en.pdf [accessed 2025-05-19]
34. High consequence infectious diseases (HCID). UK Health Security Agency. 2025. URL: https://www.gov.uk/guidance/high-consequence-infectious-diseases-hcid [accessed 2025-07-07]
35. Kraemer MUG, Scarpino SV, Marivate V, et al. Data curation during a pandemic and lessons learned from COVID-19. Nat Comput Sci 2021 Jan;1(1):9-10. [doi: 10.1038/s43588-020-00015-6] [Medline: 38217160]
36. Bhatia S, Imai N, Watson OJ, et al. Lessons from COVID-19 for rescalable data collection. Lancet Infect Dis 2023 Sep;23(9):e383-e388. [doi: 10.1016/S1473-3099(23)00121-4] [Medline: 37150186]
37. Martin K, Murray E, McHugh L, Weigel A, Alden N, Moore Z. Update to public health reporting and national notification for novel influenza A virus infection. Council of State and Territorial Epidemiologists. URL: https://cdn.ymaws.com/www.cste.org/resource/resmgr/position_statements_files_2023/24-ID-09_Novel_Influenza_A.pdf [accessed 2025-11-26]
38. Powers J, Sopko J, Avunoori M, Han A, Scarpino SV, Krugman A. Timeline: H5N1 bird flu outbreak in the US. Think Global Health. 2025. URL: https://www.thinkglobalhealth.org/article/timeline-h5n1-bird-flu-outbreak-us [accessed 2025-05-19]
39. About one health. Centers for Disease Control and Prevention. 2024. URL: https://www.cdc.gov/one-health/about/index.html [accessed 2025-05-19]
40. H5N1 influenza. US Department of Agriculture. 2025. URL: https://www.aphis.usda.gov/livestock-poultry-disease/avian/avian-influenza/hpai-detections [accessed 2025-05-19]
41. HPAI confirmed cases in livestock. US Department of Agriculture. 2025. URL: https://www.aphis.usda.gov/livestock-poultry-disease/avian/avian-influenza/hpai-detections/hpai-confirmed-cases-livestock [accessed 2025-05-19]
42. WastewaterSCAN dashboard. WastewaterSCAN. URL: https://data.wastewaterscan.org [accessed 2025-05-19]
43. About CDC's wastewater monitoring program. Centers for Disease Control and Prevention. 2025. URL: https://www.cdc.gov/nwss/about.html [accessed 2026-05-31]
44. Louis S, Mark-Carew M, Biggerstaff M, et al. Wastewater surveillance for influenza A virus and H5 subtype concurrent with the highly pathogenic avian influenza A(H5N1) virus outbreak in cattle and poultry and associated human cases - United States, May 12-July 13, 2024. MMWR Morb Mortal Wkly Rep 2024 Sep 19;73(37):804-809. [doi: 10.15585/mmwr.mm7337a1] [Medline: 39298357]
45. GenBank overview. National Center for Biotechnology Information (NCBI). 2022. URL: https://www.ncbi.nlm.nih.gov/genbank [accessed 2025-05-19]
46. GISAID. URL: https://gisaid.org/ [accessed 2025-05-19]
47. Global.health map. Global.health. 2024. URL: https://www.map.global.health/ [accessed 2025-05-19]
48. Feature lookup tables. Mapbox. 2025. URL: https://docs.mapbox.com/data/boundaries/reference/feature-lookup-tables/ [accessed 2025-05-19]
49. 2020-2024 highlights in the history of avian influenza (bird flu) timeline. Centers for Disease Control and Prevention. 2024. URL: https://www.cdc.gov/bird-flu/avian-timeline/2020s.html#:~:text=In%20February%202022%2C%20USDA/APHIS,well%20as%20in%20other%20countries [accessed 2026-04-11]
50. Highly pathogenic avian influenza A(H5N1) virus: interim recommendations for prevention, monitoring, and public health investigations. Centers for Disease Control and Prevention. 2024. URL: https://www.cdc.gov/bird-flu/prevention/hpai-interim-recommendations.html [accessed 2025-05-19]
51. CDC A(H5N1) bird flu response update November 4, 2024. Centers for Disease Control and Prevention. 2024. URL: https://www.cdc.gov/bird-flu/spotlights/h5n1-response-11012024.html [accessed 2026-05-31]

XSL•FO
RenderX

52. Peacock TP, Moncla L, Dudas G, et al. The global H5N1 influenza panzootic in mammals. Nature New Biol 2025 Jan;637(8045):304-313. [doi: 10.1038/s41586-024-08054-z] [Medline: 39317240]
53. H5N1 bird flu continues to take its toll in the United States. GISAID. 2025. URL: https://gisaid.org/resources/gisaid-in-the-news/highly-pathogenic-avian-influenza-outbreak-in-the-united-states/ [accessed 2025-09-05]
54. Garg S, Reinhart K, Couture A, et al. Highly pathogenic avian influenza A(H5N1) virus infections in humans. N Engl J Med 2025 Feb 27;392(9):843-854. [doi: 10.1056/NEJMoa2414610] [Medline: 39740051]
55. Rolfes MA, Kniss K, Kirby MK, et al. Human infections with highly pathogenic avian influenza A(H5N1) viruses in the United States from March 2024 to May 2025. Nat Med 2025 Nov;31(11):3889-3898. [doi: 10.1038/s41591-025-03905-2] [Medline: 40712649]
56. Klein B, Kraemer MU, Scarpino SV. Timeline for H5N1 in the USA during the 2024 outbreak. : GitHub; 2024. [doi: 10.5281/ZENODO.11977338]
57. Kilaru P, Hill D, Anderson K, et al. Wastewater surveillance for infectious disease: a systematic review. Am J Epidemiol 2023 Feb 1;192(2):305-322. [doi: 10.1093/aje/kwac175] [Medline: 36227259]
58. Honein MA, Olsen SJ, Jernigan DB, Daskalakis DC. Challenges and opportunities for wastewater monitoring of influenza viruses during the multistate outbreak of highly pathogenic avian influenza A(H5N1) virus in dairy cattle and poultry. Am J Public Health 2024 Dec;114(12):1309-1312. [doi: 10.2105/AJPH.2024.307860] [Medline: 39298698]
59. H5N1 bird flu virus in livestock. California Department of Food and Agriculture. 2024. URL: https://www.cdfa.ca.gov/AHFSS/Animal_Health/HPAI.html [accessed 2025-05-19]
60. Avian influenza updates. California Department of Food and Agriculture. URL: https://www.cdfa.ca.gov/ahfss/Animal_Health/Avian_Influenza.html [accessed 2026-04-11]
61. UPDATE: CDPH confirms two cases of bird flu in California. California Department of Public Health. 2024. URL: https://www.cdph.ca.gov/Programs/OPA/Pages/NR24-029.aspx [accessed 2025-05-19]
62. Zhu S, Harriman K, Liu C, et al. Human cases of highly pathogenic avian influenza A(H5N1) - California, September-December 2024. MMWR Morb Mortal Wkly Rep 2025 Mar 13;74(8):127-133. [doi: 10.15585/mmwr.mm7408a1] [Medline: 40080512]
63. Wolfe MK, Duong D, Shelden B, et al. Detection of hemagglutinin H5 Influenza A virus sequence in municipal wastewater solids at wastewater treatment plants with increases in influenza A in Spring, 2024. Environ Sci Technol Lett 2024 Jun 11;11(6):526-532. [doi: 10.1021/acs.estlett.4c00331]
64. Falender R, Radniecki TS, Kelly C, et al. Avian Influenza A(H5) Subtype in Wastewater - Oregon, September 15, 2021-July 11, 2024. MMWR Morb Mortal Wkly Rep 2025 Feb 27;74(6):102-106. [doi: 10.15585/mmwr.mm7406a5] [Medline: 40014649]
65. Sutton M, Falender R, Scholz R, et al. Avian influenza A (H5) in wastewater, July 2024 to February 2025. JAMA Netw Open 2025 Jun 2;8(6):e2517286. [doi: 10.1001/jamanetworkopen.2025.17286] [Medline: 40569599]
66. Paulos AP, Hilton SP, Shelden B, Duong D, Boehm AB, Wolfe MK. Detection of hemagglutinin H5 influenza A virus RNA and model of potential inputs in an urban California sewershed. Environ Sci Technol 2025 Aug 12;59(31):16168-16179. [doi: 10.1021/acs.est.4c14792] [Medline: 40742307]
67. Dhillon RS, Karan A, Srikrishna D. Enhancing wastewater testing for H5N1 surveillance. Lancet Infect Dis 2024 Oct;24(10):e608. [doi: 10.1016/S1473-3099(24)00503-6] [Medline: 39151442]
68. Arvanitoyannis IS, Giakoundis A. Current strategies for dairy waste management: a review. Crit Rev Food Sci Nutr 2006;46(5):379-390. [doi: 10.1080/10408390591000695] [Medline: 16891210]
69. Detections of highly pathogenic avian influenza in wild birds. US Department of Agriculture. 2026. URL: https://www.aphis.usda.gov/livestock-poultry-disease/avian/avian-influenza/hpai-detections/wild-birds?page=1 [accessed 2026-04-11]
70. U.S. Fish and Wildlife Service homepage. US Fish and Wildlife Service. URL: https://www.fws.gov/carp/partner/migratory-bird-program-administrative-flyways#:~:text=The%20U.S.%20Fish%20&%20Wildlife%20Service,with%20the%20Flyway%20Boundaries%20map [accessed 2026-04-11]
71. The full annual cycle of migratory birds. Smithsonian's National Zoo and Conservation Biology Institute. URL: https://nationalzoo.si.edu/migratory-birds/full-annual-cycle-migratory-birds [accessed 2026-04-11]
72. Youk S, Torchetti MK, Lantz K, et al. H5N1 highly pathogenic avian influenza clade 2.3.4.4b in wild and domestic birds: Introductions into the United States and reassortments, December 2021–April 2022. Virology (Auckl) 2023 Oct;587:109860. [doi: 10.1016/j.virol.2023.109860]
73. Hill NJ, Takekawa JY, Cardona CJ, et al. Cross-seasonal patterns of avian influenza virus in breeding and wintering migratory birds: a flyway perspective. Vector Borne Zoonotic Dis 2012 Mar;12(3):243-253. [doi: 10.1089/vbz.2010.0246] [Medline: 21995264]
74. Spackman E, Stallknecht DE, Slemons RD, et al. Phylogenetic analyses of type A influenza genes in natural reservoir species in North America reveals genetic variation. Virus Res 2005 Dec;114(1-2):89-100. [doi: 10.1016/j.virusres.2005.05.013] [Medline: 16039745]
75. Macken CA, Webby RJ, Bruno WJ. Genotype turnover by reassortment of replication complex genes from avian influenza A virus. J Gen Virol 2006 Oct;87(Pt 10):2803-2815. [doi: 10.1099/vir.0.81454-0] [Medline: 16963738]

XSL•FO
RenderX

76. Dugan VG, Chen R, Spiro DJ, et al. The evolutionary genetics and emergence of avian influenza viruses in wild birds. In: Perez DR, editor. PLoS Pathog 2008 May 30;4(5):e1000076. [doi: 10.1371/journal.ppat.1000076] [Medline: 18516303]
77. Chen R, Holmes EC. Frequent inter-species transmission and geographic subdivision in avian influenza viruses from wild birds. Virology (Auckl) 2009 Jan 5;383(1):156-161. [doi: 10.1016/j.virol.2008.10.015] [Medline: 19000628]
78. Pacific flyway conservation. Sacramento Audobon Society. URL: https://www.sacramentoaudubon.org/pacific-flyway-conservation [accessed 2026-04-11]
79. McDuie F, Matchett EL, Prosser DJ, et al. Pathways for avian influenza virus spread: GPS reveals wild waterfowl in commercial livestock facilities and connectivity with the natural wetland landscape. Transbound Emerg Dis 2022 Sep;69(5):2898-2912. [doi: 10.1111/tbed.14445] [Medline: 34974641]
80. Belkhiria J, Hijmans RJ, Boyce W, Crossley BM, Martínez-López B. Identification of high risk areas for avian influenza outbreaks in California using disease distribution models. PLOS ONE 2018;13(1):e0190824. [doi: 10.1371/journal.pone.0190824] [Medline: 29385158]
81. Fall migration brings the return of avian influenza in wild birds. State of California, Department of Fish and Wildlife. 2024. URL: https://wildlife.ca.gov/News/Archive/fall-migration-brings-the-return-of-avian-influenza-in-wild-birds#:~:text=December%206%2C%202024,in%20California%20in%20July%202022 [accessed 2026-04-11]
82. Branda F, Ciccozzi M, Scarpa F. Tracking the spread of avian influenza A(H5N1) with alternative surveillance methods: the example of wastewater data. Lancet Infect Dis 2024 Oct;24(10):e604-e605. [doi: 10.1016/S1473-3099(24)00498-5] [Medline: 39102833]
83. Morgan OW, Abdelmalik P, Perez-Gutierrez E, et al. How better pandemic and epidemic intelligence will prepare the world for future threats. Nat Med 2022 Aug;28(8):1526-1528. [doi: 10.1038/s41591-022-01900-5]
84. Thacker SB, Qualters JR, Lee LM, Centers for Disease Control and Prevention. Public health surveillance in the United States: evolution and challenges. MMWR Suppl 2012 Jul 27;61(3):3-9. [Medline: 22832990]
85. Agate J. Artificial intelligence methods and approaches to improve data quality in healthcare data. Artif Intell Life Sci 2025 Dec;8:100135. [doi: 10.1016/j.ailsci.2025.100135]
86. Brownstein JS, Rader B, Astley CM, Tian H. Advances in artificial intelligence for infectious-disease surveillance. N Engl J Med 2023 Apr 27;388(17):1597-1607. [doi: 10.1056/NEJMra2119215] [Medline: 37099342]
87. Fanelli A, Cescatti A, Ciscar JC, et al. Assessing the risk of diseases with epidemic and pandemic potential in a changing world. Sci Adv 2025 Jul 25;11(30):eadw6363. [doi: 10.1126/sciadv.adw6363] [Medline: 40700491]

Abbreviations

CDC: Centers for Disease Control and Prevention
HPAI: highly pathogenic avian influenza
NWSS: National Wastewater Surveillance System
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
USDA: United States Department of Agriculture
WOAH: World Organization for Animal Health

Edited by A Mavragani, T Sanchez; submitted 21.Oct.2025; peer-reviewed by FA Tobolowsky, M Sutton; revised version received 20.Apr.2026; accepted 27.Apr.2026; published 22.Jun.2026.

Please cite as:
Sopko J, Han AR, Powers J, Sauer JA, Avunoori M, Zakrzewski S, Krugman A, Dasgupta A, Sewalk K, Gertz A, Rader B, Sheldon J, Klein B, Rivera JM, Kraemer MUG, Scarpino SV, Brownstein JS
Multimodal Data Approaches for Examining the 2024-2025 Highly Pathogenic Avian Influenza Outbreak in the United States: Descriptive Study
JMIR Public Health Surveill 2026;12:e86209
URL: https://publichealth.jmir.org/2026/1/e86209
doi:10.2196/86209

© Juliana Sopko, Aimee R Han, Jacqueline Powers, Jacquelin A Sauer, Mansi Avunoori, Stanislaw Zakrzewski, Allison Krugman, Abhishek Dasgupta, Kara Sewalk, Autumn Gertz, Benjamin Rader, James Sheldon, Brennan Klein, Jessica Malaty Rivera, Moritz U G Kraemer, Samuel V Scarpino, John S Brownstein. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 22.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete

XSL•FO
RenderX

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

XSL•FO
RenderX

Predicting Tuberculosis Outcomes Using Routine Surveillance Data in Chiang Mai, Thailand: Retrospective Cohort Study

Porramat Saksaen[1], MNS; Ekkarat Boonchieng[2], PhD; Aksara Thongprachum[1], PhD; Surasak Maotheuak[3], PhD; Sineenart Chautrakarn[1], PhD; Waraporn Boonchieng[1], DrPH

[1]Faculty of Public Health, Chiang Mai University, 239 Huay Kaew Road, Muang District, Chiang Mai, Thailand
[2]Faculty of Science, Department of Computer Science, Chiang Mai University, Chiang Mai, Thailand
[3]Faculty of Education, Department of Curriculum, Teaching & Learning, Chiang Mai University, Chiang Mai, Thailand

Corresponding Author:
Waraporn Boonchieng, DrPH
Faculty of Public Health, Chiang Mai University, 239 Huay Kaew Road, Muang District, Chiang Mai, Thailand

Abstract

Background: Tuberculosis (TB) remains one of the leading causes of death from a single infectious disease worldwide. In Thailand, persistent gaps in early detection and access to TB care remain important public health challenges, particularly among populations in rural and remote areas.

Objective: This study aimed to develop and evaluate predictive models for TB outcomes using routine surveillance data to support risk stratification, thereby informing public health decision-making in Chiang Mai Province, Thailand.

Methods: A retrospective cohort study was conducted using data from 5557 TB cases registered in the National Tuberculosis Information Program in Chiang Mai Province from 2020 to 2024. Models were developed to predict treatment success, mortality, and time to treatment initiation. We evaluated model performance using the area under the receiver operating characteristic curve, the Harrell Concordance Index, and error-based metrics for time-to-treatment prediction. Scenario analyses were conducted under predefined assumptions to assess projected changes in detection-related indicators, treatment coverage, and mortality relative to provincial baseline indicators.

Results: The models demonstrated good predictive performance. Mortality was associated with HIV co-infection (hazard ratio 5.80) and was highest among older adults with HIV co-infection (hazard ratio 12.30). Treatment delays were longer among individuals living in rural or remote areas and among those without health insurance, with mean delays ranging from 8 to 18 days. These findings suggest distinct patterns in TB outcomes, with mortality more closely related to clinical vulnerability and treatment delay more closely related to health care access factors. Under modeled scenarios, the models projected that detection-related indicators would increase by 25%, treatment coverage would increase by 15%, and mortality would decrease by 20% relative to provincial baseline conditions.

Conclusions: Predictive modeling using routine TB surveillance data may support risk stratification and provide insights into treatment outcomes and delays. The findings highlight the combined roles of clinical vulnerability and health care access factors in TB outcomes and support further evaluation of data-driven approaches to inform targeted TB interventions in resource-limited settings.

(JMIR Public Health Surveill 2026;12:e86495) doi:10.2196/86495

KEYWORDS

tuberculosis; predictive modeling; treatment delay; mortality; health equity; resource-limited settings; sustainable development goals; SDG; SDG 3; SDG 10; Thailand; good health and well-being; reduced inequalities

Introduction

Tuberculosis (TB) remains one of the leading causes of death from a single infectious disease worldwide. The World Health Organization (WHO) estimates that 10.7 million people developed TB globally in 2024, of whom 8.3 million were diagnosed and notified, highlighting persistent gaps in case detection and reporting. Although global health efforts have made progress in TB diagnosis and treatment, global TB incidence declined by approximately 12% between 2015 and 2024, remaining well below the End TB Strategy target of a 50% reduction by 2025 [1].

Thailand is classified as a high-burden country for TB and TB/HIV by the WHO. Thailand has an estimated 113,000 new TB cases annually, corresponding to an incidence of 157 per 100,000 population. Of these, HIV co-infection accounts for 9,400 cases, and 13,000 deaths are attributed to TB each year [2]. Reported treatment coverage is approximately 71%,

XSL•FO
RenderX

indicating persistent gaps in access to diagnosis and care. Previous studies have shown that TB does not distribute evenly across Thailand but rather clusters in several areas, particularly along international border regions [3]. These geographic and health-system disparities highlight the need for analytical approaches that integrate epidemiological, clinical, and health service data to support more precise and equitable TB control.

Mathematical models, particularly compartmental models such as the SEIR (susceptible-exposed-infectious-recovered) framework, have been widely used to study TB epidemiology. Unlike the classical SIR model, the SEIR framework explicitly includes an exposed compartment representing individuals who are infected but not yet infectious, thereby accounting for the latent stage before progression to active TB disease [4,5]. However, classical SEIR-based approaches often rely on simplifying assumptions, such as homogeneous mixing and fixed transition rates between compartments. These approaches may have limited ability to capture individual-level heterogeneity in clinical risk, health care access, and treatment outcomes.

Researchers are increasingly applying machine learning (ML) methods to TB-related predictive tasks, such as incidence prediction, diagnosis, and treatment outcome prediction [6,7]. These approaches can exploit high-dimensional, heterogeneous data and demonstrate strong predictive performance for specific TB-related outcomes. ML-based approaches have also been incorporated into clinical decision support systems to assist with diagnostic and treatment decision-making [8,9]. However, most existing studies focus on isolated prediction tasks and do not explicitly link individual-level risk estimation with broader epidemiological disease-stage structures.

The strengths of SEIR-based epidemiological models and ML approaches are complementary. SEIR-based frameworks [10] provide a structured way to conceptualize disease progression at the population level, whereas ML approaches enable data-driven prediction of observable individual-level outcomes from clinical and surveillance data. However, there remains limited evidence on how routine TB surveillance data can be organized within an epidemiologically informed framework to generate practical predictions of treatment outcomes, mortality, and treatment delays. This gap is particularly important in settings where surveillance systems contain rich programmatic data, but practitioners do not routinely use them for risk stratification or decision support.

TB control in Thailand is primarily implemented through health care facility–based diagnosis, registration, and treatment reporting. In contrast, routine programmatic reporting remains centered on surveillance and treatment monitoring rather than risk-prediction tools. Ongoing programmatic work includes surveillance and focused interventions for key high-risk or vulnerable populations, such as people living with HIV, prisoners, migrants, older adults, and non-Thai populations [11,12]. Thailand has established the National Tuberculosis Information Program (NTIP) as a national online TB reporting system. However, there are gaps in translating surveillance data into analytical tools that can support decision-making at different levels of the health care system.

NTIP is Thailand's national TB reporting system that collects data on diagnoses, treatment registrations, and outcomes from health care facilities [13]. While NTIP provides a standardized national dataset, it reflects diagnosed and registered TB cases and does not capture unobserved infection or transmission processes. These characteristics make NTIP suitable for developing predictive models focused on routinely observed clinical and programmatic outcomes. Integration of NTIP data with analytical models that combine an epidemiologically informed structure and predictive modeling may support more systematic use of surveillance data for risk stratification and policy planning.

Therefore, this study aimed to develop and evaluate a TB surveillance framework based on an SEIR-informed conceptual structure and predictive models using NTIP data from Chiang Mai Province. Specifically, the study aimed to assess whether routinely collected surveillance data could support risk prediction, risk stratification, and public health decision-making for TB outcomes in a resource-limited setting.

Methods

Study Design

This retrospective cohort study used routine surveillance data to develop a TB surveillance framework that combines an SEIR-informed conceptual structure with individual-level predictive models. We used the SEIR framework [10,14] to guide the conceptual organization of TB disease stages. Empirical modeling focused on clinical and programmatic outcomes recorded in NTIP, including treatment success, mortality, and time to treatment initiation.

Within this framework, NTIP provided data on diagnosed and registered TB cases, treatment outcomes, deaths, and treatment initiation dates. Registered TB cases represented the observed infectious state among patients captured by the surveillance system. We operationalized recovery-related outcomes as treatment success, and we used recorded deaths to represent mortality. NTIP did not directly capture susceptible and exposed populations; therefore, we retained them as conceptual components of the SEIR-informed framework. Table 1 summarizes the observability and operationalization of SEIR-D (susceptible-exposed-infectious-recovered-death) components in NTIP data.

The study, therefore, linked an epidemiologically informed disease-stage framework with predictive modeling of observed outcomes among patients diagnosed with and registered for TB.

XSL•FO
RenderX

Table . Summarizes the observability and operationalization of SEIR-D (susceptible-exposed-infectious-recovered-death) components in the National Tuberculosis Information Program (NTIP) data[a].

Component	Epidemiological meaning	NTIP observability	Operationalization in this study
Susceptible	Population at risk of TB[b] infection	Not directly observed	Conceptual component of the SEIR-informed framework
Exposed	Individuals with TB infection before registered active TB disease	Not directly observed	Conceptual component of the SEIR-informed framework
Infectious	Diagnosed active TB cases	Observed among registered cases	Registered TB cases recorded in NTIP
Recovered	Successful treatment outcome	Observed	Treatment success recorded in NTIP
Death	Death during treatment or follow-up	Observed	Death recorded in NTIP treatment outcome data
Treatment delay	Interval from diagnosis to treatment initiation	Observed	Days between diagnosis and treatment initiation

[a]SEIR-D components were defined within an SEIR-informed conceptual framework and operationalized according to the structure of the NTIP surveillance dataset. Susceptible and exposed populations were not directly observed in NTIP and were retained as conceptual components. Empirical analyses focused on registered TB cases, treatment success, death, and time to treatment initiation. Treatment delay was considered a programmatic interval rather than a disease-state compartment.

[b]TB: tuberculosis.

Study Area and Data Sources

Data were obtained from the NTIP, including records of patients with TB registered in Chiang Mai Province between October 2020 and September 2024. The NTIP database provided individual-level data on diagnosis, treatment registration, comorbidities, and treatment outcomes.

Inclusion and Exclusion Criteria

The study population consisted of patients registered for TB treatment in NTIP during the study period. We excluded patients with missing values for key variables required for model development, including age, sex, district of residence, treatment initiation date, or treatment outcome. These variables were required to define the analytical cohort, predictor structure, and primary outcomes. The analytical dataset comprised 5557 cases after data cleaning and eligibility screening based on variable completeness. Records were excluded only when missingness involved variables essential for defining the analytical cohort, primary outcomes, or model structure; we addressed missing values in nonessential variables separately using multiple imputation.

Data Management and Preparation

The key variables in the analytical dataset were age, sex, district of residence, date of treatment initiation, and treatment outcome. Records with missing values in key analytical variables were excluded. Missing values were imputed for nonessential variables, such as selected comorbidities and health care system indicators, using multiple imputation by chained equations, which generated 20 imputed datasets to account for the uncertainty of the imputed values and to reduce potential bias associated with complete-case analysis [14-16]. The imputation model included observed demographic, clinical, and health system variables and assumed a missing-at-random mechanism conditional on these observed variables. This approach is consistent with established recommendations for handling incomplete observational data when missingness is assumed to be conditional on observed variables [14,16].

BMI, comorbidities, HIV status, insurance status, and other predictor variables were recorded at TB registration or initial clinical assessment and used as baseline predictors for subsequent recorded outcomes. A BMI of less than 18.5 kg/m^2 was defined using the BMI value recorded at registration or initial assessment. We used the Tukey fences and the isolation forest algorithm [17] to identify outliers. We set the contamination parameter of the isolation forest algorithm to 0.05, and we checked model outputs for clinical plausibility. Residence categories were defined using geographic accessibility to TB services, including the distance from residence to treatment facilities. Urban residence, representing the most accessible category, was used as the reference group for residence variables. All eligible records were retained for model development and assessed using train-test splitting and internal cross-validation.

Development of Predictive Models

We conducted predictive modeling using observable outcomes from the final analytic NTIP dataset of 5557 TB cases. We selected models according to the structure of the recorded outcomes and the statistical requirements of each prediction task. We applied logistic regression for binary treatment success [17,18], the Cox proportional hazards model for time-to-event mortality with censoring [19,20], and the log-normal accelerated failure time model for time to treatment initiation [20]. This approach allowed each model to correspond to the scale, timing, and censoring structure of each outcome under analysis. Because NTIP records diagnosed and registered TB cases rather than the full susceptible or exposed population, model construction focused on clinical and programmatic outcomes available among registered patients. The SEIR framework was used to organize disease-stage concepts, consistent with compartmental modeling approaches that distinguish susceptible, exposed, infectious, and recovered states [10,21].

XSL•FO
RenderX

A multivariable logistic regression model was used to predict treatment success as a binary outcome, selected for its interpretability and suitability for clinical decision-making [17,18]. A Cox proportional hazards model [19,20] was used to estimate the hazard of death with time-to-event data, accounting for censoring. We tested the proportional hazards assumption using Schoenfeld residuals. We fitted an accelerated failure time model with a log-normal distribution to estimate delays between diagnosis and treatment initiation; this approach estimates the effect of covariates on event time rather than on the hazard function [20].

Together, these models estimated binary outcomes, time-to-event risks, and delays within a unified analytical framework based on recorded clinical and programmatic end points, including treatment success, mortality, and diagnosis-to-treatment delay. We specified the predictive models for recorded outcomes rather than for latent or unobserved transmission states. Model performance was assessed using discrimination and calibration metrics, including the area under the receiver operating characteristic curve (AUC) for classification models [22] and concordance-based measures for survival models [23].

Model Training and Validation

We assessed model performance using a train-test split (train: 4446/5557, 80%; test: 1111/5557, 20%) and performed internal validation using 10-fold stratified cross-validation on the training set. For classification models, the AUC was used as a performance metric [22]. For survival models, the Harrell Concordance Index was used as a performance metric [23]. We assessed prediction error for the time-to-treatment model using root mean squared error and mean absolute error. Additionally, we assessed the goodness-of-fit of the logistic regression model using the Hosmer-Lemeshow test [17]. We assessed discrimination, calibration, and prediction error across outcome types and model specifications.

Management of Conflicting Outcomes

We analyzed mortality and treatment success as separate outcome-specific end points. We modeled mortality using time-to-event methods, whereas we modeled treatment success as a binary outcome. We applied this outcome-specific modeling strategy to support risk stratification across different stages of the TB care pathway. We interpreted estimates as predictive associations rather than causal effects.

Analysis of Scenarios

We performed scenario analyses to assess projected changes in detection-related indicators, treatment coverage, and mortality under predefined modeled assumptions. The baseline scenario represented the observed provincial TB program conditions prior to applying the modeled changes. Baseline values included the provincial detection-related indicator, treatment coverage, and mortality rate, which were derived from Chiang Mai provincial TB program indicators. We derived the time to treatment initiation from recorded diagnosis and treatment initiation dates in NTIP and modeled it separately using a log-normal accelerated failure time model.

We represented screening as a programmatic process related to case detection rather than as a disease-state compartment. We specified detection-related changes as proportional changes relative to provincial TB program baseline indicators. We represented changes in treatment processes using predefined relative changes in treatment coverage and mortality, which were applied to provincial baseline indicators.

In the baseline scenario, the detection-related indicator was 87.08%, treatment coverage was 71.90%, and mortality was 11.90% [12]. The scenario parameters for detection-related improvement, treatment coverage improvement, and mortality reduction were predefined assumptions rather than estimates from an observed intervention. We, therefore, interpreted scenario outcomes as model-based projections rather than observed intervention effects, consistent with the use of modeling approaches to inform public health decision-making in the absence of direct population-level evidence [24].

Statistical Analysis and Software

We conducted all analyses using R (version 4.1.2; R Foundation for Statistical Computing) and Python (version 3.11; Python Software Foundation). We defined statistical significance as $P<.05$, and we reported all estimates with 95% CI. As NTIP represents a registry of reported TB cases rather than a sampled survey dataset, we therefore did not apply survey weighting. Given the analysis's predictive and exploratory focus, we instead emphasized effect estimates, CI, and model performance metrics rather than confirmatory hypothesis testing.

Ethical Considerations

The research ethics committee of the Faculty of Public Health, Chiang Mai University (approval: ET030/2025) and the Chiang Mai Provincial Public Health Office approved the study protocol. We anonymized all data prior to analysis and retained no identifiable personal information. Data were stored in secure, encrypted systems in accordance with applicable data protection regulations.

Results

Demographic, Clinical, and Health Care Access Characteristics of the Study Participants

This study analyzed 5557 TB cases recorded in the NTIP in Chiang Mai Province between 2020 and 2024. Most participants were male (3756, 67.6% cases), and the mean age was 45.3 (SD 14.8) years. HIV co-infection was present in 556 (10.0% cases), 1000 (18.0%) cases had diabetes mellitus, and 1223 (22.0%) cases had a BMI <18.5 kg/m^2. Health care access characteristics included absence of health insurance in 445 (8.0%) cases, rural residence in 1556 (28.0%) cases, and remote residence in 445 (8.0%) cases (Table 2).

XSL•FO
RenderX

Table . Baseline demographic, clinical, and health care access characteristics of the study population (N=5557)[a].

Characteristic	Values
Male sex, n (%)	3756 (67.6)
Age (y), mean (SD)	45.3 (14.8)
HIV-positive, n (%)	556 (10.0)
Diabetes mellitus, n (%)	1000 (18.0)
BMI (<18.5 kg/m^2), n (%)	1223 (22.0)
No health insurance, n (%)	445 (8.0)
Rural residence, n (%)	1556 (28.0)
Remote residence, n (%)	445 (8.0)

[a]Baseline demographic, clinical, and health care access characteristics of National Tuberculosis Information Program–registered patients with tuberculosis in Chiang Mai Province, Thailand, 2020 - 2024. Data are presented as n (%) or mean (SD).

Predictive Model Performance

Predictive models demonstrated good performance across recorded TB outcomes. The AUC for the treatment success model was 0.88 (95% CI 0.86 - 0.90), indicating good discrimination. The mortality model showed the Harrell Concordance Index of 0.86 (95% CI 0.84 - 0.88), indicating good concordance between predicted and observed survival outcomes. We summarized prediction error for time to treatment initiation using root mean squared error and mean absolute error, with values of 8.2 days and 5.7 days, respectively. The logistic regression model was adequately calibrated, with no evidence of poor fit based on the Hosmer-Lemeshow test (*P*>.05; Table 3).

Table . Performance of predictive models for tuberculosis outcomes[a].

Outcome	Model	Performance metric
Treatment success	Logistic regression	AUC[b]=0.88 (95% CI 0.86 - 0.90)
Mortality	Cox proportional hazards model	Harrell Concordance Index=0.86 (95% CI 0.84 - 0.88)
Time to treatment initiation	Log-normal AFT[c] model	RMSE[d]=8.2 days; MAE[e]=5.7 days[f]

[a]Performance metrics of predictive models for tuberculosis outcomes in a retrospective cohort of National Tuberculosis Information Program–registered patients in Chiang Mai Province, Thailand, 2020 - 2024.

[b]AUC: area under the receiver operating characteristic curve.

[c]AFT: accelerated failure time.

[d]RMSE: root mean squared error.

[e]MAE: mean absolute error.

[f]RMSE and MAE represent prediction errors for the time to treatment initiation, expressed in days.

Treatment Success as a Recovery-Related Outcome

We modeled treatment success as the recovery-related outcome recorded in NTIP. The logistic regression model for treatment success demonstrated good discrimination, with an AUC of 0.88 (95% CI 0.86 - 0.90), indicating that routinely recorded clinical and programmatic variables are informative for classifying treatment success among registered patients with TB.

Factors Associated With Mortality

HIV infection was associated with higher mortality among patients with TB (HR 5.80; 95% CI 4.20 - 7.90; *P*<.001). Patients aged older than 65 years with HIV co-infection had the highest estimated mortality risk (HR 12.30; 95% CI 8.90 - 16.40; *P*<.001). Higher mortality was also associated with diabetes mellitus (HR 3.10; 95% CI 2.30 - 4.40; *P*<.001) and a BMI of less than 18.5 kg/m^2 (HR 2.30; 95% CI 1.70 - 3.10; *P*<.001; Table 4). We present these estimates as outcome-specific associations from the mortality model.

XSL•FO
RenderX

Table . Factors associated with mortality[a].

Factor	HR[b] (95% CI)	*P* value
HIV-positive	5.80 (4.20 - 7.90)	<.001
Age (>65 y with HIV co-infection)	12.30 (8.90 - 16.40)	<.001
Diabetes mellitus	3.10 (2.30 - 4.40)	<.001
BMI (<18.5 kg/m^2)	2.30 (1.70 - 3.10)	<.001

[a]Factors associated with mortality among patients with tuberculosis in Chiang Mai Province, Thailand, from 2020 to 2024. Hazard ratios were estimated using Cox proportional hazards models.

[b]HR: hazard ratio.

Factors Associated With Treatment Delay

Treatment delays were longer among individuals living in rural and remote areas than among those living in urban areas. Rural residence was associated with an average delay of 12 (95% CI 9 - 15) days, remote residence with 18 (95% CI 14 - 22) days, lack of health insurance with 8 (95% CI 6 - 10) days, and age >70 years with 6 (95% CI 4 - 9) days (Table 5). We present treatment-delay estimates as mean differences with 95% CI.

Table . Factors associated with treatment delays[a].

Factor	Days, mean difference (95% CI)
Rural residence	12 (9 - 15)
Remote residence	18 (14 - 22)
No health insurance	8 (6 - 10)
Age (>70 y)	6 (4 - 9)

[a]Factors associated with treatment delay among patients with tuberculosis in Chiang Mai Province, Thailand, 2020 - 2024. Estimates represent differences in the expected time to treatment initiation derived from the accelerated failure time model. Urban residence, the most accessible residence category, was used as the reference group for rural and remote residence estimates.

Simulation Results of Modeled Scenarios

We derived baseline values for scenario comparisons from Chiang Mai provincial TB program indicators. The scenario analysis did not model the redistribution of specific resources between geographic risk areas. Instead, it represented potential programmatic improvements in detection-related indicators, treatment coverage, and mortality under predefined assumptions. Under these modeled scenarios, the models projected detection-related indicators to increase by 25% (95% CI 22 - 28), treatment coverage by 15% (95% CI 12 - 18), and mortality to decrease by 20% (95% CI 16 - 24) relative to baseline conditions (Table 6). Readers should interpret these projections as scenario-based estimates rather than as observed effects of resource reallocation or implemented program changes.

Table . Projected outcomes under modeled scenarios relative to provincial baseline indicators[a].

Outcome	Provincial baseline (%)	Projected change (%)	95% CI (%)
Detection-related indicators	87.08	+25	22 - 28
Treatment coverage	71.90	+15	12 - 18
Mortality	11.90	–20	16 - 24

[a]Projected outcomes under modeled scenarios are relative to provincial baseline indicators in Chiang Mai Province, Thailand. Estimates represent model-based projections under predefined assumptions and do not reflect observed intervention effects or implemented program changes. Baseline values were obtained from provincial tuberculosis program indicators.

Discussion

Principal Findings

This study examined TB outcomes using routinely collected surveillance data and applied predictive modeling to identify factors associated with mortality and treatment delay. The findings suggest distinct patterns in TB outcomes, where mortality was more strongly associated with clinical vulnerability, while treatment delay was more closely related to health care access factors. By examining these outcomes within a unified surveillance-based analytical framework, this study highlights the importance of distinguishing between clinical risk factors and health care access barriers when designing TB control strategies, as these domains reflect distinct yet interacting pathways that influence patient outcomes.

The SEIR-informed framework provided a structure for organizing disease-stage concepts, and the predictive models focused on recorded clinical and programmatic outcomes in NTIP, including treatment success, mortality, and time to

XSL•FO
RenderX

treatment initiation. This approach linked an epidemiological disease-stage perspective with practical risk stratification using routinely collected surveillance data.

Comparison With Previous Studies

The association between HIV co-infection and higher mortality observed in this study (HR=5.80) is consistent with prior survival analyses showing reduced survival among patients with TB or HIV coinfection [25] and cohort evidence identifying HIV positivity as a predictor of mortality among patients with TB [26]. The relatively high magnitude of association observed in this study may, therefore, reflect not only HIV-related clinical vulnerability and the increased complexity of clinical management in patients with coinfection, but also contextual factors such as diagnostic timing and health system access, which previous studies have identified as contributors to diagnostic delays and poorer TB outcomes [27].

In addition to HIV co-infection, baseline metabolic and nutritional factors also appear relevant to TB outcomes. The associations observed for diabetes are consistent with data from prospective cohorts, indicating that diabetes and prediabetes are associated with adverse TB treatment outcomes [28]. Moreover, prior research has linked dysglycemia to more severe disease presentation, including higher radiographic severity and increased bacillary load, as well as increased TB lethality among patients with diabetes [29]. Low BMI may indicate nutritional vulnerability and should be interpreted in the context of a broader clinical risk profile rather than as an independent predictor, consistent with survival analyses demonstrating the significance of baseline clinical characteristics in shaping TB mortality risk [26].

In addition to these clinical risk factors, health system and access-related barriers also contributed to differences in TB outcomes. Regarding treatment delay, the finding that rural and remote populations experienced longer delays is consistent with systematic review evidence indicating that delays in diagnosis and treatment of TB continue to be important barriers to TB control [27]. One suggested approach to reduce these access gaps is active case-finding, particularly through community-based approaches [30] and screening in general or at-risk populations [31]. However, these approaches are likely to be most effective when implemented with appropriate coverage and linked to timely diagnosis and treatment initiation. These findings suggest that geographic variation in treatment delay may reflect underlying disparities in health care access and care pathways rather than solely spatial disease patterns.

Interpretation and Implications

From a methodological perspective, the results are consistent with a broader pattern in the TB modeling literature, in which different modeling approaches capture different aspects of the disease process. Researchers have applied ML methods to address several TB-related predictive tasks, including diagnostic classification with high-dimensional serological data [32], incidence prediction using meteorological and air pollution data [6], and treatment outcome prediction using routinely available clinical indicators [33]. Recent work has also examined the use of ML for TB incidence forecasting within broader epidemiological modeling frameworks [34]. Such approaches are often designed to optimize predictive accuracy; however, they may have limited interpretability regarding underlying disease mechanisms, particularly when the observed data include only diagnosed and reported cases rather than the full spectrum of transmission dynamics.

Mechanistic models, particularly SEIR-based approaches, provide a structured framework for conceptualizing disease progression and population-level disease dynamics [35] and have been extended to examine TB dynamics under different modeling assumptions [36]. Nevertheless, such models often rely on simplifying assumptions and may not fully capture real-world heterogeneity or individual-level risk. In this study, we used the SEIR framework conceptually to organize disease stages, while applying predictive models to recorded outcomes in routine surveillance data. The analysis, therefore, focused on recorded clinical and programmatic outcomes rather than population-level transmission dynamics. This approach supports alignment between theoretical disease-stage concepts and the structure of routinely available surveillance data.

Recent advances in hybrid TB modeling suggest that integrating mechanistic models with ML may provide a more balanced framework that supports predictive performance while maintaining epidemiological interpretability [37]. In this study, this integration was applied at a conceptual level, reflecting the structure and constraints of routine surveillance data. The study demonstrates a practical approach to organizing surveillance data within an epidemiologically informed framework to support the prediction of recorded TB outcomes and risk stratification.

The literature has well established that TB epidemiology exhibits spatial heterogeneity, with geographic variation observed across spatial and socioeconomic contexts [3,38]. More recently, work using satellite imagery and urban density measures has shown that spatial proxies may help identify areas for targeted active case-finding, while highlighting the need for cautious interpretation of geographic risk patterns [39]. The observed association between geographic residence and treatment delay suggests that spatial disparities in TB outcomes are attributable to differences in health care access and service availability. However, spatial clustering does not necessarily reflect transmission dynamics, and geographic variation may also be due to differences in health care access, reporting systems, population structure, and socioeconomic conditions. This distinction is important because outcomes may vary geographically due to both epidemiological and health-system factors.

From a programmatic perspective, the findings suggest that interventions targeting both early case detection and reduction of structural barriers to health care access may be needed to improve TB outcomes, rather than applying a uniform strategy across all populations. This distinction has implications for intervention design, as strategies to address clinical risk (eg, comorbidity management) may differ substantially from those intended to address access barriers (eg, service decentralization or active case-finding). This interpretation is consistent with the scenario analyses presented in this study, which show that improvements in detection and treatment processes could be

XSL•FO
RenderX

associated with changes in projected outcomes under modeled assumptions. These projections provide scenario-based estimates of possible implications for programs, and readers should interpret them in the context of the assumptions and baseline indicators used in the analysis.

Targeted case-finding strategies may help improve access to diagnosis, especially in high-risk populations. However, delays in TB care are multifactorial and influenced by both patient- and health system–level factors [27,30]. The effectiveness of these approaches may, therefore, depend on both expanded access and integration into routine care pathways. In practical terms, predictive outputs may help TB programs prioritize active follow-up, outreach screening, referral support, patient navigation, and closer monitoring for clinically vulnerable patients or those living in underserved areas. Further implementation research is needed to determine whether these outputs should guide additional resources, the redistribution of existing resources, or both. Digital health technologies and clinical decision support systems may facilitate implementation [8,40], but future studies should evaluate their practical impact in real-world clinical and public health settings.

Strengths

This study used a large dataset derived from a routine TB surveillance system, supporting the relevance of the findings to real-world programmatic settings. The analysis considered multiple outcomes, including treatment success, mortality, and time to treatment initiation, allowing for a broader assessment of the TB care pathway. In addition, the use of established statistical approaches supported risk stratification and the identification of high-risk subgroups using routinely available data. These strengths may enhance the applicability of the findings to TB control programs, although readers should interpret the results within the context of surveillance-based data.

A further strength of the study is the alignment between the analytical methods and the structure of the available surveillance data. By focusing on recorded outcomes, the framework supports practical interpretation for routine programmatic settings while retaining an epidemiologically informed structure for organizing disease-stage concepts.

Limitations

This study has several important limitations. First, the use of routinely collected surveillance data may be subject to missing data, misclassification, and variability in data quality across reporting sites. Important variables, including socioeconomic status, income, education, migrant status, health care-seeking behavior, timing of symptom onset, and distance to health care facilities, may not have been consistently recorded. Their absence may have introduced residual confounding, particularly in the analysis of treatment delay, as unmeasured social and structural barriers may partly explain the observed associations between rural or remote residence and delayed treatment initiation. The study period also overlapped with the COVID-19 pandemic, which may have affected health care-seeking behavior, TB screening activities, diagnostic access, reporting completeness, and time to treatment initiation [41].

Second, the study was based on registered TB cases in NTIP and may not capture undiagnosed cases, unreported cases, or individuals who did not access care. Readers should therefore interpret the findings as reflecting patterns among registered patients rather than the full burden and heterogeneity of TB in the underlying population. Because NTIP includes diagnosed and registered patients with TB, it cannot directly estimate the population-level effect of increased screening on undiagnosed or unreported cases. Readers should therefore interpret the assumed increase in screening or detection-related indicators in the scenario analysis as a programmatic assumption applied to provincial baseline indicators rather than as an effect estimated from the NTIP cohort itself.

Third, although we used the SEIR framework to guide the conceptual organization of disease stages, the available surveillance data did not directly observe susceptible or exposed populations, and we did not empirically estimate or validate SEIR transition parameters. Accordingly, readers should interpret the framework as SEIR-informed rather than as a fully parameterized transmission model. The observational design also limits causal inference; readers should interpret observed associations as predictive or outcome-specific rather than causal. In addition, combining an epidemiologically informed disease-stage framework with individual-level predictive models requires cautious interpretation; researchers should not directly infer population-level concepts from individual-level associations.

Finally, although we performed internal validation using train-test splitting and cross-validation, we did not conduct external validation using independent datasets. The models may perform differently in settings with varying epidemiological profiles, health care access patterns, population characteristics, reporting completeness, or data quality. The use of multiple predictors and outcome-specific models may also increase the risk of overfitting, particularly in subgroups or lower-frequency outcomes. We based the scenario analysis on predefined assumptions and provincial baseline indicators rather than observed intervention data; therefore, readers should interpret projected changes as illustrative programmatic projections. Future studies should compare performance with simpler baseline models, validate calibration and discrimination in independent datasets, examine sensitivity to scenario assumptions, and evaluate feasibility and implementation in routine TB control.

Conclusions

TB outcomes in Chiang Mai Province reflected the combined influence of clinical vulnerability and health care access barriers. Mortality was more strongly associated with clinical factors, including HIV co-infection, older age, diabetes, and low BMI. In contrast, treatment delay was more closely related to rural or remote residence and lack of health insurance. These findings suggest that TB control strategies should combine intensified clinical monitoring for high-risk patients with improved access to timely diagnosis and treatment for underserved populations.

An SEIR-informed predictive framework using routine surveillance data may support risk stratification and program planning by identifying high-risk groups and access-related

XSL•FO
RenderX

delays in the TB care pathway. The findings support practical actions such as active follow-up for patients with HIV co-infection or other clinical vulnerabilities, strengthened outreach and referral pathways in rural and remote areas, improved linkage to care for uninsured patients, and targeted prioritization of follow-up, outreach, and linkage-to-care activities.

Model-based scenario results suggest that improvements in detection-related indicators and treatment coverage could contribute to better TB outcomes under predefined assumptions. However, readers should interpret these projections cautiously because they are based on modeled scenarios rather than observed intervention effects. Before wider implementation, future studies should externally validate the predictive framework in independent datasets and evaluate it prospectively in routine TB control settings to assess feasibility, calibration, clinical usefulness, and impact on decision-making.

Acknowledgments

The authors thank the staff of the Chiang Mai Provincial Tuberculosis Control Network and the public health personnel at all levels of the health system for their cooperation in data preparation and verification. The authors declare the use of generative AI during manuscript revision. ChatGPT by OpenAI was used under full human supervision to support text generation for revised wording, proofreading and editing, and reformatting. Generative AI was not used to generate original data, conduct statistical analyses, create or modify figures, select references independently, or make independent scientific interpretations. All AI-assisted content was reviewed, edited, verified, and approved by the authors. Responsibility for the accuracy, integrity, and final content of the manuscript lies entirely with the authors. Generative AI tools are not listed as authors and do not bear responsibility for the final outcomes.

Funding

This research was supported by Chiang Mai University.

Authors' Contributions

PS conceptualized the study, performed the investigation, collected data, and drafted the original manuscript.
EB contributed to data analysis, curation, and editing.
AT conducted formal analysis and data curation, and contributed to writing and review. SM contributed to methodology development and critical review. SC was responsible for validation and manuscript review. WB supervised the overall study, provided resources, and served as the corresponding author. All authors read and approved the final version of this manuscript.

Conflicts of Interest

None declared.

Checklist 1
STROBE checklist.
[PDF File, 118 KB - publichealth_v12i1e86495_app1.pdf]

References

1. Global tuberculosis report 2025. : World Health Organization; 2025 URL: https://iris.who.int/server/api/core/bitstreams/e97dd6f4-b567-4396-8680-717bac6869a9/content [accessed 2026-06-11]
2. Tuberculosis profile: Thailand. World Health Organization. URL: https://worldhealthorg.shinyapps.io/tb_profiles/?_inputs_&tab=%22charts%22&lan=%22EN%22&iso3=%22THA%22&entity_type=%22country%22 [accessed 2026-04-29]
3. Chinpong K, Thavornwattana K, Armatrmontree P, et al. Spatiotemporal epidemiology of tuberculosis in Thailand from 2011 to 2020. Biology (Basel) 2022 May 16;11(5):755. [doi: 10.3390/biology11050755] [Medline: 35625483]
4. Side S, Mulbar U, Sidjara S, Sanusi W. A SEIR model for transmission of tuberculosis. Presented at: The 4th International Conference on Mathematical Sciences – Mathematical Sciences: Championing the Way in a Problem Based and Data Driven Society; Nov 15-17, 2016. [doi: 10.1063/1.4980867]
5. Adi YA. An investigation of Susceptible–exposed–infectious–recovered (SEIR) tuberculosis model dynamics with pseudo-recovery and psychological effect. Healthcare Analytics 2024 Dec;6:100361. [doi: 10.1016/j.health.2024.100361]
6. Tang N, Yuan M, Chen Z, et al. Machine learning prediction model of tuberculosis incidence based on meteorological factors and air pollutants. Int J Environ Res Public Health 2023 Feb 22;20(5):3910. [doi: 10.3390/ijerph20053910] [Medline: 36900920]
7. Bartl L, Zeeb M, Kälin M, et al. Machine learning-based prediction of active tuberculosis in people with HIV using clinical data. Clin Infect Dis 2025 Oct 6;81(3):521-530. [doi: 10.1093/cid/ciaf149] [Medline: 40132061]

XSL•FO
RenderX

8. Verboven L, Calders T, Callens S, et al. A treatment recommender clinical decision support system for personalized medicine: method development and proof-of-concept for drug resistant tuberculosis. BMC Med Inform Decis Mak 2022 Mar 2;22(1):56. [doi: 10.1186/s12911-022-01790-0] [Medline: 35236355]
9. Wang M, Lee C, Wei Z, Ji H, Yang Y, Yang C. Clinical assistant decision-making model of tuberculosis based on electronic health records. BioData Min 2023 Mar 16;16(1):11. [doi: 10.1186/s13040-023-00328-y] [Medline: 36927471]
10. Hethcote HW. The mathematics of infectious diseases. SIAM Rev 2000 Jan 1;42(4):599-653. [doi: 10.1137/S0036144500371907]
11. Jittimanee S, Vorasingha J, Mad-asin W, Nateniyom S, Rienthong S, Varma JK. Tuberculosis in Thailand: epidemiology and program performance, 2001-2005. Int J Infect Dis 2009 Jul;13(4):436-442. [doi: 10.1016/j.ijid.2008.07.025] [Medline: 19013094]
12. Division of Tuberculosis, Department of Disease Control. The report on the implementation of the National Tuberculosis Control Program 2016–2020. : Ministry of Public Health; 2022.
13. Division of Tuberculosis, Department of Disease Control. Report of the 6th Tuberculosis Joint International Monitoring Mission Thailand: 29 May to 6 June 2022. : Ministry of Public Health; 2022.
14. Rubin DB. Multiple Imputation for Nonresponse in Surveys: John Wiley & Sons; 1987. [doi: 10.1002/9780470316696]
15. van Buuren S. Flexible Imputation of Missing Data, 2nd edition: Chapman and Hall/CRC; 2018. [doi: 10.1201/9780429492259]
16. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? Int J Methods Psych Res 2011 Mar;20(1):40-49. [doi: 10.1002/mpr.329]
17. Hosmer DW, Lemeshow S. Applied Logistic Regression: John Wiley & Sons; 2000. [doi: 10.1002/0471722146]
18. Steyerberg EW. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating: Springer; 2009. [doi: 10.1007/978-0-387-77244-8]
19. Cox DR. Regression models and life-tables. J R Stat Soc Series B Stat Methodol 1972 Jan 1;34(2):187-202. [doi: 10.1111/j.2517-6161.1972.tb00899.x]
20. Kleinbaum DG, Klein M. Survival Analysis: A Self-Learning Text: Springer; 2012. [doi: 10.1007/978-1-4419-6646-9]
21. Keeling MJ, Rohani P. Modeling Infectious Diseases in Humans and Animals: Princeton University Press; 2008. [doi: 10.2307/j.ctvcm4gk0]
22. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982 Apr;143(1):29-36. [doi: 10.1148/radiology.143.1.7063747] [Medline: 7063747]
23. Harrell FE Jr, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. JAMA 1982 May 14;247(18):2543-2546. [doi: 10.1001/jama.1982.03320430047030] [Medline: 7069920]
24. Dowdy DW, Dye C, Cohen T. Data needs for evidence-based decisions: a tuberculosis modeler's "wish list". Int J Tuberc Lung Dis 2013 Jul;17(7):866-877. [doi: 10.5588/ijtld.12.0573] [Medline: 23743307]
25. Lelisho ME, Wotale TW, Tareke SA, et al. Survival rate and predictors of mortality among TB/HIV co-infected adult patients: retrospective cohort study. Sci Rep 2022 Nov 1;12(1):18360. [doi: 10.1038/s41598-022-23316-4] [Medline: 36319734]
26. Xie Y, Han J, Yu W, Wu J, Li X, Chen H. Survival analysis of risk factors for mortality in a cohort of patients with tuberculosis. Can Respir J 2020;2020:1654653. [doi: 10.1155/2020/1654653] [Medline: 32963642]
27. Bello S, Afolabi RF, Ajayi DT, et al. Empirical evidence of delays in diagnosis and treatment of pulmonary tuberculosis: systematic review and meta-regression analysis. BMC Public Health 2019 Jun 25;19(1):820. [doi: 10.1186/s12889-019-7026-4] [Medline: 31238906]
28. Arriaga MB, Araújo-Pereira M, Barreto-Duarte B, et al. The effect of diabetes and prediabetes on antituberculosis treatment outcomes: a multicenter prospective cohort study. J Infect Dis 2022 Feb 15;225(4):617-626. [doi: 10.1093/infdis/jiab427] [Medline: 34651642]
29. Bezerra AL, da Silva Rezende Moreira A, Isidoro-Gonçalves L, et al. Clinical, laboratory, and radiographic aspects of patients with pulmonary tuberculosis and dysglycemia and tuberculosis treatment outcomes. J Bras Pneumol 2022;48(6):e20210505. [doi: 10.36416/1806-3756/e20210505] [Medline: 36449815]
30. Burke RM, Nliwasa M, Feasey HRA, et al. Community-based active case-finding interventions for tuberculosis: a systematic review. Lancet Public Health 2021 May;6(5):e283-e299. [doi: 10.1016/S2468-2667(21)00033-5] [Medline: 33765456]
31. Bohlbro AS, Hvingelby VS, Rudolf F, Wejse C, Patsche CB. Active case-finding of tuberculosis in general populations and at-risk groups: a systematic review and meta-analysis. Eur Respir J 2021 Oct;58(4):2100090. [doi: 10.1183/13993003.00090-2021] [Medline: 33766950]
32. Rashidi HH, Dang LT, Albahra S, Ravindran R, Khan IH. Automated machine learning for endemic active tuberculosis prediction from multiplex serological data. Sci Rep 2021 Sep 9;11(1):17900. [doi: 10.1038/s41598-021-97453-7] [Medline: 34504228]
33. Wang Z, Guo Z, Wang W, et al. Prediction of tuberculosis treatment outcomes using biochemical makers with machine learning. BMC Infect Dis 2025 Feb 17;25(1):229. [doi: 10.1186/s12879-025-10609-y] [Medline: 39962412]

XSL•FO
RenderX

34. K B HM, Jose SA, Jirawattanapanit A, Mathew K. A comprehensive study on tuberculosis prediction models: integrating machine learning into epidemiological analysis. J Theor Biol 2025 Jan 21;597:111988. [doi: 10.1016/j.jtbi.2024.111988] [Medline: 39542168]
35. Biswas MHA, Paiva LT, de Pinho MDR. A SEIR model for control of infectious diseases with constraints. Math Biosci Eng 2014;11:761-784. [doi: 10.3934/mbe.2014.11.761]
36. Ucakan Y, Gulen S, Koklu K. Analysing of tuberculosis in Turkey through SIR, SEIR and BSEIR mathematical models. Math Comput Model Dyn Syst 2021 Jan 2;27(1):179-202. [doi: 10.1080/13873954.2021.1881560]
37. Ghosh A, Das P, Das SK, Das P, Upadhyay RK. Forecasting tuberculosis through mechanistic learning of transmission dynamics: insights from a case study in India. Comput Biol Med 2025 Nov;198(Pt B):111225. [doi: 10.1016/j.compbiomed.2025.111225] [Medline: 41130159]
38. Shaweno D, Karmakar M, Alene KA, et al. Methods used in the spatial analysis of tuberculosis epidemiology: a systematic review. BMC Med 2018 Oct 18;16(1):193. [doi: 10.1186/s12916-018-1178-4] [Medline: 30333043]
39. Faccin M, Geenen C, Happaerts M, Ombelet S, Migambi P, André E. Analyzing satellite imagery to target tuberculosis control interventions in densely urbanized areas of Kigali, Rwanda: cross-sectional pilot study. JMIR Public Health Surveill 2025 Apr 24;11:e68355. [doi: 10.2196/68355] [Medline: 40273403]
40. Lee Y, Raviglione MC, Flahault A. Use of digital technology to enhance tuberculosis control: scoping review. J Med Internet Res 2020 Feb 13;22(2):e15727. [doi: 10.2196/15727] [Medline: 32053111]
41. McQuaid CF, Vassall A, Cohen T, Fiekert K, White RG. The impact of COVID-19 on TB: a review of the data. Int J Tuberc Lung Dis 2021 Jun 1;25(6):436-446. [doi: 10.5588/ijtld.21.0148] [Medline: 34049605]

Abbreviations

AUC: area under the receiver operating characteristic curve
HR: hazard ratio
ML: machine learning
NTIP: National Tuberculosis Information Program
SEIR: susceptible-exposed-infectious-recovered
SEIR-D: susceptible-exposed-infectious-recovered-death
TB: tuberculosis
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 25.Oct.2025; peer-reviewed by A Girma, BA Owusu, E Lau, I Celik; revised version received 03.May.2026; accepted 30.May.2026; published 25.Jun.2026.

Please cite as:
Saksaen P, Boonchieng E, Thongprachum A, Maotheuak S, Chautrakarn S, Boonchieng W
Predicting Tuberculosis Outcomes Using Routine Surveillance Data in Chiang Mai, Thailand: Retrospective Cohort Study
JMIR Public Health Surveill 2026;12:e86495
URL: https://publichealth.jmir.org/2026/1/e86495
doi:10.2196/86495

© Porramat Saksaen, Ekkarat Boonchieng, Aksara Thongprachum, Surasak Maotheuak, Sineenart Chautrakarn, Waraporn Boonchieng. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 25.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Comparison of Measured 24-Hour Urinary Salt Excretion With Spot Urine and 24-Hour Dietary Recall Estimates Among Adolescents and Parents: Cross-Sectional Study

Sandeep Kaur[1,2,3*], MSc, PhD; Akashdeep Singh Chauhan[2,3*], MPH, PhD; Rajesh Kumar[3,4*], MBBS, MD; Manmeet Kaur[3,4*], PhD

[1]Faculty of Medicine, Imperial College London, Hammersmith Campus, London, United Kingdom
[2]Indian Institute of Public Health-Delhi, Public Health Foundation of India, New Delhi, India
[3]Department of Community Medicine and School of Public Health, Post Graduate Institute of Medical Education and Research, Chandigarh, India
[4]Health Equity Action Learnings Foundation, Chandigarh, India
[*]all authors contributed equally

Corresponding Author:
Sandeep Kaur, MSc, PhD
Faculty of Medicine, Imperial College London, Hammersmith Campus, London, United Kingdom

Abstract

Background: Hypertension, a leading cause of cardiovascular disease, is a significant public health challenge, with urban India reporting prevalence rates up to 40%. Excessive salt intake, averaging 8 to 11 g/day in India, far exceeds the World Health Organization–recommended limit of 5 g/day and is a key modifiable risk factor. While 24-hour urine collection is the gold standard for measuring 24-hour urinary excreted salt levels, its practicality in large-scale studies is limited, necessitating alternative methods such as spot urine sampling and dietary recall.

Objective: This study aimed to address current knowledge gaps by evaluating 24-hour urinary excretory salt levels across age groups and comparing the gold standard of 24-hour urinary salt excretion with estimates from spot urine samples and 24-hour dietary recalls among adolescents and their parents in Chandigarh, India.

Methods: A cross-sectional survey was conducted among 462 adolescents and 395 parents from public schools in Chandigarh, part of a cluster randomized controlled trial focused on reducing chronic disease risk factors. Twenty-four–hour urinary excreted salt levels were measured via 24-hour urine collection, spot urine samples were analyzed with the Kawasaki, Tanaka, and INTERSALT equations, and 2 nonconsecutive 24-hour dietary recalls using a correction factor of 0.9. Bland-Altman plots were used to assess agreement between methods.

Results: The mean 24-hour urinary excreted salt levels measured via 24-hour urine samples were 6.95 g/day (SD 0.24; 95% CI 6.48 - 7.42) for adolescents and 7.73 g/day (SD 0.3; 95% CI 7.18 - 8.28) for parents. Spot urine estimates showed variability, with the INTERSALT equation with potassium yielding the closest results to 24-hour measurements for parents (mean 8.12, SD 0.4 g/d; 95% CI 7.36 - 8.95) and adolescents (mean 7.97, SD 0.4 g/day; 95% CI 7.16 - 8.77). Dietary recall underestimated 24-hour urinary excreted salt levels for both groups (mean 4.15, SD 0.08 g/day for adolescents and mean 3.62, SD 0.08 g/day for parents).

Conclusions: Spot urine sampling, particularly the INTERSALT with potassium equation, offers a feasible alternative to 24-hour urine collection for estimating population-level 24-hour urinary excreted salt levels among adolescents and adults in resource-limited settings. Dietary recall alone significantly underestimates intake and should be supplemented with objective measures. These findings support the use of tailored estimation methods based on age for improved hypertension management strategies in India.

Trial Registration: Clinical Trials Registry-India CTRI/2019/09/021452; https://ctri.nic.in/Clinicaltrials/pmaindet2.php?EncHid=MzcwMzY

International Registered Report Identifier (IRRID): RR2-10.1371/journal.pone.0263584

(JMIR Public Health Surveill 2026;12:e85549) doi:10.2196/85549

KEYWORDS

hypertension; cardiovascular diseases; urinary excreted salt levels; chronic diseases; excretory salt; 24-hour urine collection; spot urine sampling; dietary recall; adolescents

XSL•FO
RenderX

Introduction

Hypertension is a major global health issue, significantly contributing to the burden of cardiovascular disease, which remains the leading cause of death worldwide [1]. In India, the prevalence of hypertension is high, with 40% of urban adults living with hypertension, compared with 17% in rural areas [2]. In Chandigarh, a union territory, the issue is particularly pronounced, affecting 45% of women and 58% of men [2].

Excessive salt intake is one of the most impactful yet modifiable risk factors for hypertension [3-5]. Globally, average salt consumption is nearly twice the maximum daily intake of 5 g/day recommended by the World Health Organization (WHO) [6]. In India, estimates indicate that people consume an average of 11 g/day, while recent national data report a mean intake of 8 g/day among adults aged 18 to 69 years [2,6]. Similar trends are seen in other countries; for instance, a study from Iran found that adolescents aged 10 to 15 years consume about 9.7 g/day, reflecting a widespread pattern of high salt intake across age groups [7]. The latest 2024 report suggests that nearly 1.4 billion people worldwide are hypertensive, underscoring the need for effective preventive and management strategies [1,8]. Studies from India and other low- and middle-income countries have shown that high salt consumption during adolescence is a risk factor for hypertension later in life [5,9].

To create effective management strategies, it is important to have clear criteria for evaluating the impact of different approaches. While the health risks linked to high salt intake are well known, accurately measuring salt consumption in the population is limited by several challenges. These difficulties make it hard to obtain reliable data on salt intake patterns, which in turn affects the development and implementation of targeted public health programs aimed at reducing health risks [10]. While the 24-hour urine collection method is the most accurate for estimating 24-hour urinary excretory salt levels, it is often impractical for large population studies due to logistical demands. Spot urine sampling provides a more feasible alternative, using formulas such as INTERSALT, Kawasaki, and Tanaka to estimate 24-hour urinary salt excretion from single urine samples or through 24-hour dietary recall [10,11]. However, the applicability and accuracy of these methods within the Indian context are not well studied, and few investigations have compared these estimates to the benchmark 24-hour urine collection [12]. Moreover, in India, previous studies on urinary salt excretion have focused exclusively on adults. To our knowledge, no study has objectively examined this important factor among adolescents using urine samples, although several studies have assessed dietary salt intake in this age group. Thus, this is the first study, to our knowledge, to objectively estimate urinary salt excretion among adolescents in the Indian context [13,14].

This study aimed to address current knowledge gaps by evaluating the 24-hour urinary excretory salt levels across age groups and comparing the gold standard of 24-hour urinary salt excretion with estimates from spot urine samples and 24-hour dietary recalls among adolescents and their parents in Chandigarh.

Methods

This study was based on a cross-sectional survey carried out as part of a cluster randomized controlled trial to assess the impact of a school-based health promotion intervention package in reducing the behavioral risk factors associated with chronic diseases [15]. The survey was conducted among adolescents studying in public schools and one of their parents in the Union Territory of Chandigarh, India, from May 2018 to September 2019.

Study Setting

The study was conducted in public schools as the Director of Public Instruction did not allow enrollment of private schools in the study. Data collected from the Department of School Education, Chandigarh, revealed that 106 public schools had eighth-grade classes, and on average, 10,790 adolescents studied in eighth grade in these schools. Most schools had 4 to 5 sections in the eighth-grade class. Each section of the class had about 30 adolescents.

Participants

On the basis of the knowledge that behaviors contributing to chronic diseases, such as a preference for salty foods, often emerge during early adolescence, this study selected assenting eighth-grade adolescents and one of their parents based on their mutual preference [16]. Inclusion criteria required that participants planned to remain in the city for at least 1 year following recruitment and that the selected parent was younger than 65 years.

Sample Size and Sampling

Details on the sample size calculation are described in the protocol paper of the broader cluster randomized controlled trial [17]. Broadly, calculations considered expected means or proportions for the control group, estimated intervention effects, and coefficient of variation values derived from adolescent data from previous studies [18-21]. To ensure sufficient power across all risk factors, the maximum calculated cluster size (n=12) was selected. Consequently, 12 clusters (schools), each with 30 adolescents, yielded an estimated total sample size of 360 adolescents and 360 parents.

Data Collection Tools

Several tools were used for data collection after receiving assent from adolescents (Multimedia Appendix 1) and consent from parents (Multimedia Appendix 2). These tools included the interview schedule for recording sociodemographic information, behavioral risk factors, and anthropometric, physiological, and biochemical measures (Multimedia Appendix 3).

A 24-hour dietary recall questionnaire was used to assess the 24-hour dietary urinary excreted salt levels. Evidence suggests that a single 24-hour dietary recall is unable to provide the usual consumption pattern of the participants. To overcome this drawback, two 24-hour dietary recalls were collected for all participants on nonconsecutive days [22].

Twenty-four–hour urine and morning second-spot urine samples were collected to estimate the level of urinary salt excretion

XSL•FO
RenderX

(g/day). To assess BMI, height was measured with a United Nations Children's Fund standardized anthropometer to the nearest 0.1 cm, and weight was measured with a portable electronic weighing scale to the nearest 0.1 kg. Blood pressure (BP) was recorded for all participants using a digital sphygmomanometer (Omron Healthcare Co Ltd).

Data Collection Methods

SK collected the dietary, anthropometric, and physiological data. Additionally, a team of 3 members, including SK and 2 field investigators, collected urine samples. The field investigators were qualified up to the graduation level, and they had received training for data collection related to urine samples.

Neither the investigator nor the data collectors were blinded to the intervention. Standard operating procedures were developed to maintain uniformity and standardization in sample collection and analysis (Multimedia Appendix 4)

After obtaining the study participants' written informed consent, SK provided questionnaires to adolescents to record their responses in the classroom setting for recording sociodemographic characteristics and 24-hour dietary recall. SK was present in the classroom and explained each question to all the adolescents before they recorded their responses. Any query raised by the adolescents was answered.

To record the two 24-hour dietary recalls, all the adolescents in the same cluster filled out their forms themselves (self-administration). In comparison, in-person interviews were carried out with parents. The first 24-hour dietary recall was recorded during the first interaction with all the participants in the schools. The second 24-hour dietary recall was collected a week later. The data collection of the 24-hour dietary recall was manually recorded using the Multiple-Pass Method [22]. After collecting sociodemographic data and dietary recalls, participants' anthropometric and BP measurements were obtained in a separate room one by one. An average of 3 readings was used for each participants' BP measurement, with a resting period of at least 5 minutes between each reading. Finally, instructions for the collection of urine samples were provided. Specific days were fixed for collecting urine samples from the home setting. The details of BMI, BP, and 24-hour and second-spot urine samples collection are provided in Multimedia Appendices 4 and 5.

During Saturday home visits, participants received instructions regarding the process of 24-hour urine collection on Sundays and second-spot urine collection on Monday morning. Urine collection containers were provided to them. They were instructed to discard the first void of the morning and then start collecting the subsequent voids for the next 24 hours (Multimedia Appendix 6). Second-spot urine samples are preferred over first morning samples because they are less influenced by overnight concentration and better reflect daytime sodium excretion. This improves the accuracy of estimating 24-hour urinary sodium compared with first morning samples [23].

Participants were advised to write the time of their voids and whether they had collected each void in the collection container or not for 24 hours using a structured proforma (Multimedia Appendix 7). The urine samples were collected on Monday mornings and analyzed for urinary salt levels on the same day.

Data Management and Analysis

The 24-hour dietary data were entered in a Microsoft Excel sheet. For the estimation of sodium intake, the Prospective Urban Rural Epidemiology study computer software was used [24]. This tool was developed explicitly for assessing the intake of Indian foods and has been previously validated in Indian settings [24]. The details of the estimation of 24-hour urinary excreted salt levels using this tool have been provided in Multimedia Appendix 8. Twenty-four–hour dietary urinary excreted salt levels from the 2 recalls of every participant were calculated separately and then averaged and multiplied by a correction factor of 0.9 to estimate their usual 24-hour urinary excreted salt levels.

For BMI categorization of adolescents into severe thinness, thinness, normal weight, overweight, and obese categories, WHO growth reference percentiles that were age and sex specific for the 5- to 19-year age group were used [25]. For parents, WHO Asia-specific BMI cutoffs were used to categorize them into underweight, normal weight, overweight, and obese categories [26].

For BP assessment among adolescents, percentiles for systolic BP and diastolic BP were assessed based on each adolescent's height, age, and sex (normal: <90th percentile; borderline-elevated BP: systolic BP or diastolic BP ≥90th percentile but <95th percentile; elevated BP: systolic BP or diastolic BP ≥95th percentile). Recent 2025 American College of Cardiology and American Heart Association guidelines were used to categorize adult participants into normal, elevated, stage 1, and stage 2 hypertension [27].

Urinary excreted salt levels were estimated from spot urine samples using a series of established estimation equations. The equations used were Tanaka, Kawasaki, and INTERSALT with potassium [28] (Multimedia Appendix 9). Twenty-four–hour urinary excreted salt levels values that were greater than 3 SDs above or below the mean were excluded from all analyses. Twenty-four–hour urinary salt excretion was estimated from 24-hour urine samples using the formula mentioned in Multimedia Appendix 9. As both 24-hour urine collections (gold standard) and spot urine–based equations estimate 24-hour urinary salt excretion, the salt intake derived from 24-hour dietary recall was adjusted using a 0.9 correction factor to approximate urinary excreted salt. This adjustment is based on the established physiological evidence that approximately 90% of ingested salt is excreted in urine under steady-state conditions [29]. Paired samples 2-tailed *t* tests were used to compare mean population 24-hour urinary excreted salt levels estimated by the 24-hour urine collections, spot urine equation-based methods, and 24-hour dietary recalls. To determine the agreement between individual-level 24-hour urinary excreted salt levels estimated from 24-hour collections and methods based on spot urine collections and dietary recall, Bland-Altman plots were used. The difference between the methods was plotted against the mean of the 2 methods. Data are presented as mean (95% CI) unless otherwise specified. Statistical analyses were conducted using R (version 4.3.0; R Foundation for Statistical Computing).

XSL•FO
RenderX

Ethical Considerations

Ethics approval was obtained from the institutional ethics committee of the Post-Graduate Institute of Medical Education and Research, Chandigarh (INT/IEC/2018/000082; January 22, 2019). Permission to conduct the study in schools was granted by the Department of Education, Chandigarh, with letter number 1296-DSE-UT-S5-11(65)11-II. As this study was part of a cluster randomized controlled trial, the study was registered in Clinical Trials Registry-India (CTRI/2019/09/021452).

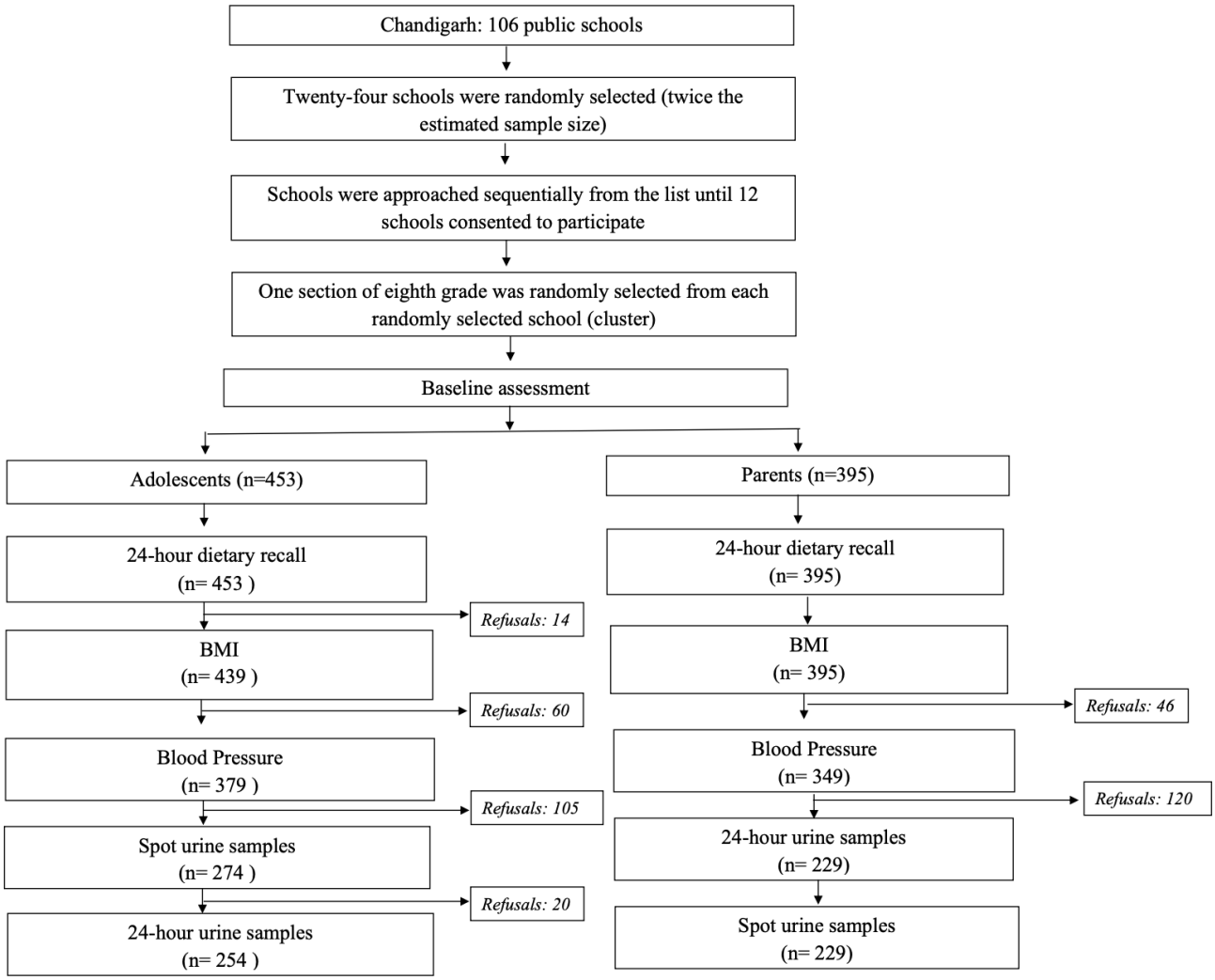
Written assent of adolescents and written consent from their parents were obtained after informing them about the study purpose, data to be collected, estimated time for data collection, the confidentiality of the data, and risks involved before their enrollment in the study (Multimedia Appendices 1 and 2).

Results

Participants

Of the 462 eligible adolescents, 453 (98.1%) consented to be part of the study (Figure 1). One parent was supposed to be enrolled in the study for each adolescent. However, 395 consenting parents enrolled in the study.

Figure 1. Sampling strategy and participation rates for various indicators among all study participants.



The mean age was 13 (SD 0.4) years for adolescents and 37 (SD 0.3) years for parents (Table 1). Of all participating adolescents, 46.1% (209/453) were girls, and of all parents, 61% (241/395) were mothers. Nearly a quarter (105/439, 23.9%) of the participating adolescents were underweight, whereas half (197/395, 49.9%) of the parents were either overweight or obese (Table 1). Among adolescents, 3.4% (13/379) had stage 2 hypertension compared with 25.5% (89/349) of parents. There were differential participation rates for each of the indicators used in this study, as shown in Figure 1.

XSL•FO
RenderX

Table . Study participants' characteristics.

Participant characteristics	Adolescents	Parents
Age (y), mean (SD; 95% CI)	13.12 (0.4; 13.05-13.19)	37.1 (0.3; 36.54-37.67)
Female, n (%)	209/453 (46.1)	241/395 (61.0)
Weight (kg), mean (SD; 95% CI)	39.4 (0.45; 38.6-40.3)	61.95 (0.63; 60.71-63.19)
BMI (kg/m^2), mean (SD; 95% CI)	17.4 (0.17; 17.1-17.8)	25.11 (0.22; 24.67-25.55)
BMI classification[a,b], n (%)		
Underweight	105/439 (23.9)	27/395 (6.8)
Overweight and obese	46/439 (10.5)	197/395 (49.9)
Blood pressure[c,d], n (%)		
Normal	346/379 (91.3)	204/349 (53.7)
Hypertension stage 1	17/379 (4.5)	56/349 (14.7)
Hypertension stage 2	13/379 (3.4)	89/349 (23.4)

[a]Adolescent BMI categories were based on World Health Organization growth reference percentiles age and sex specific for 5-19 y (2007): ≤−2 SD, severe thinness; >−2 SD to ≤−1 SD, thinness; >−1 SD to ≤+1 SD, normal weight; >+1 SD to ≤+2 SD, overweight; and >+2 SD, obesity.

[b]Parent BMI categories were defined as follows: underweight, <18.5 kg/m^2; normal weight, 18.5-24.9 kg/m^2; overweight, 25.0-29.9 kg/m^2; and obese, ≥30.0 kg/m^2.

[c]Adolescent blood pressure (BP) categories were based on age-, height-, and sex-specific percentiles: normal, <90th percentile; borderline-elevated BP, systolic BP or diastolic BP ≥90th percentile but <95th percentile; elevated BP, systolic BP or diastolic BP ≥95th percentile.

[d]Parent BP categories were defined as follows: normal, systolic BP <120 mm Hg and diastolic BP <80 mm Hg; elevated BP, systolic BP 120-129 mm Hg and diastolic BP <80 mm Hg; stage 1 hypertension, systolic BP 130-139 mm Hg or diastolic BP 80-89 mm Hg; and stage 2 hypertension, systolic BP ≥140 mm Hg or diastolic BP ≥90 mm Hg.

Urinary Excreted Salt Levels

Based on the 24-hour urine samples, the measured 24-hour urinary excreted salt level among adolescents was 6.95 g/day (SD 0.24; 95% CI 6.48-7.42), and for their parents it was 7.73 g/day (SD 0.3; 95% CI 7.18-8.28). These measured estimates were compared with estimated 24-hour urinary excreted salt levels derived from various urine equations by using spot urine samples and 24-hour dietary recall. (Multimedia Appendix 9).

Twenty-four–hour dietary recall estimates of 24-hour urinary excreted salt levels were the lowest for both adolescents (mean 4.15, SD 0.08 g/day; 95% CI 3.99-4.31) and parents (mean 3.62, SD 0.08 g/day; 95% CI 3.47, 3.77; Table 2).

Table . Population-level estimates of 24-hour urinary excreted salt among adolescents and their parents[a].

	Adolescents (g/day), mean (SD; 95% CI)	Parents (g/day), mean (SD; 95% CI)
Measured 24-hour urinary salt	6.95 (0.24; 6.48-7.42)	7.73 (0.3; 7.18-8.28)
Estimated 24-hour dietary salt	4.15 (0.08; 3.99-4.31)	3.62 (0.08; 3.47-3.77)
Estimated Kawasaki equation	12.83 (0.5; 11.92-13.74)	12.72 (0.3; 12.17-13.28)
Estimated Tanaka equation	5.79 (0.13; 5.55-6.04)	6.89 (0.12; 6.65-7.13)
Estimated INTERSALT equation with potassium	7.97 (0.4; 7.16-8.77)	8.12 (0.4; 7.36-8.95)

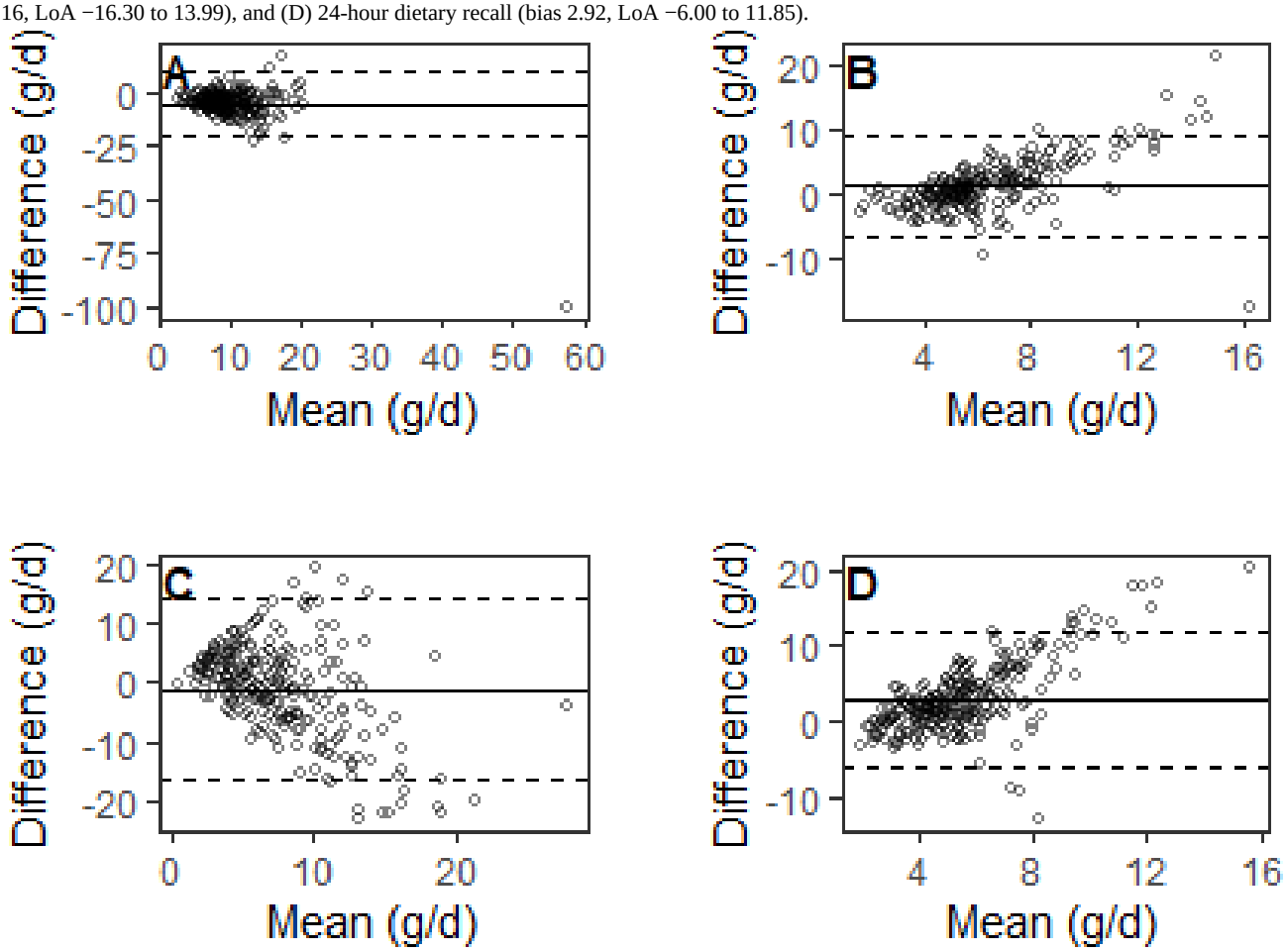
[a]Sample sizes varied by indicator and are presented in Figure 1. Equations used to estimate urinary salt excretion are provided in Multimedia Appendix 9.

Agreement Between Methods of Estimating 24-Hour Urinary Excreted Salt Levels

For population-level 24-hour urinary excreted salt level estimation, comparisons among methods revealed that among adolescents there were no significant differences between the estimates derived from 24-hour urine collections and those obtained using the Kawasaki (P<0.001) and Tanaka equations (P=0.008), as well as 24-hour dietary recall (P<0.001). The INTERSALT equation with potassium slightly underestimated the 24-hour salt levels (mean −1.16 g/day; 95% limits of agreement −16.3 to 13.99 g/day; P=.02; Figure 2).

XSL•FO
RenderX

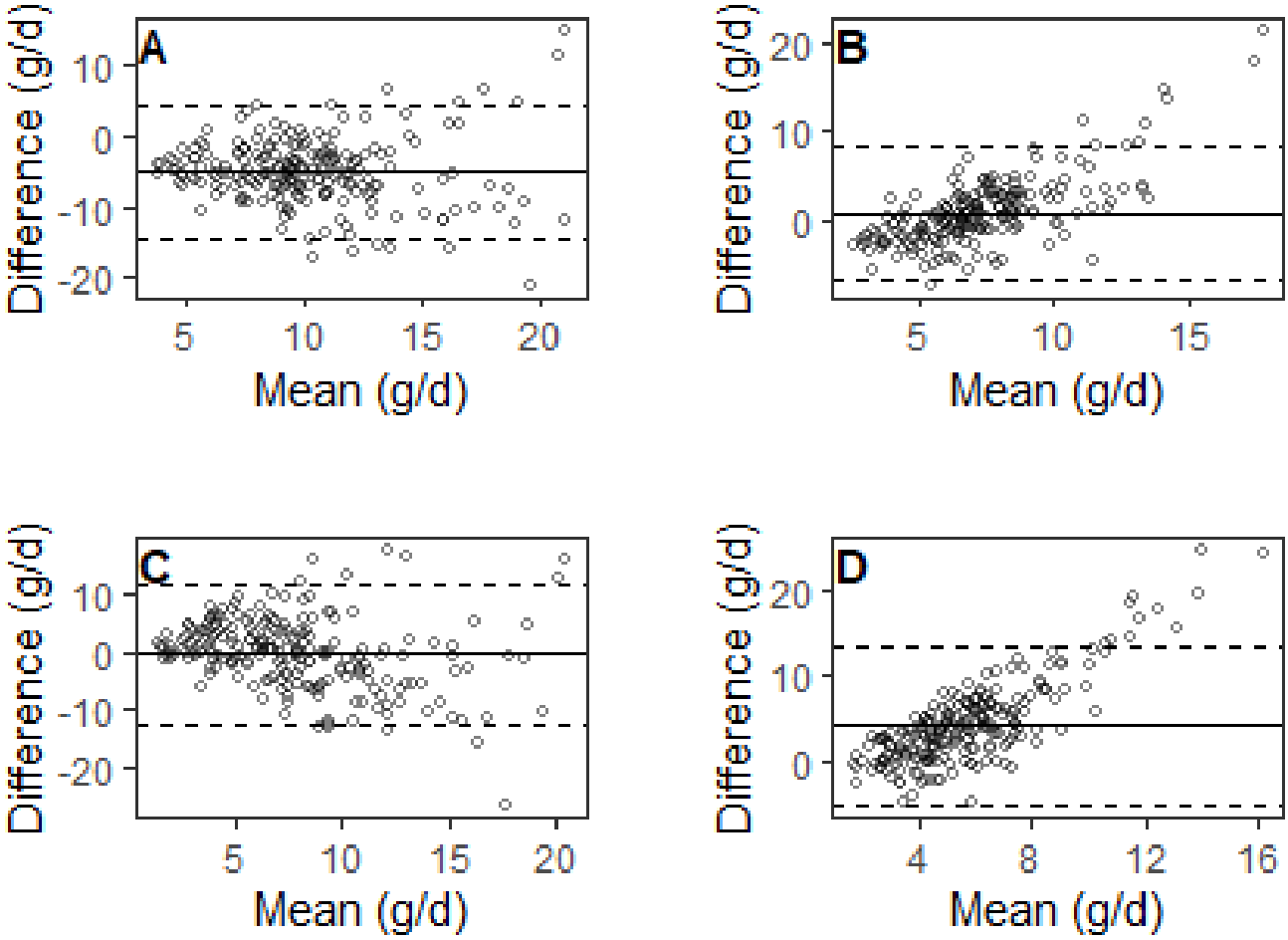
Figure 2. Bland-Altman plots comparing 24-hour urinary excretory salt estimates from 24-hour urine, spot urine samples, and 24-hour dietary recall in adolescents. Bias and 95% limits of agreement (LoA) in comparing 24-hour urinary salt excretion assessed through 24-hour urine samples (gold standard) with (A) Kawasaki (bias −5.78, LoA −21.13 to 9.57), (B) Tanaka (bias 1.29, LoA −6.52 to 9.10), (C) INTERSALT with potassium (bias −1.16, LoA −16.30 to 13.99), and (D) 24-hour dietary recall (bias 2.92, LoA −6.00 to 11.85).



Among parents, there were no significant differences between the estimates derived from 24-hour urine collections and those obtained using the Kawasaki and INTERSALT equations, as well as 24-hour dietary recall. Tanaka slightly overestimated the 24-hour salt levels (mean 0.71 g/day; 95% limits of agreement −6.82 to 8.24 g/day; *P*=.008; Figure 3).

XSL•FO
RenderX

Figure 3. Bland-Altman plots comparing 24-hour urinary excretory salt estimates from 24-hour urine samples, spot urine samples, and 24-hour dietary recall in parents. The graphs show bias and 95% limits of agreement (LoA) in assessing 24-hour urinary salt excretion through 24-hour urine samples (the gold standard) with (A) Kawasaki (bias −5.16, LoA −14.42 to 4.10), (B) Tanaka (bias 0.71, LoA −6.82 to 8.24), (C) INTERSALT with potassium (bias −0.23, LoA −12.22 to 11.75), and (D) 24-hour dietary recall (bias 4.18, LoA −5.25 to 13.60) methods.



Discussion

Key Findings

This study suggests that spot urine samples can be practical alternatives to 24-hour collections for estimating population-level 24-hour urinary excreted salt levels in Indian populations, with INTERSALT with potassium being the most suitable method. The INTERSALT with potassium equation showed closer agreement with the gold standard, 24-hour urine samples, at the population level; in contrast, the 24-hour dietary recall method tended to underestimate the urinary excreted salt levels. Adolescence is a period of significant physiological change with distinct nutritional requirements; therefore, 24-hour urine samples remain the gold standard for assessing 24-hour urinary excreted salt levels in this age group. However, understanding 24-hour urinary excreted salt levels in this group is crucial, as studies suggest that high salt consumption in early years is a major risk factor for hypertension later in life. Importantly, this is a modifiable behavior that can be addressed during early adolescence through appropriate interventions [5,9].

The mean population intake estimated in the study is comparable to the earlier studies; for instance, standard 24-hour urinary salt excretion, widely regarded as the gold standard, was measured at 7 g/day among adolescents and 8 g/day among their parents [13,30]. These findings align closely with previous studies, such as one conducted in China, where the average 24-hour urinary salt excretion for adolescents was 8.2 g/day using 24-hour urine samples [31]. Additionally, an Indian study estimated the weighted mean 24-hour urinary excreted salt levels from 24-hour urine samples to be 8.59 g/day in Delhi and 9.46 g/day in Andhra Pradesh, further supporting the validity of our findings [30]. In this study, the mean 24-hour urinary salt excretion estimated through the INTERSALT equation with potassium, based on spot urine samples, was 8.0 g/day for adolescents and 8.12 g/day for their parents. These values are consistent with Dong et al [31], where adolescent 24-hour urinary salt excretion was reported as 8.4 g/day using the same equation. Another Indian study using the INTERSALT equation with potassium yielded comparable results, with estimates of 8.13 g/day in Delhi and 8.81 g/day in Andhra Pradesh [30]. Together, these findings suggest that the INTERSALT equation with potassium can yield population-level estimates closely aligned with those obtained from 24-hour urine collections for adults, making spot urine samples a feasible option in settings where 24-hour collections are challenging.

The biochemical assessment of 24-hour urinary excreted salt levels, based on urinary measures, proves to be more accurate than dietary recall methods. In this study, dietary recalls on 2 nonconsecutive days estimated 24-hour urinary excreted salt

XSL•FO
RenderX

levels between 4 and 4.5 g/day for participants across all age groups. This underestimation is supported by the literature; for example, McLean et al [32] conducted a meta-analysis of 28 studies and confirmed that dietary recall methods consistently underestimate 24-hour urinary excreted salt levels when compared with 24-hour urine collections.

Among the various methods for estimating 24-hour urinary excreted salt levels, the INTERSALT equation with potassium applied to spot urine samples produced no statistical difference from the gold standard of 24-hour urine collection, with differences of 0.23 g/day among parents. Previous studies indicate that among urine equations, including Kawasaki, Tanaka, and INTERSALT, the INTERSALT method demonstrates the most reliable performance in estimating mean population sodium intake, minimizing bias effectively [31,33]. However, there was a significant difference between 24-hour salt levels assessed through INTERSALT with potassium for adolescents. This variation may be attributed to differences in urinary potassium levels between adolescents and adults, influenced by growth-related physiological changes, dietary habits, and hormonal fluctuations during adolescence, all of which can affect potassium excretion [34].

Strengths and Limitations

This study has several strengths and limitations. Pilot testing was conducted prior to data collection to assess feasibility and refine research instruments, helping to address potential challenges. A key strength is the use of validated tools in the Indian context, along with standardized equipment and protocols for BP assessment.

The 24-hour dietary recall is a reliable method for assessing population-level intake compared with tools such as the Food Frequency Questionnaire. Although it may not fully capture usual intake, we used recalls on 2 nonconsecutive days to improve reliability. The study was adequately powered, and the findings can inform sample size calculations for future research.

A limitation is the exclusion of private schools, which may affect representation across socioeconomic groups. However, as a large proportion of the Indian population belongs to low- and middle-income groups, the findings remain reasonably generalizable to the Indian settings.

Overall, this study supports the utility of spot urine samples, specifically with equations such as the INTERSALT equation with potassium, for accurately estimating population-level 24-hour urinary excreted salt levels in adults.

Conclusions

In settings where 24-hour urine collections are impractical, spot urine equations such as INTERSALT with potassium offer effective alternatives for estimating 24-hour urinary sodium excretion in adult populations in the Indian context. However, 24-hour dietary recall methods often underestimate 24-hour urinary excreted salt levels, and their accuracy can be significantly improved when combined with objective measurements. This combined approach is especially valuable for designing and implementing effective salt reduction interventions. As part of a larger cluster randomized controlled trial, the findings can support objective assessment of urinary salt excretion in resource-limited settings. Future interventional studies may also adopt the tools used in this study to monitor changes in behavioral risk factors, including dietary behaviors assessed through 24-hour recall.

Acknowledgments

No generative artificial intelligence tools were used in the writing, analysis, or preparation of this manuscript.

The authors would also take this opportunity to extend their special thanks to the George Institute of Global Health, Hyderabad, for providing them with the funds to conduct laboratory testing of the urine samples collected as part of this study.

Funding

MK received funding from the George Institute for Global Health, Hyderabad (project ID 5562), to support laboratory testing of the urine samples collected as part of this study. The funder had no role in the conceptualization of the study, study conduct, data analysis, interpretation of findings, or dissemination of the results.

Data Availability

The fully anonymized dataset supporting the findings of this study is available in a public repository [35].

Authors' Contributions

SK contributed to conceptualization, investigation, data curation, and writing the original draft. RK and MK contributed to conceptualization, methodology, supervision, validation, and provision of resources. ASC contributed to formal analysis. All authors contributed to writing, review, and editing of the manuscript and approved the final version for publication.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Participant informed assent form.
[PDF File, 213 KB - publichealth_v12i1e85549_app1.pdf]

Multimedia Appendix 2
Parental informed consent form.
[PDF File, 333 KB - publichealth_v12i1e85549_app2.pdf]

Multimedia Appendix 3
Data collection tool.
[PDF File, 357 KB - publichealth_v12i1e85549_app3.pdf]

Multimedia Appendix 4
Standard operating procedure for urine samples (24-hour and spot).
[PDF File, 244 KB - publichealth_v12i1e85549_app4.pdf]

Multimedia Appendix 5
Standardized tools and methods to record height, weight, and blood pressure.
[PDF File, 213 KB - publichealth_v12i1e85549_app5.pdf]

Multimedia Appendix 6
Approaches of 24-hour and spot urine sample collection.
[PDF File, 233 KB - publichealth_v12i1e85549_app6.pdf]

Multimedia Appendix 7
Twenty-four-hour urine collection: field form.
[PDF File, 214 KB - publichealth_v12i1e85549_app7.pdf]

Multimedia Appendix 8
Estimation of dietary salt.
[PDF File, 244 KB - publichealth_v12i1e85549_app8.pdf]

Multimedia Appendix 9
Estimation of urinary salt.
[PDF File, 332 KB - publichealth_v12i1e85549_app9.pdf]

References

1. Global health estimates: leading causes of DALYs. World Health Organization. URL: https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/global-health-estimates-leading-causes-of-dalys [accessed 2024-01-15]
2. National Family Health Survey (NFHS-5), 2019-21: India: volume I. : International Institute for Population Sciences; 2021 URL: https://dhsprogram.com/pubs/pdf/FR375/FR375.pdf [accessed 2026-06-17]
3. Arokiasamy P. India's escalating burden of non-communicable diseases. Lancet Glob Health 2018 Dec;6(12):e1262-e1263. [doi: 10.1016/S2214-109X(18)30448-0] [Medline: 30292427]
4. Clermont A, Rouzier V, Pierre JL, et al. High dietary sodium, measured using spot urine samples, is associated with higher blood pressure among young adults in Haiti. Glob Heart 2023;18(1):5. [doi: 10.5334/gh.1187] [Medline: 36817226]
5. Konlan KD, Lee H, Afaya A. Prevalence and risk factors of hypertension among adolescents living in Ghana. Int J Environ Health Res 2024 Mar;34(3):1431-1442. [doi: 10.1080/09603123.2023.2218280] [Medline: 37272549]
6. Diet, nutrition, and the prevention of chronic diseases: report of a joint WHO/FAO expert consultation. : World Health Organization; 2003 URL: https://iris.who.int/server/api/core/bitstreams/c4f12939-5f55-4b75-b6f9-283024956b3f/content [accessed 2026-06-17]
7. Emamian MH, Ebrahimi H, Hashemi H, Fotouhi A. Salt intake and blood pressure in Iranian children and adolescents: a population-based study. BMC Cardiovasc Disord 2021 Feb 2;21(1):62. [doi: 10.1186/s12872-021-01876-z] [Medline: 33530964]
8. Global report on hypertension 2025: high stakes: turning evidence into action. : World Health Organization; 2025 URL: https://www.who.int/publications/i/item/9789240115569 [accessed 2026-06-17]
9. Shah K, Shah H, Vijan S, Kamat S, Kirpalani D, Kirpalani A. The salt study - salt intake in adolescents today. J Hypertens 2022;40(Suppl 1):e153. [doi: 10.1097/01.hjh.0000836892.98694.9d] [Medline: 35703297]

XSL•FO
RenderX

10. Huang L, Crino M, Wu JH, et al. Mean population salt intake estimated from 24-h urine samples and spot urine samples: a systematic review and meta-analysis. Int J Epidemiol 2016 Feb;45(1):239-250. [doi: 10.1093/ije/dyv313] [Medline: 26796216]
11. Mente A, O'Donnell MJ, Dagenais G, et al. Validation and comparison of three formulae to estimate sodium and potassium excretion from a single morning fasting urine compared to 24-h measures in 11 countries. J Hypertens 2014 May;32(5):1005-1014. [doi: 10.1097/HJH.0000000000000122] [Medline: 24569420]
12. Johnson C, Praveen D, Pope A, et al. Mean population salt consumption in India: a systematic review. J Hypertens 2017 Jan;35(1):3-9. [doi: 10.1097/HJH.0000000000001141] [Medline: 27755388]
13. Petersen KS, Johnson C, Mohan S, et al. Estimating population salt intake in India using spot urine samples. J Hypertens 2017 Nov;35(11):2207-2213. [doi: 10.1097/HJH.0000000000001464] [Medline: 28697010]
14. Johnson C, Mohan S, Rogers K, et al. The association of knowledge and behaviours related to salt with 24-h urinary salt excretion in a population from North and South India. Nutrients 2017 Feb 16;9(2):144. [doi: 10.3390/nu9020144] [Medline: 28212309]
15. Kaur S, Kumar R, Lakshmi PV, Kaur M. Effectiveness of a school-based behavioural change intervention in reducing chronic disease risk factors in Chandigarh, India: a cluster-randomised controlled trial. Lancet Reg Health Southeast Asia 2024;21:100353. [doi: 10.1016/j.lansea.2024.100353] [Medline: 38312946]
16. Kaur S, Kumar R, Kaur M. School-based behaviour change intervention to reduce ultra-processed food consumption among adolescents: evidence from a cluster-randomised controlled trial in India. BMJ Glob Health 2026 Jan 9;11(1):e020799. [doi: 10.1136/bmjgh-2025-020799] [Medline: 41513306]
17. Kaur S, Kaur M, Kumar R. Health promotion intervention to prevent risk factors of chronic diseases: protocol for a cluster randomized controlled trial among adolescents in school settings of Chandigarh (India). PLoS One 2022;17(2):e0263584. [doi: 10.1371/journal.pone.0263584] [Medline: 35176046]
18. Global school-based student health survey (age 13-15). Central Board of Secondary Education. 2007. URL: https://cdn.who.int/media/docs/default-source/ncds/ncd-surveillance/data-reporting/india/gshs/2007-india-gshs-fact-sheet.pdf?sfvrsn=b96a6143_2&download=true [accessed 2026-6-23]
19. Peltzer K, Pengpid S. Fruits and vegetables consumption and associated factors among in-school adolescents in five Southeast Asian countries. Int J Environ Res Public Health 2012 Oct 11;9(10):3575-3587. [doi: 10.3390/ijerph9103575] [Medline: 23202763]
20. Kumar P, Kumar D, Ranjan A, Singh CM, Pandey S, Agarwal N. Prevalence of hypertension and its risk factors among school going adolescents of Patna, India. J Clin Diagn Res 2017 Jan;11(1):SC01-SC04. [doi: 10.7860/JCDR/2017/23886.9196] [Medline: 28274012]
21. Ha K, Chung S, Lee HS, et al. Association of dietary sugars and sugar-sweetened beverage intake with obesity in Korean children and adolescents. Nutrients 2016 Jan 8;8(1):31. [doi: 10.3390/nu8010031] [Medline: 26761029]
22. Jonnalagadda SS, Mitchell DC, Smiciklas-Wright H, et al. Accuracy of energy intake data estimated by a multiple-pass, 24-hour dietary recall technique. J Am Diet Assoc 2000 Mar;100(3):303-308. [doi: 10.1016/s0002-8223(00)00095-x] [Medline: 10719403]
23. Dougher CE, Rifkin DE, Anderson CA, et al. Spot urine sodium measurements do not accurately estimate dietary sodium intake in chronic kidney disease. Am J Clin Nutr 2016 Aug;104(2):298-305. [doi: 10.3945/ajcn.115.127423] [Medline: 27357090]
24. Miller V, Yusuf S, Chow CK, et al. Availability, affordability, and consumption of fruits and vegetables in 18 countries across income levels: findings from the Prospective Urban Rural Epidemiology (PURE) study. Lancet Glob Health 2016 Oct;4(10):e695-e703. [doi: 10.1016/S2214-109X(16)30186-3] [Medline: 27567348]
25. BMI-for-age (5-19 years). World Health Organization. URL: https://www.who.int/tools/growth-reference-data-for-5to19-years/indicators/bmi-for-age [accessed 2022-02-21]
26. WHO Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet 2004 Jan 10;363(9403):157-163. [doi: 10.1016/S0140-6736(03)15268-3] [Medline: 14726171]
27. Writing Committee Members, Jones DW, Ferdinand KC, et al. 2025 AHA/ACC/AANP/AAPA/ABC/ACCP/ACPM/AGS/AMA/ASPC/NMA/PCNA/SGIM guideline for the prevention, detection, evaluation and management of high blood pressure in adults: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. Hypertension 2025 Oct;82(10):e212-e316. [doi: 10.1161/HYP.0000000000000249] [Medline: 40811516]
28. Krishnan A, Amarchand R, Tarik M, et al. Performance of equations for estimation of 24-hour urinary sodium from morning fasting urine samples in a multisite study in India. Natl Med J India 2022;35(1):4-10. [doi: 10.25259/NMJI_352_19] [Medline: 36039619]
29. Holbrook JT, Patterson KY, Bodner JE, et al. Sodium and potassium intake and balance in adults consuming self-selected diets. Am J Clin Nutr 1984 Oct;40(4):786-793. [doi: 10.1093/ajcn/40.4.786] [Medline: 6486085]
30. Johnson C, Santos JA, Sparks E, et al. Sources of dietary salt in North and South India estimated from 24 hour dietary recall. Nutrients 2019 Feb 1;11(2):318. [doi: 10.3390/nu11020318] [Medline: 30717304]

XSL•FO
RenderX

31. Dong J, Zhao X, Dong H, Yan Y, Mi J. Can spot urine measurement be a substitute for 24-hour urine measurement to estimate sodium intake in adolescents? Pediatr Investig 2021 Mar;5(1):12-20. [doi: 10.1002/ped4.12243] [Medline: 33778422]
32. McLean R, Cameron C, Butcher E, Cook NR, Woodward M, Campbell NR. Comparison of 24‐hour urine and 24‐hour diet recall for estimating dietary sodium intake in populations: a systematic review and meta‐analysis. J Clin Hypertens (Greenwich) 2019 Dec;21(12):1753-1762. [doi: 10.1111/jch.13729] [Medline: 31769168]
33. Peng Y, Zhang Y, Li K, Liu L, Zhang S, Peng X. A new approach is needed to evaluate 24-hour urinary sodium excretion using spot urines: a validation study in a Chinese child population. J Am Heart Assoc 2020 Jan 7;9(1):e014575. [doi: 10.1161/JAHA.119.014575] [Medline: 31865890]
34. Gonçalves C, Abreu S, Padrão P, et al. Sodium and potassium urinary excretion and dietary intake: a cross-sectional analysis in adolescents. Food Nutr Res 2016;60:29442. [doi: 10.3402/fnr.v60.29442] [Medline: 27072344]
35. Kaur S. Comparison of measured 24-hour salt excretion with spot urine and 24-hour dietary 1 recall estimates among adolescents and parents: a cross-sectional study from north 2 India. Figshare. 2025. URL: https://figshare.com/articles/dataset/Comparison of Measured 24hour Salt Excretion with Spot Urine and 24hour Dietary 1 Recall Estimates among Adolescents and Parents A Cross-Sectional Study from North 2 India/28622999/1?file=53098892 [accessed 2026-06-23]

Abbreviations

BP: blood pressure
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 10.Oct.2025; peer-reviewed by A Clermont, A Olayonwa; revised version received 30.Apr.2026; accepted 11.Jun.2026; published 30.Jun.2026.

Please cite as:
Kaur S, Chauhan AS, Kumar R, Kaur M
Comparison of Measured 24-Hour Urinary Salt Excretion With Spot Urine and 24-Hour Dietary Recall Estimates Among Adolescents and Parents: Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e85549
URL: https://publichealth.jmir.org/2026/1/e85549
doi:10.2196/85549

© Sandeep Kaur, Akashdeep Singh Chauhan, Rajesh Kumar, Manmeet Kaur. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 30.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Comparison of In Vitro Metrics With Real-World Risk of Drug-Induced Parkinsonism Due to Antipsychotic Drugs: Retrospective Cohort Study

Woo-Taek Lim[1*]; Hyun Woo Lee[2,3*], PharmD; Seungyeon Kim[4], PhD; Kwangsoo Kim[5], PhD; Yong Min Ahn[6], MD, PhD; Minseok Hong[7], MD; Yun Mi Yu[2,3*], PhD; Ha Young Jang[1], PharmD, PhD

[1]College of Pharmacy, Gachon University, 191, Hambangmoe-ro, Yeonsu-gu, Incheon, Republic of Korea

[2]Department of Pharmaceutical Medicine and Regulatory Sciences, Colleges of Medicine and Pharmacy, Yonsei University, Incheon, Republic of Korea

[3]Department of Pharmacy and Yonsei Institute of Pharmaceutical Sciences, College of Pharmacy, Yonsei University, Incheon, Seoul, Republic of Korea

[4]College of Pharmacy, Dankook University, Cheonan, Republic of Korea

[5]Division of Clinical Bioinformatics, Biomedical Research Institute, Seoul National University Hospital, Seoul, Republic of Korea

[6]Department of Psychiatry, Seoul National University Hospital, Seoul, Republic of Korea

[7]Department of Psychiatry, Uijeongbu Eulji Medical Center, Eulji University, Uijeongbu, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Ha Young Jang, PharmD, PhD
College of Pharmacy, Gachon University, 191, Hambangmoe-ro, Yeonsu-gu, Incheon, Republic of Korea

Abstract

Background: Drug-induced parkinsonism (DIP) predominantly occurs due to antipsychotic drugs (APDs) blocking dopamine D2 receptors (D_2Rs). However, in vitro assays often fail to fully reflect real-world variability in clinical outcomes.

Objective: This study aimed to evaluate whether in vitro pharmacological metrics correspond to real-world risk of DIP associated with APD use.

Methods: For 8 commonly used APDs, key in vitro parameters—including inhibition constants (K_i) of D_2Rs and the serotonin 2A receptor, reversal rate (K_r) of D_2Rs, and blood-brain barrier (BBB) penetration rate—were compiled to construct 6 composite DIP risk metrics. The real-world DIP risk was assessed using the Seoul National University Hospital common data model (2002 - 2021). APD users were matched 1:1 to selective serotonin reuptake inhibitor users using propensity score matching, and Cox proportional hazard regression was performed to estimate the hazard ratios (HRs) for DIP risk. Correlation between each in vitro metric and real-world DIP risk was evaluated using logarithmic regression models.

Results: Among 44,664 patients from 8 matched cohorts, haloperidol showed the highest DIP risk (HR=4.56, 95% CI 2.29 - 9.07), whereas aripiprazole exhibited the lowest risk (HR=2.11, 95% CI 1.56 - 2.86). Metric 4 (pK_r × BBB penetration rate) exhibited the strongest correlation with real-world DIP risk (R^2=0.95). The correlation decreased when aripiprazole, a partial D_2R agonist, was included in the analysis (R^2=0.58).

Conclusions: Integrating receptor-binding kinetics with BBB penetration may provide an in vitro framework that reflects real-world variation in DIP risk among D_2R-antagonizing APDs. These findings support the relevance of combining kinetic and central nervous system exposure parameters for early safety evaluation.

(JMIR Public Health Surveill 2026;12:e81876) doi:10.2196/81876

KEYWORDS

drug-induced parkinsonism; antipsychotic drug; binding affinity; dopamine D2 receptor; blood-brain barrier

Introduction

The pharmacological action of drugs generally occurs through receptor binding, and side effects are often the consequences of these interactions [1]. To understand how drugs function and anticipate their potential side effects, extensive in vitro experiments on drug-receptor interactions are conducted from the early stages of drug development. However, relying on in vitro analysis, often referred to as "test tube experiments," makes it difficult to predict how often side effects will occur in real-world patients [2,3]. This translational gap has highlighted the need for mechanistically informed models that integrate

XSL•FO
RenderX

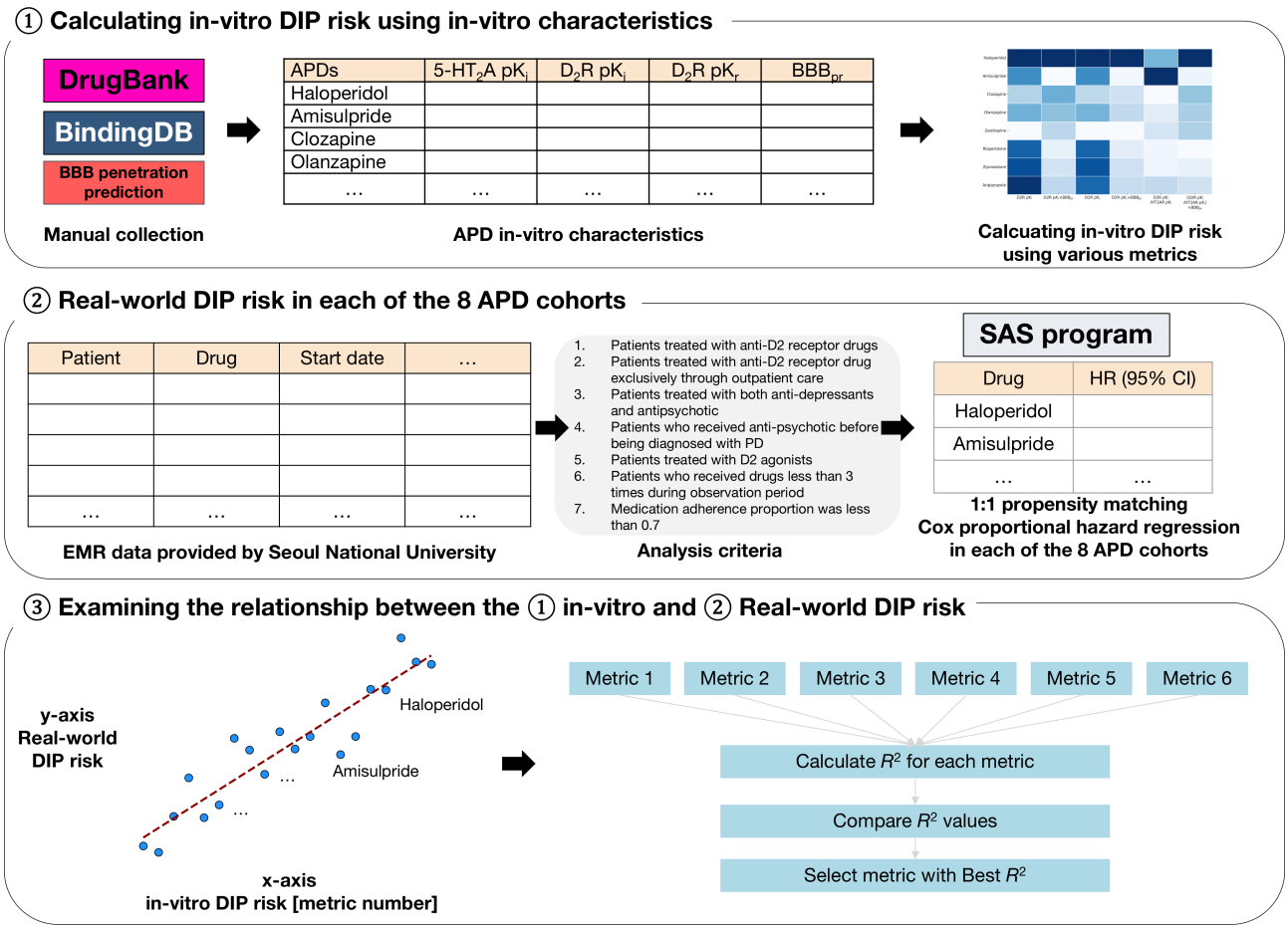
binding kinetics and pharmacokinetic factors to better approximate real-world safety outcomes [3].

Drug-induced parkinsonism (DIP), predominantly caused by antipsychotic drugs (APDs), is one of the most common forms of secondary parkinsonism and remains a clinically significant dose-limiting adverse effect [4]. DIP arises primarily from dopamine D2 receptor (D_2R) blockade within the nigrostriatal pathway, leading to symptoms of movement disorders such as muscle stiffness, slow movements, and tremors [5-10]. Because these symptoms can impair quality of life and daily functioning, understanding the pharmacologic determinants that underlie variability in DIP risk across APDs has become an important clinical and research priority [11,12].

In efforts to explain interdrug differences, the inhibition constant (K_i) and dissociation constant (K_d) have been widely used as a key measure for D_2R blockade; lower values indicate stronger binding to D_2Rs at equivalent drug concentrations [13]. Sykes et al [14] further expanded this framework by demonstrating that receptor-rebinding kinetics—reflecting the probability that a ligand re-engages adjacent receptors after dissociation—can also account for differences in extrapyramidal symptom liability. Other studies have implicated the serotonin 2A receptor ($5\text{-}HT_{2A}R$) in DIP modulation, with newer APDs often exhibiting a higher $5\text{-}HT_{2A}R$–to-D_2R affinity ratio associated with lower DIP risk [15-17]. While binding kinetics contributes to understanding variation in APD-related neurological syndromes, additional pharmacological factors also appear to play important roles. A recent meta-analysis demonstrated that the dose of APDs and D_2R occupancy correlate with extrapyramidal symptom onset, indicating that in vivo exposure should be considered alongside in vitro parameters [18,19]. Furthermore, blood-brain barrier (BBB) permeability, which regulates central nervous system (CNS) drug exposure and can change with certain clinical conditions, may also influence the clinical expression of D_2R blockade [20].

These prior observations collectively suggest that DIP risk reflects an interplay among receptor-binding kinetics, CNS pharmacokinetics, and dose-response. However, despite extensive research on DIP, its precise mechanism and the quantitative translation between in vitro pharmacological parameters and real-world DIP risk remain unclear [6,7,21]. Therefore, this study aimed to evaluate whether specific in vitro pharmacological metrics correspond to variation in real-world DIP risk, with a particular focus on D_2R-antagonizing APDs. First, we derived multiple candidate in vitro DIP risk metrics based on receptor-binding and pharmacokinetic parameters of APDs (Figure 1). Second, using longitudinal real-world data (RWD), we estimated the DIP risk associated with commonly used APDs. Finally, we assessed the extent to which each in vitro metric correlated with the observed real-world DIP risk.

Figure 1. Study overview. $5\text{-}HT_{2A}R$: serotonin 2A receptor; APD: antipsychotic drug; BBB_{pr}: blood-brain barrier penetration rate; D_2R: dopamine D2 receptor; DIP: drug-induced parkinsonism; EMR: electronic medical record; HR: hazard ratio; PD: Parkinson disease.



XSL•FO
RenderX

Methods

Study Design and Overview

This study consisted of two components: (1) an assessment of in vitro pharmacological parameters of APDs related to DIP and (2) a retrospective cohort study evaluating the effect of APDs on the risk of DIP using the common data model (CDM) of Seoul National University Hospital. We subsequently examined the relationship between in vitro metrics and the real-world DIP risk. This study was reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology guidelines [22].

In Vitro DIP Risk Metrics

The key parameters of the APDs used to calculate the in vitro DIP risk were the pK_r values for D_2Rs and 5-HT_{2A}R collected from DrugBank and BindingDB [23-25]. The receptor reversal rate (K_r) values for D_2Rs as suggested by Sykes et al [14] were also obtained. Considering that the primary site of action for APDs is within the CNS, the BBB penetration rate of each APD was considered an adjustment factor [26]. The 6 main metrics for the in vitro DIP risk were as follows:

- Metric 1: pK_i for D_2Rs
- Metric 2: (pK_i for D_2Rs) × BBB penetration rate
- Metric 3: pK_r for D_2Rs
- Metric 4: (pK_r for D_2Rs) × BBB penetration rate
- Metric 5: (pK_i for D_2Rs) / (pK_i for 5-HT_{2A}R)
- Metric 6: (pK_i for D_2Rs) / (pK_i for 5-HT_{2A}R) × BBB penetration rate

All the data were collected through a manual search conducted by the research team. When multiple values were reported, the geometric mean was used as an integrated value.

RWD Source

The CDM of Seoul National University Hospital included longitudinal data on patient demographics, diagnostic information (such as Parkinson disease and other comorbidities), and prescription details (including prescribed drugs, prescription dates, dosages, and duration of use) from 2002 to 2021. The CDM includes standardized fields for these domains, so conventional item-level missingness is minimal; however, care and prescriptions received outside the hospital are not captured and may result in incomplete ascertainment of medication exposure and clinical events, which was considered when interpreting the results.

Study Population

Patients were recruited into the study cohort based on prescription records. The experimental group consisted of patients prescribed APDs, including haloperidol, olanzapine, quetiapine, risperidone, amisulpride, aripiprazole, clozapine, and ziprasidone. The active comparator group, which served as the control, consisted of patients treated with selective serotonin reuptake inhibitors (SSRIs), including citalopram and escitalopram, fluoxetine, paroxetine, and sertraline, which are known to have minimal D_2R binding affinity (K_i of approximately 10,000 nM) [23-25].

The index date was defined as the date of the first outpatient prescription of an APD or an SSRI during the study period. To maintain methodological consistency, only outpatient prescriptions of the oral formulations of the study drugs were included in the analysis. Patients were excluded based on the following criteria: (1) concurrent prescriptions of APDs and SSRIs; (2) diagnosis of Parkinson disease prior to the index date as the study focused on identifying new-onset DIP attributable to drug exposure; (3) use of D_2R agonists within 1 year prior to the index date; (4) fewer than 3 prescriptions of the study drugs after the index date because such limited exposure is unlikely to represent sustained treatment and may not meaningfully influence DIP onset (in our setting, outpatient prescriptions of APDs and SSRIs are usually written for short durations [approximately 2 - 4 weeks], so at least 3 prescriptions typically correspond to 2 to 3 months of continuous therapy and help avoid misclassifying sporadic or trial use as ongoing treatment, consistent with previous register-based studies evaluating psychotropic and antipsychotic medication exposure [27,28]); and (5) poor medication adherence, defined as a proportion of days covered of less than 0.7 as inconsistent medication use could confound assessments of DIP risk [29].

Outcome Assessment

The onset of DIP was defined by the presence of diagnostic codes for DIP (G21.1, G21.2, G21.8, and G21.9) based on the *International Classification of Diseases, 10th Revision* [30]. To enhance diagnostic specificity, patients were considered to have developed DIP only if they received prescriptions for D_2R agonists or anticholinergic agents (standard treatments for DIP) within 60 days of initial diagnosis.

Covariates

A total of 28 covariates were selected based on the presence of comorbidities or concurrent medication history within 1 year of the start date of medication use. The 11 comorbidities included chronic obstructive pulmonary disease, dementia, diabetes mellitus, dyslipidemia, end-stage renal disease, gout, hypertension, liver disease, osteoarthritis, osteoporosis, and stroke. The list of *International Classification of Diseases, 10th Revision* codes for comorbidities is shown in Table S1 in Multimedia Appendix 1. The 17 concurrent medications included renin-angiotensin-aldosterone system inhibitors, such as angiotensin-converting enzyme inhibitors and angiotensin receptor blockers; alpha-glucosidase inhibitors; anticonvulsants; anxiolytics; beta-blockers; calcium channel blockers; dipeptidyl peptidase-4 inhibitors; erythropoiesis-stimulating agents; glucagonlike peptide-1 receptor agonists; insulin; iron; loop diuretics; meglitinides; metformin; sodium-glucose cotransporter 2 inhibitors; statins; and sulfonylureas. A detailed list of concurrent medications is provided in Table S2 in Multimedia Appendix 1.

Statistical Analysis

Statistical analyses were performed for the treated population. Patients were followed up on until the earliest events of DIP onset, the last day the patient took their prescribed medication + 30 days, or the end of the 1-year study period. Each APD user was matched 1:1 to an SSRI user, and the distribution of the

XSL•FO
RenderX

propensity score was inspected [31]. The matching variables included age, sex, comorbidities, and concurrent medications. A standardized difference of >0.1 was regarded as a sign of imbalance [32]. The baseline characteristics were summarized using descriptive statistics. Cox proportional hazard regression was used to estimate the hazard ratio (HR) and 95% CI for the risk of DIP associated with APD use. A dose-response analysis was performed by stratifying patients according to their average daily exposure based on the defined daily dose (DDD) [33]. Patients were categorized into 3 groups (<0.5 the DDD, 0.5 - 1.5 the DDD, and ≥1.5 the DDD) to contextualize the in vitro correlation within a clinical dose-response pattern.

Correlation Analysis

To explore the relationship between the in vitro and real-world DIP risks, we correlated the 6 in vitro metrics with the HR estimated from the real-world cohort. Coefficients of determination (R^2) were calculated to quantify the explanatory strength of each association, and an R^2 value of 0.7 was used as a descriptive threshold for a strong relationship [34]. The primary analysis focused on APDs with D_2R antagonist properties, whereas aripiprazole—a partial agonist—was examined separately in an exploratory analysis to reflect its distinct pharmacologic profile [35].

For each drug i, the association between the clinical outcome—expressed as an HR—and the corresponding in vitro pharmacological metric (metric 1 - 6) was modeled using a logarithmic regression:

$$HRi=\beta 0+\beta 1\ln(Xi)+\varepsilon i$$

In this expression, Xi represents the in vitro metric for the ith drug. As each HR estimate was accompanied by a 95% CI, weighted least squares was used to incorporate the varying uncertainty across drugs, allowing the regression to account for the differing precision of each HR estimate. A 95% confidence band for the fitted regression line was constructed based on the SE of the mean predicted HR values:

$$CI95\%(x)=\hat{HR}(x)\pm t0.975,df\cdot SE\hat{HR}(x)$$

In this equation, $\hat{HR}(x)$ is the predicted HR at a given in vitro metric value x; t0.975,df is the 2-tailed critical value from the Student t distribution at a 95% confidence level, with df representing the df; and $SE\hat{HR}(x)$ is the SE of the estimated mean HR at x, derived from the variance-covariance matrix of the weighted least squares model. The upper and lower confidence limits were visualized as a shaded band around the fitted trend line for each metric. All analyses were performed using SAS (version 9.4; SAS Institute) and Python (version 3.12.12; Python Software Foundation).

Sensitivity Analysis

A sensitivity analysis was conducted in which the outcome definition was modified by removing the requirement for anticholinergic prescriptions within 60 days of initial DIP diagnosis, allowing for the assessment of potential misclassification. In a second analysis, we recalculated the in vitro metrics using pK_i values extracted exclusively from the single-source dataset reported by Sykes et al [14]. Because this recalculation affected only the pK_i-dependent metrics, the additional evaluation was performed to determine whether the observed in vitro–clinical relationships were robust to variability in pK_i data sources.

Ethical Considerations

This study was approved by the institutional review boards of Gachon University Gil Hospital (1044396-202312-HR-230-01) and Seoul National University Hospital (E-2409-042-1569). The requirement for informed consent was waived by both review boards because the study involved a retrospective analysis of fully anonymized data, and no identifiable personal information was accessed. All procedures adhered to applicable local and national regulations regarding the protection of personal information, privacy, and confidentiality. As this study involved only secondary analysis of existing anonymized data, no compensation was provided to participants. All personal information was encrypted to ensure that the individuals could not be identified, and access to the dataset was restricted to authorized investigators in accordance with institutional data governance policies.

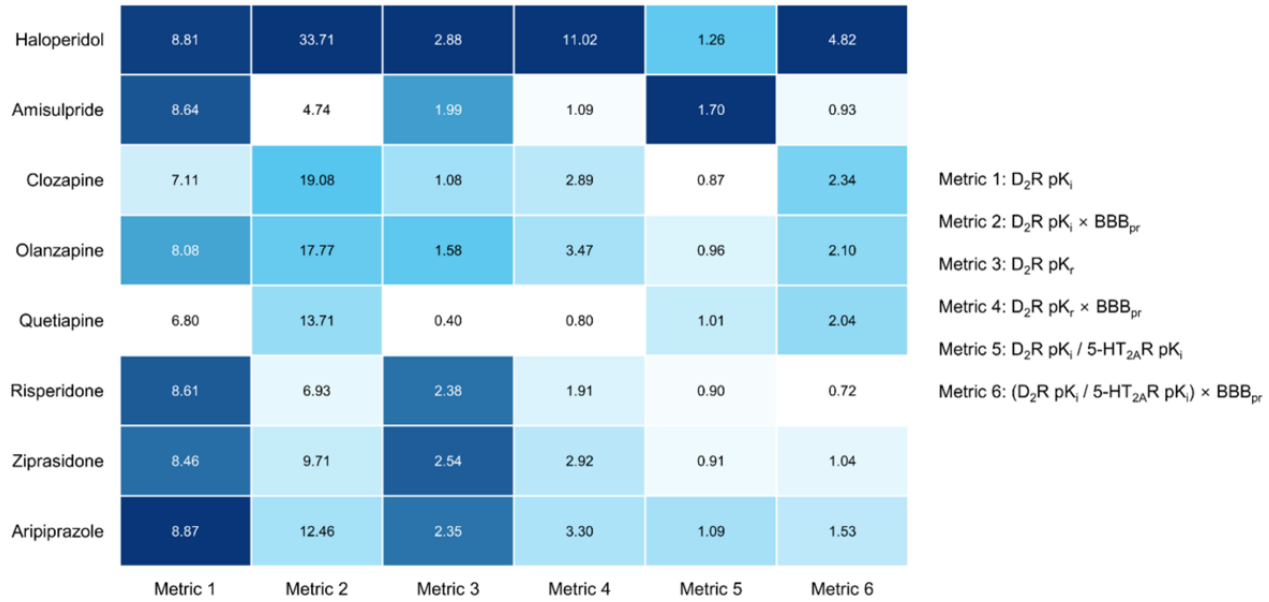
Results

In Vitro DIP Risks

The in vitro pharmacological characteristics of APDs used to derive the 6 in vitro DIP risk metrics are summarized in Table S3 in Multimedia Appendix 1, and corresponding metric values are visualized in Figure 2. Haloperidol showed consistently high values across metrics, ranking first in 4 of the 6 metrics. In contrast, quetiapine exhibited the lowest overall metric values. Amisulpride ranked lower on most measures but held first place in metric 5 (D_2R pK_i/5-$HT_{2A}R$ pK_i = 1.70), reflecting its particularly low affinity for 5-$HT_{2A}R$. Clozapine showed a mixed pattern: although it had relatively low D_2R pK_r, its BBB penetration rate–adjusted values were higher, placing it second in both metric 2 (D_2R pK_i × BBB penetration rate = 19.08) and metric 6 (D_2R pK_i/5-$HT_{2A}R$ pK_i × BBB penetration rate = 2.34). Olanzapine, risperidone, and ziprasidone generally exhibited low to intermediate levels across metrics. Aripiprazole showed high values for several metrics, ranking first in metric 3 (D_2R pK_i=8.87) despite its distinct partial agonist mechanism.

XSL•FO
RenderX

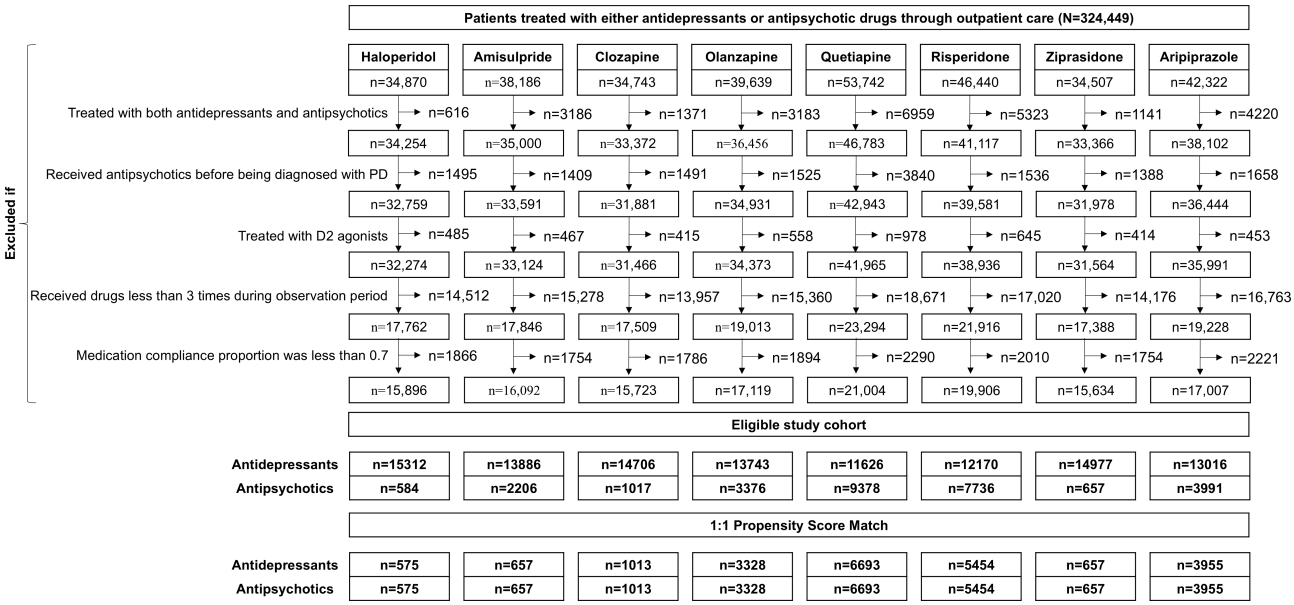
Figure 2. Calculated in vitro risk metrics for drug-induced parkinsonism. The intensity of the blue color represents the relative magnitude of the antipsychotic drug's metric within the same in vitro measurement. 5-HT_{2A}R: serotonin 2A receptor; BBB_{pr}: blood-brain barrier penetration rate; D_2R: dopamine D2 receptor.



Patient Characteristics

A total of 324,449 patients who were exclusively prescribed APDs or SSRIs during outpatient visits were included in the cohort (Figure 3). After eligibility assessment, the final eligible cohort included 109,436 SSRI users and 28,945 APD users. Table 1 presents the baseline characteristics of the haloperidol cohort as a representative example, and those of the remaining cohorts are shown in Tables S4-S10 in Multimedia Appendix 1. Before matching, SSRI users were generally older, whereas APD users showed higher anticonvulsant use. After 1:1 propensity score matching, the final matched cohort included the following numbers for each drug: 575 for haloperidol, 657 for amisulpride, 1013 for clozapine, 3328 for olanzapine, 6693 for quetiapine, 5454 for risperidone, 657 for ziprasidone, and 3955 for aripiprazole. Differences in age and anticonvulsant use, along with other minor discrepancies, were substantially reduced, resulting in well-balanced matched cohorts. Standardized differences for all covariates were below 0.1 across all study cohorts. The median follow-up period was 287 (IQR 142 - 366) days across all participants, with a median of 298 (IQR 159 - 366) days in the SSRI group and 272 (IQR 131 - 364) days in the APD group. The median time to DIP was 71 (IQR 28 - 171) days for the entire dataset, 90 (IQR 39 - 194) days for the SSRI group, and 63 (IQR 23 - 157) days for the APD group.

Figure 3. Flowchart of real-world data analysis. PD: Parkinson disease.



XSL•FO
RenderX

Table . Baseline characteristics of the haloperidol cohort.

Variable	Before matching			After matching		
	SSRI[a] (n=15,312)	Haloperidol (n=584)	Standardized difference	SSRI (n=575)	Haloperidol (n=575)	Standardized difference
Sex (male), n (%)	6789 (44.3)	362 (62.0)	0.3	361 (62.8)	356 (61.9)	−0.01
Age (y), mean (SD)	44.0 (20.7)	39.0 (22.7)	−0.2	41.4 (22.4)	38.6 (22.6)	−0.1
Comorbidities, n (%)						
COPD[b]	76 (0.5)	6 (1.0)	0.06	5 (0.9)	5 (0.9)	0
Dementia	1074 (7.0)	21 (3.6)	−0.1	18 (3.1)	21 (3.7)	0.02
DM[c]	695 (4.5)	38 (6.5)	0.08	33 (5.7)	33 (5.7)	0
Dyslipidemia	725 (4.7)	17 (2.9)	−0.09	21 (3.7)	17 (3.0)	−0.03
ESRD[d]	53 (0.3)	9 (1.5)	0.1	5 (0.9)	8 (1.4)	0.04
Gout	25 (0.2)	2 (0.3)	0.03	1 (0.2)	2 (0.3)	0.03
Hypertension	978 (6.4)	27 (4.6)	−0.07	27 (4.7)	26 (4.5)	−0.008
Liver disease	7 (0.0)	12 (2.1)	0.2	4 (0.7)	3 (0.5)	−0.02
Osteoarthritis	329 (2.1)	6 (1.0)	−0.09	4 (0.7)	6 (1.0)	0.03
Osteoporosis	197 (1.3)	10 (1.7)	0.03	8 (1.4)	9 (1.6)	0.01
Stroke	1412 (9.2)	47 (8.0)	−0.04	50 (8.7)	45 (7.8)	−0.03
Concurrent medications, n (%)						
ACEIs[e] and ARBs[f]	1223 (8.0)	52 (8.9)	0.03	53 (9.2)	50 (8.7)	−0.01
BBs[g]	2545 (16.6)	136 (23.3)	0.1	135 (23.5)	129 (22.4)	−0.02
CCBs[h]	1269 (8.3)	74 (12.7)	0.1	68 (11.8)	67 (11.7)	−0.005
Anticonvulsants	2084 (13.6)	138 (23.6)	0.2	153 (26.6)	138 (24.0)	−0.06
Anxiolytics	8643 (56.4)	286 (49.0)	−0.1	231 (40.2)	279 (48.5)	0.1
ESAs[i]	74 (0.5)	13 (2.2)	0.1	8 (1.4)	12 (2.1)	0.05
Iron	14 (0.1)	2 (0.3)	0.05	0 (0.0)	2 (0.3)	0.08
Loop diuretics	302 (2.0)	58 (9.9)	0.3	41 (7.1)	50 (8.7)	0.05
Other diuretics	743 (4.9)	72 (12.3)	0.2	57 (9.9)	63 (11.0)	0.03
Statins	1403 (9.2)	54 (9.2)	0.003	52 (9.0)	49 (8.5)	−0.01
AGIs[j]	55 (0.4)	2 (0.3)	−0.003	0 (0.0)	2 (0.3)	0.08
DPP4[k] inhibitors	179 (1.2)	26 (4.5)	0.2	20 (3.5)	21 (3.7)	0.009
GLP-1[l] agonists	2 (0.0)	0 (0.0)	−0.01	0 (0.0)	0 (0.0)	0
Insulin	247 (1.6)	41 (7.0)	0.2	31 (5.4)	34 (5.9)	0.02
Meglitinides	28 (0.2)	2 (0.3)	0.03	2 (0.3)	2 (0.3)	0
Metformin	441 (2.9)	34 (5.8)	0.1	29 (5.0)	30 (5.2)	0.007
SGLT2[m] inhibitors	29 (0.2)	3 (0.5)	0.05	2 (0.3)	3 (0.5)	0.02
Sulfonylurea	291 (1.9)	19 (3.3)	0.08	15 (2.6)	16 (2.8)	0.01

[a]SSRI: selective serotonin reuptake inhibitor.
[b]COPD: chronic obstructive pulmonary disease.
[c]DM: diabetes mellitus.
[d]ESRD: end-stage renal disease.
[e]ACEI: angiotensin-converting enzyme inhibitor.
[f]ARB: angiotensin II receptor blocker.

XSL•FO
RenderX

[g]BB: beta-blocker.
[h]CCB: calcium channel blocker.
[i]ESA: erythropoiesis-stimulating agent.
[j]AGI: alpha-glucosidase inhibitor.
[k]DPP4: dipeptidyl peptidase-4.
[l]GLP-1: glucagonlike peptide-1.
[m]SGLT2: sodium-glucose cotransporter-2.

Real-World DIP Risks

The HRs and 95% CIs for DIP across various medications are summarized in Table 2. The typical APD, haloperidol, had the highest DIP risk (HR=4.56, 95% CI 2.29 - 9.07). Among the atypical APDs, clozapine (HR=3.59, 95% CI 2.33 - 5.52), olanzapine (HR=3.53, 95% CI 2.68 - 4.66), risperidone (HR=3.16, 95% CI 2.45 - 4.06), and ziprasidone (HR=3.04, 95% CI 1.68 - 5.50) showed relatively higher risks. Lower HRs were observed with amisulpride (HR=2.36, 95% CI 1.67 - 3.34), quetiapine (HR=2.21, 95% CI 1.76 - 2.78), and aripiprazole (HR=2.11, 95% CI 1.56 - 2.86). Overall, among the medications analyzed, haloperidol exhibited the highest DIP risk, whereas aripiprazole and quetiapine were associated with the lowest risks.

Table . Hazard ratios of drug-induced parkinsonism associated with antipsychotic drugs vs selective serotonin reuptake inhibitors.

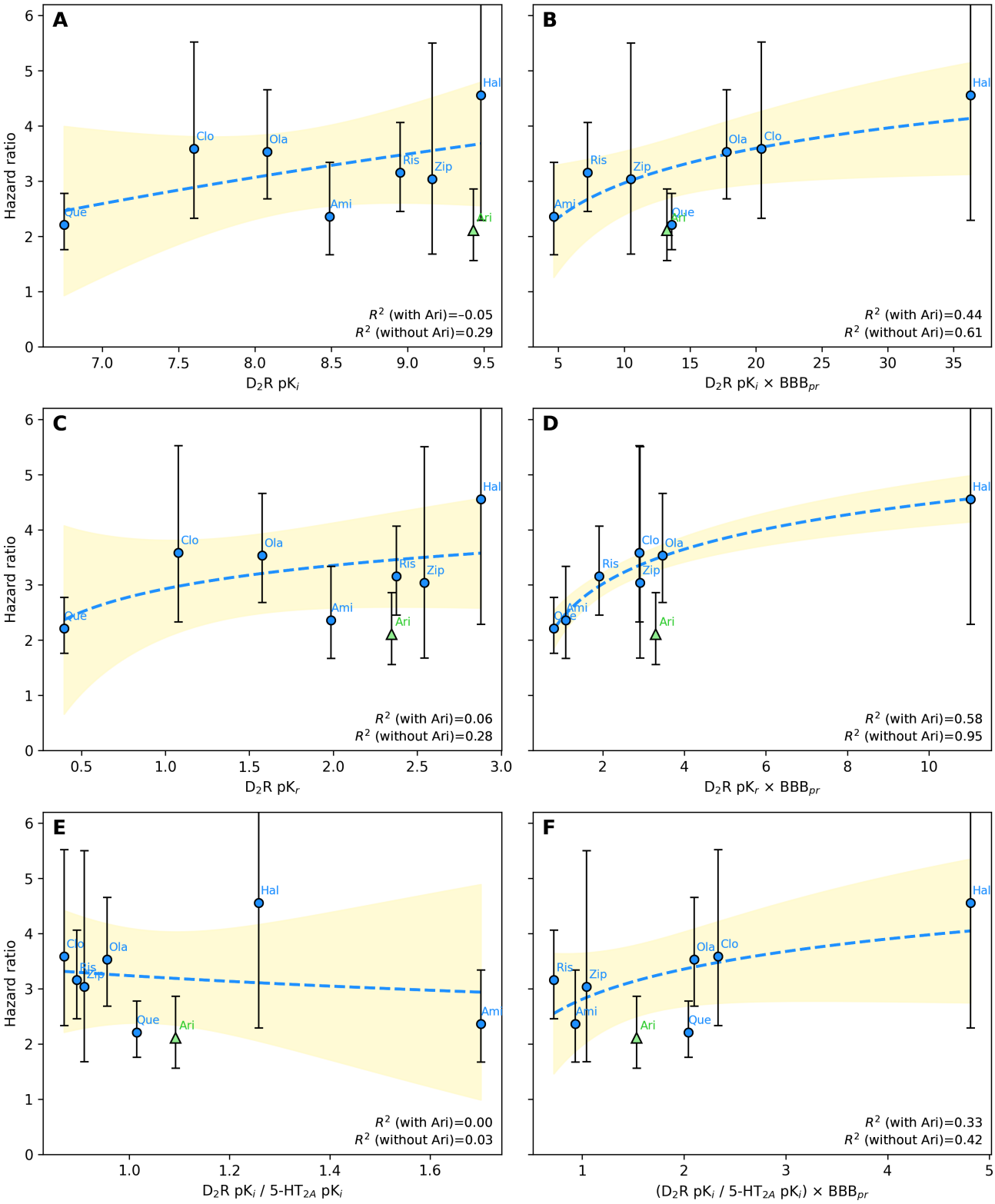
Antipsychotic drug	Hazard ratio (95% CI)
Haloperidol	4.56 (2.29 - 9.07)
Amisulpride	2.36 (1.67 - 3.34)
Clozapine	3.59 (2.33 - 5.52)
Olanzapine	3.53 (2.68 - 4.66)
Quetiapine	2.21 (1.76 - 2.78)
Risperidone	3.16 (2.45 - 4.06)
Ziprasidone	3.04 (1.68 - 5.50)
Aripiprazole	2.11 (1.56 - 2.86)

Correlation Between In Vitro and Real-World DIP Risks

Among the 6 in vitro metrics evaluated, metric 4 (D_2R pK_r × BBB penetration rate) demonstrated the strongest correlation with real-world DIP risk (R^2=0.95), whereas metric 5 (D_2R pK_i/5-$HT_{2A}R$ pK_i) showed the weakest correlation (R^2=0.03; Figure 4). Incorporating BBB penetration markedly improved the explanatory strength across all metric pairs (metric 1 vs metric 2, metric 3 vs metric 4, and metric 5 vs metric 6). Aripiprazole, a partial D_2R agonist, deviated from the pattern observed among D_2R antagonists. When aripiprazole was included in the analysis for metric 4, the correlation coefficient decreased substantially from 0.95 to 0.58, reflecting its fundamentally different pharmacological mechanism. Across all 6 metrics, stratification by DDD suggested a trend toward higher DIP risk with increasing exposure (Figure 5). For metric 4, the stratified curves showed the clearest parallel pattern across exposure groups, although the explanatory strength declined within strata.

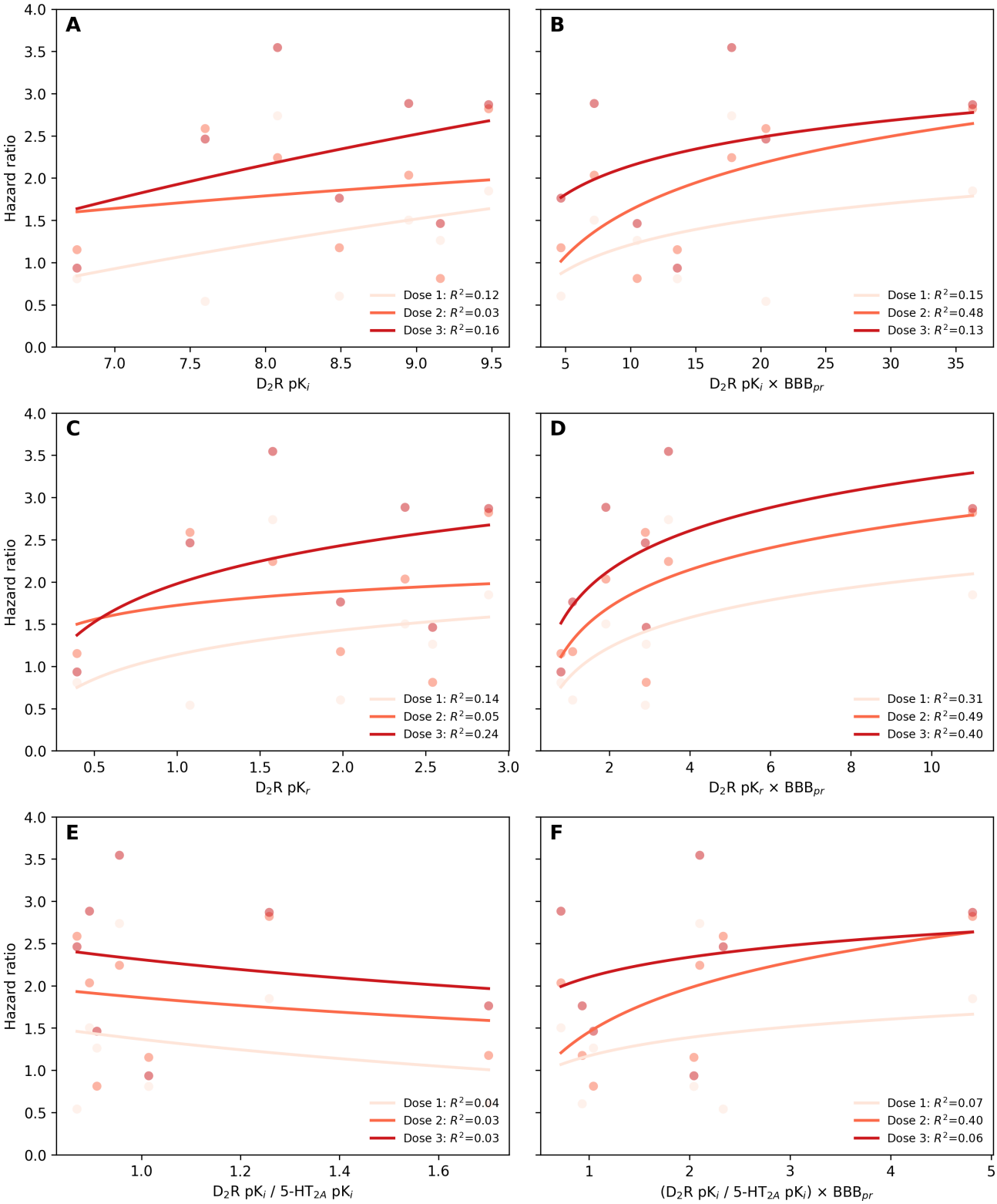
XSL•FO
RenderX

Figure 4. Correlation analysis of in vitro metrics with clinical risk of drug-induced parkinsonism: (A) dopamine D2 receptor (D_2R) pK_i (metric 1); (B) D_2R pK_i × blood-brain barrier penetration rate (BBB_{pr}; metric 2); (C) D_2R pK_r (metric 3); (D) D_2R pK_r × BBB_{pr} (metric 4); (E) D_2R pK_i/serotonin 2A receptor (5-$HT_{2A}R$) pK_i (metric 5); and (F) D_2R pK_i/5-$HT_{2A}R$ pK_i × BBB_{pr} (metric 6). The vertical error bars represent the 95% CIs of the hazard ratios. The blue dashed line and yellow-shaded region indicate the logarithmic regression and 95% confidence band, respectively. The coefficient of determination (R^2) is shown for models calculated with and without aripiprazole (Ari). Ami: amisulpride; Clo: clozapine; Hal: haloperidol; Ola: olanzapine; Ris: risperidone; Zip: ziprasidone.



XSL•FO
RenderX

Figure 5. Dose-response analysis of in vitro metrics with clinical risk of drug-induced parkinsonism: (A) dopamine D2 receptor (D_2R) pK_i (metric 1); (B) D_2R pK_i × blood-brain barrier penetration rate (BBB_{pr}; metric 2); (C) D_2R pK_r (metric 3); (D) D_2R pK_r × BBB_{pr} (metric 4); (E) D_2R pK_i/serotonin 2A receptor (5-$HT_{2A}R$) pK_i (metric 5); and (F) D_2R pK_i/5-$HT_{2A}R$ pK_i × BBB_{pr} (metric 6). Each curve represents the fitted logarithmic regression within a dose stratum. Dose 1 (<0.5 the defined daily dose [DDD]), dose 2 (0.5 - 1.5 the DDD), and dose 3 (≥1.5 the DDD) are shown in light, intermediate, and dark red, respectively. The coefficient of determination (R^2) is shown for models calculated without aripiprazole.



Sensitivity Analysis

In the sensitivity analysis using an outcome definition without anticholinergic confirmation, all overall correlation patterns were preserved, and metric 4 continued to show the strongest correlation, with R^2=0.83 (Figure S1 in Multimedia Appendix 1). When the in vitro metrics were recalculated using pK_i values derived from a single-study source, the overall correlation

XSL•FO
RenderX

pattern remained similar (Table S11 in Multimedia Appendix 1).

Discussion

Principal Findings

In this study, we demonstrated that in vitro–derived DIP risk was closely related to the risk of DIP due to the use of APDs observed in real-world clinical settings. Previous studies have evaluated the inhibitory effects of APDs on D_2Rs or 5-HT_{2A}R through in vitro experiments [14,36-38]. However, such in vitro receptor affinity measures alone often fail to reliably predict the frequency of clinical adverse events such as DIP [24,25]. Conversely, RWD studies have assessed the impact of APDs on the incidence of DIP. However, most of these studies have often been limited by the evaluation of only a small subset of medications or lack of appropriate active comparator groups [30,39-43]. Consequently, numerous randomized controlled trials comparing the risk of DIP among patients taking APDs have been conducted over time, resulting in significant expenditure of time and resources [44-51]. A notable strength of this study is that it provides a comprehensive comparison of DIP risks associated with 8 commonly used APDs within a single institutional cohort using robust matching techniques to minimize confounding and demonstrating the relationship between in vitro pharmacological metrics and observed clinical outcomes. These findings suggest that integrating in vitro pharmacological data with clinical evidence may help generate hypotheses for future safety evaluation frameworks, particularly in early exploratory stages of drug development.

Consistent with previous literature, APDs significantly increased the risk of DIP compared to SSRIs, which served as active comparators [14,40,52]. Haloperidol, a typical APD, exhibited the highest DIP risks in both the in vitro metrics and the real-world analysis, aligning with its strong D_2R antagonist properties. In contrast, atypical APDs showed lower HRs, with aripiprazole exhibiting the lowest DIP risk despite its relatively high receptor affinity and kinetic parameters (pK_i, pK_r, and pK_r × BBB penetration rate). Sensitivity analyses supported the robustness of the overall associations: using an outcome definition without anticholinergic confirmation yielded correlation patterns that were directionally similar, and recalculating the in vitro metrics with pK_i values derived from a single-study source produced a correlation structure comparable to that of the main analysis. Although the DDD-based stratification was exploratory, the generally increasing DIP risk across exposure levels offers qualitative support for a dose-response pattern consistent with prior pharmacological understanding of D_2R blockade [18,40]. This pattern was most apparent for metric 4, which showed a broadly similar ranking across dose strata. However, the reduced explanatory strength within strata suggests that dose-based subgrouping introduces analytic instability, likely reflecting loss of covariate balance and reduced variability after stratification. These observations indicate that, while dose may influence the in vitro–clinical relationship, larger datasets with broader dosing distributions and more detailed exposure metrics, such as treatment duration and receptor occupancy, will be required to more robustly characterize dose-dependent effects.

Aripiprazole deviated from the general correlation pattern, presenting a lower HR than that expected from metric 4 (D_2R pK_r × BBB penetration rate), with similar discrepancies across other metrics. This divergence is most likely attributable to the unique pharmacodynamic profile of aripiprazole as a partial D_2R agonist [35]. Unlike full antagonists, aripiprazole displays intrinsic activity at the D_2Rs, enabling it to maintain a degree of dopaminergic signaling even at high receptor occupancy. A number of positron emission tomography studies have demonstrated that therapeutic doses of aripiprazole achieve high D_2R occupancy without producing corresponding extrapyramidal adverse effects, indicating that occupancy does not translate directly into functional blockade for partial agonists [53-55]. Additionally, aripiprazole exhibits functional selectivity and stabilizes D_2R signaling rather than fully suppressing it, which reduces the likelihood of inducing movement disorders despite strong binding [35]. Such pharmacological nuances highlight that future extensions of this framework may require incorporating measures of functional efficacy, such as intrinsic activity, partial agonist efficacy, or occupancy-response weighting, to more accurately model APDs whose clinical effects diverge from antagonist-based predictions.

Another key finding of this study was that K_r better reflects the DIP risks observed in clinical settings than K_i. This observation aligns with that made in a previous study by Sykes et al [14], which highlighted the importance of rebinding kinetics in DIP. Our study extends these findings by integrating BBB penetration rate, which improved the strength of the correlation across all metrics. Among the models tested, the combination of D_2R pK_r and BBB penetration rate (metric 4) showed the strongest correlation with clinical DIP risk. Notably, the correlation coefficient increased from 0.28 (pK_r alone) to 0.95 (pK_r × BBB penetration rate), indicating that accounting for CNS exposure meaningfully shifted the strength of the association. Conversely, the metrics that incorporated D_2R K_i did not sufficiently reflect the variations in real-world DIP risk. Taken together, these findings suggest that CNS accessibility and rebinding kinetics capture important mechanistic features underlying observed differences among APDs.

Several studies have illustrated the relevance of incorporating BBB penetration rate into risk models. Amisulpride and risperidone, both of which have BBB penetration rate values below 1 (0.55 and 0.81, respectively), appeared below the trend line in unadjusted models, indicating a lower-than-expected DIP risk despite high in vitro receptor affinity (Figures 4A and 4C). This discrepancy is likely due to the limited BBB penetration. When BBB penetration rate–adjusted metrics were applied (Figures 4B and 4D), both drugs aligned closer to the trend line. A similar trend was observed for olanzapine and clozapine, which showed comparable HRs in the clinical data. Although olanzapine exhibited a 46% higher pK_r value than clozapine, this difference was reduced to 20% after adjusting for BBB penetration rate, suggesting that the superior BBB penetration of clozapine may compensate for its lower binding

XSL•FO
RenderX

affinity, highlighting the importance of accounting for BBB permeability in DIP risk prediction.

Limitations

Our study had several limitations. First, the in vitro indicators presented herein are aggregate estimates that are not intended for individual patient prediction. Future research could develop individualized risk prediction models that incorporate these indicators. Second, because most patients in the cohort were treated with relatively low APD doses (≤1.0 the DDD), dose-response characterization was inherently limited, and the relationship between the in vitro metrics and clinical risk across the full dosing spectrum remains unvalidated. Future studies encompassing a wider dosing distribution will be needed to more fully delineate dose-dependent patterns. Third, while the metrics demonstrated correlation with real-world DIP risk among D_2R antagonists, the analysis could not cover all APDs owing to insufficient prescribing frequency, and the current framework does not incorporate intrinsic efficacy, limiting its applicability to partial agonists such as aripiprazole. Therefore, expansion to additional APDs and methodological extensions that account for intrinsic efficacy will be essential for assessing generalizability. Finally, this study used a single-institution CDM database; prescribing patterns and clinical characteristics outside this setting were not captured, and residual confounding inherent to the observational nature may remain despite rigorous propensity score matching. In addition, the identification of DIP onset, even with an enhanced operational definition, may still allow for potential misclassification. Multicenter studies will be valuable to further evaluate the generalizability of the proposed framework. Given these constraints, the findings should be interpreted as exploratory associations rather than definitive causal inferences.

Conclusions

In conclusion, this study demonstrated that combining receptor-binding kinetics with BBB penetration provides robust in vitro metrics that strongly correlate with the real-world clinical risk of DIP. These findings underscore the importance of integrating receptor kinetics and neuropharmacokinetics with real-world evidence, offering a conceptual foundation that may support more mechanistically informed approaches to future pharmacovigilance and adverse event predictions. Ultimately, such integrative methodologies may help refine early safety assessment and improve decision-making across the drug development continuum, potentially reducing the cost and duration of clinical trials.

Funding

This study was supported by a National Research Foundation of Korea grant funded by the Korean government (RS-2023-00280552).

Data Availability

The data generated or analyzed during this study are not publicly available due to institutional restrictions but are available from the corresponding author upon reasonable request and with permission from Seoul National University Hospital.

Authors' Contributions

HYJ and YMY conceptualized, supervised the study, and served as corresponding authors. HYJ, YMY, HWL, and WTL contributed to data acquisition, analysis, and interpretation. All authors participated in drafting and critically revising the manuscript. All authors have approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Tables S1-S11 and Figure S1.
[DOC File, 671 KB - publichealth_v12i1e81876_app1.DOC]

Checklist 1
STROBE checklist.
[PDF File, 217 KB - publichealth_v12i1e81876_app2.pdf]

References

1. Lounkine E, Keiser MJ, Whitebread S, et al. Large-scale prediction and testing of drug activity on side-effect targets. Nature New Biol 2012 Jun 10;486(7403):361-367. [doi: 10.1038/nature11159] [Medline: 22722194]
2. Scannell JW, Blanckley A, Boldon H, Warrington B. Diagnosing the decline in pharmaceutical R&D efficiency. Nat Rev Drug Discov 2012 Mar 1;11(3):191-200. [doi: 10.1038/nrd3681] [Medline: 22378269]
3. Rock BM, Foti RS. Pharmacokinetic and drug metabolism properties of novel therapeutic modalities. Drug Metab Dispos 2019 Oct;47(10):1097-1099. [doi: 10.1124/dmd.119.088708] [Medline: 31399505]

XSL•FO
RenderX

4. Shin HW, Chung SJ. Drug-induced parkinsonism. J Clin Neurol 2012 Mar;8(1):15-21. [doi: 10.3988/jcn.2012.8.1.15] [Medline: 22523509]
5. Ginovart N, Kapur S. Role of dopamine D(2) receptors for antipsychotic activity. Handb Exp Pharmacol 2012;2012(212):27-52. [doi: 10.1007/978-3-642-25761-2_2] [Medline: 23129327]
6. Miyamoto S, Miyake N, Jarskog LF, Fleischhacker WW, Lieberman JA. Pharmacological treatment of schizophrenia: a critical review of the pharmacology and clinical effects of current and future therapeutic agents. Mol Psychiatry 2012 Dec;17(12):1206-1227. [doi: 10.1038/mp.2012.47] [Medline: 22584864]
7. Meltzer HY. Update on typical and atypical antipsychotic drugs. Annu Rev Med 2013;64:393-406. [doi: 10.1146/annurev-med-050911-161504] [Medline: 23020880]
8. Leucht S, Corves C, Arbter D, Engel RR, Li C, Davis JM. Second-generation versus first-generation antipsychotic drugs for schizophrenia: a meta-analysis. Lancet 2009 Jan 3;373(9657):31-41. [doi: 10.1016/S0140-6736(08)61764-X] [Medline: 19058842]
9. Farde L, Nordström AL, Wiesel FA, Pauli S, Halldin C, Sedvall G. Positron emission tomographic analysis of central D1 and D2 dopamine receptor occupancy in patients treated with classical neuroleptics and clozapine. Relation to extrapyramidal side effects. Arch Gen Psychiatry 1992 Jul;49(7):538-544. [doi: 10.1001/archpsyc.1992.01820070032005] [Medline: 1352677]
10. Caron MG, Beaulieu M, Raymond V, et al. Dopaminergic receptors in the anterior pituitary gland. Correlation of [3H]dihydroergocryptine binding with the dopaminergic control of prolactin release. J Biol Chem 1978 Apr 10;253(7):2244-2253. [doi: 10.1016/S0021-9258(17)38065-1] [Medline: 416027]
11. Weng J, Zhang L, Yu W, et al. Risk factors, clinical correlates, and social functions of Chinese schizophrenia patients with drug-induced parkinsonism: a cross-sectional analysis of a multicenter, observational, real-world, prospective cohort study. Front Pharmacol 2023;14:1077607. [doi: 10.3389/fphar.2023.1077607] [Medline: 36937864]
12. Feldman M, Marmol S, Margolesky J. Updated perspectives on the management of drug-induced parkinsonism (dip): insights from the clinic. Ther Clin Risk Manag 2022;18:1129-1142. [doi: 10.2147/TCRM.S360268] [Medline: 36573102]
13. Seeman P, Tallerico T. Antipsychotic drugs which elicit little or no parkinsonism bind more loosely than dopamine to brain D2 receptors, yet occupy high levels of these receptors. Mol Psychiatry 1998 Mar;3(2):123-134. [doi: 10.1038/sj.mp.4000336] [Medline: 9577836]
14. Sykes DA, Moore H, Stott L, et al. Extrapyramidal side effects of antipsychotics are linked to their association kinetics at dopamine D2 receptors. Nat Commun 2017 Oct 2;8(1):763. [doi: 10.1038/s41467-017-00716-z] [Medline: 28970469]
15. Janssen PA, Niemegeers CJ, Awouters F, Schellekens KH, Megens AA, Meert TF. Pharmacology of risperidone (R 64 766), a new antipsychotic with serotonin-S2 and dopamine-D2 antagonistic properties. J Pharmacol Exp Ther 1988 Feb;244(2):685-693. [Medline: 2450200]
16. Meltzer HY, Matsubara S, Lee JC. Classification of typical and atypical antipsychotic drugs on the basis of dopamine D-1, D-2 and serotonin2 pKi values. J Pharmacol Exp Ther 1989 Oct;251(1):238-246. [Medline: 2571717]
17. Roth BL, Sheffler DJ, Kroeze WK. Magic shotguns versus magic bullets: selectively non-selective drugs for mood disorders and schizophrenia. Nat Rev Drug Discov 2004 Apr;3(4):353-359. [doi: 10.1038/nrd1346] [Medline: 15060530]
18. Siafis S, Wu H, Wang D, et al. Antipsychotic dose, dopamine D2 receptor occupancy and extrapyramidal side-effects: a systematic review and dose-response meta-analysis. Mol Psychiatry 2023 Aug;28(8):3267-3277. [doi: 10.1038/s41380-023-02203-y] [Medline: 37537284]
19. Kapur S, Zipursky R, Jones C, Remington G, Houle S. Relationship between dopamine D(2) occupancy, clinical response, and side effects: a double-blind PET study of first-episode schizophrenia. Am J Psychiatry 2000 Apr;157(4):514-520. [doi: 10.1176/appi.ajp.157.4.514] [Medline: 10739409]
20. Luptáková D, Vallianatou T, Nilsson A, et al. Neuropharmacokinetic visualization of regional and subregional unbound antipsychotic drug transport across the blood-brain barrier. Mol Psychiatry 2021 Dec;26(12):7732-7745. [doi: 10.1038/s41380-021-01267-y] [Medline: 34480089]
21. Perkins DO. Predictors of noncompliance in patients with schizophrenia. J Clin Psychiatry 2002 Dec;63(12):1121-1128. [doi: 10.4088/jcp.v63n1206] [Medline: 12523871]
22. Elm EV, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ 2007 Oct 20;335(7624):806-808. [doi: 10.1136/bmj.39335.541782.AD]
23. Knox C, Wilson M, Klinger CM, et al. DrugBank 6.0: the DrugBank knowledgebase for 2024. Nucleic Acids Res 2024 Jan 5;52(D1):D1265-D1275. [doi: 10.1093/nar/gkad976] [Medline: 37953279]
24. Gilson MK, Liu T, Baitaluk M, Nicola G, Hwang L, Chong J. BindingDB in 2015: a public database for medicinal chemistry, computational chemistry and systems pharmacology. Nucleic Acids Res 2016 Jan 4;44(D1):D1045-D1053. [doi: 10.1093/nar/gkv1072] [Medline: 26481362]
25. Liu T, Lin Y, Wen X, Jorissen RN, Gilson MK. BindingDB: a web-accessible database of experimentally determined protein-ligand binding affinities. Nucleic Acids Res 2007 Jan;35(Database issue):D198-D201. [doi: 10.1093/nar/gkl999] [Medline: 17145705]

XSL•FO
RenderX

26. Shaker B, Lee J, Lee Y, et al. A machine learning-based quantitative model (LogBB_Pred) to predict the blood-brain barrier permeability (logBB value) of drug compounds. Bioinformatics 2023 Oct 3;39(10):btad577. [doi: 10.1093/bioinformatics/btad577] [Medline: 37713469]
27. Jennum P, Baandrup L, Iversen HK, Ibsen R, Kjellberg J. Mortality and use of psychotropic medication in patients with stroke: a population-wide, register-based study. BMJ Open 2016 Mar 8;6(3):e010662. [doi: 10.1136/bmjopen-2015-010662] [Medline: 26956165]
28. Prior A, Laursen TM, Larsen KK, et al. Post-stroke mortality, stroke severity, and preadmission antipsychotic medicine use--a population-based cohort study. PLoS ONE 2014;9(1):e84103. [doi: 10.1371/journal.pone.0084103] [Medline: 24416196]
29. Choudhry NK, Shrank WH, Levin RL, et al. Measuring concurrent adherence to multiple related medications. Am J Manag Care 2009 Jul;15(7):457-464. [Medline: 19589013]
30. Jeong S, Cho H, Kim YJ, Ma HI, Jang S. Drug-induced parkinsonism: a strong predictor of idiopathic Parkinson's disease. PLoS ONE 2021;16(3):e0247354. [doi: 10.1371/journal.pone.0247354] [Medline: 33647030]
31. Parsons LS, editor. Reducing bias in a propensity score matched-pair sample using greedy matching techniques. Presented at: Proceedings of the 26th Annual SAS Users Group International Conference; Apr 22-25, 2001 URL: https://support.sas.com/resources/papers/proceedings/proceedings/sugi26/p214-26.pdf [accessed 2026-01-01]
32. Data matching - optimal and greedy. NCSS Statistical Software. URL: https://www.ncss.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Data_Matching-Optimal_and_Greedy.pdf
33. ATC/DDD index 2025. Norwegian Institute of Public Health. 2025. URL: https://atcddd.fhi.no/atc_ddd_index [accessed 2025-01-13]
34. Moore DS. The Basic Practice of Statistics, 5th edition: W.H. Freeman and Company; 2010. URL: https://ndl.ethernet.edu.et/bitstream/123456789/39040/1/David%20S.%20Moore.pdf [accessed 2026-01-21]
35. Shapiro DA, Renock S, Arrington E, et al. Aripiprazole, a novel atypical antipsychotic drug with a unique and robust pharmacology. Neuropsychopharmacology 2003 Aug;28(8):1400-1411. [doi: 10.1038/sj.npp.1300203] [Medline: 12784105]
36. Martinot JL, Paillère-Martinot ML, Poirier MF, Dao-Castellana MH, Loc'h C, Mazière B. In vivo characteristics of dopamine D2 receptor occupancy by amisulpride in schizophrenia. Psychopharmacology (Berl) 1996 Mar;124(1-2):154-158. [doi: 10.1007/BF02245616] [Medline: 8935811]
37. Kapur S, Seeman P. Does fast dissociation from the dopamine d(2) receptor explain the action of atypical antipsychotics?: a new hypothesis. Am J Psychiatry 2001 Mar;158(3):360-369. [doi: 10.1176/appi.ajp.158.3.360] [Medline: 11229973]
38. McCormick PN, Kapur S, Graff-Guerrero A, Raymond R, Nobrega JN, Wilson AA. The antipsychotics olanzapine, risperidone, clozapine, and haloperidol are D2-selective ex vivo but not in vitro. Neuropsychopharmacology 2010 Jul;35(8):1826-1835. [doi: 10.1038/npp.2010.50] [Medline: 20410873]
39. d'Errico A, Strippoli E, Vasta R, Ferrante G, Spila Alegiani S, Ricceri F. Use of antipsychotics and long-term risk of parkinsonism. Neurol Sci 2022 Apr;43(4):2545-2553. [doi: 10.1007/s10072-021-05650-z] [Medline: 34652577]
40. Marras C, Herrmann N, Anderson GM, Fischer HD, Wang X, Rochon PA. Atypical antipsychotic use and parkinsonism in dementia: effects of drug, dose, and sex. Am J Geriatr Pharmacother 2012 Dec;10(6):381-389. [doi: 10.1016/j.amjopharm.2012.11.001] [Medline: 23217531]
41. Sharma A, Sorrell JH. Aripiprazole-induced parkinsonism. Int Clin Psychopharmacol 2006 Mar;21(2):127-129. [doi: 10.1097/01.yic.0000185025.40656.30] [Medline: 16421466]
42. Tarsy D, Baldessarini RJ, Tarazi FI. Effects of newer antipsychotics on extrapyramidal function. CNS Drugs 2002;16(1):23-45. [doi: 10.2165/00023210-200216010-00003] [Medline: 11772117]
43. Wickremaratchi M, Morris HR, Ali IM. Aripiprazole associated with severe exacerbation of Parkinson's disease. Mov Disord 2006 Sep;21(9):1538-1539. [doi: 10.1002/mds.21025] [Medline: 16817207]
44. Crespo-Facorro B, Pérez-Iglesias R, Mata I, et al. Effectiveness of haloperidol, risperidone and olanzapine in the treatment of first-episode non-affective psychosis: results of a randomized, flexible-dose, open-label 1-year follow-up comparison. J Psychopharmacol 2011 Jun;25(6):744-754. [doi: 10.1177/0269881110388332] [Medline: 21292922]
45. Cuesta MJ, Sánchez-Torres AM, de Jalón EG, et al. Spontaneous parkinsonism is associated with cognitive impairment in antipsychotic-naive patients with first-episode psychosis: a 6-month follow-up study. Schizophr Bull 2014 Sep;40(5):1164-1173. [doi: 10.1093/schbul/sbt125] [Medline: 24072809]
46. Kahn RS, Fleischhacker WW, Boter H, et al. Effectiveness of antipsychotic drugs in first-episode schizophrenia and schizophreniform disorder: an open randomised clinical trial. Lancet 2008 Mar 29;371(9618):1085-1097. [doi: 10.1016/S0140-6736(08)60486-9] [Medline: 18374841]
47. Lieberman JA, Phillips M, Gu H, et al. Atypical and conventional antipsychotic drugs in treatment-naive first-episode schizophrenia: a 52-week randomized trial of clozapine vs chlorpromazine. Neuropsychopharmacology 2003 May;28(5):995-1003. [doi: 10.1038/sj.npp.1300157] [Medline: 12700715]
48. Möller HJ, Riedel M, Jäger M, et al. Short-term treatment with risperidone or haloperidol in first-episode schizophrenia: 8-week results of a randomized controlled trial within the German Research Network on Schizophrenia. Int J Neuropsychopharmacol 2008 Nov;11(7):985-997. [doi: 10.1017/S1461145708008791] [Medline: 18466670]

XSL•FO
RenderX

49. Robinson DG, Woerner MG, Napolitano B, et al. Randomized comparison of olanzapine versus risperidone for the treatment of first-episode schizophrenia: 4-month outcomes. Am J Psychiatry 2006 Dec;163(12):2096-2102. [doi: 10.1176/ajp.2006.163.12.2096] [Medline: 17151160]
50. Sanger TM, Lieberman JA, Tohen M, Grundy S, Beasley C, Tollefson GD. Olanzapine versus haloperidol treatment in first-episode psychosis. Am J Psychiatry 1999 Jan;156(1):79-87. [doi: 10.1176/ajp.156.1.79] [Medline: 9892301]
51. Schooler N, Rabinowitz J, Davidson M, et al. Risperidone and haloperidol in first-episode psychosis: a long-term randomized trial. Am J Psychiatry 2005 May;162(5):947-953. [doi: 10.1176/appi.ajp.162.5.947] [Medline: 15863797]
52. Kim S, Suh HS. Treatment changes and prognoses in patients with incident drug-induced parkinsonism using a Korean nationwide healthcare claims database. J Clin Med 2023;12(8):2860. [doi: 10.3390/jcm12082860]
53. Kegeles LS, Slifstein M, Frankle WG, et al. Dose-occupancy study of striatal and extrastriatal dopamine D2 receptors by aripiprazole in schizophrenia with PET and [18F]fallypride. Neuropsychopharmacology 2008 Dec;33(13):3111-3125. [doi: 10.1038/npp.2008.33] [Medline: 18418366]
54. Mamo D, Graff A, Mizrahi R, Shammi CM, Romeyer F, Kapur S. Differential effects of aripiprazole on D(2), 5-HT(2), and 5-HT(1A) receptor occupancy in patients with schizophrenia: a triple tracer PET study. Am J Psychiatry 2007 Sep;164(9):1411-1417. [doi: 10.1176/appi.ajp.2007.06091479] [Medline: 17728427]
55. Yokoi F, Gründer G, Biziere K, et al. Dopamine D2 and D3 receptor occupancy in normal humans treated with the antipsychotic drug aripiprazole (OPC 14597): a study using positron emission tomography and [11C]raclopride. Neuropsychopharmacology 2002 Aug;27(2):248-259. [doi: 10.1016/S0893-133X(02)00304-4] [Medline: 12093598]

Abbreviations

5-HT$_{2A}$R: serotonin 2A receptor
APD: antipsychotic drug
BBB: blood-brain barrier
CDM: common data model
CNS: central nervous system
D$_2$R: dopamine D2 receptor
DDD: defined daily dose
DIP: drug-induced parkinsonism
HR: hazard ratio
RWD: real-world data
SSRI: selective serotonin reuptake inhibitor
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Mavragani, T Sanchez; submitted 05.Aug.2025; peer-reviewed by AS Mutluer, M Ali; revised version received 09.Dec.2025; accepted 30.Dec.2025; published 28.Jan.2026.

Please cite as:
Lim WT, Lee HW, Kim S, Kim K, Ahn YM, Hong M, Yu YM, Jang HY
Comparison of In Vitro Metrics With Real-World Risk of Drug-Induced Parkinsonism Due to Antipsychotic Drugs: Retrospective Cohort Study
JMIR Public Health Surveill 2026;12:e81876
URL: https://publichealth.jmir.org/2026/1/e81876
doi:10.2196/81876

© Woo-Taek Lim, Hyun Woo Lee, Seungyeon Kim, Kwangsoo Kim, Yong Min Ahn, Minseok Hong, Yun Mi Yu, Ha Young Jang. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 28.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Texting to Recruit Unassisted Tobacco Users at a Large Safety Net Health System Into Quitline Service for a Medicaid Value-Based Care Program: Pragmatic Randomized Controlled Trial

Cindy V Valencia[1], MPH, PhD; Melanie S Dove[2], MPH, ScD; Karen Kim[3], MD; Paul Giboney[3], MD; Hal F Yee Jr[3], MD, PhD; Carrie A Kirby[4], MS; Christopher M Anderson[4]; Shu-Hong Zhu[4], PhD; Elisa K Tong[5], MA, MD

[1]Center for Healthcare Policy Research, University of California, Davis, 4150 V Street, Suite 2400, Sacramento, CA, United States

[2]Department of Public Health Sciences, University of California, Davis, Sacramento, CA, United States

[3]Los Angeles County Department of Health Services, Los Angeles, CA, United States

[4]Moores Cancer Center, University of California, San Diego, La Jolla, CA, United States

[5]Department of Internal Medicine and Center for Healthcare Policy Research, University of California, Davis, 4150 V Street, Suite 2400, Sacramento, CA, United States

Corresponding Author:
Elisa K Tong, MA, MD
Department of Internal Medicine and Center for Healthcare Policy Research, University of California, Davis, 4150 V Street, Suite 2400, Sacramento, CA, United States

Abstract

Background: A growing body of research supports the efficacy of text messaging programs to help tobacco users quit, but texting as a strategy for recruiting tobacco users into other evidence-based cessation services, such as quitline coaching, is less well understood. Texting to offer treatment could increase use of cessation resources, an important consideration for health systems trying to improve their quality metric performance on tobacco assessment and counseling.

Objective: The aim of this study is to compare the effects of text messages offering free quitline coaching or free nicotine patches on engagement with quitline services by patients identified in electronic health records as unassisted tobacco users.

Methods: Participants (N=4171) were adult patients of Los Angeles County Department of Health Services who had had a clinical visit in the past 12 months, were willing to receive text messages, and were identified as unassisted tobacco users (ie, those with no documented advice or assistance to quit in the past 24 months). They included 3139 English speakers and 1032 Spanish speakers. Participants were stratified by language, then randomly assigned to one of two groups. Group 1 received a text offering free quitline coaching. Group 2 received a text offering free nicotine patches. The texts were sent in April and May 2022. Outcome measures were the proportions calling the quitline and receiving evidence-based cessation treatments. Data were analyzed in 2025.

Results: Overall, 1.5% (61/4171) of participants called the quitline, including 1.3% (28/2086) in Group 1 and 1.6% (33/2085) in Group 2, and 0.5% (21/4171) received treatment, including 0.4% (8/2086) in Group 1 and 0.6% (13/2085) in Group 2. There was no significant difference either in calls (P=.52) or receiving treatment (P=.29). However, Spanish speakers were significantly more likely to call than English speakers, 2.3% (24/1032) vs 1.2% (37/3139), respectively (P=.008), and engaged in treatment at approximately twice the rate of English speakers, 0.9% (9/1032) vs 0.4% (12/3139), although the latter difference was nonsignificant (P=.07).

Conclusions: A single text was effective in connecting unassisted tobacco users with evidence-based quitline services for both English and Spanish speakers, but especially the latter. A population health approach of reaching out to tobacco-using patients outside clinical visits can meaningfully supplement provider referral. Research is needed on ways to increase the population health impact of this strategy.

Trial Registration: ClinicalTrials.gov NCT07102290; https://www.clinicaltrials.gov/study/NCT07102290

(JMIR Public Health Surveill 2026;12:e83269) doi:10.2196/83269

KEYWORDS

quitline; tobacco cessation; Spanish; safety net; clinic; text

XSL•FO
RenderX

Introduction

The US Preventive Services Task Force urges health systems to support clinicians in treating tobacco use and calls on them to promote quitline utilization [1]. Quitlines are widely available and provide evidence-based cessation counseling [2]. Most partner with health systems to encourage clinicians to ask their patients if they use tobacco, advise those who do to quit, and connect them to quitlines via electronic referral (e-referral) [3,4]. Some quitlines provide additional services, including nicotine replacement therapy (NRT) and mobile health (mHealth) interventions [5].

mHealth interventions such as mobile apps and text messaging programs have received increasing attention due to their potential to expand the reach of tobacco cessation services. The approach is promising, given how widespread cell phone ownership is. In the United States, 98% of adults have a cell phone, including the same high percentage of White, Black, Latino, and English-speaking Asian adults [6]. In 2024, Americans exchanged nearly 2.2 trillion text messages [7]. Moreover, there is evidence that texting can improve the use of health care services. For example, sending appointment reminders via text increases attendance [8]. Text messages offering cessation treatment could potentially increase use of such services, an important consideration for health systems trying to improve their quality metric performance on tobacco assessment and counseling [9].

While there is growing evidence that text messaging programs can help tobacco users quit [10-12], less research has been done on texting as a way to recruit tobacco users into other evidence-based cessation services [13,14]. Researchers in the United Kingdom found that calls plus texts drove more recruitment for a text messaging trial than calls alone [15] and that sending a reminder about the trial plus a text drove more recruitment than the reminder alone [16]. A New York study reported that multimodal outreach including interactive voice response calls, emails, postcards, and texts led to a 6-fold increase in quitline reengagement [17]. Another New York study found that sending smokers identified in the electronic health record (EHR) 5‑7 texts resulted in 10.0% accepting quitline referral [13]. Researchers in Washington, DC, found that sending emergency department patients screened for tobacco use and willingness to receive up to 4 texts promoting a cessation program led 63.8% to opt in [14]. Illinois researchers reported that sending safety net health center patients identified as smokers in the EHR a letter plus 5 texts outperformed e-referral in driving quitline utilization [18]. Researchers in Vietnam found that frequent texting significantly increased quitline utilization [19]. All of these studies involved multiple text messages, additional modalities, or both.

This study, a pragmatic randomized control trial, implemented a text message campaign in a safety net health system promoting cessation in patients identified as unassisted tobacco users in the EHR. We compared two treatment offers and their uptake by English- and Spanish-speaking patients. The hypotheses were that texting would be a feasible means of recruiting unassisted tobacco users into quitline services, that offering free NRT would attract more enrollees than offering free coaching, and that recruiting by text would work equally well for English and Spanish speakers.

Methods

Study Participants and Intervention

Participants were patients of the Los Angeles County Department of Health Services (LADHS), the second largest municipal health system in the United States [20]. LADHS serves about 750,000 unique patients annually [20] and participates in a state Medicaid value-based care program to improve quality metric performance, including on “Tobacco Assessment and Counseling” [9]. Eligible patients were identified by the LADHS population health team using data extraction from the EHR for quality metric reporting. Eligibility criteria were being ≥18 years old, having had a clinical visit in the past 12 months, being willing to receive text messages, and being an unassisted tobacco user (ie, one with no documented advice or assistance to quit in the past 24 months). Of 4171 patients meeting these criteria, 3139 were English speakers and 1032 were Spanish speakers.

Participants were stratified by language, then randomly assigned to one of two groups. Randomization blocks were generated by the quitline, Kick It California, while participants were assigned and text messages were sent by LADHS. The messages were in English or Spanish according to the participant’s language preference. English speakers in Group 1 received a text reading, “Hello [first name], Kick It California works with your provider to help you quit smoking at no cost. A Quit Coach can double your chances of quitting. For more info call [toll-free phone number].” English speakers in Group 2 received a text reading, “Hello [first name], Kick It California works with your provider to help you quit smoking at no cost. You may be able to get nicotine patches mailed to your home at no cost. For more info call [toll-free phone number].” Spanish speakers in each group received a translated version of the message for that group. Four dedicated toll-free phone numbers were used, each of which was only available to members of a single group who spoke either English or Spanish. These steps were taken to enable tracking of incoming calls and to provide callers with linguistically appropriate services. The texts were sent in April and May 2022.

Staff of the quitline who handled incoming calls were blind to the allocation of participants and offered coaching to all who enrolled. Coaching followed a structured, multicall protocol previously proven effective [21,22] and described elsewhere [23]. Quitline staff also offered a free 2-week supply of nicotine patches to all who enrolled unless NRT was contraindicated, regardless of which text message they received. The patches were mailed directly to participants’ homes. To ensure equitable access, a second text message offering free patches was sent to all participants in June 2022.

Measures

Demographic information obtained from the EHR included age, sex, race/ethnicity, and language. The outcome measures were the proportions of participants calling the quitline and receiving

XSL•FO
RenderX

evidence-based cessation treatments, which could include NRT, quitline coaching, or both. Calls to the quitline and treatments received were attributed to the initial text message only if they took place prior to when the second text message was sent. For those who received coaching, the number of follow-up sessions was also tracked.

Analysis

Descriptive statistics were used to describe patient characteristics by treatment condition. The proportions of participants calling the quitline and receiving treatment (NRT, coaching, or both) were calculated for each treatment and language group. Differences between treatment groups were assessed using χ^2 tests for categorical variables or the Fisher exact test in the event of small cell sizes. Analyses were conducted in 2025 using SAS (version 9.4; SAS Institute Inc).

Ethical Considerations

The study was approved as exempt from human subjects review by the University of California Davis institutional review board (application number 1301241 - 6) and the LADHS institutional review board (application number 2018-12-785). Because the institutional review board determination was that the study was not research, informed consent was not required. Data were deidentified to safeguard participant information. Participants were not offered compensation. The trial was registered on ClinicalTrials.gov (NCT07102290).

Results

Table 1 shows the demographic characteristics, by treatment group, of LADHS patients identified as unassisted tobacco users in the EHR who were randomized into the study. There were no significant differences between groups with regard to age, sex, race/ethnicity, or language, suggesting that the groups were equivalent at baseline.

Table . Characteristics of text message recipients in the Los Angeles County Department of Health Services system by randomized condition (N=4171).

	Group 1, free coaching (n=2086), %	Group 2, free nicotine replacement therapy (n=2085), %	*P* value[a]
Age (years)			
18-24	2.2	2.3	.49
25-44	31.5	30.4	
45-64	52.9	52.2	
>64	13.5	15.1	
Sex			
Male	63.1	60.4	.13
Female	36.9	39.6	
Other	0.05	0.05	
Race/ethnicity			
American Indian/Alaska Native	0.34	0.58	.25
Asian	2.5	3.0	
Black	25.9	25.1	
White	6.5	7.1	
Hispanic/Latino	43.9	44.1	
Non-Hispanic other[b]	19.3	17.8	
Unknown/blank/refused	1.6	2.4	
Language			
English	75.3	75.3	.99
Spanish	24.7	24.8	

[a]*P* value from χ^2 test for all variables except sex, which used the Fisher exact test because of cell sizes <5.

[b]Non-Hispanic other includes Pacific Islander, other race, and multiple races.

Figure 1 summarizes the flow of participants through the study. Of 4171 LADHS patients identified in the EHR as unassisted tobacco users, half (n=2086) were randomly assigned to Group 1 and received a text message offering free quitline coaching. The other half (n=2085) were randomly assigned to Group 2 and received a text message offering free NRT. A total of 61 participants responded to the text they received by calling the quitline, including 28 in Group 1 (1.1%) and 33 in Group 2 (1.6%). Of these, 21 participants engaged in treatment, including 8 in Group 1 (0.4%) and 13 in Group 2 (0.6%). Most of these (n=19) received NRT, including 6 in Group 1 (0.3%) and 13 in Group 2 (0.6%), while 11 received a coaching session to help

XSL•FO
RenderX

them prepare to quit, including 5 in Group 1 (0.2%) and 6 in Group 2 (0.3%). Eight participants also completed at least one follow-up session to help them stay on track with quitting, completing 2.1 follow-up sessions on average (not shown in figure).

Figure 1. Flow of Los Angeles County Department of Health Services text message recipients through the study.

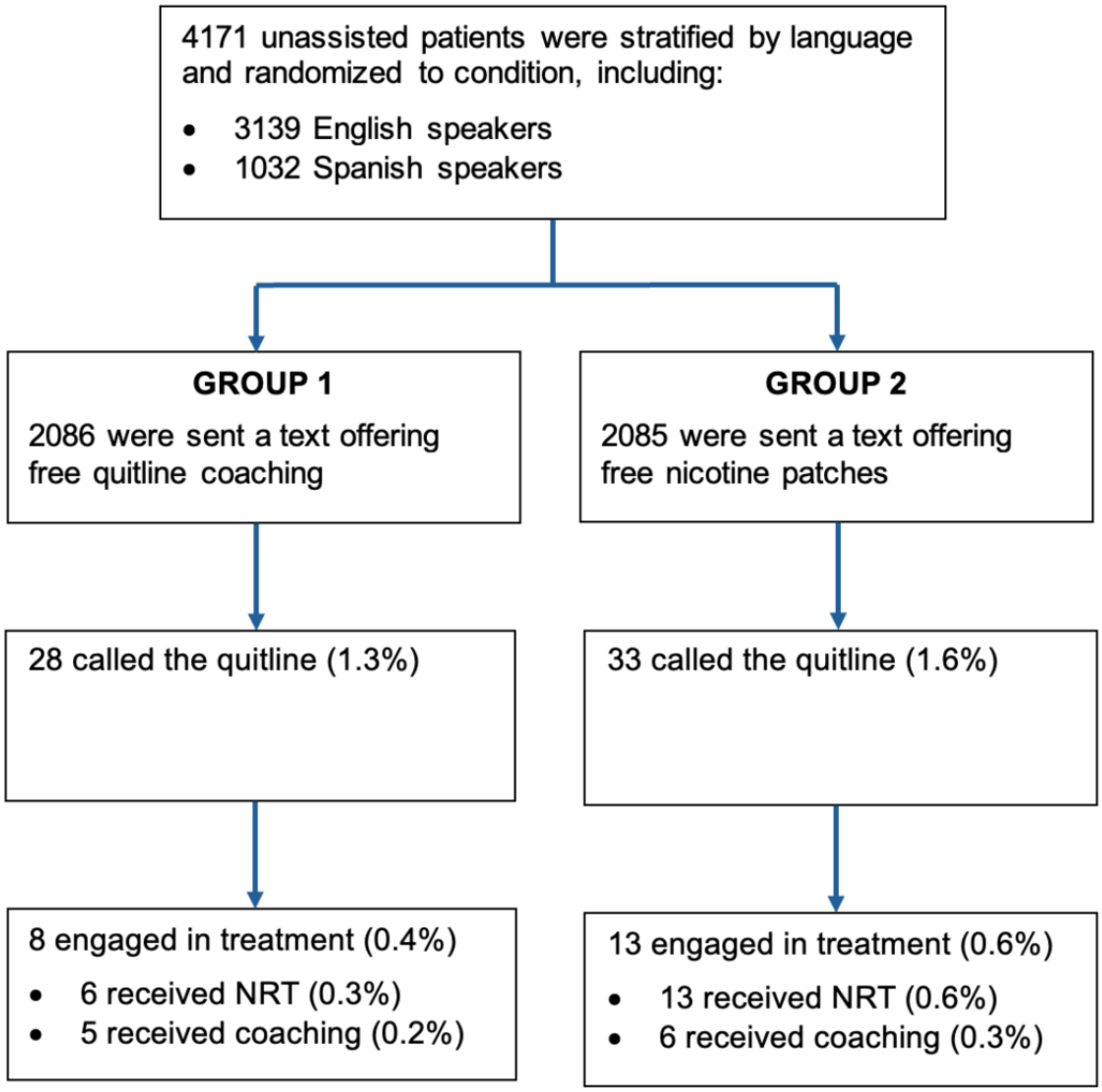


Table 2 shows in more detail the effects of the intervention on calls to the quitline by treatment condition and language. There was no significant difference by treatment condition, with 1.3% (28/2086) of Group 1 participants and 1.6% (33/2085) of Group 2 participants calling the quitline (*P*=.52), nor were there between-group differences within languages. Among English speakers, 1.1% (17/1570) in Group 1 called the quitline vs 1.3% (20/1569) in Group 2 (*P*=.62), and among Spanish speakers, 2.1% (11/516) in Group 1 called the quitline vs 2.5% (13/516) in Group 2 (*P*=.68). However, when examined by the stratification variable, language, there was a significant difference, with Spanish speakers twice as likely as English speakers to call the quitline, 2.3% (24/1032) versus 1.2% (37/3139), respectively (*P*=.008).

Table . Calls to the quitline by Los Angeles County Department of Health Services text message recipients by randomized condition and language.

Language	Group 1, free coaching (n=2086), % (n/N)	Group 2, free nicotine replacement therapy (n=2085), % (n/N)	*P* value[a]	Total (N=4171), % (n/N)	*P* value[b]
English	1.1 (17/1570)	1.3 (20/1569)	.62	1.2 (37/3139)	.008
Spanish	2.1 (11/516)	2.5 (13/516)	.68	2.3 (24/1032)	
Total	1.3 (28/2086)	1.6 (33/2085)	.52	1.5 (61/4171)	

[a]*P* value from χ^2 test comparing Group 1 to Group 2 for each language category.

[b]*P* value from χ^2 test comparing English to Spanish.

XSL•FO
RenderX

Table 3 shows the effects of the intervention on treatments received. Overall, 21 of 4171 or 0.5% of participants received evidence-based treatment, including 0.4% (8/2086) in Group 1 and 0.6% (13/2085) in Group 2. As with calls to the quitline, there was no significant difference by treatment condition in this outcome (*P*=.29), nor were there between-group differences within languages. Among English speakers, 0.3% (5/1570) in Group 1 engaged in treatment versus 0.4% (7/1569) in Group 2 (*P*=.58), and among Spanish speakers, 0.6% (3/516) in Group 1 engaged in treatment versus 1.2% (6/516) in Group 2 (*P*=.51). Spanish speakers engaged in treatment at approximately twice the rate of English speakers, but the difference was nonsignificant, 0.9% (9/1032) versus 0.4% (12/3139), respectively (*P*=.07).

Table . Los Angeles County Department of Health Services text message recipients' receipt of tobacco cessation treatment by randomized condition and language.

Language	Group 1, free coaching (n=2086), % (n/N)	Group 2, free nicotine replacement therapy (n=2085), % (n/N)	*P* value[a]	Total (N=4171), % (n/N)	*P* value[b]
English	0.3 (5/1570)	0.4 (7/1569)	.58	0.4 (12/3139)	.07
Spanish	0.6 (3/516)	1.2 (6/516)	.51	0.9 (9/1032)	
Total	0.4 (8/2086)	0.6 (13/2085)	.29	0.5 (21/4171)	

[a]*P* value from Fisher exact χ^2 test comparing Group 1 to Group 2 for each language category.

[b]*P* value from Fisher exact χ^2 test comparing English to Spanish.

Discussion

Principal Findings

This pragmatic trial tested a strategy of sending unassisted tobacco users in a safety net health system a text message promoting quitline services: either free coaching or free NRT. Prior to the trial, clinicians in the system could e-refer patients to the quitline, but use of this mechanism tends to be low [24,25]. The trial offered an opportunity to test a population health strategy of reaching out to tobacco users outside clinical visits, targeting those who had not recently received help to quit. The intervention resulted in 1.5% of identified tobacco users calling the quitline and 0.5% receiving one or more evidence-based treatments. The latter rate is comparable to the treatment reach rate of US quitlines, which is less than 1% annually [5]. Moreover, it resulted from a low-intensity outreach effort—a single text.

The offers of free coaching and free NRT were equally effective at driving enrollment. This result was unexpected, as previous research found that mailers sent to Medicaid households were much more effective at driving quitline utilization when they offered free NRT than when they offered more generic help [26]. By targeting unassisted tobacco users, the texts may have reached an audience more concerned with obtaining any help to quit than a particular kind of help. The texts may also have done a better job "selling" the coaching than the NRT, as the coaching offer included the message, "a Quit Coach can double your chances of quitting." The NRT offer, in contrast, made no explicit attempt to boost positive expectancy. In any case, there was no discernible difference in the two offers' ability to motivate enrollment.

In contrast, there was a notable difference by language: Spanish speakers called the quitline and engaged in treatment at twice the rates of English speakers. These are encouraging results, given that Latino tobacco users are less likely to receive clinical advice to quit than non-Latino White tobacco users [27]. Latinos and Spanish speakers often lack knowledge about available cessation resources, hold misconceptions about cessation treatment, and may view tobacco use as a personal weakness, rather than as a condition to be treated [28-30]. Moreover, there is evidence that while Spanish speakers face barriers accessing digital health information, such as through patient portals, text messaging actually facilitates their access [31]. This may help explain why receiving a language-concordant text from their health system appears to have increased engagement with the quitline among Spanish-speaking patients. Regardless, the result suggests that population health strategies like the one tested in this study could help address disparities in the receipt of clinical advice.

The study adds to previous research on the use of texting to recruit patients into cessation treatment. Several studies have examined texting either alone [13,14,19] or as part of a multimodal outreach approach [15-18]. All studies in the former category involved multiple texts. Taken together, these studies have shown that texting can be effective in recruiting tobacco users into cessation service in the context of multitext or multimodal campaigns. This study adds to this body of research by confirming that recruiting tobacco users via text is feasible and showing that even a single text can increase recruitment.

The lesson we draw from this finding is not that organizations using multiple text messages to recruit tobacco users into evidence-based cessation treatment or using text messages in combination with other modalities should consider downgrading to a single-text format, but rather that this low-cost, almost universally available outreach tool is robust enough to measurably increase recruitment even with a single dose. It speaks to the value of quitlines partnering with health systems on text messaging campaigns to reach tobacco-using patients outside of clinical visits to offer help with quitting. Such population health strategies have the potential to meaningfully supplement direct health care provider treatment and referral [32] and to improve quality metric performance for safety net health systems in a state Medicaid value-based care program without additional staffing costs or burden [9]. Complementary strategies may further increase the effectiveness of text

XSL•FO
RenderX

messaging as a recruitment tool, including using it to reach out to additional groups of patients, using opt-out approaches [33], increasing the number and variety of messages sent [13], using dynamic tailoring [34], offering incentives [26], and offering other mHealth services [14].

The public health significance of this low-cost approach is that other health systems can quickly and widely adopt it to improve population-based tobacco quality metrics, which would have the additive effect of increasing the treatment reach of evidence-based resources such as quitlines. Starting in 2026, community cancer centers accredited by the American College of Surgeons Commission on Cancer, which care for 70% of the nation's patients with cancer, have a new standard to screen every patient newly diagnosed with cancer for smoking and to offer a referral within 30 days [35]. Also in 2026, health plans have a new population-based metric for "Tobacco Use Screening and Cessation Intervention" [36]. Low-cost proactive outreach strategies like the one described in this study can help cancer centers, health plans, provider networks, and other health care entities reach beyond clinical encounters to offer tobacco treatment to their members and patients who may need it.

To our knowledge, this study is the first to show that texting may be a more effective recruitment tool among Spanish-speaking tobacco users than among English-speaking tobacco users. This also speaks to the value of quitlines and health systems partnering to reach underserved patients [27]. All state quitlines in the United States serve both English and Spanish speakers; therefore, all health systems could potentially partner with their quitline on bilingual text messaging campaigns to reduce disparities in access to evidence-based cessation services [37].

Limitations

This study had limitations. First, only unassisted tobacco users were included in the study; the response and engagement rates may have differed if assisted tobacco users were included. Second, there was no means to identify patients with incorrect phone numbers in the participant list, and there was no way to know whether text messages were received unless patients contacted the quitline. Third, it is unknown whether participants enrolling in the quitline benefited from the service, as the trial was not powered or designed to assess quitting outcomes. Fourth, the study included only one safety net health system, limiting generalizability. Finally, the trial was conducted a little more than two years into the COVID-19 pandemic, and it is unknown to what extent COVID-19–related stressors such as social, economic, and psychosocial distress may have affected patients' willingness to respond to a text message offering help to quit.

Conclusion

This study showed that a single text message can effectively connect unassisted tobacco users to evidence-based cessation services for both English and Spanish speakers, but especially the latter. It showed that a population health approach of reaching out to tobacco-using patients outside clinical visits can meaningfully supplement provider referral. This population health strategy is being disseminated to and adapted by other safety net health systems and state Medicaid managed care plans in a statewide learning collaborative [38,39]. Future research should examine ways of increasing the population health impact of the strategy.

Acknowledgments

The authors thank Yue-Lin Zhuang, who assisted with the analysis.

Funding

The study was funded by the Tobacco-Related Disease Research Program (Community Practice-Based Implementation Research Award number 28CP-0039 and Research Award number T33IR6713). The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

The datasets generated or analyzed in this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Checklist 1
CONSORT checklist.
[PDF File, 248 KB - publichealth_v12i1e83269_app1.pdf]

References

1. Fiore MC, Jaén CR, Baker TB, et al. Treating Tobacco Use and Dependence: 2008 Update Clinical Practice Guideline: US Department of Health and Human Services. Public Health Service; 2008. URL: https://www.ncbi.nlm.nih.gov/books/NBK63952/ [accessed 2026-02-20]

XSL•FO
RenderX

2. Matkin W, Ordóñez-Mena JM, Hartmann-Boyce J. Telephone counselling for smoking cessation. Cochrane Database Syst Rev 2019 May 2;5(5):CD002850. [doi: 10.1002/14651858.CD002850.pub4] [Medline: 31045250]
3. Vidrine JI, Shete S, Cao Y, et al. Ask-Advise-Connect: a new approach to smoking treatment delivery in health care settings. JAMA Intern Med 2013 Mar 25;173(6):458-464. [doi: 10.1001/jamainternmed.2013.3751] [Medline: 23440173]
4. Piñeiro B, Vidrine DJ, Wetter DW, et al. Implementation of Ask-Advise-Connect in a safety net healthcare system: quitline treatment engagement and smoking cessation outcomes. Transl Behav Med 2020 Feb 3;10(1):163-167. [doi: 10.1093/tbm/iby108] [Medline: 30476236]
5. Mason K, editor. FY24 annual survey of state quitlines: results. North American Quitline Consortium. 2025 May 15. URL: https://www.naquitline.org/page/survey [accessed 2026-02-20]
6. Mobile fact sheet. Pew Research Center. 2025 Nov 20. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2025-11-29]
7. Annual survey highlights. CTIA. URL: https://www.ctia.org/news/2025-annual-survey-highlights [accessed 2025-11-29]
8. Gurol-Urganci I, de Jongh T, Vodopivec-Jamsek V, Atun R, Car J. Mobile phone messaging reminders for attendance at healthcare appointments. Cochrane Database Syst Rev 2013 Dec 5;2013(12):CD007458. [doi: 10.1002/14651858.CD007458.pub3] [Medline: 24310741]
9. Improving quality and reducing disparities through the Quality Incentive Pool (QIP). : California Health Care Safety Net Institute, California Association of Public Hospitals and Health Systems; 2021 Apr 28 URL: https://caph.org/wp-content/uploads/2021/04/caph_qip_final-4.28.21.pdf [accessed 2026-02-20]
10. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y, Dobson R. Mobile phone text messaging and app-based interventions for smoking cessation. Cochrane Database Syst Rev 2019 Oct 22;10(10):CD006611. [doi: 10.1002/14651858.CD006611.pub5] [Medline: 31638271]
11. Fang YE, Zhang Z, Wang R, et al. Effectiveness of eHealth smoking cessation interventions: systematic review and meta-analysis. J Med Internet Res 2023 Jul 28;25:e45111. [doi: 10.2196/45111] [Medline: 37505802]
12. Puljević C, Meciar I, Holland A, et al. Systematic review and meta-analysis of text messaging interventions to support tobacco cessation. Tob Control 2025 Apr 1;34(2):228-238. [doi: 10.1136/tc-2023-058323] [Medline: 38448226]
13. Krebs P, Sherman SE, Wilson H, et al. Text2Connect: a health system approach to engage tobacco users in quitline cessation services via text messaging. Transl Behav Med 2020 Feb 3;10(1):292-301. [doi: 10.1093/tbm/ibz033] [Medline: 32011721]
14. Abroms LC, Wu KC, Krishnan N, et al. A pilot randomized controlled trial of text messaging to increase tobacco treatment reach in the emergency department. Nicotine Tob Res 2021 Aug 18;23(9):1597-1601. [doi: 10.1093/ntr/ntab036] [Medline: 33684207]
15. Free C, Hoile E, Robertson S, Knight R. Three controlled trials of interventions to increase recruitment to a randomized controlled trial of mobile phone based smoking cessation support. Clin Trials 2010 Jun;7(3):265-273. [doi: 10.1177/1740774510367687] [Medline: 20484492]
16. Free CJ, Hoile E, Knight R, Robertson S, Devries KM. Do messages of scarcity increase trial recruitment? Contemp Clin Trials 2011 Jan;32(1):36-39. [doi: 10.1016/j.cct.2010.09.002] [Medline: 20840874]
17. Carlini B, Miles L, Doyle S, Celestino P, Koutsky J. Using diverse communication strategies to re-engage relapsed tobacco quitline users in treatment, New York State, 2014. Prev Chronic Dis 2014;12:E179. [doi: 10.5888/pcd12.150191]
18. Hitsman B, Matthews PA, Papandonatos GD, et al. An EHR-automated and theory-based population health management intervention for smoking cessation in diverse low-income patients of safety-net health centers: a pilot randomized controlled trial. Transl Behav Med 2022 Oct 7;12(9):892-899. [doi: 10.1093/tbm/ibac026] [Medline: 36205472]
19. Sadasivam RS, Nagawa CS, Wijesundara JG, et al. Peer texting to promote quitline use and smoking cessation among rural participants in Vietnam: randomized clinical trial. Int J Public Health 2024;69:1606941. [doi: 10.3389/ijph.2024.1606941] [Medline: 38651035]
20. Who we are. Los Angeles County Health Services. URL: https://dhs.lacounty.gov/who-we-are [accessed 2025-12-04]
21. Zhu SH, Stretch V, Balabanis M, Rosbrook B, Sadler G, Pierce JP. Telephone counseling for smoking cessation: effects of single-session and multiple-session interventions. J Consult Clin Psychol 1996;64(1):202-211. [doi: 10.1037/0022-006X.64.1.202]
22. Zhu SH, Anderson CM, Tedeschi GJ, et al. Evidence of real-world effectiveness of a telephone quitline for smokers. N Engl J Med 2002 Oct 3;347(14):1087-1093. [doi: 10.1056/NEJMsa020660] [Medline: 12362011]
23. Zhu S, Tedeschi GJ, Anderson CM, Pierce JP. Telephone counseling for smoking cessation: what's in a call? J Couns Dev 1996 Nov 12;75(2):93-102. [doi: 10.1002/j.1556-6676.1996.tb02319.x]
24. Hood-Medland EA, Stewart SL, Nguyen H, et al. Health system implementation of a tobacco quitline eReferral. Appl Clin Inform 2019 Aug;10(4):735-742. [doi: 10.1055/s-0039-1697593] [Medline: 31578046]
25. Tong EK, Zhu SH, Anderson CM, et al. Implementation, maintenance, and outcomes of an electronic referral to a tobacco quitline across five health systems. Nicotine Tob Res 2023 May 22;25(6):1135-1144. [doi: 10.1093/ntr/ntad008] [Medline: 36977494]
26. Anderson CM, Kirby CA, Tong EK, Kohatsu ND, Zhu SH. Effects of offering nicotine patches, incentives, or both on quitline demand. Am J Prev Med 2018 Dec;55(6 Suppl 2):S170-S177. [doi: 10.1016/j.amepre.2018.07.007] [Medline: 30454671]

XSL•FO
RenderX

27. Valencia CV, Dove M, Tong EK. Factors associated with receipt of smoking cessation advice and assistance by health professionals among Latino and Non-Latino White smokers with Medicaid insurance in California. JAMA Netw Open 2022 Jan 4;5(1):e2144207. [doi: 10.1001/jamanetworkopen.2021.44207] [Medline: 35044467]
28. Bock BC, Niaura RS, Neighbors CJ, Carmona-Barros R, Azam M. Differences between Latino and non-Latino White smokers in cognitive and behavioral characteristics relevant to smoking cessation. Addict Behav 2005 May;30(4):711-724. [doi: 10.1016/j.addbeh.2004.08.017] [Medline: 15833576]
29. Levinson AH, Borrayo EA, Espinoza P, et al. An exploration of Latino smokers and the use of pharmaceutical aids. Am J Prev Med 2006;31(2):167-171. [doi: 10.1016/j.amepre.2006.03.022]
30. Piñeiro B, Díaz DR, Monsalve LM, et al. Systematic transcreation of self-help smoking cessation materials for Hispanic/Latino smokers: improving cultural relevance and acceptability. J Health Commun 2018;23(4):350-359. [doi: 10.1080/10810730.2018.1448487] [Medline: 29533167]
31. Higashi RT, Thakur B, Repasky EC, et al. Digital health technology use among Spanish speakers in the US: a scoping review. JAMA Netw Open 2025 May 1;8(5):e2510386. [doi: 10.1001/jamanetworkopen.2025.10386] [Medline: 40372754]
32. Valencia CV, Dove MS, Cummins SE, et al. A proactive outreach strategy using a local area code to refer unassisted smokers in a safety net health system to a quitline: a pragmatic randomized trial. Nicotine Tob Res 2023 Jan 1;25(1):43-49. [doi: 10.1093/ntr/ntac156] [Medline: 36103393]
33. Richter KP, Catley D, Gajewski BJ, et al. The effects of opt-out vs opt-in tobacco treatment on engagement, cessation, and costs: a randomized clinical trial. JAMA Intern Med 2023 Apr 1;183(4):331-339. [doi: 10.1001/jamainternmed.2022.7170] [Medline: 36848129]
34. Graham AL, Papandonatos GD, Jacobs MA, et al. Optimizing text messages to promote engagement with internet smoking cessation treatment: results from a factorial screening experiment. J Med Internet Res 2020 Apr 2;22(4):e17734. [doi: 10.2196/17734] [Medline: 32238338]
35. Standard 5.9: smoking cessation for patients with cancer. American College of Surgeons, Commission on Cancer. 2025 Jul. URL: https://accreditation.facs.org/accreditationdocuments/CoC/Resources/SmokingCessation/Standard5.9.pdf [accessed 2025-12-04]
36. Kolinski B. Tobacco cessation HEDIS measure planned for MY 2026. NCQA. 2025 Jun 12. URL: https://www.ncqa.org/blog/tobacco-cessation-hedis-measure-planned-for-my-2026 [accessed 2025-12-04]
37. Quitline profiles. North American Quitline Consortium. URL: https://www.naquitline.org/page/quitlineprofiles [accessed 2025-12-04]
38. Kaslow AA, Romano PS, Schwarz E, Shaikh U, Tong EK. Building and scaling-up California Quits: supporting health systems change for tobacco treatment. Am J Prev Med 2018 Dec;55(6 Suppl 2):S214-S221. [doi: 10.1016/j.amepre.2018.07.045] [Medline: 30454676]
39. California Quits. CA Quits: advancing tobacco treatment with Medi-Cal for health systems change. UC David Health. 2025 Nov. URL: https://health.ucdavis.edu/chpr/programs/ca-quits/resources/data-reports [accessed 2026-02-20]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
e-referral: electronic referral
EHR: electronic health record
LADHS: Los Angeles County Department of Health Services
mHealth: mobile health
NRT: nicotine replacement therapy

Edited by A Mavragani, T Sanchez; submitted 30.Aug.2025; peer-reviewed by B Yang, Y Fang; revised version received 23.Dec.2025; accepted 24.Dec.2025; published 27.Feb.2026.

Please cite as:
Valencia CV, Dove MS, Kim K, Giboney P, Yee Jr HF, Kirby CA, Anderson CM, Zhu SH, Tong EK
Texting to Recruit Unassisted Tobacco Users at a Large Safety Net Health System Into Quitline Service for a Medicaid Value-Based Care Program: Pragmatic Randomized Controlled Trial
JMIR Public Health Surveill 2026;12:e83269
URL: https://publichealth.jmir.org/2026/1/e83269
doi:10.2196/83269

© Cindy Valencia, Melanie Dove, Karen Kim, Paul Giboney, Hal F Yee Jr, Carrie Kirby, Christopher Anderson, Shu-Hong Zhu, Elisa K Tong. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 27.Feb.2026. This is

XSL•FO
RenderX

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

XSL•FO
RenderX

Multiyear Wastewater Monitoring of High-Risk Substances Surrounding an Annual Electronic Dance Music Festival, Las Vegas, Nevada, USA, 2023-2025: Surveillance Study

Daniel Gerrity[1,2*], PhD; Casey A Barber[1*], PhD; Rebecca A Trenholm[1], MS; Brett J Vanderford[1], BS; Katherine C Crank[1], PhD; Ching-Lan Chang[3,4,5], BS; Michael A Moshi[3,4,5], BS; Anil T Mangla[6,7], PhD; Cassius Lockett[6], PhD; Edwin C Oh[3,4,5,8,9], PhD

[1]Applied Research and Development Center, Southern Nevada Water Authority, Las Vegas, NV, United States

[2]Department of Civil and Environmental Engineering and Construction, Howard R. Hughes College of Engineering, University of Nevada, Las Vegas, 4505 S Maryland Pkwy, Mail Stop 4005, Las Vegas, NV, United States

[3]Laboratory of Neurogenetics and Precision Medicine, College of Sciences, University of Nevada, Las Vegas, Las Vegas, NV, United States

[4]Neuroscience Interdisciplinary Ph.D. Program, University of Nevada, Las Vegas, NV, United States

[5]Center for Water Intelligence and Community Health, University of Nevada, Las Vegas, Las Vegas, NV, United States

[6]Southern Nevada Health District, Las Vegas, NV, United States

[7]School of Public Health, University of Nevada, Las Vegas, NV, United States

[8]Department of Brain Health, School of Integrated Health Sciences, University of Nevada, Las Vegas, Las Vegas, NV, United States

[9]Department of Internal Medicine, Kirk Kerkorian School of Medicine, University of Nevada, Las Vegas, Las Vegas, NV, United States

[*]these authors contributed equally

Corresponding Author:
Daniel Gerrity, PhD
Applied Research and Development Center, Southern Nevada Water Authority, Las Vegas, NV, United States

Abstract

Background: Prior studies involving music festivals (including on-site surveys, drug checking, and wastewater-based analyses) suggest that recreational substance use is common among festivalgoers, though preferred substances and usage may be region- and event-specific. However, multiyear, event-specific wastewater surveillance data in the United States are limited.

Objective: This study aimed to characterize concentrations of high-risk substances (HRS) and their metabolites in wastewater surrounding an annual electronic dance music (EDM) festival in Las Vegas, Nevada, USA. Using a progressive sampling approach over 3 consecutive years (2023 - 2025), we assessed trends and examined comparisons by sample type, day, and location to inform situational awareness and future public health surveillance efforts and risk mitigation strategies.

Methods: Wastewater samples were collected from utility access points at the festival site (2024 - 2025), off-site in the surrounding community (2023 - 2024), and at the downstream wastewater treatment plant (WWTP; 2023 - 2025) during the festival week and weekend. For comparison, samples from a nonfestival weekend were collected off-site and at the WWTP in 2023. Target analytes consisted of 24 parent compounds and metabolites, including stimulants, opioids, and sedatives, which were analyzed by liquid chromatography–tandem mass spectrometry (LC-MS/MS) with isotope dilution. Analyte selection prioritized substances associated with overdose risk and with prior festival-related surveillance.

Results: Cocaine (and 3 of 4 measured metabolites), methamphetamine, amphetamine, 11-nor-9-carboxy-Δ-9-tetrahydrocannabinol (THC-COOH), and tramadol were detected in 100% of festival weekend samples; 3,4-methylenedioxymethamphetamine (MDMA; ecstasy) was detected (at concentrations of up to nearly 2 mg/L) in all but 1 festival site sample collected prior to the start of the event in 2025. Festival-associated signals were most pronounced for stimulant-related analytes, which were also detected in WWTP and off-site samples during the festival. Compared with nonfestival weeks, MDMA concentrations were approximately 10 to 30 times higher at off-site locations and nearly 100 times higher at the downstream WWTP during the festival. Opioid-related concentrations were comparable at the WWTP and the festival site, although there were notable detections of norfentanyl and relatively high concentrations of tramadol in some festival site samples. MDMA- and cocaine-related analyte concentrations peaked during the final festival days in 2024 and 2025; the 2023 samples highlighted the presence of HRS in local wastewater during nonfestival weekends.

Conclusions: Event-focused, multiyear wastewater surveillance detected stimulants and psychoactive substances, opioids, and Δ-9-tetrahydrocannabinol (THC) associated with an annual EDM festival weekend. Findings are consistent with previous wastewater-based studies of music and dance festivals, with MDMA concentrations among the highest reported in similar

XSL•FO
RenderX

surveillance efforts. Codetection of opioid- and stimulant-related analytes suggests varied usage preferences, polysubstance use, and/or unintentional adulteration (eg, lacing). These findings demonstrate the value of wastewater monitoring as a complementary, population-level tool to support public health preparedness, harm-reduction strategies, and attendee safety during large-scale events. These data also inform wastewater utilities about expected influent composition during special events, with relevance for treatment operations and water reuse applications.

(JMIR Public Health Surveill 2026;12:e102679) doi:10.2196/102679

KEYWORDS

high-risk substances; HRS; opioid; 3,4-methylenedioxymethamphetamine; MDMA; wastewater monitoring; event-based wastewater surveillance; wastewater-based epidemiology; WBE; substance use prevention; public health

Introduction

Large events, including music festivals, can present unique public health and safety challenges, particularly if attendees choose to engage in high-risk health behaviors such as recreational substance use. Surveys suggest that substance use is common among festivalgoers [1,2], with 24% of respondents in one study reporting an intention to use drugs at music festivals [3]. Researchers have previously investigated underlying reasons for adverse outcomes associated with festival-specific substance use behaviors, identifying a complex blend of factors such as first-time use among inexperienced attendees, new or unfamiliar drug sources (ie, drugs purchased from strangers), crowding, heat, exhaustion, and more [4]. The presence of adulterated or counterfeit drugs in this setting further exacerbates potential health risks among festivalgoers; a review of 6 studies involving voluntary drug-checking stations at festival sites found adulterants in 11% to 55% of tested samples [5]. These data sources may still underestimate the scope of this problem, as most event-related surveys, interviews, and drug-checking studies rely on voluntary participation and self-reporting of stigmatized (and often illicit) substance use behaviors.

In the search for a complementary, noninvasive tool for characterizing event-related recreational substance use, researchers have explored wastewater monitoring and wastewater-based epidemiology (WBE) of high-risk substances (HRS) at music festivals. Wastewater results from festivals in Australia [6,7], Brazil [8], and 6 European countries [9] suggest event- and region-specific differences in the types and amounts of drugs present. Lai et al [6] noted consistency in per capita consumption estimates for conventional and emerging illicit drugs (based on daily, on-site, and composite wastewater samples) across 2 years of a recurring summer music festival in Australia, suggesting that attendees may exhibit consistent year-over-year substance use patterns surrounding an event.

Las Vegas, Nevada, USA, is an international entertainment destination with multiple large annual events. Year-round activities draw tourists and locals alike, but one of the largest annual events is an electronic dance music (EDM) festival that has attracted more than 520,000 annual attendees in recent years [10]. During the festival weekend in 2022, surface water in the Las Vegas Wash, which primarily conveys discharged effluent from local wastewater treatment plants (WWTPs), was analyzed for the presence of pharmaceuticals and recreational drugs that had persisted through wastewater treatment, as a means of assessing ecological and environmental risks [11]. That study identified a unique fingerprint of community substance use associated with the festival weekend, including 3,4-methylenedioxymethamphetamine (MDMA; ecstasy), methamphetamine, cocaine, and ketamine. In fact, the median MDMA concentration in the Las Vegas Wash during the festival was 46 times higher than the study control (ie, nonfestival period) [11]. Elevated concentrations of drugs and their metabolites have important implications for Las Vegas Wash ecology and for understanding substance use in the local community during the event. However, due to the efficacy of wastewater treatment in attenuating concentrations of some trace organics [12,13], including HRS [14], direct monitoring of untreated wastewater (rather than effluent-impacted surface water) is preferable for WBE efforts.

Building on these local findings [11], this study characterizes HRS occurrence in site-specific and community-level raw wastewater surrounding the festival weekend over 3 consecutive years (2023, 2024, and 2025). Specifically, this study examines wastewater-based assessments of HRS over time and across sampling locations, festival days, and sample types (grab vs composite). This work further characterizes the wastewater-based HRS profile of the festival weekend for both the corresponding wastewater utility (to understand event-related changes in influent wastewater composition and to potentially inform wastewater treatment and reuse operations) and public health and safety personnel (to inform future event-related public health surveillance activities, responses, and resource allocations).

Methods

Festival Information

During each year of the study period (2023 - 2025), the festival was held over a single mid-May weekend at a large outdoor venue within the same sewershed as the casino-resort corridor known as the Las Vegas Strip. Off-site promotional events catering to attendees began as early as the preceding Tuesday, but the on-site campground opened on Thursday. The on-site festival was then held nightly on Friday, Saturday, and Sunday, closing at 5:30 AM each morning. The festival concluded on Monday morning, with the campground closing later that day.

Wastewater Sampling

Sampling Locations and Sample Handling

Festival-related wastewater sampling expanded each year of the study, from pilot sampling in 2023 to more comprehensive

XSL•FO
RenderX

sampling before, during, and after the festival weekend in 2024 and 2025. Similar sampling schemes (pre-event, during, and postevent) have been described in previous event-based wastewater monitoring studies [15]. Sample locations and types by year and collection day are shown in Table 1. An aliquot (80 mL) of each collected sample was transferred to two 40-mL amber glass vials (containing 50 mg/L of ascorbic acid for oxidant quenching and 1 g/L of sodium azide for biological preservation) and then stored at 4°C until processing and analysis. A map of these sampling locations in the context of Las Vegas area sewersheds is shown in Figure 1. Although the site-specific sampling efforts required mapping of wastewater flows originating from contributing buildings, sites, or complexes, it was not possible to determine each sample's specific points of origin *within* those contributing buildings, sites, or complexes, thereby maintaining individual privacy.

Table . Wastewater sampling locations by day and year during the festival week and weekend. Samples were composite unless otherwise noted[a].

Year	Festival week and weekend sample collection days					
	Thursday	Friday	Saturday	Sunday	Monday	Tuesday
2023[b]	—[c]	—	• Strip	• Bars	• WWTP[d,e] • Airport	—
2024	—	• Festival[d] • WWTP • Airport	• Festival[d] • WWTP • Strip • Bars	• Festival[d] • WWTP	• Festival[d] • WWTP (n=2)[f]	—
2025	• Festival • WWTP	• Festival • WWTP	• Festival • WWTP	• Festival • WWTP	• Festival • WWTP	• Festival (n=2)[f] • WWTP (n=2)[f]

[a]Off-site sample locations include airport, bars, and Las Vegas Strip.

[b]Additional nonfestival-weekend samples from the Las Vegas Strip, bars, airport, and WWTP (n=1 per site; not shown here) were collected in June and July 2023 for comparison.

[c]No samples collected.

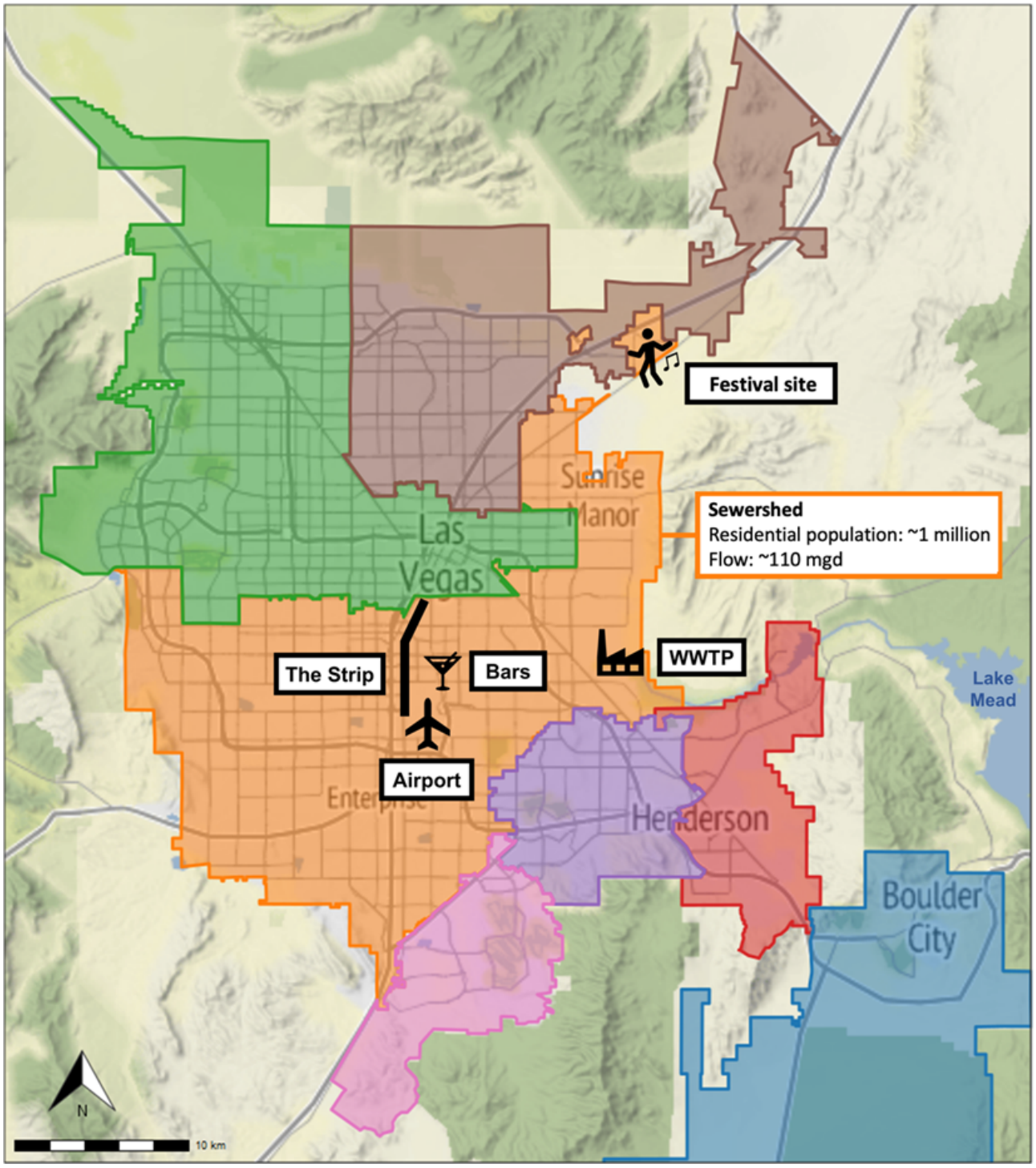
[d]Grab sample only.

[e]WWTP: wastewater treatment plant.

[f]One composite and one grab sample were collected for sample-type comparison.

XSL•FO
RenderX

Figure 1. Map of the Las Vegas area sewersheds and wastewater sampling locations. mgd: million gallons per day; WWTP: wastewater treatment plant.



WWTP Sampling

Raw influent wastewater samples were collected each year from the WWTP downstream of the event venue by collaborating utility partners. This WWTP treats approximately 110 million gallons per day (mgd) and serves the casino-resort corridor of the Las Vegas Strip (with an average weekly visitor volume of approximately 800,000 throughout the year), in addition to the international airport, a large university, and a residential population of approximately 1 million. In 2023, 1 grab sample was collected at the WWTP after the final night of the festival (on Monday morning, May 22, 2023), and again 5 and 7 weeks later (nonfestival Mondays of June 26 and July 10, 2023, respectively). In 2024, 24-hour volume-proportional (flow-paced) influent composite samples were collected each morning during the festival (starting Friday and ending Monday, with collections representing the preceding 24 h), with a corresponding grab sample collected for comparison on Monday morning (as in 2023). Finally, the 2025 WWTP sampling expanded the 2024 sampling scheme to include composite sampling on Thursday (prefestival) and Tuesday mornings

XSL•FO
RenderX

(postfestival), with another grab sample collected for comparison on Tuesday.

Community (Off-Site) Sampling

Time-proportional composite samples were collected using an autosampler (AS950, Hach) from utility access points (ie, manholes) receiving wastewater from the international airport, the casino-resort corridor of the Las Vegas Strip, and a cluster of bars and nightclubs near the Strip. Sampling details associated with these "off-site" manhole locations were described previously [16], and the days of the week represented by these samples are detailed in Table 1 and Multimedia Appendix 1. In 2023, samples from the airport, the Las Vegas Strip, and bars were collected during the festival weekend (in May) and the 2 nonfestival comparison weekends (in late June and early July). Similar sentinel-site sampling was performed in 2024 but not in 2025. Previous event-based HRS wastewater studies similarly included samples from the surrounding community [6] and nonfestival weekends [9] for comparison. No flow data were available for these sampling locations.

Festival-Site Sampling

In response to preliminary community-scale results from 2023, festival site-specific sampling was initiated in 2024 and repeated in 2025. Wastewater was collected by the collaborating wastewater utility from a manhole receiving wastewater only from the festival venue and campground (ie, mobile showers and restrooms discharging to the sewer collection system); these flows ultimately travel to the downstream WWTP. In 2024, a field-deployable autosampler was not yet available for use by the collaborating utility, so grab samples were collected daily from Friday to Monday morning (n=4) to complement the samples collected at the downstream WWTP. In 2025, the sampling period was extended from Thursday (prefestival) to Tuesday (postfestival), and, in lieu of grab samples (except for a supplemental Tuesday grab), 24-hour time-proportional composite samples were collected using an autosampler (CA101, C.E.C. Analytics). Therefore, sample type is a potential confounding variable for direct comparisons between 2024 and 2025 concentrations. The festival attracted more than 520,000 attendees each year, but no corresponding wastewater flow data were available for the festival-site sampling.

Sample Analysis

Target analytes in this study included 24 parent compounds and metabolites (Table S1 in Multimedia Appendix 2). As in our concurrent campus-based wastewater monitoring campaign [17], nonopioid substances included cocaine and its metabolites (benzoylecgonine [BZE], ecgonine methyl ester [EME], ecgonine, and norcocaine), methamphetamine (and amphetamine), MDMA (and 3,4-methylenedioxyamphetamine [MDA]), Δ-9-tetrahydrocannabinol (THC; also 11-hydroxy-Δ-9-tetrahydrocannabinol [THC-OH] and 11-nor-9-carboxy-Δ-9-tetrahydrocannabinol [THC-COOH]), and the tranquilizer xylazine. Natural (heroin, morphine, and codeine), semisynthetic (hydrocodone and oxycodone), and synthetic (fentanyl, methadone, and tramadol) opioid targets were also included. Relevant opioid metabolites included 6-acetylmorphine for heroin, norfentanyl for fentanyl, and 2-ethylidene-1,5-dimethyl-3,3-diphenylpyrrolidine (EDDP) for methadone. Fentanyl and xylazine were added to the analyte panel in 2025.

Analyses involved direct injection of aqueous samples after a 10-fold dilution to reduce matrix interference; samples were diluted further if concentrations fell outside the calibration range. All compounds were analyzed by liquid chromatography–tandem mass spectrometry (LC-MS/MS) with isotope dilution quantification according to previously published methods [18,19]. Analyses used a CTC Autosampler (CTC Analytics) and a 1260 LC Binary Pump (Agilent). Target compounds and their isotopically labeled analogs were then analyzed using positive electrospray ionization (ESI) in multiple reaction monitoring (MRM) mode on a 6500 QTRAP mass spectrometer (SCIEX). Additional analytical details for all target compounds, including method reporting limits (MRLs) and MRM transitions, are included in Text S1 in Multimedia Appendix 2. Field and analytical blanks (consisting of Milli-Q water treated as samples) were negative for all target compounds.

Statistical Analysis

Wilcoxon signed-rank tests were conducted in R using RStudio. The first analysis compared site-specific analyte concentrations between 2024 and 2025 (ie, festival 2024 vs festival 2025 and WWTP 2024 vs WWTP 2025), and data were limited to sampling days that overlapped in both years. Note that the festival comparison was potentially confounded by the difference in sample type (grab in 2024 and composite in 2025). The second analysis compared same-day concentrations at the festival site and the WWTP across both 2024 and 2025. Any concentrations below the MRL were omitted from the statistical analyses, along with their paired samples.

Ethical Considerations

The University of Nevada, Las Vegas Institutional Review Board reviewed this project and determined it to be exempt from human participant research regulations and from the requirement for informed consent, according to federal regulations and university policy.

Results

Detection Frequencies and Analyte Concentrations by Year and Sampling Location

Principal Findings

Analyte detection frequencies and corresponding concentrations were shared electronically with the collaborating wastewater utility and the local health authority as soon as the data were finalized. Detection frequencies for each analyte by year and sampling location during the festival weekend are shown in Table 2, with the airport, bars, and Las Vegas Strip sampling sites combined as "off-site" locations.

Cocaine (and 3 of 4 metabolites), methamphetamine, amphetamine, THC-COOH, and tramadol were detected (above the MRL) in 100% of samples across all years and all locations during the festival week and weekend (excluding the 2023

XSL•FO
RenderX

nonfestival comparison weeks). This finding is consistent with previous wastewater surveillance studies in Southern Nevada; for example, methamphetamine and THC typically exhibit some of the highest per capita consumption rates based on Southern Nevada's community-scale wastewater concentrations [19]. Additionally, amphetamine has been one of the most frequently detected compounds in the sewer collection system of a local university campus [17]. In 2010, elevated cocaine consumption was linked to a major sporting event weekend based on wastewater samples collected from the same Southern Nevada WWTP in this study [14]—a finding also observed in 2023 [19]. Here, norcocaine—the one exception among the cocaine metabolites—was detected in only a subset of the samples collected upstream of the WWTP, specifically in all 3 samples from the bars in 2023 (festival and nonfestival weekends) and in 1 sample from the festival site in 2025 (grab sample only). Neither heroin, THC, nor xylazine (analyzed only in 2025) was detected above the MRL in any samples in this study, and the major heroin metabolite acetylmorphine was not detected at all in 2025 or at the WWTP in 2024.

MDMA was detected in all festival-impacted samples (including off-site, festival site, and WWTP samples), except for a 2025 festival site composite collected *before* the campground was officially opened for the weekend. In the previous community-scale study, MDMA occurrence was rare in neighboring Southern Nevada sewersheds (20%±8% detection frequencies in the previous study), compared to the WWTP in this study (92% detection frequency previously) [19]. MDMA occurrence in influent wastewater at this WWTP is presumably driven primarily by contributions from the Las Vegas Strip, but the high MDMA detection frequency across all festival weekend samples (including those from the airport) highlights the impact of the event across the city. Although detected in 100% of festival weekend off-site samples, MDMA was below the MRL in both airport samples and in one of the Las Vegas Strip samples during nonfestival weeks in 2023, and it was also below the MRL in all samples collected at a local university campus in 2025 [17]. Similarly, MDA (a major metabolite of MDMA) is rarely observed in local wastewater samples; for example, the only detection of MDA in the previous sewershed-level wastewater study in Southern Nevada was associated with an unrelated music festival in September 2022 [19]. Thus, the MDA detection frequencies in this study, primarily associated with the bars and festival samples, are notable.

Although stimulant-related analytes were most frequently detected, several opioid-related analytes (eg, tramadol, codeine, hydrocodone, oxycodone, norfentanyl, and morphine) were present in more than 50% of samples across the study years and locations. Some detection frequencies varied by location; for example, codeine, hydrocodone, and norfentanyl detection frequencies were lower at the festival site compared with off-site (2024) and the WWTP (2024 and 2025). Both festival and off-site samples yielded fewer detections of EDDP compared to the WWTP samples. The co-occurrence of opioids and stimulants may reflect a variety of substance use behaviors, preferences, polysubstance use, and/or unintentional adulteration (ie, lacing).

Analyte concentrations by year and sample location during the festival week and weekend are shown in Multimedia Appendix 2 (Figure S1), and concentrations are listed in Multimedia Appendix 1. Within each site, and despite the difference in sample type at the festival site (ie, grab in 2024 and composite in 2025), concentrations of each analyte were not significantly different between 2024 and 2025 (all Wilcoxon signed-rank $P>.05$). Notably, several stimulant-related (cocaine metabolites, amphetamine, methamphetamine, MDA, and MDMA) and THC-related (THC-COOH and THC-OH) concentrations observed in the festival site wastewater exceeded those in samples from the WWTP or off-site locations by more than an order of magnitude (Figure S1, Multimedia Appendix 2). In fact, stimulant-related analytes had some of the highest observed wastewater concentrations, with BZE reaching as high as 587,000 ng/L at the festival site. This was exceeded only by MDMA in festival site samples, which reached 1,280,000 ng/L in 2024 and 1,860,000 ng/L in 2025; these concentrations are among the highest festival-related wastewater MDMA concentrations reported in the literature.

XSL•FO
RenderX

Table . Detection frequency (%) across all samples (grab and composite) by year and location during the festival week and weekend[a].

Analyte	Festival week and weekend detection frequencies (% of N samples above MRL[b])						
	2023: sample locations; sample types		2024: sample locations; sample types			2025: sample locations; sample types	
	Off-site[c] (N=3); composite (n=3), %	WWTP[d] (N=1); grab (n=1), %	Off-site[c] (N=3); composite (n=3), %	Festival (N=4); grab (n=4), %	WWTP (N=5); composite (n=4) and grab (n=1), %	Festival (N=7); composite (n=6) and grab (n=1), %	WWTP (N=7); composite (n=6) and grab (n=1), %
Cocaine	100	100	100	100	100	100	100
BZE[e,f]	100	100	100	100	100	100	100
Ecgonine[e]	100	100	100	100	100	100	100
EME[e,g]	100	100	100	100	100	100	100
Norcocaine[e]	33	0	0	0	0	14	0
Codeine	100	100	100	75	100	57	100
Hydrocodone	100	100	100	75	100	29	0
Oxycodone	100	100	100	100	100	86	100
Fentanyl[h]	—[i]	—	—	—	—	0	14
Norfentanyl[e]	100	100	100	75	100	43	100
Heroin	0	0	0	0	0	0	0
Acetylmorphine[e]	67	100	67	50	0	0	0
MDMA[j]	100	100	100	100	100	86	100
MDA[k]	100	100	100	100	80	71	57
Methadone	0	100	33	0	0	14	0
EDDP[e,l]	33	100	67	25	100	43	100
Methamphetamine	100	100	100	100	100	100	100
Amphetamine	100	100	100	100	100	100	100
Morphine	100	100	100	100	100	86	100
THC[m]	0	0	0	0	0	0	0
THC-COOH[e,n]	100	100	100	100	100	100	100
THC-OH[e]	0	100	0	100	0	86	29
Tramadol	100	100	100	100	100	100	100
Xylazine[h]	—	—	—	—	—	0	0

[a] An analyte was "detected" if its concentration was above the MRL.

[b]MRL: method reporting limit.

[c]Off-site locations in 2023 - 2024 are airport, bars, and Las Vegas Strip. The 2023 nonfestival-weekend samples are excluded here.

[d]WWTP: wastewater treatment plant.

[e]These analytes are metabolites only.

[f]BZE: benzoylecgonine.

[g]EME: ecgonine methyl ester.

[h]Samples were analyzed for fentanyl and xylazine in 2025 only.

[i]Not available.

[j]MDMA: 3,4-methylenedioxymethamphetamine.

XSL•FO
RenderX

[k]MDA: 3,4-methylenedioxyamphetamine.

[l]EDDP: 2-ethylidene-1,5-dimethyl-3,3-diphenylpyrrolidine.

[m]THC: tetrahydrocannabinol.

[n]THC-COOH: 11-nor-9-carboxy-Δ-9-tetrahydrocannabinol.

2023 Pilot Year: Festival vs Nonfestival Weeks

Analyte concentrations in the 2023 samples are shown in Multimedia Appendix 2 (Figure S2) by sample location and collection event (festival weekend in May vs nonfestival weekends in June and July). Although no 2023 sampling occurred at the festival site, MDMA and MDA concentrations were notably higher in the festival weekend samples across all sites (bars, the Las Vegas Strip, the airport, and the WWTP). In fact, the airport-related detections of MDMA and MDA occurred only during the festival weekend, highlighting potential contributions from visitors traveling to Las Vegas for festival-related activities. MDMA concentrations at the bars and the Las Vegas Strip were approximately 10 and 30 times higher, respectively, during the festival compared with nonfestival sampling events, and nearly 100 times higher at the WWTP (treating relatively consistent community flows of ~110 mgd). Along with the approximate 10-fold change in wastewater MDMA concentration at the bars, the norfentanyl concentration was nearly 2 times higher during the festival week, highlighting potential risks of fentanyl use, adulteration, or contamination [20], and the need for increased awareness of these possibilities [21].

Concentrations of other substances (eg, methamphetamine, amphetamine, and THC metabolites) were similar or even higher on nonfestival weekends; a similar observation for methamphetamine was observed for a baseline (higher concentration) vs special event weekend (lower concentration) in an earlier study conducted in Las Vegas [14]. Regardless of the week, wastewater concentrations at the bars were consistently higher than those at the Las Vegas Strip, the airport, and the WWTP for cocaine (and multiple metabolites) and methamphetamine. Importantly, these results highlight the widespread consumption of HRS across Southern Nevada during festival and nonfestival weekends.

2024-2025: Day of Festival Week

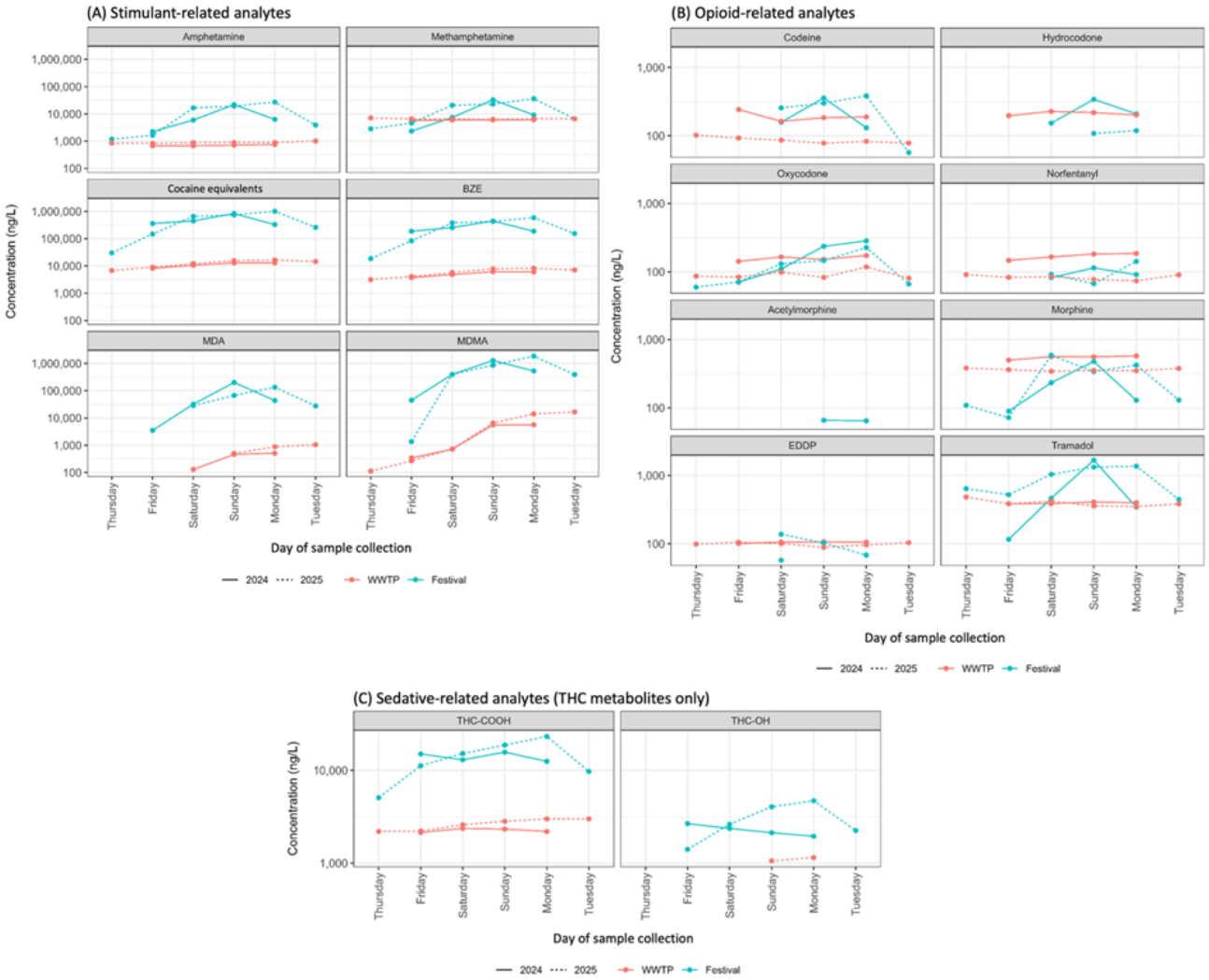
Analyte concentrations by year and day of wastewater sample collection (Friday-Monday in 2024; Thursday-Tuesday in 2025) and sample location (festival vs WWTP) are shown in Figure 2; analytes not detected at these sites in either year (heroin, THC, and xylazine) and analytes with singular detections (fentanyl, methadone, and norcocaine) are not shown. Gerrity et al [19] previously documented the instability of cocaine-related compounds in wastewater but demonstrated that the overall mass balance of cocaine and its metabolites was conserved over 60 days of hold time. Thus, the cocaine-related compounds are aggregated in Figure 2 as “cocaine equivalents” (after accounting for differences in molecular weight), with BZE also shown for comparison.

For the analytes detected at least twice, concentrations tended to peak in samples collected on either Sunday or Monday, surrounding the culmination of the festival. Increasing concentrations of stimulants (including amphetamine, methamphetamine, MDMA, MDA, BZE, and cocaine equivalents in Figure 2A) were particularly evident over the main festival days at both sites. In 2025, festival-site MDMA concentrations increased rapidly from below the MRL before the campground was opened to 1370 ng/L after the first full day of the festival, and ultimately peaked at 1,860,000 ng/L at the end of the weekend (in the Sunday-Monday sample); residual high concentrations persisted through the Tuesday sample collections.

Although each analyte’s concentration within the sampling sites did not differ significantly between 2024 and 2025, concentrations across both years were significantly different between the festival site and the WWTP for amphetamine, cocaine equivalents, BZE, ecgonine, MDMA, and THC-COOH (all Wilcoxon signed-rank P=.008). As expected due to significant dilution from municipal wastewater flows, stimulant concentrations were often lower at the WWTP relative to the festival site. That being said, BZE more than doubled, and MDMA increased more than 100-fold at the WWTP during the 2025 festival week. THC-COOH and THC-OH were also higher in the festival samples compared to the WWTP samples (Figure 2C). Although there were some day-to-day variations, opioid-related concentrations were generally more similar or even higher at the WWTP relative to the festival (Figure 2B), but some opioids (eg, oxycodone, morphine, and tramadol) exhibited increasing trends at the festival site over the first few days of the festival.

XSL•FO
RenderX

Figure 2. Concentrations in wastewater from the festival site (blue) and WWTP (red) by day of collection in 2024 (solid line) and 2025 (dashed line) for (A) stimulants, (B) opioids, and (C) sedatives (THC-COOH and THC-OH only). Note: 2024 festival samples are grab samples; others are composite. Analytes not detected in either year (heroin, THC, and xylazine) and analytes with singular detections (fentanyl, methadone, and norcocaine) are not shown. BZE: benzoylecgonine; EDDP: 2-ethylidene-1,5-dimethyl-3,3-diphenylpyrrolidine; MDA: 3,4-methylenedioxyamphetamine; MDMA: 3,4-methylenedioxy-methamphetamine; THC: tetrahydrocannabinol; THC-COOH: 11-nor-9-carboxy-Δ-9-tetrahydrocannabinol; WWTP: wastewater treatment plant.

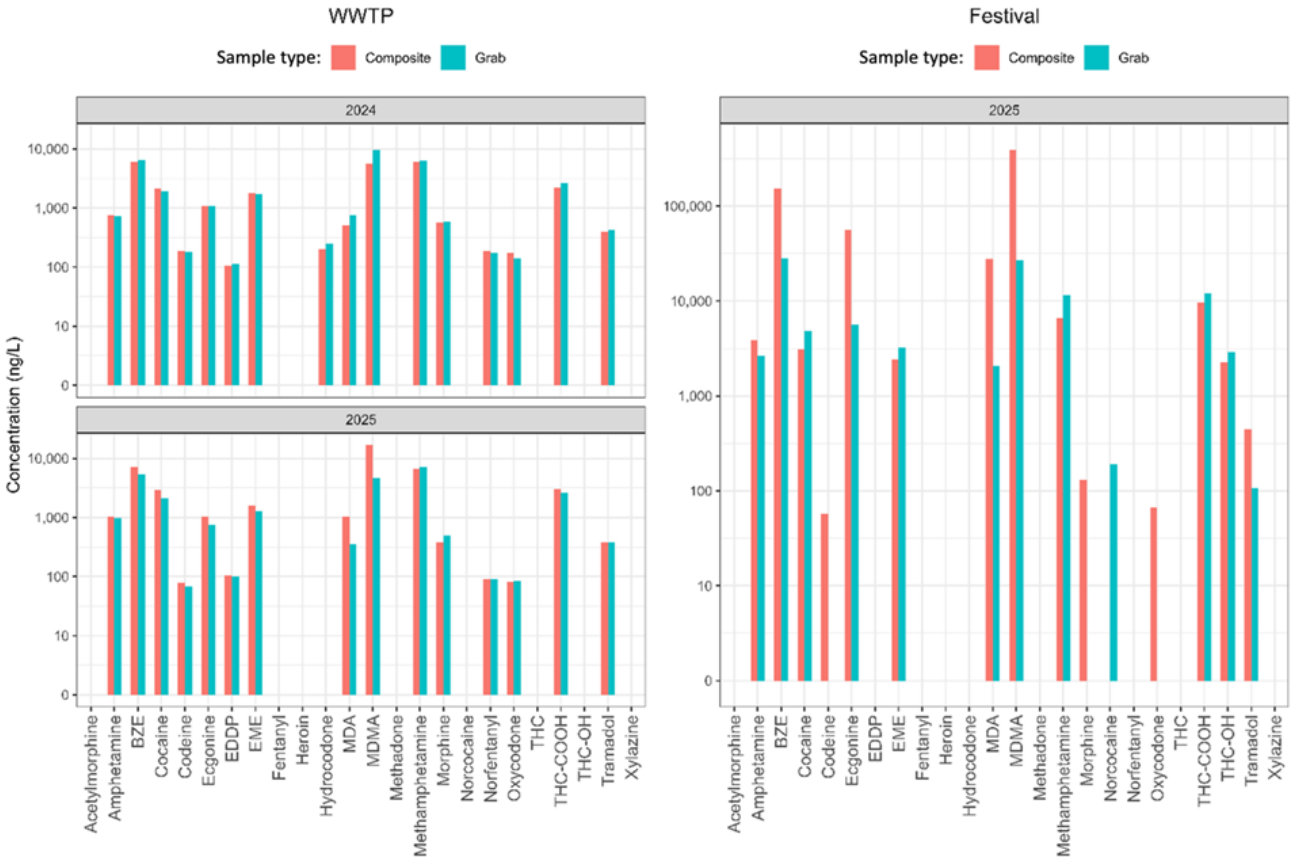


2024-2025: Grab vs Composite Sampling

As shown in Figure 3, concentrations were similar between grab and composite samples collected at the WWTP in 2024. In 2025, WWTP composite sample concentrations were higher than the corresponding grab samples for the analytes expected to be impacted by the festival, including some cocaine-related compounds, MDA, and MDMA. These differences between 2024 (Monday collections) and 2025 (Tuesday collections) may have been driven by the day of the week, with the Tuesday grab likely to be less affected by the festival weekend. Similarly, the 2025 festival composite sample (Monday-Tuesday) yielded higher concentrations of BZE (5.5-fold), ecgonine (9.9-fold), MDA (13.4-fold), MDMA (14.6-fold), and tramadol (4.2-fold) compared to the corresponding Tuesday grab sample. Notably, codeine, morphine, and oxycodone were detected in the composite but not in the corresponding grab sample, while only norcocaine was detected in the grab sample and not in the corresponding composite sample.

XSL•FO
RenderX

Figure 3. Comparison of analyte concentrations in corresponding composite (red) and grab (blue) samples for the WWTP (2024 and 2025; left) and the festival site (2025 only; right). BZE: benzoylecgonine; EDDP: 2-ethylidene-1,5-dimethyl-3,3-diphenylpyrrolidine; EME: ecgonine methyl ester; MDA: 3,4-methylenedioxyamphetamine; MDMA: 3,4-methylenedioxy-methamphetamine; THC: tetrahydrocannabinol; WWTP: wastewater treatment plant.



Discussion

Principal Findings

This multiyear surveillance study analyzed wastewater from multiple locations to characterize the presence and concentration of substances (and their metabolites) surrounding an annual EDM festival. The findings in 2023 reaffirmed the presence of HRS in raw community wastewater during festival and nonfestival time points, with notable differences by weekends and sample locations. The findings from 2024 to 2025 further illustrated the prevalence of stimulants (cocaine, methamphetamine, and notably MDMA) in festival-site wastewater, although opioid-related analytes were also detected there and in the community. Detections related to THC were not unexpected, given the legality of recreational marijuana products in the state of Nevada and prior literature indicating its frequency in festival-related WBE studies [22].

The study festival (with more than 520,000 annual attendees) is large in comparison to other events included in prior wastewater-based studies of substance use. The 2010 to 2011 Australian festival described by Lai et al [6], for example, had daily attendees between 12,100 and 19,700; Bijlsma et al [9] sampled from 6 European music festivals with a total combined attendance of 465,000. A recent systematic review [22] included 23 WBE studies on HRS, with 6200 to 600,000 attendees; only 2 studies reported attendance in excess of 500,000 people.

Although this study included just 3 sequential years, the results offer a preliminary wastewater-based profile of substances and metabolites present at the festival site, namely stimulants (eg, MDMA, cocaine, and methamphetamine), as well as select opioids. Lai et al [6] similarly noted consistency in festival attendance and wastewater-based HRS profiles each year, in line with findings that event-related WBE findings may vary by attendee characteristics and time of year [23]. Elsewhere, researchers have documented wastewater-based concentrations rising to a peak during the event [23], a pattern also observed here on the final days of the festival and potentially reflecting accumulation and/or intentional discarding at the end of the event [8]. Multiple WBE studies have highlighted wastewater-based profiles of substance use surrounding holidays [23-25], during weekends compared to weekdays [24], and among nightclub, music event, and festival attendees in general [26-28], with some variation by region and music genre [22]. Several results in this study indicated concurrent detections of select analytes (eg, MDMA and norfentanyl at the bars), highlighting another potential opportunity for multifaceted public health responses to address intentional (or unintentional) polydrug use [29] amid the recent wave of polysubstance use and related deaths in the US overdose crisis [30].

MDMA and Other Stimulants

Across analytical approaches, geographic locations, and event types, MDMA was the most consistently detected substance in the Cainamisir et al [22] review of 23 wastewater studies of

XSL•FO
RenderX

psychoactive substances at music festivals. The maximum MDMA concentrations in this study (1,280,000 ng/L and 1,860,000 ng/L), however, exceeded concentrations reported previously [22], including those from a neighboring Southern Nevada sewershed during a separate music festival in the fall of 2022 (1100 ng/L) [19] and from a festival in Thailand (1267 ng/L) [31]. An analysis of pooled wastewater from a festival in Belgium, however, reported a higher concentration (7580 ng/mL=7,580,000 ng/L) [32]. The accumulation of MDMA over a festival weekend (with a blood half-life of 8 h; 95% clearance within 40 h [33]) may help explain the concentrations in this study. Importantly, prior work has observed that MDMA concentrations tend to be higher on weekends, even in the absence of a large event or festival [34,35]. Recent wastewater findings in Europe also emphasized higher MDMA loads during weekends compared to weekdays (occurring in more than 75% of cities studied), though overall MDMA wastewater loads in European cities decreased by nearly 16% from 2024 to 2025 [36].

Methamphetamine and cocaine (and its metabolites) were consistently present over the festival weekends (with 100% detection frequencies), but the 2023 nonfestival weekend concentrations suggest community use independent of the event. A separate local wastewater surveillance study conducted in 2022 to 2023 [19] similarly concluded that methamphetamine and cocaine were among the highest per capita consumption estimates across all Las Vegas sewersheds. Wastewater detections of cocaine (the parent compound) have been reported in similar festival-based studies [22], albeit with slightly lower concentrations (eg, a maximum of 8914 ng/L in Brazil [8] vs 10,000 ng/L in this study). Here, we used a mass-balance approach to visualize total cocaine equivalents, but focusing on one of the more abundant cocaine metabolites, specifically BZE, may be sufficient for public health and event-related wastewater monitoring of cocaine use (Figure 2). Cainamisir et al [22] highlighted geographic variability in wastewater-based indicators of methamphetamine presence and consumption at festivals elsewhere, though the methamphetamine detections in this study align with the established presence of the substance in Southern Nevada [37].

Opioid HRS Targets

Although stimulants were detected most frequently in this study, several opioids were also present during the festival week and weekend, both on-site and in the broader community. Norfentanyl, for example, was detected in both festival and WWTP wastewater, though more consistently at the WWTP; this aligns with the prior local findings [19] and could reflect some medical use. Other festival-related WBE studies have reported detections of norfentanyl, though it was not the most common analyte; in one study, norfentanyl was detected at only 2 of 6 festivals sampled [7]. Generally, festival-related opioid concentrations in this study mirrored those at the downstream WWTP, suggesting a less-pronounced event-associated use compared with the stimulants discussed earlier. However, the co-occurrence of some opioids and stimulants in site-specific wastewater samples may reflect intentional or unintentional polysubstance use (ie, through lacing) [29], although this cannot be confirmed with wastewater alone.

The absence of heroin (the parent compound) detections was not unexpected, given previous work suggesting its potential instability in the presence of the quenching agent, ascorbic acid, used here [19]. Heroin and its metabolite acetylmorphine are known to be subject to rapid degradation in wastewater [38], though acetylmorphine has been detected during previous hold-time experiments [19], in other local studies [17], and here in 2023 (n=4 off-site samples and 2 WWTP samples) and in 2024 (n=2 off-site samples and 2 festival-site samples).

Limitations

Several elements of this study may limit and/or confound the interpretation of results and their broader contextualization in the fields of WBE and public health. First, due to practical and logistical limitations, the wastewater sampling framework varied each year of the study, complicating year-over-year comparisons. Multiday sampling in 2024 and 2025 offered more granular detail about the festival week and weekend, and the limited comparisons between grab and composite samples highlighted the potential effect of sample type on analyte detections (eg, codeine, morphine, oxycodone, and norcocaine) and on concentrations (eg, MDMA). Generally, composite samples are preferred over grab samples in the WBE literature, given concerns about potentially nonrepresentative slugs of wastewater that may underestimate or overestimate the presence of a target analyte in wastewater [39]. However, the results of this study show that even limited grab sampling offers valuable information about HRS presence and use surrounding an event.

Analyte source, consumption, metabolism, excretion, stability, and dilution in wastewater are also key factors that affect interpretation of wastewater results. Others have addressed metabolism with correction factors to back-calculate consumption estimates [19,40]. This would be a valuable addition to future festival-specific WBE studies for HRS, although corresponding flow measurements would be needed to calculate mass loadings. Wastewater analyses also do not distinguish between the various individuals present at a festival, potentially including attendees, artists, vendors, staff, and volunteers. Supplemental analyses to distinguish between licit and illicit consumption of select substances (eg, methamphetamine) have been established [41], but were not performed here. This means, for example, that a portion of the amphetamine and methamphetamine concentrations observed here may be due to medically prescribed use, rather than illicit use. Furthermore, these wastewater detections may partially reflect intentional discarding (ie, dumping) of substances in festival restrooms and subsequent in-sewer degradation during and/or at the end of the event, as has been noted previously [8]. Although wastewater monitoring does not identify specific sources (eg, original user or lacing vs polysubstance use vs intentional discarding), these findings still complement existing public health surveillance activities and indicate the presence of these substances on-site and in the community surrounding the event.

Results were not normalized to wastewater flow rate, a fecal indicator, or another analyte (eg, caffeine or sucralose). As described earlier, normalization would have facilitated consumption calculations and/or captured differences in human

XSL•FO
RenderX

urine or fecal load within these samples [19]. Site-specific consumption estimates could be calculated with site-level flow data and corrections for metabolism or excretion or molecular weight equivalence [19,40]. Importantly, upstream flows may be highly variable throughout the day and across the event.

This study included a selection of conventional prescription and recreational substances of public health significance, but a variety of other targets could be included in future work. Several published event-focused studies have used wastewater to characterize the presence of less-regulated novel psychoactive substances (NPS) [7,9]. Wastewater may be particularly valuable in tracking the emergence of new substance use trends, including NPS and misuse of existing substances (eg, xylazine), in tandem with other sources such as online discussion forums [42]. Future wastewater-based efforts might also include medetomidine, naloxone (despite potential instability in wastewater [43]), and ketamine, which may be increasingly common [26] and co-occurring with fentanyl [44]. Others have previously characterized the community wastewater-based occurrence of ketamine in Europe [24]; ketamine was also found in the Las Vegas Wash during the 2022 festival weekend [11].

At least one similar WBE study endeavored to relate wastewater HRS results to clinical outcomes or other datasets, although these analyses were beyond the scope of this study. This previous festival-based work measured wastewater HRS concentrations along with self-reported substance use behaviors [7], but links between festival HRS wastewater data and health outcomes associated with substance use (eg, overdoses, health care encounters, and deaths) are less clear in the literature. This does not, however, diminish the serious health outcomes that have been reported among festivalgoers [4,45], including critical care admissions at a tertiary care hospital in Las Vegas associated with the festival weekend [46,47]. In a 2023 case series of intensive care unit (ICU) admissions, substances reportedly used included MDMA (6/6 patients), cocaine (n=2), fentanyl (n=1), lysergic acid diethylamide (LSD; n=1), THC (n=1), and/or alcohol [1,46]. A recent local retrospective cohort study also found that drug or alcohol overdose patients related to the festival in this study were younger and more likely to be admitted to the ICU than nonfestival-related drug or alcohol overdose patients, although the 2 groups did not differ significantly in other characteristics, drug use patterns, or ultimate health outcomes [47]. In 2025, the Clark County coroner's office reported that MDMA contributed to 2 deaths in the Las Vegas area during the festival weekend [48].

Implications for Public Health Response

These results offer valuable insights for public health and clinical situational awareness [22] of substance use during the festival week or weekend, complementing other surveillance activities, emergency responses, and the ultimate reporting of health care use and serious clinical outcomes. Selected results, such as the relatively few heroin-related detections in this study, may reflect an evolving landscape of substance use behaviors among attendees. Palamar et al [26] described changes in substances of choice in New York City nightlife and dance club settings from 2017 to 2022, with survey responses indicating decreased use of heroin, nonmedical prescription opioids, and cocaine (among other stimulants) in the post–COVID-19 era, accompanied by increased use of ketamine, psilocybin, poppers, and other NPS. Together with ongoing surveys, online forum monitoring [42], drug checking, and traditional public health or clinical surveillance, event-based wastewater surveillance may inform continued investment in prevention and future responses or interventions, both on-site and in the community, to support successful events while prioritizing attendee and community health and safety. Text S2 in Multimedia Appendix 2 provides additional discussion and recommendations on how event-based wastewater surveillance can inform on-site and community-level public health activities, including activities that have been implemented specifically for the music festival in question.

Implications for Utilities

Event-related wastewater monitoring also offers insight into expected changes in the composition of influent wastewater flows surrounding the event. Although in-sewer mixing, dilution, and degradation likely explain the lower analyte concentrations in the WWTP samples compared with the festival site, the WWTP still experienced significant increases (10- to 100-fold) in influent concentrations of select analytes (eg, MDMA and MDA) over the event weekend. In addition to changes in flow driven by large influxes of people associated with special events, these compositional changes have implications for wastewater treatment processes, water reuse, and downstream environmental surface water quality [11,22,49], highlighting how wastewater monitoring can be informative and mutually beneficial for utilities and public health entities.

Conclusions

Event-related wastewater monitoring is an evolving subtopic in the field of WBE. Several studies have already explored targeted wastewater monitoring for pathogens such as SARS-CoV-2 at the Olympics [50] and other mass gatherings [15], but WBE has also been applied to substance use at music festivals in countries like Australia and Thailand. In Las Vegas, this approach has now been extended to characterize the use of HRS at the community level and in treated wastewater effluent surrounding events and festivals, as has been done elsewhere [11,14,17,19]. Here, wastewater surveillance surrounding a large EDM festival in 3 sequential years at on-site and off-site locations further explored WBE as a complementary tool for public health monitoring of substance use or misuse. Stimulant concentrations were notably elevated during event weekends, with high peak concentrations of MDMA and cocaine-related analytes, but opioid-related detections were also observed. These findings may inform both public health surveillance and response activities and wastewater utility operations.

Wastewater monitoring should complement, not replace, existing public health surveillance and response structures. Multifaceted public health surveillance systems can integrate wastewater surveillance for both pathogens and HRS to expand event-related preparedness, including interventions to prioritize the health and safety of attendees on-site and in the community. Future research may explore additional HRS targets, other non-EDM festivals and events, trends over time, and alignment with other data sources (including health outcomes). Local public health

XSL•FO
RenderX

agencies, community stakeholders (eg, health care systems, nonprofit organizations, and local government), and other tourism and event destinations similarly stand to benefit from data-driven understanding of the public health implications of events and mass gatherings.

Acknowledgments

We thank the Clark County Water Reclamation District (CCWRD) for its contributions to this study. Specifically, we thank Carlton Utter, Adrian Romero, Alycia Ybarra, and LaToya Blanche of CCWRD for helping collect representative samples. The authors attest that there was no use of generative AI technology in the generation of text, figures, or other informational content of the manuscript.

Funding

This project was supported in part by the Fund for a Resilient Nevada (FRN), which is administered by the Nevada Department of Human Services (DHS) to address the impacts, risks, and harms of opioid use in the State of Nevada. The funder had no involvement in the study design, data collection, analysis, interpretation, or writing of the manuscript. These findings are solely the responsibility of the authors and do not necessarily represent the official views of FRN or Nevada DHS.

Data Availability

Study data are included in Multimedia Appendix 1.

Authors' Contributions

Conceptualization: BJV, DG, ECO
Data curation: CAB, RAT, DG
Formal analysis: CAB, RAT, DG
Funding acquisition: ECO, DG
Investigation: CAB, RAT, CLC, MAM, DG
Methodology: RAT, BJV, DG
Project administration: ECO, DG
Resources: RAT, BJV, ECO
Supervision: BJV, DG
Validation: RAT, BJV
Visualization: CAB, KCC, DG
Writing – original draft: CAB
Writing – review & editing: RAT, BJV, KCC, CLC, MAM, ECO, ATM, CL, DG

Conflicts of Interest

None declared.

Multimedia Appendix 1
Wastewater concentrations of analytes and corresponding sample information.
[XLSX File, 54 KB - publichealth_v12i1e102679_app1.xlsx]

Multimedia Appendix 2
Supplemental analytical methods, compound and mass spectrometry parameters, wastewater concentration data, and additional discussion of on-site and community-level public health activities.
[PDF File, 779 KB - publichealth_v12i1e102679_app2.pdf]

References

1. Jenkinson R, Bowring A, Dietze P, Hellard M, Lim MSC. Young risk takers: alcohol, illicit drugs, and sexual practices among a sample of music festival attendees. J Sex Transm Dis 2014;2014(1):357239. [doi: 10.1155/2014/357239] [Medline: 26316974]
2. Day N, Criss J, Griffiths B, et al. Music festival attendees' illicit drug use, knowledge and practices regarding drug content and purity: a cross-sectional survey. Harm Reduct J 2018 Jan 5;15(1):1. [doi: 10.1186/s12954-017-0205-7] [Medline: 29304871]
3. Douglass CH, Raggatt M, Wright CJC, et al. Alcohol consumption and illicit drug use among young music festival attendees in Australia. Drugs Educ Prev Policy 2022 Mar 4;29(2):175-181. [doi: 10.1080/09687637.2021.1889978]

XSL•FO
RenderX

4. Palamar JJ, Sönmez İ. A qualitative investigation exploring why dance festivals are risky environments for drug use and potential adverse outcomes. Harm Reduct J 2022 Feb 5;19(1):12. [doi: 10.1186/s12954-022-00598-5] [Medline: 35120530]
5. Palamar JJ, Fitzgerald ND, Keyes KM, Cottler LB. Drug checking at dance festivals: a review with recommendations to increase generalizability of findings. Exp Clin Psychopharmacol 2021 Jun;29(3):229-235. [doi: 10.1037/pha0000452] [Medline: 33600199]
6. Lai FY, Thai PK, O'Brien J, et al. Using quantitative wastewater analysis to measure daily usage of conventional and emerging illicit drugs at an annual music festival. Drug Alcohol Rev 2013 Nov;32(6):594-602. [doi: 10.1111/dar.12061] [Medline: 23782033]
7. Brett J, Siefried KJ, Healey A, et al. Wastewater analysis for psychoactive substances at music festivals across New South Wales, Australia in 2019-2020. Clin Toxicol (Phila) 2022 Apr;60(4):440-445. [doi: 10.1080/15563650.2021.1979233] [Medline: 34542387]
8. Gomes CS, de Jesus Soares Freire D, de Souza Ramos Pontes Moura H, et al. Wastewater surveillance to assess cocaine and methylenedioxymethamphetamine use trends during a major music festival in Brazil. Drug Test Anal 2025 Jan;17(1):88-100. [doi: 10.1002/dta.3682] [Medline: 38544438]
9. Bijlsma L, Celma A, Castiglioni S, et al. Monitoring psychoactive substance use at six European festivals through wastewater and pooled urine analysis. Sci Total Environ 2020 Jul 10;725:138376. [doi: 10.1016/j.scitotenv.2020.138376] [Medline: 32298891]
10. Under the electric sky: the history of EDC Las Vegas. The Las Vegas Sun. 2025 May 8. URL: https://lasvegassun.com/news/2025/may/08/under-the-electric-sky-the-history-of-edc-las-vega/ [accessed 2026-08-22]
11. Sims DB, Monk JR, Garner MC, et al. Occurrence, source, and ecological risk of pharmaceuticals and illicit drugs in surface water in Las Vegas, USA. Int J Environ Sci Technol 2024 Aug;21(12):7865-7880. [doi: 10.1007/s13762-024-05522-x]
12. Achermann S, Falås P, Joss A, et al. Trends in micropollutant biotransformation along a solids retention time gradient. Environ Sci Technol 2018 Oct 16;52(20):11601-11611. [doi: 10.1021/acs.est.8b02763] [Medline: 30208701]
13. Gerrity D, Neyestani M. Impacts of solids retention time and antibiotic loading in activated sludge systems on secondary effluent water quality and microbial community structure. Water Environ Res 2019 Jun;91(6):546-560. [doi: 10.1002/wer.1060] [Medline: 30693624]
14. Gerrity D, Trenholm RA, Snyder SA. Temporal variability of pharmaceuticals and illicit drugs in wastewater and the effects of a major sporting event. Water Res 2011 Nov 1;45(17):5399-5411. [doi: 10.1016/j.watres.2011.07.020] [Medline: 21920575]
15. Brighton K, Fisch S, Wu H, Vigil K, Aw TG. Targeted community wastewater surveillance for SARS-CoV-2 and Mpox virus during a festival mass-gathering event. Sci Total Environ 2024 Jan 1;906:167443. [doi: 10.1016/j.scitotenv.2023.167443] [Medline: 37793442]
16. Barber CA, Chang CL, Moshi MA, et al. Operationalizing SARS-CoV-2 wastewater monitoring to assess traveler health in Las Vegas, Nevada, USA. IJID Reg 2025 Jun;15:100619. [doi: 10.1016/j.ijregi.2025.100619] [Medline: 40176854]
17. Gerrity D, Barber CA, Trenholm RA, et al. Piloting wastewater-based monitoring on a university campus to inform public health surveillance and response for opioids and other high-risk substances. J Water Health 2026 Jan;24(1):65-80. [doi: 10.2166/wh.2026.182] [Medline: 41627797]
18. Gerrity D, Papp K, Dickenson E, et al. Characterizing the chemical and microbial fingerprint of unsheltered homelessness in an urban watershed. Sci Total Environ 2022 Sep 20;840:156714. [doi: 10.1016/j.scitotenv.2022.156714] [Medline: 35709998]
19. Gerrity D, Crank K, Oh EC, Quinones O, Trenholm RA, Vanderford BJ. Wastewater surveillance of high risk substances in Southern Nevada: sucralose normalization to translate data for potential public health action. Sci Total Environ 2024 Jan 15;908:168369. [doi: 10.1016/j.scitotenv.2023.168369] [Medline: 37951274]
20. Cruz SL, Bencomo-Cruz M, Medina-Mora ME, Vázquez-Quiroz F, Fleiz-Bautista C. First drug-checking study at an electronic festival and fentanyl detection in the central region of Mexico. Harm Reduct J 2023 Dec 6;20(1):174. [doi: 10.1186/s12954-023-00905-8] [Medline: 38053148]
21. Palamar JJ, Fitzgerald ND, Cottler LB. Shifting awareness among electronic dance music party attendees that drugs may contain fentanyl or other adulterants. Int J Drug Policy 2021 Nov;97:103353. [doi: 10.1016/j.drugpo.2021.103353] [Medline: 34242960]
22. Cainamisir R, Zeng X, Himmerich SB, Himmerich H. Wastewater analyses for psychoactive substances at music festivals: a systematic review. Behav Sci (Basel) 2025 Dec 3;15(12):1672. [doi: 10.3390/bs15121672] [Medline: 41464015]
23. Rousis N, Bade R, Romero-Sánchez I, et al. Festivals following the easing of COVID-19 restrictions: prevalence of new psychoactive substances and illicit drugs. Environ Int 2023 Aug;178:108075. [doi: 10.1016/j.envint.2023.108075] [Medline: 37399770]
24. Van Wichelen N, Boogaerts T, Quireyns M, et al. Ketamine, a new (or old) kid on the block: a comprehensive three-year spatio-temporal study in Belgium through wastewater-based epidemiology. Water Res 2025 May 15;276:123269. [doi: 10.1016/j.watres.2025.123269] [Medline: 39954459]
25. Bade R, Rousis N, Adhikari S, et al. Three years of wastewater surveillance for new psychoactive substances from 16 countries. Water Res X 2023 May 1;19:100179. [doi: 10.1016/j.wroa.2023.100179] [Medline: 37143710]

XSL•FO
RenderX

26. Palamar JJ, Le A, Cleland CM, Keyes KM. Trends in drug use among nightclub and festival attendees in New York City, 2017-2022. Int J Drug Policy 2023 May;115:104001. [doi: 10.1016/j.drugpo.2023.104001] [Medline: 36934660]
27. Palamar JJ, Abukahok N, Acosta P, Walton SE, Stang B, Krotulski AJ. Exposures to synthetic cathinones, fentanyl, and xylazine among nightclub attendees in New York City, 2024. Drug Alcohol Depend 2025 Oct 1;275:112792. [doi: 10.1016/j.drugalcdep.2025.112792] [Medline: 40706442]
28. Fernández-Calderón F, Cleland CM, Palamar JJ. Polysubstance use profiles among electronic dance music party attendees in New York City and their relation to use of new psychoactive substances. Addict Behav 2018 Mar;78:85-93. [doi: 10.1016/j.addbeh.2017.11.004] [Medline: 29128711]
29. Ciccarone D. The Donovan Memorial Lecture: 'Fentanyl-Plus': a new era of fentanyl polydrug combinations in the North American overdose crisis. J Med Toxicol 2026 Jan;22(1):49-62. [doi: 10.1007/s13181-025-01104-6] [Medline: 41214415]
30. Ciccarone D. The rise of illicit fentanyls, stimulants and the fourth wave of the opioid overdose crisis. Curr Opin Psychiatry 2021 Jul 1;34(4):344-350. [doi: 10.1097/YCO.0000000000000717] [Medline: 33965972]
31. Jiang JJ, Lee CL, Fang MD, Tu BW, Liang YJ. Impacts of emerging contaminants on surrounding aquatic environment from a youth festival. Environ Sci Technol 2015 Jan 20;49(2):792-799. [doi: 10.1021/es503944e] [Medline: 25495157]
32. Geuens M, Van Hoofstadt K, Hoogmartens O, Van den Eede N, Sabbe M. Pooled urine analysis at a Belgian music festival: trends in alcohol consumption and recreational drug use. Prehosp Disaster Med 2022 Dec;37(6):806-809. [doi: 10.1017/S1049023X22001285] [Medline: 36148611]
33. Kalant H. The pharmacology and toxicology of “ecstasy” (MDMA) and related drugs. CMAJ 2001 Oct 2;165(7):917-928. [Medline: 11599334]
34. Gushgari AJ, Driver EM, Steele JC, Halden RU. Tracking narcotics consumption at a Southwestern U.S. university campus by wastewater-based epidemiology. J Hazard Mater 2018 Oct 5;359:437-444. [doi: 10.1016/j.jhazmat.2018.07.073] [Medline: 30059885]
35. González-Mariño I, Zuccato E, Santos MM, Castiglioni S. Monitoring MDMA metabolites in urban wastewater as novel biomarkers of consumption. Water Res 2017 May 15;115:1-8. [doi: 10.1016/j.watres.2017.01.063] [Medline: 28254532]
36. Wastewater analysis and drugs—a European multi-city study. : European Union Drugs Agency (EUDA); 2026 Mar 18 URL: https://www.euda.europa.eu/sites/default/files/pdf/2757_en.pdf?576293 [accessed 2026-04-29]
37. 2025 Southern Nevada community health assessment report. : Southern Nevada Health District (SNHD); 2025 URL: https://www.healthysouthernnevada.org/content/sites/snhd/2025_CHA/CHA_Full_Report-Binder1_FINAL.pdf [accessed 2025-05-19]
38. Verovšek T, Krizman-Matasic I, Heath D, Heath E. Investigation of drugs of abuse in educational institutions using wastewater analysis. Sci Total Environ 2021 Dec 10;799:150013. [doi: 10.1016/j.scitotenv.2021.150013] [Medline: 34602307]
39. Gerrity D, Papp K, Pecson BM. Pathogen peak “averaging” in potable reuse systems: lessons learned from wastewater surveillance of SARS-CoV-2. ACS ES T Water 2022 Nov 11;2(11):1863-1870. [doi: 10.1021/acsestwater.1c00378] [Medline: 37566355]
40. Puljević C, Tscharke B, Wessel EL, et al. Characterising differences between self-reported and wastewater-identified drug use at two consecutive years of an Australian music festival. Sci Total Environ 2024 Apr 15;921:170934. [doi: 10.1016/j.scitotenv.2024.170934] [Medline: 38360330]
41. Verovšek T, Heath D, Heath E. Enantiomeric profiling of amphetamines in wastewater using chiral derivatisation with gas chromatographic-tandem mass spectrometric detection. Sci Total Environ 2022 Aug 20;835:155594. [doi: 10.1016/j.scitotenv.2022.155594] [Medline: 35490814]
42. Online monitoring. National Drug Early Warning System (NDEWS). 2025. URL: https://ndews.org/novel-surveillance/web-surveillance/ [accessed 2025-06-11]
43. Ahmed F, Li J, O’Brien JW, et al. In-sewer stability of selected analgesics and their metabolites. Water Res 2021 Oct 1;204:117647. [doi: 10.1016/j.watres.2021.117647] [Medline: 34536687]
44. NDEWS Weekly Briefing Issue 239—This Week’s Focus: Ketamine & Fentanyl. National Drug Early Warning System (NDEWS). 2025. URL: https://ndews.org/newsletter/ndews-weekly-briefing-issue-239/ [accessed 2025-07-22]
45. Polkinghorne BG, Massey PD, Durrheim DN, Byrnes T, MacIntyre CR. Prevention and surveillance of public health risks during extended mass gatherings in rural areas: the experience of the Tamworth Country Music Festival, Australia. Public Health 2013 Jan;127(1):32-38. [doi: 10.1016/j.puhe.2012.09.014] [Medline: 23141111]
46. Chen C, Shih J, Sutton J, Cabugao P, Romero A. 116: Electric Daisy Carnival and associated complication rates in ICU patients: a case series. Crit Care Med 2024;52(1):S33-S33. [doi: 10.1097/01.ccm.0000998656.49825.c6]
47. Ren J, Batra K, Goel V, Jensen T, Tran S, Kioka M. Critical care utilization associated with an electronic dance music festival in Las Vegas, Nevada: a retrospective cohort study of the Electric Daisy Carnival. Prev Med Rep 2026 Apr;64:103439. [doi: 10.1016/j.pmedr.2026.103439] [Medline: 41852618]
48. Torres-Cortez R. MDMA contributed to two EDC weekend deaths in Las Vegas, coroner says. Las Vegas Review-Journal. 2025 Aug 25. URL: https://www.reviewjournal.com/news/politics-and-government/clark-county/mdma-contributed-to-two-edc-weekend-deaths-in-las-vegas-coroner-says-3429887/ [accessed 2026-04-27]

XSL•FO
RenderX

49. Aberg D, Chaplin D, Freeman C, Paizs B, Dunn C. The environmental release and ecosystem risks of illicit drugs during Glastonbury Festival. Environ Res 2022 Mar;204(Pt B):112061. [doi: 10.1016/j.envres.2021.112061] [Medline: 34543637]
50. Kitajima M, Murakami M, Iwamoto R, Katayama H, Imoto S. COVID-19 wastewater surveillance implemented in the Tokyo 2020 Olympic and Paralympic Village. J Travel Med 2022 May 31;29(3):taac004. [doi: 10.1093/jtm/taac004] [Medline: 35134222]

Abbreviations

BZE: benzoylecgonine
EDDP: 2-ethylidene-1,5-dimethyl-3,3-diphenylpyrrolidine
EDM: electronic dance music
EME: ecgonine methyl ester
ESI: electrospray ionization
HRS: high-risk substances
ICU: intensive care unit
LC-MS/MS: liquid chromatography–tandem mass spectrometry
LSD: lysergic acid diethylamide
MDA: 3,4-methylenedioxyamphetamine
MDMA: 3,4-methylenedioxy-methamphetamine
MGD: million gallons per day
MRL: method reporting limit
MRM: multiple reaction monitoring
NPS: novel psychoactive substances
THC: Δ-9-tetrahydrocannabinol
THC-COOH: carboxy-Δ-9-tetrahydrocannabinol
WBE: wastewater-based epidemiology
WWTP: wastewater treatment plant

Edited by A Mavragani, T Sanchez; submitted 27.May.2026; peer-reviewed by S Fu, T Nefau; revised version received 10.Aug.2026; accepted 11.Aug.2026; published 08.Sep.2026.

Please cite as:
Gerrity D, Barber CA, Trenholm RA, Vanderford BJ, Crank KC, Chang CL, Moshi MA, Mangla AT, Lockett C, Oh EC
Multiyear Wastewater Monitoring of High-Risk Substances Surrounding an Annual Electronic Dance Music Festival, Las Vegas, Nevada, USA, 2023-2025: Surveillance Study
JMIR Public Health Surveill 2026;12:e102679
URL: https://publichealth.jmir.org/2026/1/e102679
doi:10.2196/102679

© Daniel Gerrity, Casey A Barber, Rebecca A Trenholm, Brett J Vanderford, Katherine C Crank, Ching-Lan Chang, Michael A Moshi, Anil T Mangla, Cassius Lockett, Edwin C Oh. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 8.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Performance of an Electronic Sentinel Surveillance System for Bacterial Gastroenteritis: Mixed Methods Evaluation Study

Inmyung Song[1], PhD; Yerin Heo[1], BSc; Minsol Jo[1], MPH; Jung-yeon Park[1], MPH; Sukhyun Ryu[2], MD, PhD; Dong-Sook Kim[1], PhD

[1]College of Nursing and Health, Kongju National University, Gongjudaehakro 56, Gongju, Chungcheongnam-do, Republic of Korea

[2]Department of Preventive Medicine, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea

Corresponding Author:
Dong-Sook Kim, PhD
College of Nursing and Health, Kongju National University, Gongjudaehakro 56, Gongju, Chungcheongnam-do, Republic of Korea

Abstract

Background: Bacterial gastroenteritis remains a substantial public health burden. South Korea has operated an electronic sentinel surveillance system (SSS) since 2000; in 2023, a total of 211 sentinel hospitals conducted surveillance for 11 notifiable types of bacterial gastroenteritis. The system has not undergone a comprehensive evaluation.

Objective: This study evaluated the SSS's operational performance and reporting completeness, assessed the population-based geographic allocation of sentinel sites, and explored concordance between SSS reports and National Health Insurance (NHI) claim records.

Methods: We conducted a mixed methods evaluation using a 2024 survey of staff at the 211 sentinel hospitals, an interview with the Korea Disease Control and Prevention Agency (KDCA) operator, KDCA operational reporting records, nationwide NHI claim records, and regional population statistics. The primary evaluation year was 2023; 2021 was a nonconsecutive descriptive comparator because these were the 2 complete annual extracts available in the analytic dataset. We calculated delayed reporting, submission completeness, and on-time completeness. Exploratory Pearson correlations compared paired weekly all-pathogen totals (52 weeks per year), annual surveillance target, and diagnosis code category counts within each age group (11 paired categories), and actual vs population-proportional sentinel allocations across 17 regions.

Results: Of 211 sentinel hospitals, 180 (85.3%) responded. In total, 90% (162/180) rated reporting as simple or adequate; responses on goal awareness and overall satisfaction were predominantly neutral. The delayed reporting proportion was 2.4% (255/10,652) in 2021 and 1.1% (122/10,972) in 2023; submission completeness was 99.9% (10,652/10,660) and 100% (10,972/10,972), respectively; and on-time completeness was 97.5% (10,397/10,660) and 98.9% (10,850/10,972), respectively. The exploratory weekly SSS-NHI correlation was 0.899 in 2021 and 0.937 in 2023. Age-stratified annual cross-category correlations ranged from 0.665 to 0.969 in 2021 and from 0.939 to 0.983 in 2023. Correlations between actual and population-proportional sentinel allocations were 0.946 in 2021 and 0.951 in 2023.

Conclusions: The SSS demonstrated favorable operational performance, with low reporting delay and near-complete submission. The claim comparison showed strong ecological comovement but cannot establish sensitivity or diagnostic equivalence because the data sources use different case definitions and counting units. Priorities are clearer communication of operational goals, standardized operator training, and future validation using patient- or episode-level data linked to laboratory results.

(JMIR Public Health Surveill 2026;12:e102373) doi:10.2196/102373

KEYWORDS

bacterial gastroenteritis; sentinel surveillance; public health surveillance; electronic reporting; claim data; data linkage; evaluation; South Korea; notifiable diseases; foodborne illness

Introduction

Gastroenteritis is caused by a wide range of pathogens, including bacteria, viruses, and parasites, that infect the gastrointestinal tract [1]. The causal pathogen is often not identified or reported [2], yet bacterial enteric pathogens are responsible for a large share of the burden, accounting for an estimated 20% to 40% of diarrhea cases worldwide [3,4]. The downstream consequences are striking: enteric infectious diseases caused an estimated 1.27 million deaths worldwide in 2023, with diarrheal disease accounting for most of these deaths (1.11 million) [5], whereas in high-income settings, acute gastroenteritis drives substantial outpatient and inpatient costs [6,7].

XSL•FO
RenderX

South Korea is no exception to this burden. Health insurance claims for 18 foodborne pathogens recorded more than 300,000 cases annually between 2008 and 2012 [8]. Among Korean children between 2004 and 2019, bacterial pathogens were responsible for 3% to 20% of acute gastroenteritis cases, with *Escherichia coli* and *Salmonella* spp. predominating [9]. By 2023, bacterial gastroenteritis accounted for half of all gastroenteritis cases reported to the Korea Disease Control and Prevention Agency (KDCA); *Salmonella* spp., *Campylobacter* spp., and enteropathogenic *E. coli* were the most frequently notified pathogens [10]. The case for close monitoring is likely to intensify: South Korea's summers are growing warmer and longer, which favors bacterial transmission [11], and rising international travel has made surveillance of previously unreported pathogen genotypes increasingly important [9].

South Korea introduced a sentinel surveillance and electronic reporting system in 2000 [12] targeting infectious diseases with high incidence but low severity [13]. By 2023, a total of 11 types of bacterial gastroenteritis were classified as group 4 notifiable diseases and had to be reported to the health authority within 7 days of diagnosis through the sentinel surveillance system (SSS) [13]. Today, 211 sentinel hospitals submit weekly case reports to the KDCA through this nationwide network. Surveillance systems for gastrointestinal diseases vary widely worldwide in design and data source, ranging from laboratory-based and sentinel-clinician networks to syndromic and administrative claim–based approaches; a recent systematic review found such systems to be validated for routine use in at least 10 countries [14]. A previous evaluation looked at the SSS for hand, foot, and mouth disease [15], but to our knowledge, no comprehensive evaluation has yet examined the SSS specifically for bacterial gastroenteritis. Closing this gap matters: routine surveillance for this disease group needs to be assessed on its own terms, and concrete strategies for improvement need to be identified.

To address this gap, we conducted a mixed methods evaluation of South Korea's electronic SSS for bacterial gastroenteritis with 2023 as the primary evaluation year and 2021 as a nonconsecutive descriptive comparator. The objectives were to (1) assess simplicity, acceptability, timeliness, submission completeness, and on-time completeness; (2) examine whether sentinel site allocation across 17 regions was proportional to population size; and (3) explore whether weekly all-pathogen totals and annual age-stratified surveillance target or diagnosis code category patterns in SSS reports were concordant with National Health Insurance (NHI) claim records.

Methods

Study Design

This mixed methods surveillance system evaluation combined a cross-sectional survey and structured operator interview with retrospective analysis of operational reporting records, NHI claim records, and regional population statistics. The primary evaluation year was 2023, with 2021 used only as a nonconsecutive descriptive comparator. The operational and external comparison components were analyzed separately so that claim data were not treated as a reference standard for laboratory-confirmed SSS cases.

Data Sources

Overview

From July 29, 2024, to August 30, 2024, we emailed a semistructured questionnaire to the staff member responsible for case notification at each of the 211 sentinel hospitals. Items covered the perceived simplicity of reporting, awareness of SSS operational goals, overall satisfaction, and mean weekly reporting time. From these 211 hospitals, we received 180 (85.3%) completed questionnaires.

We conducted a structured interview with the KDCA operator responsible for routine SSS operation. Topics included reporting compliance, ease of data collection and dissemination, recurrent data quality problems, operator workload, and whether formal training or standardized operating materials were provided. Separately, KDCA operational extracts supplied the number of designated sentinel sites, facilities with no laboratory-confirmed case notifications for more than 6 months (an indicator distinct from missing scheduled weekly submissions), and weekly on-time and delayed submission counts for 2021 and 2023.

Nationwide NHI claim records were obtained from the Health Insurance Review and Assessment Service for January 2021 to December 2021 and January 2023 to December 2023. Variables included *International Classification of Diseases, 10th Revision* (*ICD-10*), diagnosis code and position; outpatient or inpatient status; age group; treating institution region; service date; and billed laboratory test procedures. Official resident registration population counts for the 17 administrative regions in the corresponding years were obtained from the Korean Statistical Information Service and were used only to calculate population-proportional expected numbers of sentinel sites [16].

Table 1 lists the 11 types of gastroenteritis that are notifiable through the SSS, with their causative pathogens and diagnostic criteria [17].

XSL•FO
RenderX

Table . Laboratory-confirmed types of bacterial gastroenteritis included in South Korea's electronic sentinel surveillance system (SSS) and national SSS case counts in 2023.

Diseases	Causative pathogen	Diagnostic criteria	KDCA[a]-reported cases in 2023, n
Salmonellosis	Nontyphoidal *Salmonella* spp.	Isolation and identification of nontyphoidal *Salmonella* spp. from specimens (stool and rectal swabs)	3540
Vibrio parahaemolyticus gastroenteritis	*V parahaemolyticus*	Isolation and identification of *V parahaemolyticus* from specimens (stool and rectal swabs)	101
ETEC[b]	ETEC	Isolation and identification of *Escherichia coli* possessing heat-labile toxin genes (*LT*) or heat-stable toxin genes (*ST*) from specimens (stool and rectal swabs)	481
EIEC[c]	EIEC	Isolation and identification of *E. coli* possessing the invasive factor gene (*ipaH*) from specimens (stool and rectal swabs)	55
EPEC[d]	EPEC	Isolation and identification of *E. coli* possessing adherence factor genes (*eaeA* and *bfpA*) from specimens (stool and rectal swabs)	1963
Campylobacteriosis	*Campylobacter jejuni* and *Campylobacter coli*	Isolation and identification of *Campylobacter* spp. from specimens (stool, rectal swabs, and vomitus)	3167
Clostridium perfringens enteritis	*C. perfringens*	Detection of *C. perfringens* at levels of $\geq 10^6$ CFU[e] per gram in specimens (stool and vomitus) or isolation and identification of *C. perfringens* possessing enterotoxin-specific genes (*cpa* and *cpe*) from specimens (stool, rectal swabs, and vomitus)	423
Staphylococcus aureus intoxication	*S. aureus*	Isolation and identification of *S. aureus* possessing enterotoxin genes from specimens (stool, rectal swabs, and vomitus)	160
Bacillus cereus gastroenteritis	*B. cereus*	Isolation and identification of *B. cereus* possessing toxin genes (*hblC*, *nheA*, *entFM*, *cytK2*, *becT*, and *CER*) from specimens (stool, rectal swabs, and vomitus)	42
Yersiniosis	*Yersinia enterocolitica*	Isolation and identification of *Y. enterocolitica* from specimens (stool, rectal swabs, and vomitus)	169
Listeriosis	*Listeria monocytogenes*	Isolation and identification of *L. monocytogenes* from specimens (stool and rectal swabs)	10

[a]KDCA: Korea Disease Control and Prevention Agency.

[b]ETEC: enterotoxigenic *E. coli*.

[c]EIEC: enteroinvasive *E. coli*.

[d]EPEC: enteropathogenic *E. coli*.

[e]CFU: colony-forming unit.

Case Identification in NHI Claim Data

To identify NHI claim records corresponding to the 11 SSS targets, we used *ICD-10* codes A02, A04, A05, and A32 and their pathogen-specific subcodes where available. These diagnosis code categories correspond to salmonellosis, other bacterial intestinal diseases, bacterial foodborne intoxications, and listeriosis. They do not establish microbiological identification of the causative pathogen. Some parent codes encompass more than one organism; therefore, category-level analyses compared SSS surveillance targets with corresponding

XSL•FO
RenderX

diagnosis code categories and should not be interpreted as one-to-one pathogen matching.

Counts were extracted at the claim record level rather than at the unique patient or episode level; repeated outpatient or inpatient claims for the same person could therefore be included. To increase coding specificity, we retained primary or secondary diagnosis records accompanied by a billed laboratory test procedure and excluded follow-up records without an active diagnosis code. A billed test indicates only that testing was ordered or claimed; the result is not available in NHI data. Consequently, the algorithm may include empirically treated or miscoded records and is not equivalent to the laboratory-confirmed SSS case definition. We did not use these data to calculate incidence, case count ratios, capture fractions, or diagnostic accuracy. Throughout the manuscript, "claim records" denotes NHI observations, and "cases" denotes SSS reports.

Study Period

The primary evaluation year was 2023 because it was the most recent complete year in the project dataset and was closest to the 2024 survey and operator interview. The analytic dataset also contained a complete 2021 extract, which was retained as a nonconsecutive historical comparator; 2022 was not included in the extract available to this study. We did not treat 2021 as an unaffected prepandemic baseline. Comparisons between 2021 and 2023 are descriptive snapshots, and no continuous temporal trend or causal effect of the COVID-19 pandemic was inferred.

Evaluation Framework

Overview

Our evaluation used the KDCA guidelines [18], which adapt the US Centers for Disease Control and Prevention framework for evaluating public health surveillance systems [19]. The framework had previously been applied to another South Korean SSS [15]. We assessed operational performance (simplicity, acceptability, and timeliness), reporting completeness (submission completeness and on-time completeness), and geographic allocation. Because sensitivity requires a valid denominator or reference standard and NHI claims do not provide one, sensitivity was not estimated. Weekly and age-stratified claim comparisons were analyzed as a separate exploratory external concordance component. Table 2 shows the operational definitions and data sources.

XSL•FO
RenderX

Table . Operational definitions and data sources for the mixed methods evaluation of South Korea's electronic bacterial gastroenteritis sentinel surveillance system (SSS) using 2021 and 2023 operational and claim data and a 2024 survey and operator interview.

Domains	Attribute	Operational definition	Source
Performance	Simplicity	Percentage of surveyed sentinel staff rating the reporting process as simple or adequate	2024 sentinel hospital survey
Performance	No case notifications for >6 mo	Percentage of designated sentinel hospitals with no laboratory-confirmed bacterial gastroenteritis case notifications for >6 mo; distinct from missing scheduled weekly submissions	KDCA[a] operational data
Performance	Acceptability	Distribution of responses on awareness of SSS operational goals	2024 sentinel hospital survey
Performance	Acceptability	Distribution of responses on overall satisfaction with SSS operation	2024 sentinel hospital survey
Performance	Timeliness	Delayed reports divided by all submitted reports; B/(A + B)	KDCA operational data
Reporting completeness	Submission completeness	All submitted reports divided by expected reports; (A + B)/C	KDCA operational data
Reporting completeness	On-time completeness	Reports submitted within 7 d divided by expected reports; A/C	KDCA operational data
Exploratory external comparison	Weekly temporal concordance	Pearson correlation between weekly all-pathogen SSS cases and NHI[b] claim records (n=52 wk per y); not a sensitivity measure	KDCA and NHI claims
Exploratory external comparison	Age-stratified cross-category concordance	Pearson correlation between annual counts across 11 SSS surveillance targets or corresponding diagnosis code categories within each age group; not one-to-one pathogen validation	KDCA and NHI claims
Network allocation	Geographic allocation	Pearson correlation between population-proportional expected and actual sentinel site counts across 17 regions	KDCA and KOSIS[c] population data

[a]KDCA: Korea Disease Control and Prevention Agency.

[b]NHI: National Health Insurance.

[c]KOSIS: Korean Statistical Information Service.

Simplicity

Simplicity captures the ease of structuring and operating the surveillance system [18]. We checked it in 3 aspects: case reporting, data collection, and dissemination of analytic results. Sentinel hospital personnel rated the simplicity of the case reporting process. The KDCA operator was asked about ease of data collection and result dissemination. Each item used a 3-point scale: simple, adequate, or complex.

Acceptability

Acceptability reflects how willing stakeholders are to take part in operations, reporting, and data use [18]. We measured it in 3 ways. The first was the proportion of sentinel hospitals with no laboratory-confirmed bacterial gastroenteritis case notifications for 6 months or longer in a year divided by the total number of sentinel hospitals multiplied by 100. This indicator concerned the absence of notified cases and was distinct from scheduled weekly report submission; a zero-case weekly submission counted toward completeness. The second was awareness of operational goals, asked as follows: "Are you aware of the operational goals of the surveillance system?" (responses: "yes," "neutral," or "no"). The third was overall satisfaction, asked as follows: "What is your overall satisfaction with the operation of the surveillance system?" (responses: "good," "neutral," or "poor").

Timeliness

The KDCA expects sentinel hospitals to report cases within 7 days of diagnosis [13]. Broadly, timeliness refers to the time needed for reporting, analysis, and feedback within the SSS [18]. We measured it as the delayed reporting proportion: reports submitted after the 7-day window divided by the total number of reports submitted (the sum of on-time and delayed reports).

XSL•FO
RenderX

Completeness

If A denotes scheduled weekly submissions received within 7 days, B denotes delayed scheduled weekly submissions, C denotes the expected number of scheduled weekly submissions (number of designated sentinel hospitals multiplied by 52 weeks), and weekly submissions could report one or more cases or zero cases, we calculated submission completeness as (A + B)/C and on-time completeness as A/C. The delayed reporting proportion, our timeliness measure, was B/(A + B). These indicators were not arithmetic complements: timeliness was conditional on submissions that were received, whereas completeness used all expected submissions as the denominator. Reporting both forms of completeness separated failure to submit from delay among submitted weekly reports.

Exploratory External Concordance

For exploratory external comparison, we aggregated SSS cases and NHI claim records across all 11 target categories by epidemiologic week and calculated a Pearson correlation using 52 paired weekly totals within each year. This ecological correlation asks whether the 2 data streams comove over weeks; it does not determine whether they identify the same patients or diagnoses. In a separate age-stratified analysis, we compared annual counts across 11 SSS surveillance target or corresponding diagnosis code categories within each age group (n=11 paired categories per panel). Because these correlations are sensitive to the highest-incidence categories, some parent *ICD-10* codes encompass more than one organism, and the claim definition is not laboratory confirmed, they were interpreted descriptively rather than as validation metrics.

Geographic Allocation

For each of South Korea's 17 administrative regions, the population-proportional expected number of sentinel hospitals was calculated as the total number of sentinel sites multiplied by the region's share of the national resident population. Pearson correlation coefficients compared expected and actual site counts across regions for 2021 and 2023. This indicator assessed whether the network's allocation followed population distribution; it did not evaluate regional disease incidence, regional reporting completeness, or geographic representativeness of reported cases.

Statistical Analysis

Survey and operational data were summarized using frequencies, proportions, and means. Pearson correlation coefficients were calculated separately for weekly external concordance (52 paired weeks per year), age-stratified annual cross-category concordance (11 paired categories per age group and year), and geographic allocation (17 paired regions per year). Weekly coefficients were treated as descriptive because consecutive observations may be temporally autocorrelated and shared seasonality was not modeled. *P* values are reported descriptively and were not used to classify system performance; no adjustment for multiple exploratory correlations was applied. Analyses were conducted in SAS (version 9.4; SAS Institute).

Reporting Guideline

We used the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist for cross-sectional studies as a reporting aid for the survey and observational claim components [20]. A completed checklist is provided as Checklist 1. Because this study also included surveillance operational records and a single-operator interview, the STROBE checklist does not cover every component of the mixed methods evaluation.

Ethical Considerations

The institutional review board of Kongju National University approved this study (approval number KNU_IRB_2024-072). Because we used deidentified, aggregated sentinel hospital data; an institutional interview with the KDCA operator; and anonymized NHI claim data, individual informed consent was waived. All data handling complied with the relevant Korean regulations on personal information and statistical data.

Results

Survey Results From Sentinel Hospitals

Of the 180 responding sentinel hospitals, 18 (10%) rated the reporting process as simple, 144 (80%) rated it as adequate, and 18 (10%) rated it as complex. Thus, 90% (162/180) rated reporting as simple or adequate (Table 3).

Responses on operational goal awareness and satisfaction were predominantly neutral. For awareness, of the 180 respondents, 55 (30.6%) answered "yes," 97 (53.9%) were neutral, and 28 (15.6%) answered "no." For overall satisfaction, 4.4% (n=8) rated the system as good, 79.4% (n=143) rated it as neutral, and 16.1% (n=29) rated it as poor (Table 3). These marginal distributions were analyzed separately; no within-respondent association among the 3 survey items was tested.

Table . Responses to the 2024 survey of staff responsible for bacterial gastroenteritis notification at 85.3% (180/211) of the sentinel hospitals in South Korea (n=180)[a].

Items	Positive response, n (%)	Neutral or "adequate" response, n (%)	Negative response, n (%)
Rating of the reporting process	Simple: 18 (10)	Adequate: 144 (80)	Complex: 18 (10)
Awareness of the goals of the SSS[b]	Yes: 55 (30.6)	Neutral: 97 (53.9)	No: 28 (15.6)
Overall satisfaction with the SSS	Good: 8 (4.4)	Neutral: 143 (79.4)	Poor: 29 (16.1)

[a]Mean time taken to report cases per week 30.6 (SD 92.7) minutes.

[b]SSS: sentinel surveillance system.

XSL•FO
RenderX

KDCA Operational Data

In 2023, a total of 2.8% (6/211) of the sentinel hospitals reported no laboratory-confirmed bacterial gastroenteritis cases for more than 6 months. This did not indicate missing scheduled weekly submissions; zero-case weekly submissions were counted in the completeness measures. The KDCA operator also reported that no formal standardized training program was provided for the role in 2023.

The delayed reporting proportion was 2.4% (255/10,652) of submitted weekly reports in 2021 and 1.1% (122/10,972) in 2023. Submission completeness was 99.9% (10,652/10,660) of expected weekly submissions in 2021 and 100% (10,972/10,972) in 2023. On-time completeness was 97.5% (10,397/10,660) and 98.9% (10,850/10,972) submissions in 2021 and 2023, respectively (Table 4).

Table . Operational reporting indicators, sentinel-reported cases, and National Health Insurance (NHI) claim record counts for 11 bacterial gastroenteritis diagnosis categories in South Korea (2021 and 2023)[a].

Indicators	2021	2023
Sentinels in the surveillance system, n	205	211
Facilities with no laboratory-confirmed case notifications for >6 mo, n/N (%)	3/205 (1.5)	6/211 (2.8)
On-time scheduled weekly submissions received within 7 d (denoted as A), n	10,397	10,850
Delayed scheduled weekly submissions (denoted as B), n	255	122
Delayed reporting proportion (calculated as B/[A + B]), %	2.4	1.1
Expected scheduled weekly submissions (denoted as C [calculated as sentinel sites × 52 wk]), n	10,660	10,972
Submission completeness ([A + B]/C), %	99.9	100
On-time completeness (A/C), %	97.5	98.9
Total NHI claim records for the 11 bacterial gastroenteritis diagnosis categories, n	36,572	39,376
Health care facilities with NHI claim records for the 11 diagnosis categories, n	1030	1183
NHI claim records for patients aged 0 - 6 y, n	8045	6307
Sentinel-reported cases for patients aged 0 - 6 y, n	1333	1370

[a]Weekly submissions may include zero-case reports. The >6 month indicator denotes no laboratory-confirmed case notifications and is distinct from missing weekly submissions. NHI values are claim record counts and are not deduplicated by patient or episode; they must not be interpreted as incidence or directly compared with sentinel surveillance system case counts as if the counting units were equivalent.

Exploratory External Concordance and Geographic Allocation

Using 52 paired weekly all-pathogen totals per year, the exploratory Pearson correlation between SSS-reported cases and NHI claim records was 0.899 in 2021 and 0.937 in 2023 (Figure 1). These coefficients describe ecological weekly comovement and do not measure diagnostic agreement or surveillance sensitivity.

In the annual age-stratified cross-category analyses, each coefficient was based on 11 paired surveillance target or diagnosis code categories. The coefficients ranged from 0.665 to 0.969 in 2021 and from 0.939 to 0.983 in 2023. For children aged 0 to 6 years, the coefficients were 0.969 in 2021 and 0.982 in 2023; the largest 2023 coefficient was 0.983 in the group aged 7 to 12 years. The coefficients differed numerically between the 2 years, particularly in older age groups, but the nonconsecutive years and small number of category-level observations do not support a trend or causal interpretation. Full panels are shown in Multimedia Appendices 1 and 2.

The correlation between population-proportional expected and actual sentinel site counts across South Korea's 17 administrative regions was 0.946 in 2021 and 0.951 in 2023 ($P<.001$ in both cases; Table 5), indicating close alignment of site allocation with population distribution.

XSL•FO
RenderX

Figure 1. Weekly temporal concordance between total National Health Insurance (NHI) claim records carrying 1 of 11 bacterial gastroenteritis diagnosis codes and laboratory-confirmed cases reported through South Korea's sentinel surveillance system shown separately for 2021 (A) and 2023 (B). Each point represents 1 epidemiologic week (n=52 per year); the line is the fitted linear relationship, and the shaded area is its 95% CI. The correlation is ecological and descriptive; shared seasonality and serial autocorrelation were not modeled, and the coefficient does not indicate diagnostic agreement or sensitivity. KDCA: Korea Disease Control and Prevention Agency.

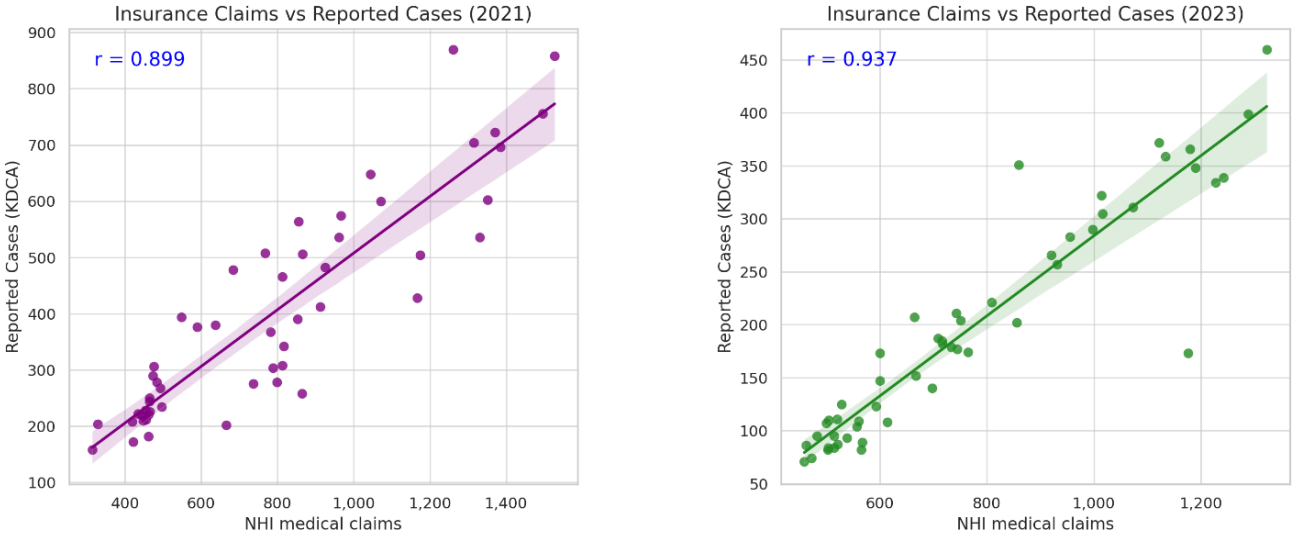


Table . Summary of operational performance reporting completeness, exploratory external concordance, and geographic allocation for South Korea's electronic bacterial gastroenteritis sentinel surveillance system (2021 and 2023)[a].

Domains	Attribute	2021	2023	Interpretation
Performance	Reporting process rated as simple or adequate	—[b]	90% (162/180)	Predominantly simple or adequate
Performance	Awareness of operational goals; response distribution	—	Yes: 30.6% (55/180); neutral: 53.9% (97/180); no: 15.6% (28/180)	Predominantly neutral
Performance	Overall satisfaction; response distribution	—	Good: 4.4% (8/180); neutral: 79.4% (143/180); poor: 16.1% (29/180)	Predominantly neutral
Performance	Delayed reporting proportion	2.4% (255/10,652)	1.1% (122/10,972)	Low
Reporting completeness	Submission completeness	99.9% (10,652/10,660)	100% (10,972/10,972)	Near complete
Reporting completeness	On-time completeness	97.5% (10,397/10,660)	98.9% (10,850/10,972)	High
Exploratory external comparison	Weekly SSS[c]-NHI[d] concordance (Pearson *r*; n=52 wk)	0.899	0.937	Strong descriptive ecological correlation
Exploratory external comparison	Age-stratified cross-category concordance (Pearson *r* for age of 0 - 6 y; n=11 categories)	0.969	0.982	Strong exploratory correlation
Network allocation	Population-proportional vs actual sentinel site allocation (Pearson *r*; n=17 regions)	0.946	0.951	Strong alignment

[a]Neutral survey responses are reported separately and were not combined with positive responses. Correlation coefficients are descriptive. The sentinel surveillance system–National Health Insurance coefficients do not establish sensitivity, diagnostic agreement, one-to-one pathogen matching, or equivalence of counting units; the geographic coefficient assesses population-based site allocation only.

[b]No corresponding survey data were available for 2021; the survey was conducted in 2024 to evaluate the 2023 surveillance system.

[c]SSS: sentinel surveillance system.

[d]NHI: National Health Insurance.

XSL•FO
RenderX

Discussion

Principal Findings

Relative to the study objectives, the electronic SSS showed favorable operational performance. In 2023, a total of 90% (162/180) of respondents rated reporting as simple or adequate, 1.1% (122/10,972) of submitted weekly reports were delayed, submission completeness was 100% (10,972/10,972), and on-time completeness was 98.9% (10,850/10,972). Actual sentinel site allocation also closely followed population-proportional allocation across 17 regions. The exploratory external comparisons showed strong weekly ecological comovement between SSS cases and NHI claim records and generally strong annual cross-category correlations within age groups.

Acceptability findings were less favorable and more difficult to interpret. Only 30.6% (55/180) of respondents clearly reported awareness of the SSS's operational goals, whereas 53.9% (97/180) selected "neutral" and 15.6% (28/180) selected "no." Overall satisfaction was also mostly neutral (143/180, 79.4%), with 4.4% (8/180) rating it as good and 16.1% (29/180) rating it as poor. Because the awareness item did not state the operational goals, neutral or negative responses may reflect either limited awareness or ambiguity in the question. Nevertheless, clearer communication and routine onboarding are warranted. The absence of a standardized training program for the KDCA operator further supports the need for role-specific training and written operating procedures.

Comparison With Prior Work

Systematic evaluation of surveillance systems can identify operational gaps and improve the information available to public health agencies. Although established frameworks exist, evaluations differ in the attributes assessed, data sources used, and analytic methods applied [19,21,22]. Our approach combined the US Centers for Disease Control and Prevention framework [19] with the KDCA-adapted framework [18], as in a previous evaluation of South Korea's hand, foot, and mouth disease SSS [15]. Electronic notification systems have also been evaluated in other settings [23]. The present system likewise performed favorably for reporting simplicity, timeliness, and completeness while also revealing opportunities to improve stakeholder engagement and operator training.

Timeliness is a central surveillance attribute because delayed notification can reduce the usefulness of information for public health action [22]. NHI claims are not a substitute for real-time reporting because billing and data access processes introduce delay. The SSS, in contrast, requires notification within 7 days of diagnosis [13]. The 1.1% (122/10,972) delayed reporting proportion in 2023, therefore, indicates strong adherence to the specified reporting window. This metric does not measure onset-to-detection delay or outbreak detection speed, and its interpretation depends on the common 7-day threshold rather than pathogen-specific urgency [22,24]. Automated outbreak detection systems use different end points and require separate evaluation [25].

The external comparison requires more cautious interpretation. Surveillance sensitivity is the proportion of true community cases captured, which cannot be estimated without a credible denominator or reference standard. Other studies have compared gastroenteritis surveillance with web-based symptom reports or medication sales [26,27], and Korean claim data have been used to complement passive surveillance [28]. In our study, the weekly correlation used 52 ecological observations per year and indicates comovement of aggregate counts only. Consecutive weeks may be temporally autocorrelated, and shared seasonal patterns were not modeled; the coefficient should therefore be interpreted as descriptive comovement rather than an independent observation validation estimate. The age-stratified panels used 11 annual category-level observations and are particularly sensitive to leverage from high-incidence categories. Some *ICD-10* parent codes encompass more than one organism, further limiting pathogen-specific interpretation. Moreover, NHI claim records are not laboratory confirmed and were not deduplicated by patient or episode. Differences between the 2021 and 2023 coefficients, therefore, should not be interpreted as evidence that sensitivity changed. Validation of capture and diagnostic accuracy would require patient- or episode-level linkage to laboratory results or another microbiologically confirmed reference source. Geographic allocation should also be distinguished from case representativeness: our regional analysis assessed where sentinel hospitals were located relative to population, not whether reported disease incidence was representative.

The results support the operational value of a nationally coordinated electronic reporting platform, but they do not establish early outbreak detection performance, which was not evaluated. Potentially transferable design features include population-informed sentinel site allocation, standardized laboratory case definitions, and centralized electronic reporting. Site placement can influence outbreak detection and access to surveillance facilities [29,30]. Periodic comparison with administrative data may be useful as a quality assurance signal when interpreted as ecological consistency rather than as validation of cases or estimation of underreporting.

Public Health and Informatics Implications

Electronic sentinel reports and administrative claims serve different functions. The SSS provides the timely, laboratory-based notifications required for routine surveillance. Claim data become available later and reflect reimbursement coding rather than confirmed infection. Their appropriate role in this context is retrospective quality assurance: marked divergence in weekly or category-specific patterns could prompt targeted review of coding, laboratory practice, or reporting workflows. Claim data alone cannot identify which source is correct, quantify underreporting, or replace laboratory-based surveillance.

Three practical improvements follow from the evaluation. First, routine feedback to sentinel hospitals could display reporting completeness and carefully labeled ecological concordance indicators, with alerts used to trigger review rather than grade performance automatically. Second, concise digital onboarding and standardized training materials should be provided to both

XSL•FO
RenderX

sentinel personnel and KDCA operators. Third, future evaluations should obtain patient- or episode-level claims and link them, where legally and technically feasible, to laboratory-confirmed surveillance records. Such validation would directly address the case definition and duplicate use limitations identified in this study.

Limitations

The principal limitation concerns the external comparator. NHI claims record diagnosis codes and billed procedures, not laboratory results, whereas SSS cases require laboratory confirmation. Claims may therefore include false-positive coding, empiric treatment, or unconfirmed testing. Accordingly, the correlations are interpreted only as exploratory ecological comparisons and not as a proxy for sensitivity.

Claims were counted at the record level and were not deduplicated by patient or clinical episode. Repeated use could alter weekly or category-specific patterns, particularly if revisit frequency varies by pathogen, age, or year. We did not compare absolute count levels, calculate incidence, or estimate capture fractions. Patient- or episode-level analysis is required before claims can be used for those purposes.

The survey response rate was 85.3% (180/211), but the characteristics of respondents were not formally compared with those of nonrespondents. Simplicity and acceptability estimates may therefore be affected by nonresponse and social desirability bias. We also did not test within-respondent associations among reporting simplicity, goal awareness, and satisfaction; the 3 marginal distributions should be interpreted separately. The operator interview involved 1 person, so it was used as contextual evidence rather than as a basis for generalizable quantitative inference.

The correlation analyses also have scope limitations. Weekly concordance was assessed in only 2 nonconsecutive years, so no continuous trend or pandemic effect can be inferred. Consecutive weekly observations may be temporally autocorrelated, and high coefficients may partly reflect shared seasonal patterns; seasonality and serial correlation were not modeled. Each age-stratified panel contained 11 category-level observations, several clustered near zero, and Pearson correlations were sensitive to high-incidence categories; multiple exploratory correlations were not adjusted for multiplicity. Some *ICD-10* parent codes encompass more than one organism, so the category-level comparisons do not imply one-to-one pathogen matching. The geographic analysis compared population-proportional and actual site allocation across 17 regions but did not compare regional disease incidence patterns.

Finally, our timeliness metric used a common 7-day threshold and did not measure the interval from symptom onset to detection or public health action. We did not evaluate outbreak detection algorithms, detection delay, flexibility, positive predictive value, stability, or cost-effectiveness [31,32]. These attributes require additional data and should be addressed in future evaluations.

Conclusions

South Korea's electronic SSS for bacterial gastroenteritis demonstrated favorable operational performance, including low reporting delay, near-complete submission, and population-aligned sentinel site allocation. Exploratory correlations with NHI claim records showed ecological comovement. Clearer communication of system goals, standardized operator training, and validation using deduplicated patient- or episode-level data linked to laboratory results are the most important next steps. These distinctions are relevant to other jurisdictions considering administrative data as a supplement to rather than a replacement for laboratory-based electronic surveillance.

Acknowledgments

The authors thank the staff of the participating sentinel hospitals and the Korea Disease Control and Prevention Agency personnel who contributed survey and operational information. During revision, the authors used ChatGPT (OpenAI) for language editing, structural organization, and internal consistency checks. The tool was not used to generate or analyze study data. All outputs were critically reviewed, verified, and revised by the authors, who take full responsibility for the manuscript.

Funding

This work was supported by the research program funded by the Korea Disease Control and Prevention Agency (grant 2024-04-019). The funder had no involvement in the study design; data collection, analysis, and interpretation; manuscript preparation; or decision to submit the work for publication.

Data Availability

The aggregated survey and Korea Disease Control and Prevention Agency–reported data supporting the findings of this study are presented within the article and its multimedia appendices. The National Health Insurance medical claims data are not publicly available because of restrictions imposed by the Health Insurance Review and Assessment Service (HIRA) of the Republic of Korea, but they can be requested from the HIRA through a formal application process for legitimate research purposes.

XSL•FO
RenderX

Authors' Contributions

IS and DSK conceptualized and designed the study. YH, MJ, JP, and SR contributed to data collection and analysis. IS drafted the original manuscript. YH, MJ, JP, SR, and DSK reviewed and edited the manuscript. All authors had full access to all data in the study and accept final responsibility for the decision to submit the manuscript for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Exploratory annual cross-category concordance between National Health Insurance claim records and sentinel-reported bacterial gastroenteritis cases within 6 age groups in South Korea in 2021.
[DOCX File, 305 KB - publichealth_v12i1e102373_app1.docx]

Multimedia Appendix 2
Exploratory annual cross-category concordance between National Health Insurance claim records and sentinel-reported bacterial gastroenteritis cases within 6 age groups in South Korea in 2023.
[DOCX File, 307 KB - publichealth_v12i1e102373_app2.docx]

Checklist 1
STROBE checklist.
[DOCX File, 35 KB - publichealth_v12i1e102373_app3.docx]

References

1. Nic Fhogartaigh C, Edgeworth JD. Bacterial gastroenteritis. Medicine 2009 Nov;37(11):586-593. [doi: 10.1016/j.mpmed.2009.08.006]
2. Scallan E, Griffin PM, McLean HQ, Mahon BE. Hospitalisations due to bacterial gastroenteritis: a comparison of surveillance and hospital discharge data. Epidemiol Infect 2018 Jun;146(8):954-960. [doi: 10.1017/S0950268818000882] [Medline: 29655383]
3. Fleckenstein JM, Matthew Kuhlmann F, Sheikh A. Acute bacterial gastroenteritis. Gastroenterol Clin North Am 2021 Jun;50(2):283-304. [doi: 10.1016/j.gtc.2021.02.002] [Medline: 34024442]
4. Barrett J, Fhogartaigh CN. Bacterial gastroenteritis. Medicine 2017 Nov;45(11):683-689. [doi: 10.1016/j.mpmed.2017.08.002]
5. GBD 2023 Diarrhoeal Disease and Enteric Infectious Diseases Collaborators. Global burden of enteric infectious diseases, diarrhoeal diseases, and corresponding aetiologies, 1990–2023: a systematic analysis for the Global Burden of Disease Study 2023. Lancet Infect Dis 2026 Jun 2. [doi: 10.1016/S1473-3099(26)00194-5] [Medline: 42229499]
6. Dickerson JF, Salas SB, Donald J, et al. Economic burden of acute gastroenteritis among members of integrated healthcare delivery system, United States, 2014-2016. Emerg Infect Dis 2024 May;30(5):968-973. [doi: 10.3201/eid3005.230356] [Medline: 38666613]
7. Barker SF, Zomer E, O'Toole J, et al. Cost of gastroenteritis in Australia: a healthcare perspective. PLoS One 2018;13(4):e0195759. [doi: 10.1371/journal.pone.0195759] [Medline: 29649285]
8. Park MS, Kim YS, Lee SH, Kim SH, Park KH, Bahk GJ. Estimating the burden of foodborne disease, South Korea, 2008-2012. Foodborne Pathog Dis 2015 Mar;12(3):207-213. [doi: 10.1089/fpd.2014.1858] [Medline: 25622301]
9. Ryoo E. Causes of acute gastroenteritis in Korean children between 2004 and 2019. Clin Exp Pediatr 2021 Jun;64(6):260-268. [doi: 10.3345/cep.2020.01256] [Medline: 32972055]
10. Infectious disease statistics. Korea Disease Control and Prevention Agency. URL: https://dportal.kdca.go.kr/pot/is/st/gstrnftn.do [accessed 2024-11-22]
11. Sung J, Cheong HK, Kwon HJ, Kim JH. Pathogen-specific response of infectious gastroenteritis to ambient temperature: national surveillance data in the Republic of Korea, 2015-2019. Int J Hyg Environ Health 2022 Mar;240:113924. [doi: 10.1016/j.ijheh.2022.113924] [Medline: 35065521]
12. Park O, Choi BY. Introduction and evaluation of communicable disease surveillance in the Republic of Korea [Article in Korean]. J Prev Med Public Health 2007 Jul;40(4):259-264. [doi: 10.3961/jpmph.2007.40.4.259] [Medline: 17693727]
13. Infectious disease surveillance system. Korea Disease Control and Prevention Agency. URL: https://www.kdca.go.kr/eng/4354/subview.do [accessed 2026-08-11]
14. Adedire O, Love NK, Hughes HE, Buchan I, Vivancos R, Elliot AJ. Early detection and monitoring of gastrointestinal infections using syndromic surveillance: a systematic review. Int J Environ Res Public Health 2024 Apr 17;21(4):489. [doi: 10.3390/ijerph21040489] [Medline: 38673400]

XSL•FO
RenderX

15. Kim BI, Achangwa C, Cho S, et al. The hand, foot, and mouth disease sentinel surveillance system in South Korea: retrospective evaluation study. JMIR Public Health Surveill 2024 Jul 23;10:e59446. [doi: 10.2196/59446] [Medline: 39045828]
16. Resident population by city, county, and district. Korean Statistical Information Service. URL: https://kosis.kr/statHtml/statHtml.do?tblId=DT_1B040A3&language=en&orgId=101 [accessed 2024-11-22]
17. Infectious disease guidelines [Article in Korean]. Korea Disease Control and Prevention Agency. URL: https://tinyurl.com/y5ck32xp [accessed 2026-08-14]
18. Yu SH. Research on sample surveillance system evaluation and quality management measures. : PRISM; 2023 URL: https://www.prism.go.kr/homepage/asmt/popup/1790387-202300062 [accessed 2026-08-11]
19. German RR, Lee LM, Horan JM, et al. Updated guidelines for evaluating public health surveillance systems: recommendations from the Guidelines Working Group. MMWR Recomm Rep 2001 Jul 27;50(RR-13):1-35. [Medline: 18634202]
20. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med 2007 Oct 16;4(10):e296. [doi: 10.1371/journal.pmed.0040296] [Medline: 17941714]
21. Watkins RE, Eagleson S, Hall RG, Dailey L, Plant AJ. Approaches to the evaluation of outbreak detection methods. BMC Public Health 2006 Oct 24;6:263. [doi: 10.1186/1471-2458-6-263] [Medline: 17059615]
22. Jajosky RA, Groseclose SL. Evaluation of reporting timeliness of public health surveillance systems for infectious diseases. BMC Public Health 2004 Jul 26;4:29. [doi: 10.1186/1471-2458-4-29] [Medline: 15274746]
23. Ibrahim W, Himatt S, Heikal S, et al. Enhancing the notification system for surveillance of infectious diseases in Qatar during the FIFA World Cup 2022: project overview. BMC Public Health 2024 Feb 27;24(1):625. [doi: 10.1186/s12889-024-18016-9] [Medline: 38413899]
24. Yoo HS, Park O, Park HK, et al. Timeliness of national notifiable diseases surveillance system in Korea: a cross-sectional study. BMC Public Health 2009 Mar 31;9:93. [doi: 10.1186/1471-2458-9-93] [Medline: 19331696]
25. Stachel A, Pinto G, Stelling J, et al. Implementation and evaluation of an automated surveillance system to detect hospital outbreak. Am J Infect Control 2017 Dec 1;45(12):1372-1377. [doi: 10.1016/j.ajic.2017.06.031] [Medline: 28844384]
26. Pini A, Merk H, Carnahan A, et al. High added value of a population-based participatory surveillance system for community acute gastrointestinal, respiratory and influenza-like illnesses in Sweden, 2013-2014 using the web. Epidemiol Infect 2017 Apr;145(6):1193-1202. [doi: 10.1017/S0950268816003290] [Medline: 28137317]
27. Pelat C, Boëlle PY, Turbelin C, Lambert B, Valleron AJ. A method for selecting and monitoring medication sales for surveillance of gastroenteritis. Pharmacoepidemiol Drug Saf 2010 Oct;19(10):1009-1018. [doi: 10.1002/pds.1965] [Medline: 20712024]
28. Jung J, Im JH, Ko YJ, et al. Complementing conventional infectious disease surveillance with national health insurance claims data in the Republic of Korea. Sci Rep 2019 Jun 19;9(1):8750. [doi: 10.1038/s41598-019-45409-3] [Medline: 31217476]
29. Buckeridge DL. Outbreak detection through automated surveillance: a review of the determinants of detection. J Biomed Inform 2007 Aug;40(4):370-379. [doi: 10.1016/j.jbi.2006.09.003] [Medline: 17095301]
30. Nikolay B, Salje H, Sturm-Ramirez K, et al. Evaluating hospital-based surveillance for outbreak detection in Bangladesh: analysis of healthcare utilization data. PLoS Med 2017 Jan;14(1):e1002218. [doi: 10.1371/journal.pmed.1002218] [Medline: 28095468]
31. Freeman R, Charlett A, Hopkins S, et al. Evaluation of a national microbiological surveillance system to inform automated outbreak detection. J Infect 2013 Nov;67(5):378-384. [doi: 10.1016/j.jinf.2013.07.021] [Medline: 23876330]
32. Drewe JA, Hoinville LJ, Cook AJ, Floyd T, Stärk KD. Evaluation of animal and public health surveillance systems: a systematic review. Epidemiol Infect 2012 Apr;140(4):575-590. [doi: 10.1017/S0950268811002160] [Medline: 22074638]

Abbreviations

ICD-10: International Classification of Diseases, 10th Revision
KDCA: Korea Disease Control and Prevention Agency
NHI: National Health Insurance
SSS: sentinel surveillance system
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 25.May.2026; peer-reviewed by I Friesema, SV Ahn; revised version received 21.Jul.2026; accepted 28.Jul.2026; published 26.Aug.2026.

Please cite as:
Song I, Heo Y, Jo M, Park JY, Ryu S, Kim DS
Performance of an Electronic Sentinel Surveillance System for Bacterial Gastroenteritis: Mixed Methods Evaluation Study
JMIR Public Health Surveill 2026;12:e102373
URL: https://publichealth.jmir.org/2026/1/e102373
doi:10.2196/102373

© Inmyung Song, Yerin Heo, Minsol Jo, Jung-yeon Park, Sukhyun Ryu, Dong-Sook Kim. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 26.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Age-Related Patterns in Health Literacy and Health Information–Seeking Behavior in a Korean Population: Cross-Sectional Study

Da Hae Kwon[1], PhD; Young Dae Kwon[2,3], MD, PhD

[1]Incheon Public Health Policy Institute, Incheon, Republic of Korea

[2]Department of Humanities and Social Medicine, College of Medicine, The Catholic University of Korea, 222 Banpo-Daero, Seocho-gu, Seoul, Republic of Korea

[3]Catholic Institute for Public Health and Healthcare Management, The Catholic University of Korea, Seoul, Republic of Korea

Corresponding Author:
Young Dae Kwon, MD, PhD
Department of Humanities and Social Medicine, College of Medicine, The Catholic University of Korea, 222 Banpo-Daero, Seocho-gu, Seoul, Republic of Korea

Abstract

Background: Health information–seeking behavior (HISB) is a critical component of personal health management. Health literacy (HL) shapes whether and how people seek health information, but its role across age groups remains underexplored. Korea's highly digitalized information environment and rapidly aging population provide an important context for examining age-group differences in HL and HISB.

Objective: This study examined age-group differences in the association of HL with 2 related stages of HISB—HISB experience and, conditional on engagement, source-based HISB type—and assessed whether the roles of HL, education, and health-related factors differed across age groups.

Methods: Cross-sectional data from the 2020‑2021 Korea Health Panel and 2021 HL Supplementary Survey (N=7910) were analyzed. HL was measured using the European Health Literacy Survey Questionnaire-16-item (HLS-EU-Q16). HISB type was classified as source-based "active seeking" or "passive acquisition" according to the respondents' top-ranked information source. Age-stratified multivariable logistic regression was used for HISB experience and type, and sensitivity analyses examined age×HL interactions and potential selection into the HISB type sample.

Results: Mean HL scores were lower in older age groups (14.10, SD 2.94 in ages 19‑44 years vs 8.34, SD 4.65 in ≥65 years). HISB prevalence was highest in the youngest group (1072/1608, 66.7%) and lowest in the oldest group (1052/3487, 30.2%). Among those who engaged in HISB, 92.7% (993/1072) of the youngest group and 48.2% (507/1052) of the oldest group were classified in the active-seeking category. Primary age-stratified and interaction analyses showed that HL was significantly associated with HISB experience and type among adults aged 45‑64 and ≥65 years, but not among those aged 19‑44 years, whereas education and health-related factors were more salient among younger adults. The HL-HISB type association was sensitive to modeling selection.

Conclusions: Age-group patterns were evident in HISB experience and source-based type. HL was a more prominent correlate among middle-aged and older adults, while education and health-related factors were more salient among younger adults. Age-tailored strategies that also consider structural, motivational, and digital-access factors alongside HL are needed.

(JMIR Public Health Surveill 2026;12:e86992) doi:10.2196/86992

KEYWORDS

health information–seeking behavior; health literacy; age differences; logistic models; Korea; internet use

Introduction

Background

Health information–seeking behavior (HISB) is a fundamental component of personal health management and a recognized social determinant of health. In an increasingly digital environment, individuals are expected to engage proactively with diverse sources to make informed decisions on prevention, diagnosis, and treatment [1-3]. Health information seeking has been linked to increased health knowledge, improved decision-making, healthier behaviors, and favorable health outcomes [1-8].

Health literacy (HL) is critical in shaping individuals' HISB [1,5]. HL affects whether individuals seek information, how

XSL•FO
RenderX

actively they engage with it, and whether they can evaluate and apply what they find [6,8]. Those with higher HL tend to use more diverse and credible information sources, demonstrate greater initiative in information seeking, and report higher confidence in managing their health [9,10]. Conversely, individuals with low HL may encounter barriers in navigating health systems, communicating with providers, or even recognizing the need for additional information [7,11-13].

Among multiple dimensions of HISB, the type of behavior, particularly the distinction between active information seeking and passive information acquisition, has been used as an important analytical lens [5,7,8]. At the conceptual level, active seeking refers to purposeful, goal-directed efforts to find health information, whereas passive acquisition refers to incidental exposure without deliberate information seeking [5,7,14]. This distinction is analytically useful because active and passive forms of information engagement may differ in their associations with health-related outcomes. HL is thus considered a predictor of whether people seek health information and how they do so [8].

Age-group differences in HL and HISB have been widely documented: older adults generally have lower HL, which may reflect educational, cognitive, and technological factors [13-16], and, despite having greater health needs, are less likely to actively seek information [17-19]. Although typically higher in HL, younger adults may rely on less vetted sources or exhibit superficial engagement with information [20]. These contrasting patterns underscore the importance of examining age-specific dynamics in the HL-HISB relationship. They can be interpreted through the lens of information-seeking frameworks such as the Comprehensive Model of Information Seeking (CMIS), which emphasizes that information seeking is shaped by individual characteristics, perceived health-related relevance, and information-source factors [21]. This perspective is useful for understanding why HL, education, and perceived health needs may play different roles in active versus passive HISB across age groups [21,22].

Objectives

Previous studies have primarily focused on HL or HISB, either separately or within specific subgroups [23,24]. Few have analyzed their relationship using nationwide adult panel data [15,25,26], and even fewer have explored age-specific differences, despite clear theoretical and empirical reasons to do so [27]. Building on the distinction between active information seeking and passive information acquisition [5,7], this study addresses this gap by examining not only whether HL is associated with HISB, but also whether the correlates of HISB type differ across age groups. We focus on whether HL, educational attainment, and perceived health needs are differentially associated with active versus passive HISB across age groups. The South Korean context, characterized by rapid population aging alongside a highly digitalized health information environment, provides a relevant setting for examining these age-specific dynamics. Using data derived from a nationally designed panel, we investigate (1) how HISB patterns vary by age group, (2) whether HL is associated with an individual's likelihood of engaging in HISB, and (3) whether the source-based HISB type (active vs passive) varies according to HL within each age group.

Methods

Data

This cross-sectional study used data from the 2020 - 2021 Korea Health Panel (KHP) and 2021 HL Supplementary Survey, which were collected within a nationally designed panel framework. The KHP, jointly administered by the Korea Institute for Health and Social Affairs and the National Health Insurance Service, is an annual longitudinal survey that assesses patterns of health care utilization, expenditures, and key determinants, including individual health conditions and socioeconomic characteristics. The sampling frame was derived from the 2016 Korean Population and Housing Census for the second panel cohort, introduced in 2020, to ensure broad national coverage.

The KHP data structure combines cross-sectional and longitudinal elements, with repeated annual measurements collected via face-to-face interviews and self-reported questionnaires. As a longitudinal panel dataset, the KHP can capture both between-individual variation at specific time points and within-individual changes over time. However, for this study, only between-individual variation at a single time point was examined. A single cross-sectional wave was used to align with the HL survey period and to assess age-group differences in HL and health behavior.

The HL Supplementary Survey was conducted between March and July 2021 among respondents who had completed the KHP. This supplemental module included a validated and standardized tool for assessing HL among respondents from nationally designed KHP. Approximately 9530 adults aged 19 years or older completed the HL module. To ensure completeness and consistency, 1620 individuals were excluded due to missing responses on key variables, including health care utilization (n=1482), private health insurance status (n=95), high-risk drinking (n=38), smoking status (n=4), and type of health coverage (n=1). The final analytic sample consisted of 7910 adults. A flow diagram of the sample selection process is shown in Figure S1 in Multimedia Appendix 1.

Operational Definitions of HISB Experience and Type

To assess HISB in greater detail, participants were first asked whether they had sought any health-related information during the past 12 months. This binary (yes or no) variable defined the overall presence of HISB. Among those who responded affirmatively, the type of HISB was categorized as either active seeking or passive acquisition based solely on the respondent's top-ranked information source from a predefined list of 9 options. We treated this source as the respondent's primary health information source. Drawing on prior literature, we classified the sources by the mode of access typically associated with each source [3,14,26,28-31]. Information sources such as health care professionals, government websites, hospital websites, internet portal searches, and targeted searches on platforms like YouTube (Google) were assigned to the active-seeking category [3,14,28,29]. In contrast, sources such as television, radio, newspapers, magazines, books, and

XSL•FO
RenderX

interpersonal channels such as family or friends were classified under the passive acquisition category [26,29-31]. The classification scheme used to determine HISB type is summarized in Table 1. The KHP HL Supplement identified respondents' primary health information source but did not assess whether the reported source was intentionally accessed or merely encountered incidentally. Accordingly, this classification should be interpreted as a source-based proxy for active seeking and passive acquisition rather than a direct measure of behavioral intentionality. For example, some sources, including YouTube, newspapers and magazines, and interpersonal channels, may involve either active searching or passive exposure depending on the context of use.

Table . Classification of health information sources into health information–seeking behavior (HISB) types among Korean adults: a cross-sectional analysis of the 2020 - 2021 Korea Health Panel and 2021 Health Literacy Supplementary Survey[a].

Type	Sources
Active information–seeking category	• Health care professionals (doctors, nurses, pharmacists, etc) [29] • Government public sector channels (website, YouTube, and SNS[b]) [3,14] • Hospital or public health center channels (website, YouTube, and SNS) [3,14] • Internet portals [14] • YouTube (various sources) [28]
Passive information acquisition category	• Television [26,29,31] • Radio [26,29] • Newspapers, magazines, or books (print and online) [29] • Family, friends, colleagues, and acquaintances [29,30]

[a]HISB type was classified based on the respondent's top-ranked information source and represents a source-based proxy rather than a direct measure of intentionality.

[b]SNS: social network service.

Variables and Measurements

HL was measured using the European Health Literacy Survey Questionnaire-16-item (HLS-EU-Q16), a validated instrument adapted for the Korean population. This tool comprises 16 items covering health care, disease prevention, and health promotion domains. Respondents were asked to rate each item based on perceived difficulty using a four-point Likert scale: "very difficult," "difficult," "easy," and "very easy," with an additional "do not know" option. In this study, responses were dichotomized into "easy" (1 point) and "difficult" (0 points), while "do not know" responses were treated as missing. The total score ranged from 0 to 16, with higher scores indicating greater HL. The overall HL score was considered missing if a respondent had any missing values. For descriptive purposes, HL scores were categorized as inadequate (0 - 8), problematic (9-12), or adequate (13-16) [32], but in multivariable analyses, HL was used as a continuous variable.

A range of covariates was selected based on prior research linking them to HL and HISB and included to account for potential confounding. Sociodemographic variables comprised gender, age group, marital status, educational attainment, household income, employment status, type of health coverage, and private health insurance enrollment [26,33-35]. Age was categorized into three groups: 19 - 44 years (young adults), 45 - 64 years (middle-aged adults), and 65 years and older (older adults). Given the cross-sectional design, these age groups were interpreted as reflecting age-cohort differences rather than pure physiological or cognitive aging effects, because chronological age and birth cohort could not be disentangled in the current data [36,37]. Educational attainment was grouped into elementary school or less, middle or high school, and college or above. Household income was equivalized by dividing total monthly household income by the square root of the number of household members. Employment status was categorized as economically active or inactive. Health coverage included National Health Insurance and Medical Aid. Health behavior variables included current smoking, high-risk drinking, and walking [25,33,34]. Current smoking was defined as the use of conventional cigarettes or nicotine-containing electronic cigarettes at the time of the survey. High-risk drinking was defined as consuming 7 or more drinks per occasion for men (5 or more for women) at least twice a week during the prior year. Walking behavior was defined as walking for 30 minutes or more at least 5 days per week. Health status and health care utilization variables included self-rated health (SRH; on a five-point Likert scale), number of chronic diseases (categorized as none, one, or two or more), health-related quality of life (assessed using the EQ-5D index), disability status, number of outpatient visits, hospitalization experience, and whether the respondent had a usual source of care [15,26,33,35]. Unless otherwise specified, all covariates reflected values as of December 31, 2020.

Statistical Analysis

Descriptive statistics summarized participants' demographic, behavioral, and health-related characteristics. Chi-square tests were used to examine the distribution of key variables across age groups, including health information–seeking experience and behavior type. The KHP provides survey weights for the full-panel cohort, but none were calibrated specifically for the restricted analytic sample of respondents who completed the 2021 HL Supplementary Survey and had complete data on all analytic variables. Applying the full-panel weights to this complete-case subsample could yield weighted estimates that

XSL•FO
RenderX

do not accurately represent either the original target population or the final analytic sample. We therefore retained unweighted analyses as the primary approach.

Primary inferential analyses distinguished two related stages of HISB: whether respondents reported any HISB experience, and, among those who did, the source-based type of HISB. This reflected the outcome data structure, in which HISB experience was observed for the full analytic sample, whereas HISB type was defined only for respondents who reported HISB experience. First, multivariable logistic regression was used to examine the association between HL and the likelihood of seeking health information in the past 12 months (yes or no). Second, among respondents who had engaged in information seeking, an additional multivariable logistic regression model was used to assess the association between HL and the source-based HISB type (active seeking vs passive acquisition). These stage-specific models constituted the primary analyses because they addressed complementary inferential objectives: HISB engagement in the full analytic sample and HISB type conditional on engagement. For each stage, models were estimated for the relevant overall sample and separately within each age group (19‑44, 45‑64, and ≥65 years).

To examine whether the association between HL and each outcome differed across age groups, two pooled multivariable logistic regression models incorporating age×HL interaction terms were estimated in a separate sensitivity analysis: one for HISB experience in the full analytic sample and one for HISB type among respondents who had engaged in HISB. The overall significance of the age×HL interaction terms was evaluated using joint Wald χ^2 tests, and age-specific associations between HL and each outcome were estimated from the pooled interaction models. Separately, potential selection into the HISB type analytic sample was evaluated using a separate Heckman-type 2-step sensitivity analysis. Covariate coding for this sensitivity analysis was harmonized with that used in the primary HISB type model. In the first step, a probit model was used to estimate the probability of selection into the HISB type analytic sample. In the second step, the inverse Mills ratio (IMR) derived from the first-step probit model was included in a logistic regression model comparing active seeking with passive acquisition. Although the original Heckman model was developed for continuous second-stage outcomes, related extensions have been applied to binary second-stage outcomes [38-40]. All analyses were conducted using Statistical Analysis System (SAS) version 9.4 (SAS Institute Inc.), and statistical significance was determined using a two-tailed *P* value threshold of .05.

Ethical Considerations

The study was conducted in accordance with the principles outlined in the Declaration of Helsinki. The study was approved by the Institutional Review Board of the Catholic University of Korea (MC24ZASI0109), and informed consent was waived because publicly available secondary data were used. To ensure privacy and confidentiality, all data were fully anonymized and deidentified prior to access. Participant compensation was not applicable because this study involved a secondary analysis of survey data.

Results

Participant Characteristics

Among the 7910 participants, the largest proportion (3487/7910, 44.1%) was aged 65 years and older, followed by 35.6% (2815/7910) aged 45‑64 years and 20.3% (1608/7910) aged 19‑44 years. Education level differed significantly by age, with nearly half (1734/3487, 49.7%) of the oldest group having completed elementary school or below, compared with just 0.3% (4/1608) in the youngest group. Age differences were also observed in private health insurance enrollment rates, which were approximately 90% (4015/4423) among participants aged 19‑64 years but dropped to 53.1% (1852/3487) among those aged 65 years or older. A similar pattern appeared in economic activity: 68.1% (1095/1608) of individuals aged 19‑44 years were economically active, compared to 42.4% (1479/3487) of the oldest group. Health behaviors and SRH also varied with age. The oldest adults reported the highest rate of walking (1697/3487, 48.7%) and the lowest prevalence of current smoking (334/3487, 9.6%) and high-risk drinking (142/3487, 4.1%). Older age groups reported poorer SRH, with a lower proportion reporting good health and a higher proportion with multiple chronic conditions.

HISB Prevalence and Type

HISB also showed notable age-group differences. While 66.7% (1072/1608) of adults aged 19‑44 years had sought health information, only 30.2% (1052/3487) of those aged 65 years and older had done so. Furthermore, among those who had sought information, 51.8% (545/1052) of the oldest group was classified in the passive acquisition category, compared with only 7.3% (78/1072) of the youngest group. The overall mean HL score was 10.96 (SD 4.79), and mean scores were lower in older age groups (Table 2).

XSL•FO
RenderX

Table . General characteristics by age group among Korean adults: a cross-sectional analysis of the 2020 - 2021 Korea Health Panel and 2021 Health Literacy Supplementary Survey.

Variable	Age group (years)			Chi-square or *F* test (*df*)	Chi-square or *F* test (*df*)
	All (N=7910)	19 - 44 (n=1608)	45 - 64 (n=2815)		
Sex, n (%)				2.33 (2)[a]	2.33 (2)[a]
Male	3369 (42.6)	658 (40.9)	1208 (42.9)		
Female	4541 (57.4)	950 (59.1)	1607 (57.1)		
Marital status, n (%)				2004.23 (4)[a,b]	2004.23 (4)[a,b]
Married	5715 (72.3)	1004 (62.4)	2277 (80.9)		
Divorced, separated, or widowed	1481 (18.7)	45 (2.8)	415 (14.7)		
Not married	714 (9.0)	559 (34.8)	123 (4.4)		
Education level, n (%)				3508.29 (4)[a,b]	3508.29 (4)[a,b]
Elementary school or below	2001 (25.3)	4 (0.3)	263 (9.3)		
Middle or high school	3442 (43.5)	356 (22.1)	1618 (57.5)		
College or above	2467 (31.2)	1248 (77.6)	934 (33.2)		
Health care coverage, n (%)				49.29 (2)[a,b]	49.29 (2)[a,b]
National Health Insurance	7561 (95.6)	1581 (98.3)	2701 (96.0)		
Medical Aid	349 (4.4)	27 (1.7)	114 (4.1)		
Private health insurance, n (%)				1447.63 (2)[a,b]	1447.63 (2)[a,b]
Yes	5867 (74.2)	1487 (92.5)	2528 (89.8)		
No	2043 (25.8)	121 (7.5)	287 (10.2)		
Economic activity, n (%)				576.82 (2)[a,b]	576.82 (2)[a,b]
Yes	4543 (57.4)	1095 (68.1)	1969 (70.0)		
No	3367 (42.6)	513 (31.9)	846 (30.1)		
Current smoking, n (%)				121.34 (2)[a,b]	121.34 (2)[a,b]
Yes	1146 (14.5)	294 (18.3)	518 (18.4)		
No	6764 (85.5)	1314 (81.7)	2297 (81.6)		
High-risk drinking, n (%)				194.50 (2)[a,b]	194.50 (2)[a,b]
Yes	723 (9.1)	223 (13.9)	358 (12.7)		
No	7187 (90.9)	1385 (86.1)	2457 (87.3)		
Walking, n (%)				7.20 (2)[a,c]	7.20 (2)[a,c]
Yes	3740 (47.3)	718 (44.7)	1325 (47.1)		
No	4170 (52.7)	890 (55.4)	1490 (52.9)		
Self-rated health, n (%)				683.35 (8)[a,c]	683.35 (8)[a,c]
Very good	246 (3.1)	88 (5.5)	76 (2.7)		
Good	2511 (31.7)	723 (45.0)	990 (35.2)		

XSL•FO
RenderX

Variable	Age group (years)			Chi-square or *F* test (*df*)	Chi-square or *F* test (*df*)
	All (N=7910)	19 - 44 (n=1608)	45 - 64 (n=2815)		
Fair	3448 (43.6)	645 (40.1)	1369 (48.6)		
Poor	1590 (20.1)	147 (9.1)	361 (12.8)		
Very poor	115 (1.5)	5 (0.3)	19 (0.7)		
Number of chronic diseases, n (%)				3060.36 (4)[a,b]	
0	2739 (34.6)	1280 (79.6)	1146 (40.7)		
1	1764 (22.3)	267 (16.6)	816 (29.0)		
≥2	3407 (43.1)	61 (3.8)	853 (30.3)		
Disability, n (%)				223.74 (2)[a,b]	
Yes	594 (7.5)	27 (1.7)	137 (4.9)		
No	7316 (92.5)	1581 (98.3)	2678 (95.1)		
Hospital admission, n (%)				113.14 (2)[a,b]	
Yes	1142 (14.4)	146 (9.1)	332 (11.8)		
No	6768 (85.6)	1462 (90.9)	2483 (88.2)		
Usual source of care, n (%)				771.53 (2)[a,b]	
Yes	5015 (63.4)	639 (39.7)	1634 (58.1)		
No	2895 (36.6)	969 (60.3)	1181 (42.0)		
Health information seeking experience, n (%)				790.75 (2)[a,b]	
Yes	3768 (47.6)	1072 (66.7)	1644 (58.4)		
No	4142 (52.4)	536 (33.3)	1171 (41.6)		
HISB[d] type[e], n (%)				572.96 (2)[a,b]	
Active information-seeking category	2785 (73.9)	993 (92.7)	1285 (78.2)		
Passive information-acquisition category	982 (26.1)	78 (7.3)	359 (21.8)		
Monthly household income (1000 KRW)[g], mean (SD)	2,323.41 (1,713.49)	2,976.28 (1,677.64)	2,764.25 (1,842.94)	528.33 (2, 7907)[b,f]	
Health-related quality of life (EQ-5D), mean (SD)	0.94 (0.10)	0.98 (0.05)	0.96 (0.07)	423.88 (2, 7907)[b,f]	
Number of outpatient visits, mean (SD)	19.81 (23.02)	10.10 (13.60)	15.93 (19.81)	411.89 (2, 7907)[b,f]	

XSL•FO
RenderX

Variable	Age group (years)			Chi-square or *F* test (*df*)	Chi-square or *F* test (*df*)
	All (N=7910)	19 - 44 (n=1608)	45 - 64 (n=2815)		
Health literacy level, mean (SD)	10.96 (4.79)	14.10 (2.94)	12.43 (4.06)	1,337.46 (2, 7907)[b,f]	

[a]Chi-square test.

[b]*P*<.001.

[c]*P*<.05.

[d]HISB: health information–seeking behavior.

[e]Among participants who sought health information (n=3767), excluding 1 respondent who selected “other” as the only source.

[f]*F* test.

[g]KRW: Korean Won.

Factors Associated With HISB Experience by Age Group

In the first stage of the primary analysis, a logistic regression analysis was used to examine factors associated with HISB experience. HL was significantly associated with HISB experience; participants with higher HL scores had higher odds of seeking health information (odds ratio [OR] 1.139, 95% CI 1.123 - 1.154; *P*<.001). Other variables significantly associated with HISB experience included age, gender, marital status, education, income, private health insurance, current smoking status, health-related quality of life, and having a usual source of care. Compared with adults aged 65 years and older, those aged 19 - 44 years had more than twice the odds of seeking health information (OR 2.072, 95% CI 1.690 - 2.541), and those aged 45 - 64 years also had higher odds (OR 1.728, 95% CI 1.502 - 1.988). Conversely, men, individuals with lower education levels, current smokers, and those with higher EQ-5D scores had lower odds of seeking health information. Age-stratified logistic regression analyses showed different patterns of association across age groups. Among participants aged 19 - 44 years, information seeking was more likely among women, married participants, those with two or more chronic conditions, and those with a usual source of care. Significant correlates of HISB experience in the 45 - 64 age group included gender, education level, household income, current smoking, SRH, health-related quality of life, having a usual source of care, and HL. Among adults aged 65 years and older, education level, private health insurance, current smoking, number of outpatient visits, and HL were significantly associated with HISB experience (Table 3).

XSL•FO
RenderX

Table . Factors associated with health information–seeking experience by age group among Korean adults: a cross-sectional logistic regression analysis of the 2020 - 2021 Korea Health Panel and 2021 Health Literacy Supplementary Survey.

Variables	Age group (years)			
	All (N=7910)	19 - 44 (n=1608)	45 - 64 (n=2815)	65+ (n=3487)
	OR[a] (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Age (ref[b]=65+)				
19 - 44	2.072[c] (1.690 - 2.541)	—[d]	—	—
45 - 64	1.728[e] (1.502 - 1.988)	—	—	—
Sex (ref=female)				
Male	0.753[c] (0.667 - 0.850)	0.632[e] (0.489 - 0.816)	0.633[c] (0.514 - 0.778)	0.918 (0.759 - 1.112)
Marital status (ref=not married)				
Married	1.655[c] (1.364 - 2.008)	1.798[f] (1.422 - 2.274)	0.877 (0.572 - 1.346)	0.834 (0.357 - 1.949)
Divorced, separated, or widowed	1.598[f] (1.264 - 2.019)	1.234 (0.618 - 2.464)	1.085 (0.682 - 1.727)	0.784 (0.334 - 1.838)
Education level (ref=college or above)				
Elementary school or below	0.342[c] (0.283 - 0.412)	0.161 (0.015 - 1.734)	0.245[c] (0.174 - 0.345)	0.448[c] (0.330 - 0.609)
Middle or high school	0.616 (0.540 - 0.704)	0.676 (0.517 - 0.885)	0.552 (0.456 - 0.669)	0.656 (0.494 - 0.871)
Monthly household income (in Korean won)	1.000[e] (1.000 - 1.001)	1.000 (0.999 - 1.001)	1.001[f] (1.000 - 1.001)	1.001 (1.000 - 1.001)
Health care coverage (ref=NHI)[g]				
Medical aid	0.838 (0.636 - 1.104)	1.515 (0.608 - 3.775)	0.615 (0.377 - 1.005)	0.805 (0.549 - 1.181)
Private health insurance (ref=no)				
Yes	1.217[e] (1.063 - 1.394)	1.522[f] (1.002 - 2.314)	1.004 (0.752 - 1.341)	1.208[f] (1.017 - 1.435)
Economic activity (ref=no)				
Yes	0.948 (0.846 - 1.063)	1.017 (0.790 - 1.310)	0.945 (0.771 - 1.158)	0.908 (0.764 - 1.080)
Current smoking (ref=no)				
Yes	0.772[e] (0.661 - 0.902)	0.872 (0.639 - 1.190)	0.775[f] (0.612 - 0.983)	0.652[e] (0.485 - 0.876)
High-risk drinking (ref=no)				
Yes	0.885 (0.741 - 1.057)	0.935 (0.678 - 1.289)	0.857 (0.664 - 1.105)	0.915 (0.611 - 1.372)
Walking (ref=no)				
Yes	1.054 (0.953 - 1.166)	1.048 (0.840 - 1.308)	1.082 (0.918 - 1.274)	0.986 (0.837 - 1.162)
Self-rated health (ref=very poor)				
Very good	0.961 (0.549 - 1.683)	5.075 (0.597 - 43.146)	0.436[f] (0.140 - 1.353)	0.965 (0.429 - 2.175)
Good	1.045 (0.641 - 1.706)	5.284 (0.652 - 42.921)	0.435 (0.155 - 1.218)	1.163 (0.623 - 2.169)
Fair	1.370[e] (0.847 - 2.215)	6.106 (0.756 - 49.293)	0.657 (0.237 - 1.824)	1.296 (0.711 - 2.362)
Poor	1.125 (0.702 - 1.803)	6.717 (0.814 - 55.429)	0.470 (0.168 - 1.310)	1.117 (0.627 - 1.991)
Number of chronic diseases (ref=0)				
1	1.097 (0.945 - 1.274)	1.144 (0.831 - 1.574)	0.954 (0.741 - 1.178)	1.229 (0.889 - 1.698)
≥2	1.044 (0.887 - 1.228)	3.353[**] (1.411 - 7.972)	0.936 (0.741 - 1.184)	1.212 (0.892 - 1.646)
Health-related quality of life (EQ-5D)	0.303[e] (0.162 - 0.569)	0.155 (0.013 - 1.857)	0.043[c] (0.011 - 0.160)	0.531 (0.235 - 1.202)
Disability (ref =no)				
Yes	0.854 (0.691 - 1.056)	0.423 (0.165 - 1.083)	0.687 (0.455 - 1.036)	0.941 (0.726 - 1.219)

XSL•FO
RenderX

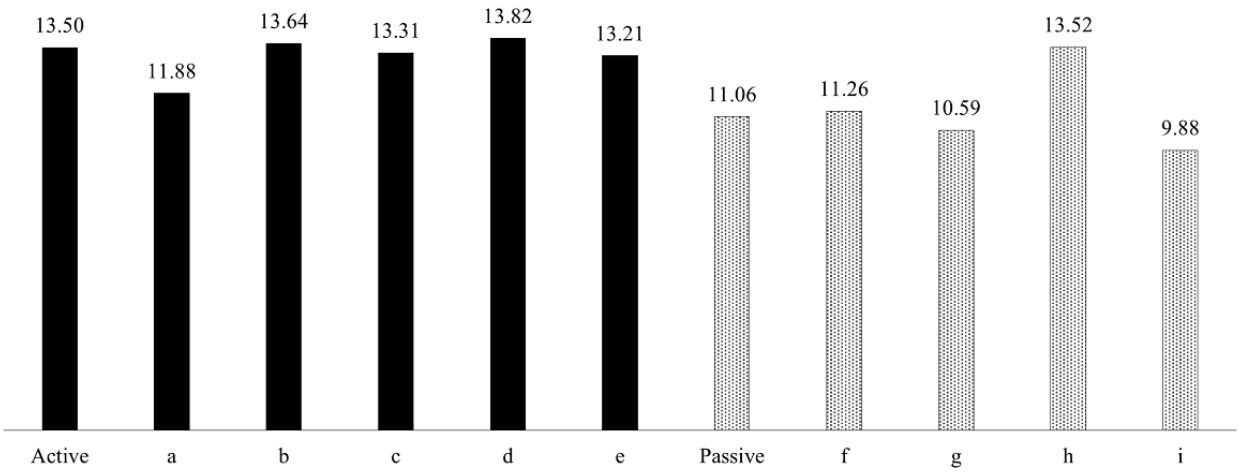
Variables	Age group (years)			
	All (N=7910)	19 - 44 (n=1608)	45 - 64 (n=2815)	65+ (n=3487)
	OR[a] (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Hospital admission (ref =no)				
Yes	0.995 (0.858 - 1.154)	0.943 (0.628 - 1.416)	0.962 (0.741 - 1.247)	0.997 (0.810 - 1.226)
Number of outpatient visits	1.002 (1.000 - 1.005)	1.007 (0.995 - 1.020)	0.999 (0.994 - 1.003)	1.004[f] (1.000 - 1.007)
Usual source of care (ref =no)				
Yes	1.518[c] (1.347 - 1.711)	1.819[c] (1.426 - 2.320)	1.579[c] (1.311 - 1.900)	1.214 (0.980 - 1.503)
Health literacy	1.139[c] (1.123 - 1.154)	1.011 (0.972 - 1.051)	1.138[c] (1.113 - 1.163)	1.178[c] (1.154 - 1.202)

[a]OR: odds ratio.
[b]ref: reference.
[c]$P<.001$.
[d]Not applicable.
[e]$P<.01$.
[f]$P<.05$.
[g]NHI: National Health Insurance.

HL Scores by HISB Type and Information Source

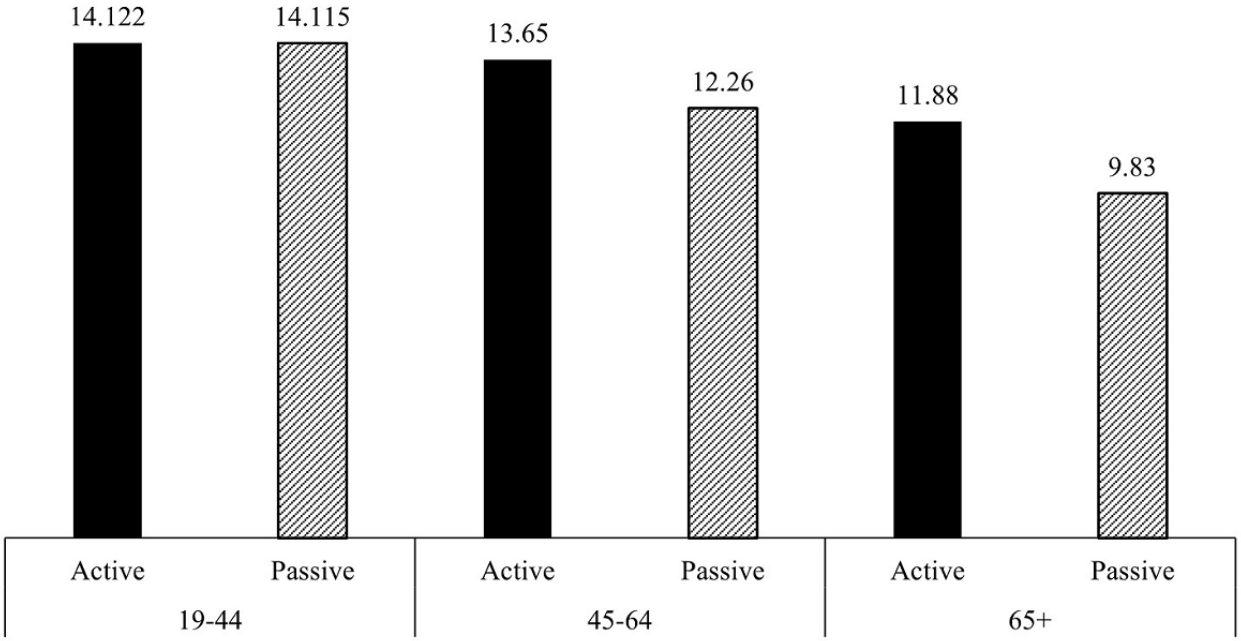
Figures 1 and 2 illustrate differences in HL according to HISB type and information source. Overall, individuals classified in the active-seeking category had higher mean HL scores (mean 13.50, SD 3.18) than those in the passive acquisition category (mean 11.06, SD 4.17), with the highest mean score observed among users of internet portals (13.82, SD 2.94). Within the passive acquisition category, printed media users had the highest mean HL score (13.52, SD 2.99), whereas those relying on interpersonal sources (eg, family or acquaintances) had the lowest mean score (9.88, SD 4.26; Figure 1). Across age groups, mean HL scores were higher among individuals classified in the active-seeking category than among those in the passive acquisition category. However, the difference was smallest among adults aged 19 - 44 years and was virtually absent in this group (14.122 vs 14.115; Figure 2).

Figure 1. Health literacy scores by source-based health information–seeking behavior type and source among Korean adults: a cross-sectional analysis of the 2020 - 2021 Korea Health Panel and 2021 health literacy supplementary survey. Active: active information seeking: (a) health care professionals (doctors, nurses, pharmacists, etc); (b) government or public sector channels (website, YouTube, and social network service); (c) hospital or public health channels (website, YouTube, social network service); (d) internet portals; e: YouTube (various sources). Passive: passive information acquisition: (f) television; (g) radio; (h) newspapers, magazines, or books (print and online); (i) family, friends, colleagues, and acquaintances.



XSL•FO
RenderX

Figure 2. Health literacy scores by age group and source-based health information–seeking behavior type among Korean adults: a cross-sectional analysis of the 2020 - 2021 Korea Health Panel and 2021 Health Literacy Supplementary Survey. Active: active information seeking; passive: passive information acquisition; age groups (years): 19 - 44, 45 - 64, and 65+.



Factors Associated With HISB Type by Age Group

In the second stage of the primary analysis, logistic regression showed that participants with higher HL scores had higher odds of classification in the active-seeking category (OR 1.105, 95% CI 1.080-1.131; $P<.001$). Other significant correlates of HISB type included age, education level, SRH, number of chronic diseases, and health-related quality of life. Participants with higher education, fewer chronic diseases, and poorer SRH had higher odds of being classified as active-seeking. Compared with adults aged 65 years and older, individuals aged 19 - 44 years had more than six times the odds of active-seeking classification (OR 6.072, 95% CI 4.374 - 8.428), while those aged 45 - 64 years had more than twice the odds (OR 2.582, 95% CI 2.096 - 3.180). Age-stratified models showed further differences. In the youngest group, education level, SRH, and disability were significantly associated with HISB type, while HL was not. In this group, individuals with disabilities had higher odds of classification in the active-seeking category. Among adults aged 65 years and older, higher HL and higher EQ-5D scores were significant correlates of classification in the active-seeking category (Table 4).

XSL•FO
RenderX

Table . Factors associated with source-based health information–seeking behavior (HISB) type by age group among Korean adults: a cross-sectional logistic regression analysis of the 2020 - 2021 Korea Health Panel and 2021 Health Literacy Supplementary Survey[a].

Variables	Age group (years)			
	All (n=3767)	19 - 44 (n=1071)	45 - 64 (n=1644)	65+ (n=1052)
	OR[b] (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Age group, years (ref[c]=65+)				
19 - 44	6.072[f] (4.374 - 8.428)	—[g]	—	—
45 - 64	2.582[d] (2.096 - 3.180)	—	—	—
Sex (ref=female)				
Male	1.017 (0.836 - 1.238)	0.711 (0.410 - 1.233)	0.807 (0.596 - 1.092)	1.329 (0.980 - 1.802)
Marital status (ref=unmarried)				
Married	0.973 (0.796 - 1.189)	1.108 (0.659 - 1.863)	0.802 (0.575 - 1.119)	0.927 (0.675 - 1.274)
Education level (ref=college or above)				
High school or below	0.639[f] (0.520 - 0.784)	0.555[f] (0.320 - 0.963)	0.592[e] (0.447 - 0.783)	0.757 (0.510 - 1.122)
Monthly household income	1.000 (1.000 - 1.001)	0.999 (0.998 - 1.001)	1.000 (0.999 - 1.001)	1.001 (1.000 - 1.002)
Private health insurance (ref=no)				
Yes	1.025 (0.820 - 1.280)	0.718 (0.231 - 2.229)	1.008 (0.649 - 1.568)	1.016 (0.767 - 1.348)
Economic activity (ref=no)				
Yes	0.885 (0.736 - 1.064)	0.921 (0.531 - 1.599)	0.785 (0.584 - 1.056)	0.969 (0.735 - 1.277)
Current smoking (ref=no)				
Yes	0.952 (0.724 - 1.252)	1.198 (0.593 - 2.421)	1.116 (0.754 - 1.650)	0.686 (0.419 - 1.123)
High-risk drinking (ref=no)				
Yes	0.991 (0.729 - 1.348)	0.596 (0.322 - 1.105)	1.252 (0.819 - 1.915)	0.982 (0.505 - 1.907)
Walking (ref=no)				
Yes	1.065 (0.902 - 1.257)	1.129 (0.698 - 1.827)	0.881 (0.690 - 1.126)	1.262 (0.968 - 1.644)
Self-rated health (ref=poor)				
Good	0.921 (0.700 - 1.211)	0.320[d] (0.102 - 0.999)	0.947 (0.608 - 1.476)	0.972 (0.647 - 1.460)
Fair	0.785[d] (0.617 - 0.999)	0.430 (0.142 - 1.302)	0.757 (0.509 - 1.124)	0.795 (0.564 - 1.121)
Number of chronic diseases (ref=0)				
1	0.775 (0.609 - 0.986)	1.318 (0.670 - 2.595)	0.644 (0.467 - 0.887)	0.962 (0.576 - 1.605)
≥2	0.660[e] (0.513 - 0.850)	1.257 (0.393 - 4.020)	0.551[e] (0.392 - 0.776)	0.821 (0.506 - 1.332)
Health-related quality of life (EQ-5D; ref=<1)				
1	1.324[e] (1.104 - 1.589)	1.578 (0.878 - 2.837)	1.196 (0.916 - 1.562)	1.396[d] (1.046 - 1.865)
Disability (ref=no)				
Yes	0.812 (0.576 - 1.144)	0.217[d] (0.049 - 0.961)	0.811 (0.437 - 1.505)	0.877 (0.571 - 1.348)
Hospital admission (ref=no)				
Yes	1.118 (0.879 - 1.421)	1.292 (0.538 - 3.100)	1.172 (0.789 - 1.742)	1.043 (0.743 - 1.465)
Number of outpatient visits	1.002 (0.998 - 1.006)	0.990 (0.978 - 1.003)	1.002 (0.994 - 1.010)	1.004 (0.999 - 1.009)
Usual source of care (ref=no)				
Yes	1.081 (0.888 - 1.316)	0.862 (0.522 - 1.425)	1.065 (0.806 - 1.408)	1.293 (0.909 - 1.840)
Health literacy	1.105[f] (1.080 - 1.131)	0.999 (0.915 - 1.091)	1.111[f] (1.073 - 1.151)	1.122[f] (1.084 - 1.161)

XSL•FO
RenderX

[a]Active seeking was coded as 1 and passive acquisition as 0. HISB type was classified based on the respondent's top-ranked information source.
[b]OR: odds ratio.
[c]ref: reference.
[d]*P*<.05.
[e]*P*<.01.
[f]*P*<.001.
[g]Not applicable.

Sensitivity Analyses

In pooled multivariable logistic regression models including age×HL interaction terms, the joint test of the interaction terms was statistically significant for both HISB experience (Wald χ^2_2=57.84; *P*<.001) and HISB type (Wald χ^2_2=7.79; *P*=.02). Age-specific estimates showed that HL was significantly associated with HISB experience and classification in the active-seeking category among adults aged 45 - 64 and ≥65 years, but not among those aged 19 - 44 years. These findings were consistent with the age-stratified primary analyses (Table S1 in Multimedia Appendix 2).

In a separate sensitivity analysis, the Heckman-type 2-step analysis showed that respondents with higher HL scores had higher odds of classification in the active-seeking category in the primary conditional logistic model among those who engaged in HISB (OR 1.105, 95% CI 1.080 - 1.130; *P*<.001). After the IMR was included in the second-stage logistic model, the IMR was statistically significant (OR 0.147, 95% CI 0.062 - 0.348; *P*<.001), whereas the estimate for HL was attenuated and no longer statistically significant (OR 0.995, 95% CI 0.945 - 1.048; *P*=.86). These findings indicate that the estimated association between HL and HISB type was sensitive to accounting for potential selection into the HISB type analytic sample (Table S2 in Multimedia Appendix 3).

Discussion

Principal Findings

This study demonstrated a distinct age-group pattern in HL, with higher mean HL scores among younger adults and lower mean scores among older adults. Given the cross-sectional design, this pattern should be interpreted as reflecting a combination of age and cohort differences rather than as evidence of HL decline caused by aging alone [36]. Differences in cognitive and sensory functions [16] and communication barriers in health care settings may partly contribute to the lower HL scores observed among older adults [24]. At the same time, cohort-specific differences in educational opportunities and formative exposure to digital technologies may also have contributed to distinct historical and technological environments [41]. These factors may also shape whether individuals engage in HISB and, among those who do, the source-based type of HISB observed in this study.

Table 5 provides an integrative summary of the main age-group patterns observed in this study. The matrix shows that the main correlates of HISB varied across age groups, with education level and health-related needs more salient among younger adults and HL emerging as a more prominent correlate among middle-aged and older adults, particularly for classification in the active-seeking category in the primary analyses. These contrasts may reflect both age-related differences in health needs and cohort-specific educational and information environments [36,37], consistent with recent evidence that information-source preferences differ across generations [41]. Viewed through the CMIS, these findings can be understood as age-group differences in the relative importance of individual antecedents, perceived health-related salience, and information-source factors rather than as a direct test of the full model [21,22]. In younger adults, education and perceived health needs may represent key antecedents and salience-related factors associated with information seeking, whereas HL may represent a more prominent cognitive and experiential resource among middle-aged and older adults, potentially supporting the recognition of information needs, evaluation of available channels, and engagement with health information (Table 5).

XSL•FO
RenderX

Table . Summary matrix of main correlates of health information–seeking behavior by age group among Korean adults: a cross-sectional analysis of the 2020 - 2021 Korea Health Panel and 2021 Health Literacy Supplementary Survey.

Age group (years)	Observed HISB[a] pattern	Main correlates	Interpretive summary
19 - 44	Predominantly active seeking	Education level; health-related needs[b]	HISB was common and active, but HL[c] was not significantly associated with HISB.
45 - 64	Predominantly active seeking, with transitional features	HL; Education level; health-related needs[b]	HL was a significant correlate, alongside education and health-related needs[b].
65+	Higher reliance on passive acquisition	HL; health-related quality of life	HL was significantly associated with classification in the active-seeking category.

[a]HISB: health information–seeking behavior.

[b]Health-related needs refer to health status indicators such as self-rated health, number of chronic diseases, disability status, or health-related quality of life, which were associated with HISB in different ways across age groups.

[c]HL: health literacy.

Building on these patterns, this analysis examined how HL, sociodemographic factors, and health-related factors were associated with the HISB experience and source-based HISB type across age groups. In the primary age-stratified analyses, HL was significantly associated with both HISB experience and classification in the active-seeking category among adults aged 45 years and older, but not among those aged 19 - 44 years. The significant age ×HL interaction further supported heterogeneity in these associations across age groups. However, in the Heckman-type sensitivity analysis, the association between HL and HISB type was attenuated after accounting for potential selection into the HISB type analytic sample. Taken together, these findings suggest that HL may play a more prominent role in HISB among middle-aged and older adults, while the independent association between HL and HISB type should be interpreted cautiously because it was sensitive to the modeling of selection. This age-group pattern may reflect differences in cognitive processing needs, reliance on health care services, or familiarity with different information channels. The latter may also reflect cohort-specific exposure to changing media and digital environments [41]. In this context, HL may function not merely as a set of skills or knowledge but as a cognitive asset that facilitates engagement with, retention of, and application of health information [41,42]. This interpretation is consistent with a recent scoping review identifying eHealth literacy as an important individual resource for online HISB, but also showing that its role is embedded within broader educational, socioeconomic, health care, and digital-environmental conditions [43]. Prior evidence also points to a close relationship between HL and HISB, although its direction remains uncertain. Information seeking was more strongly associated with healthy lifestyle behaviors among individuals with higher HL [9], while older adults who regularly sought health content online had higher HL [10].

Among adults aged 19 - 44 years, health information seeking was widespread, and more than 90% (993/1072) of seekers were classified in the active-seeking category. Despite this high engagement, HL was not significantly associated with either HISB experience or source-based HISB type in this group. Instead, the correlates differed by outcome: gender, marital status, multiple chronic conditions, and having a usual source of care were associated with HISB experience, whereas education level, SRH, and disability were associated with HISB type. While HL and education are often correlated, education, but not HL, was associated with classification in the active-seeking category, possibly reflecting broader educational resources, health awareness, or confidence in navigating information sources rather than specific information-processing skills [44]. Additionally, among those who engaged in HISB, individuals with better SRH had lower odds of classification in the active-seeking category, suggesting that perceived health-related need may be relevant to source-based information seeking [45,46]. These findings suggest that health-related needs and sociodemographic factors may be more salient correlates than HL among younger adults. However, because passive acquisition was uncommon in this age group, estimates for HISB type should be interpreted cautiously, and the nonsignificant association with HL should not be interpreted as evidence that HL is unimportant. A similar caveat applies to the predominance of classification in the active-seeking category among younger adults: this source-based classification does not directly capture intentionality, and frequent digital health information seeking does not necessarily translate into greater health empowerment or effective health management [47]. Therefore, interventions for this age group may benefit from emphasizing preventive care and future health risks while also strengthening the appraisal of digital health information and connections between digital information-seeking environments and professional guidance, including offline or telehealth-based support [47,48].

In the 45 - to 64-year age group, the primary age-stratified analyses showed that HL and education were significantly associated with both HISB experience and source-based HISB type, whereas the number of chronic diseases was associated with HISB type. In the HISB type model, higher HL and education were associated with higher odds of classification in the active-seeking category [14,49], whereas a greater number of chronic diseases was associated with lower odds of such classification [2,50,51]. This finding contrasts with prior research and may reflect a gap between the need for information

XSL•FO
RenderX

and the capacity to engage, potentially due to information overload, limited digital skills, or cognitive fatigue among individuals managing multiple health conditions [17,19]. These barriers are likely to be more pronounced among socioeconomically disadvantaged groups, where resource constraints can hinder the meaningful use of information [18]. In this context, HL may serve as an important resource for navigating information-access and processing challenges. Thus, strengthening HL, including digital HL, and ensuring access to accessible, trustworthy offline resources, such as community clinics and health centers, could support effective engagement with health information [52,53]. Tailored interventions are particularly relevant for middle-aged adults with limited incomes and lower education levels, who often face greater barriers to effective health communication [44].

Adults aged 65 years and older exhibited markedly lower levels of HISB, both in prevalence and in the proportion classified in the active-seeking category, compared with younger groups. In the primary age-stratified HISB type model, higher HL and EQ-5D scores were significantly associated with higher odds of classification in the active-seeking category. The association with EQ-5D suggests that overall functional health may be relevant to the capacity to engage with health information [2,51]. These findings indicate that HISB among older adults may be shaped not merely by perceived health needs, but also by the availability of cognitive and structural resources [14,50]. Consistent with the CMIS-informed interpretation introduced above, HL can be understood as an important resource for navigating cognitive, functional, and channel-related barriers to active engagement with health information [22,54]. At the same time, the higher proportion classified in the passive-acquisition category among older adults warrants attention, as prior media research has shown that health-related news may frame health problems in ways that emphasize individual responsibility over broader social determinants [55]. Such framing may be especially consequential for older adults with lower HL, who may have fewer opportunities or resources to verify, compare, or contextualize health information [54,56]. Therefore, interventions for older adults could combine efforts to strengthen HL with the reduction of access barriers through user-friendly technologies and community-based education tailored to their needs [53].

Taken together, these findings point to the value of strategies tailored to the needs and information environments of different age groups to support equitable and effective engagement with health information [48], consistent with recent evidence that health information source preferences differ across generations [41]. For older adults, who showed lower HISB prevalence and a higher proportion classified in the passive-acquisition category, interventions should not simply encourage more digital information seeking [54]. Rather, health systems could provide accessible proactive communication methods, such as short message service health alerts or automated reminders, while reducing digital barriers through user-friendly national health portals, community-based digital education, and peer-based technology mentoring programs [54,56]. These approaches may offer trusted entry points for older adults to gradually engage with digital health information while maintaining support from the community and health care professionals [48]. For younger adults, among whom most seekers were classified in the active-seeking category, the priority should shift from increasing information seeking itself to strengthening their ability to assess the credibility of, interpret, and apply the information they encounter [47]. Linking commonly used digital platforms with professional guidance, such as telehealth services or verified clinician-led information channels, may support more informed and effective use of health information [47,48]. As digital health environments increasingly include generative AI tools for health-related inquiries, tailored digital HL education may also be important for helping users critically evaluate AI-generated health information and reducing the risk of emerging informational inequalities across age groups [57].

Limitations

This study has several limitations. First, because of its cross-sectional design, causal inferences cannot be drawn, and the temporal relationships between HL and HISB remain unclear. Because chronological age and birth cohort cannot be disentangled in cross-sectional data, the observed age-group pattern should be interpreted as potentially reflecting both age and cohort differences, rather than as evidence of physiological or cognitive changes attributable to aging alone [36]. For example, the lower prevalence of HISB and the higher proportion classified in the passive-acquisition category among older adults may partly reflect cohort differences in formative exposure to the internet and digital technologies, rather than aging alone [47].

Second, the analysis was limited to respondents who participated in the HL supplement of the KHP and had complete data for the variables included in the analyses, a subsample of the broader national cohort. As appropriate weights specifically constructed for this restricted analytic subsample were not available, the analyses were conducted without applying survey weights. Therefore, the reported estimates should be interpreted as unweighted associations within the analytic sample rather than as nationally representative population estimates, and their generalizability may be limited.

Third, unmeasured confounding may persist despite adjustment for several socioeconomic and baseline health status variables. In particular, the KHP dataset did not include detailed measures of physical access to digital devices, internet affordability, or perceived accessibility and usability of online health information channels, all of which may influence both HL and HISB, particularly among older adults and other groups with limited digital resources [22,47,54,56]. Other potentially relevant factors, such as digital self-efficacy, prior experience with online health services, trust in specific information sources, and cognitive functioning, were also not directly measured. A recent scoping review among individuals with diabetes identified multilevel determinants of online HISB spanning individual, interpersonal, health care, and digital-environmental domains, further underscoring the possibility that relevant contextual factors were not fully captured in the present analysis [43].

Fourth, HISB type was operationalized as a source-based proxy using the respondent's highest-ranked information source. This approach did not directly assess intentionality and could not

XSL•FO
RenderX

capture the simultaneous use of multiple sources, changes in information-seeking mode across situations, or variation in how actively individuals engaged with a given source. This binary classification into active seeking and passive acquisition may not fully capture the complexity of real-world HISB.

Fifth, the HISB type analysis was conditional on respondents having engaged in HISB, creating the possibility of selection into the analytic sample. In the Heckman-type sensitivity analysis, the estimated association between HL and HISB type was attenuated after the IMR was included. Accordingly, the primary findings regarding HL and HISB type should be interpreted with caution because they were sensitive to modeling potential selection.

Sixth, passive acquisition was uncommon among adults aged 19‑44 years. The relatively small number of observations in this outcome category may have reduced the precision and stability of estimates in the youngest age group. Consequently, the nonsignificant association between HL and HISB type in this group should not be interpreted as evidence that no association exists.

Seventh, although HL was treated as a focal independent variable in this study, its role may be more nuanced, potentially functioning as a mediator or moderator in broader health engagement pathways. The analyses also could not evaluate reciprocal relationships in which engagement with health information may itself contribute to HL over time. Additionally, the assumption that classification in the active-seeking category is inherently beneficial warrants caution, as active engagement with health information may sometimes lead to information overload or exposure to misinformation [52].

Eighth, the dataset did not allow us to account for potentially important age-varying factors, including eHealth literacy and trust in specific media or information sources, that may shape the association between HL and HISB. Among younger adults, the ability and confidence to navigate online platforms may be more salient than general HL, whereas older adults' greater reliance on passive acquisition may partly reflect trust in familiar sources such as television or interpersonal networks [29,41]. Older adults' engagement with digital health services may also be constrained by technology-related anxiety, privacy concerns, and concerns that digitally mediated care could weaken interpersonal connections, which may reinforce a preference for in-person care [58]. The analyses did not capture patient activation or discussions of online health information with health care providers, which may influence whether information seeking translates into patient-centered communication [59]. Future studies should incorporate these digital, psychological, and interpersonal factors to clarify the mechanisms underlying age-group differences in HL and HISB. Finally, data were collected during the COVID-19 pandemic, and changes in health service accessibility and communication patterns may have influenced HL and information-seeking behaviors.

Conclusions

This study identified clear age-group patterns in the association between HL and HISB among Korean adults, underscoring that the role of HL and source-based HISB type differed across age groups. In the primary analyses, higher HL was associated with HISB experience and classification in the active-seeking category among middle-aged and older adults, whereas educational attainment and health-related needs were more salient among younger adults. However, because the association between HL and HISB type was sensitive to accounting for potential selection into the HISB type analytic sample, this specific association should be interpreted cautiously. At the same time, the broader age-group patterns observed across HISB experience and source-based HISB type highlight the value of designing strategies tailored to the needs and information environments of different age groups to improve access to and effective use of health information. Rather than focusing solely on strengthening HL, future interventions should also consider structural, motivational, and digital access factors to promote equitable engagement with health information across age groups.

Acknowledgments

During the revision process, the authors used generative AI (ChatGPT) to assist with language editing, the restructuring of selected passages, and the refinement of table, figure, and text presentation. ChatGPT was not used to generate or analyze data, conduct statistical analyses, interpret the results independently, or create references. It was not treated as a source of evidence; all factual statements, interpretations, citations, and reference details were reviewed and verified by the authors against the underlying data and original sources. All AI-assisted content was critically reviewed, revised, and finalized by the authors, who take full responsibility for the accuracy, integrity, and final content of the manuscript.

Funding

The authors declared no financial support was received for this work.

Data Availability

The raw data supporting this study's findings are publicly available from the Korea Health Panel Survey with permission. However, access to processed data is limited because the Korea Health Panel Survey bylaws prohibit the free sharing of data between individuals. The processed data is available upon reasonable request and with the permission of the National Health Insurance Services and the Korea Institute for Health and Social Affairs

XSL•FO
RenderX

Authors' Contributions

Conceptualization: DHK (lead), YDK (equal).
Formal analysis: DHK.
Methodology: DHK (lead), YDK (equal).
Writing – original draft: DHK.
Writing – review and editing: YDK.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Flow diagram of sample selection among Korean adults.
[DOCX File, 72 KB - publichealth_v12i1e86992_app1.docx]

Multimedia Appendix 2
Sensitivity analysis of age×health literacy interactions for health information–seeking experience and behavior type among Korean adults.
[DOCX File, 17 KB - publichealth_v12i1e86992_app2.docx]

Multimedia Appendix 3
Sensitivity analysis comparing logistic models of HISB type with and without the inverse Mills ratio among Korean adults who engaged in HISB. HISB: health information–seeking behavior.
[DOCX File, 18 KB - publichealth_v12i1e86992_app3.docx]

References

1. Mengiste M, Ahmed MH, Bogale A, Yilma T. Information-seeking behavior and its associated factors among patients with diabetes in a resource-limited country: a cross-sectional study. Diabetes Metab Syndr Obes 2021;14:2155-2166. [doi: 10.2147/DMSO.S289905] [Medline: 34040402]
2. Levy H, Janke A. Health literacy and access to care. J Health Commun 2016;21 Suppl 1(Suppl):43-50. [doi: 10.1080/10810730.2015.1131776] [Medline: 27043757]
3. Neter E, Brainin E. eHealth literacy: extending the digital divide to the realm of health information. J Med Internet Res 2012 Jan 27;14(1):e19. [doi: 10.2196/jmir.1619] [Medline: 22357448]
4. Gedefaw A, Yilma TM, Endehabtu BF. Information seeking behavior about cancer and associated factors among university students, Ethiopia: a cross-sectional study. Cancer Manag Res 2020;12:4829-4839. [doi: 10.2147/CMAR.S259849] [Medline: 32606972]
5. Kelly S, Eldredge SA, Dalton ED, Miller LE. Health-information behavior: an initial validity portfolio for active and passive measures. Commun Res Rep 2014 Apr;31(2):171-182. [doi: 10.1080/08824096.2014.907145]
6. Ishikawa H, Yano E. Patient health literacy and participation in the health-care process. Health Expect 2008 Jun;11(2):113-122. [doi: 10.1111/j.1369-7625.2008.00497.x] [Medline: 18494956]
7. Lambert SD, Loiselle CG. Health information seeking behavior. Qual Health Res 2007 Oct;17(8):1006-1019. [doi: 10.1177/1049732307305199] [Medline: 17928475]
8. D. Wilson T. Human information behavior. ISI 2000;3(2):049-056. [doi: 10.28945/576]
9. Jia W, Sun M, Huang G, Payton B, Yu W. The indirect effects of health information seeking on health lifestyle: health literacy matters. Health Educ Behav 2025 Feb;52(1):113-121. [doi: 10.1177/10901981241278587] [Medline: 39308078]
10. Lam MK, Lam LT. Health information-seeking behaviour on the internet and health literacy among older Australians. Electron J Health Inform 2012;7(2):e15 [FREE Full text]
11. Wister AV, Malloy-Weir LJ, Rootman I, Desjardins R. Lifelong educational practices and resources in enabling health literacy among older adults. J Aging Health 2010 Sep;22(6):827-854. [doi: 10.1177/0898264310373502] [Medline: 20595098]
12. Davis TC, Wolf MS, Bass PF 3rd, et al. Low literacy impairs comprehension of prescription drug warning labels. J Gen Intern Med 2006 Aug;21(8):847-851. [doi: 10.1111/j.1525-1497.2006.00529.x] [Medline: 16881945]
13. Scott TL, Gazmararian JA, Williams MV, Baker DW. Health literacy and preventive health care use among Medicare enrollees in a managed care organization. Med Care 2002 May;40(5):395-404. [doi: 10.1097/00005650-200205000-00005] [Medline: 11961474]
14. Wigfall LT, Friedman DB. Cancer information seeking and cancer-related health outcomes: a scoping review of the health information national trends survey literature. J Health Commun 2016 Sep;21(9):989-1005. [doi: 10.1080/10810730.2016.1184358] [Medline: 27466828]

XSL•FO
RenderX

15. Eronen J, Paakkari L, Portegijs E, Saajanaho M, Rantanen T. Assessment of health literacy among older Finns. Aging Clin Exp Res 2019 Apr;31(4):549-556. [doi: 10.1007/s40520-018-1104-9] [Medline: 30578457]
16. Safeer RS, Keenan J. Health literacy: the gap between physicians and patients. Am Fam Physician 2005 Aug 1;72(3):463-468. [Medline: 16100861]
17. Shiferaw KB, Tilahun BC, Endehabtu BF, Gullslett MK, Mengiste SA. E-health literacy and associated factors among chronic patients in a low-income country: a cross-sectional survey. BMC Med Inform Decis Mak 2020 Aug 6;20(1):181. [doi: 10.1186/s12911-020-01202-1] [Medline: 32762745]
18. McCloud RF, Okechukwu CA, Sorensen G, Viswanath K. Beyond access: barriers to internet health information seeking among the urban poor. J Am Med Inform Assoc 2016 Nov;23(6):1053-1059. [doi: 10.1093/jamia/ocv204] [Medline: 27206459]
19. Heckman BW, Mathew AR, Carpenter MJ. Treatment burden and treatment fatigue as barriers to health. Curr Opin Psychol 2015 Oct 1;5:31-36. [doi: 10.1016/j.copsyc.2015.03.004] [Medline: 26086031]
20. Bartlett JC. Consumer health information literacy and information behavior of young adults. In: Jean BS, editor. Roles and Responsibilities of Libraries in Increasing Consumer Health Literacy and Reducing Health Disparities: Emerald Publishing; 2021:103-118. [doi: 10.1108/S0065-283020200000047005]
21. Johnson JD, Donhoue WA, Atkin CK, Johnson S. A comprehensive model of information seeking: tests focusing on a technical organization. Sci Commun 1995;16(3):274-303. [doi: 10.1177/1075547095016003003]
22. Xu Y, Wang M, Bao L, Cheng Z, Li X. A cross-sectional study based on the Comprehensive Model of Information seeking: which factors influence health information-seeking behavior in patients with periodontitis. BMC Oral Health 2024 Oct 28;24(1):1307. [doi: 10.1186/s12903-024-05068-x] [Medline: 39468496]
23. Ghaddar SF, Valerio MA, Garcia CM, Hansen L. Adolescent health literacy: the importance of credible sources for online health information. J Sch Health 2012 Jan;82(1):28-36. [doi: 10.1111/j.1746-1561.2011.00664.x] [Medline: 22142172]
24. Gray NJ, Klein JD, Noyce PR, Sesselberg TS, Cantrill JA. Health information-seeking behaviour in adolescence: the place of the internet. Soc Sci Med 2005 Apr;60(7):1467-1478. [doi: 10.1016/j.socscimed.2004.08.010] [Medline: 15652680]
25. Sansom-Daly UM, Lin M, Robertson EG, et al. Health literacy in adolescents and young adults: an updated review. J Adolesc Young Adult Oncol 2016 Jun;5(2):106-118. [doi: 10.1089/jayao.2015.0059] [Medline: 26859721]
26. Cutilli CC. Health literacy in geriatric patients: an integrative review of the literature. Orthop Nurs 2007;26(1):43-48. [doi: 10.1097/00006416-200701000-00014] [Medline: 17273109]
27. Chesser A, Burke A, Reyes J, Rohrberg T. Navigating the digital divide: a systematic review of eHealth literacy in underserved populations in the United States. Inform Health Soc Care 2016;41(1):1-19. [doi: 10.3109/17538157.2014.948171] [Medline: 25710808]
28. Lim MSC, Molenaar A, Brennan L, Reid M, McCaffrey T. Young adults' use of different social media platforms for health information: insights from web-based conversations. J Med Internet Res 2022 Jan 18;24(1):e23656. [doi: 10.2196/23656] [Medline: 35040796]
29. Longo DR, Ge B, Radina ME, et al. Understanding breast-cancer patients' perceptions: health information-seeking behaviour and passive information receipt. Comm Health 2009 May 1;2(2):184-206. [doi: 10.1179/175380609790912779]
30. Soroya SH, Farooq A, Mahmood K, Isoaho J, Zara SE. From information seeking to information avoidance: understanding the health information behavior during a global health crisis. Inf Process Manag 2021 Mar;58(2):102440. [doi: 10.1016/j.ipm.2020.102440] [Medline: 33281273]
31. Brodie M, Foehr U, Rideout V, et al. Communicating health information through the entertainment media. Health Aff (Millwood) 2001;20(1):192-199. [doi: 10.1377/hlthaff.20.1.192] [Medline: 11194841]
32. Bergman L, Nilsson U, Dahlberg K, Jaensson M, Wångdahl J. Validity and reliability of the arabic version of the HLS-EU-Q16 and HLS-EU-Q6 questionnaires. BMC Public Health 2023 Feb 10;23(1):304. [doi: 10.1186/s12889-023-15226-5] [Medline: 36765302]
33. Jayasinghe UW, Harris MF, Parker SM, et al. The impact of health literacy and life style risk factors on health-related quality of life of Australian patients. Health Qual Life Outcomes 2016 May 4;14:68. [doi: 10.1186/s12955-016-0471-1] [Medline: 27142865]
34. Sun X, Shi Y, Zeng Q, et al. Determinants of health literacy and health behavior regarding infectious respiratory diseases: a pathway model. BMC Public Health 2013 Mar 22;13:261. [doi: 10.1186/1471-2458-13-261] [Medline: 23521806]
35. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. Ann Intern Med 2011 Jul 19;155(2):97-107. [doi: 10.7326/0003-4819-155-2-201107190-00005] [Medline: 21768583]
36. Rohrer JM. Thinking clearly about age, period, and cohort effects. Adv Methods Pract Psychol Sci 2025 Apr:1-20. [doi: 10.1177/25152459251342750]
37. Glenn ND. Distinguishing age, period, and cohort effects. In: Mortimer JT, Shanahan MJ, editors. Handbook of the Life Course: Springer; 2003:465-476 URL: https://link.springer.com/book/10.1007/b100507 [accessed 2026-08-31]
38. Sartori AE. An estimator for some binary-outcome selection models without exclusion restrictions. Polit anal 2003;11(2):111-138. [doi: 10.1093/pan/mpg001]

XSL•FO
RenderX

39. Sarma PK. Adoption and impact of super granulated urea (guti urea) technology on farm productivity in Bangladesh: a Heckman two-stage model approach. Environ Chall 2021 Dec;5:100228. [doi: 10.1016/j.envc.2021.100228]
40. Galimard JE, Chevret S, Curis E, Resche-Rigon M. Heckman imputation models for binary or continuous MNAR outcomes and MAR predictors. BMC Med Res Methodol 2018 Aug 31;18(1):90. [doi: 10.1186/s12874-018-0547-1] [Medline: 30170561]
41. Zabala JF, Sablik M, Finical GR, Keeton VF, Bell JF. COVID-19 Information sources and vaccination status among Californian adults by generation using the 2022 California health interview survey: cross-sectional study. JMIR Public Health Surveill 2026 Feb 17;12:e85904. [doi: 10.2196/85904] [Medline: 41701926]
42. Meppelink CS, Smit EG, Diviani N, Van Weert JCM. Health literacy and online health information processing: unraveling the underlying mechanisms. J Health Commun 2016;21(sup2):109-120. [doi: 10.1080/10810730.2016.1193920] [Medline: 27668318]
43. Tiao PH, Chan HY, Huang YM. Navigating diabetes management in the digital era: scoping review of online health information-seeking behavior. J Med Internet Res 2026 Jun 16;28:e82081. [doi: 10.2196/82081] [Medline: 42303245]
44. Kontos E, Blake KD, Chou WYS, Prestin A. Predictors of eHealth usage: insights on the digital divide from the Health Information National Trends Survey 2012. J Med Internet Res 2014 Jul 16;16(7):e172. [doi: 10.2196/jmir.3117] [Medline: 25048379]
45. Eastin MS, Guinsler NM. Worried and wired: effects of health anxiety on information-seeking and health care utilization behaviors. Cyberpsychol Behav 2006 Aug;9(4):494-498. [doi: 10.1089/cpb.2006.9.494] [Medline: 16901253]
46. Dutta-Bergman MJ. Developing a profile of consumer intention to seek out additional information beyond a doctor: the role of communicative and motivation variables. Health Commun 2005;17(1):1-16. [doi: 10.1207/s15327027hc1701_1] [Medline: 15590339]
47. Jiao W, Chang A, Ho M, Lu Q, Liu MT, Schulz PJ. Predicting and empowering health for generation z by comparing health information seeking and digital health literacy: cross-sectional questionnaire study. J Med Internet Res 2023 Oct 30;25:e47595. [doi: 10.2196/47595] [Medline: 37902832]
48. Dom NC, Dapari R, Kamal HA, Harun QN, Salleh SA, Aljaafre AF. Exploring community willingness and barriers to digital solutions and training for dengue prevention: a cross-sectional study in major urban areas of Malaysia. BMC Public Health 2025 Apr 15;25(1):1402. [doi: 10.1186/s12889-025-22677-5] [Medline: 40234776]
49. Jiang S, Street RL. Pathway linking internet health information seeking to better health: a moderated mediation study. Health Commun 2017 Aug;32(8):1024-1031. [doi: 10.1080/10410236.2016.1196514] [Medline: 27464036]
50. Paige SR, Krieger JL, Stellefson ML. The influence of eHealth literacy on perceived trust in online health communication channels and sources. J Health Commun 2017 Jan;22(1):53-65. [doi: 10.1080/10810730.2016.1250846] [Medline: 28001489]
51. Pluye P, Grad R, Repchinsky C, et al. Four levels of outcomes of information - seeking: a mixed methods study in primary health care. J Am Soc Inf Sci Tec 2013 Jan;64(1):108-125. [doi: 10.1002/asi.22793]
52. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. J Med Internet Res 2006 Jun 16;8(2):e9. [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
53. Dewalt DA, Berkman ND, Sheridan S, Lohr KN, Pignone MP. Literacy and health outcomes: a systematic review of the literature. J Gen Intern Med 2004 Dec;19(12):1228-1239. [doi: 10.1111/j.1525-1497.2004.40153.x] [Medline: 15610334]
54. Zhao YC, Zhao M, Song S. Online health information seeking behaviors among older adults: systematic scoping review. J Med Internet Res 2022 Feb 16;24(2):e34790. [doi: 10.2196/34790] [Medline: 35171099]
55. Chang A, Schulz PJ, Jiao W, Liu MT. Obesity-related communication in digital Chinese news from mainland China, Hong Kong, and Taiwan: automated content analysis. JMIR Public Health Surveill 2021 Nov 23;7(11):e26660. [doi: 10.2196/26660] [Medline: 34817383]
56. Kim JE, Calderon SJ, Kim M. Digital health literacy on diabetes knowledge among older adults in the United States. J Community Health Nurs 2026;43(2):153-167. [doi: 10.1080/07370016.2025.2593395] [Medline: 41294338]
57. Xian X, Chang A, Xiang YT, Liu MT. Debate and dilemmas regarding generative AI in mental health care: scoping review. Interact J Med Res 2024 Aug 12;13:e53672. [doi: 10.2196/53672] [Medline: 39133916]
58. Wang W, Ma X, Liu MTC, Chang A. Trust barriers and vulnerabilities in older adults' telemedicine adoption: scoping review. Interact J Med Res 2026 Jul 14;15:e84818. [doi: 10.2196/84818] [Medline: 42447448]
59. Lai YK, Ye JF, Yan C, Zhang L, Zhao X, Liu MTC. From online to offline: how different sources of online health information seeking affect patient-centered communication in Chinese older adults? The roles of patient activation and patient-provider discussion of online health information. Health Commun 2025 Aug;40(9):1684-1695. [doi: 10.1080/10410236.2024.2419194] [Medline: 39434584]

Abbreviations

CMIS: Comprehensive Model of Information Seeking
HISB: health information–seeking behavior
HL: health literacy
HLS-EU-Q16: European Health Literacy Survey Questionnaire-16-item

XSL•FO
RenderX

IMR: inverse Mills ratio
KHP: Korea Health Panel
OR: odds ratio
SAS: Statistical Analysis System
SRH: self-rated health

Edited by A Mavragani, T Sanchez; submitted 03.Nov.2025; peer-reviewed by A Kirbis, MT Liu; revised version received 29.Jul.2026; accepted 31.Jul.2026; published 23.Sep.2026.

Please cite as:
Kwon DH, Kwon YD
Age-Related Patterns in Health Literacy and Health Information–Seeking Behavior in a Korean Population: Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e86992
URL: https://publichealth.jmir.org/2026/1/e86992
doi:10.2196/86992

© Da Hae Kwon, Young Dae Kwon. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 23.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Characteristics Influencing Adherence to SARS-CoV-2 Serial Antigen Testing Following a Federal Safety Communication: Cohort Study

Micajah Z McGarity[1*], MSE, PhD; Catherine C Pollack[1*], PhD; Elizabeth A Currier[1], MPH, MBA; Andrew C Godoshian[2], MPH; Zachary R Smith[2,3], PhD; Marcus Siddall[2], PhD; Gregg M Tavolacci[4], BS; Reynolds M Salerno[5], PhD; Thomas C Tsai[6], MD, MPH; Michael F Iademarco[7*], MD, MPH; Joseph D Miller[2], MBA, PhD

[1]Johns Hopkins University Applied Physics Laboratory, Laurel, MD, United States

[2]Increasing Community Access to Testing, Treatment, and Response Program, Centers for Disease Control and Prevention, Atlanta, GA, United States

[3]Walgreens Co, Deerfield, IL, United States

[4]eTrueNorth, Fort Worth, TX, United States

[5]Office of Laboratory Systems and Response, Centers for Disease Control and Prevention, Atlanta, GA, United States

[6]Department of Health Policy and Management, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA, United States

[7](retired), United States Public Health Service, Washington, DC, United States

[*]these authors contributed equally

Corresponding Author:
Joseph D Miller, MBA, PhD
Increasing Community Access to Testing, Treatment, and Response Program
Centers for Disease Control and Prevention
1600 Clifton Rd NE
Atlanta, GA, 30329
United States
Phone: 1 404 645 3355
Email: faz9@cdc.gov

Abstract

Background: Throughout the COVID-19 pandemic, regular population testing for SARS-CoV-2 was crucial for disease monitoring and management. Initially, reverse transcriptase–polymerase chain reaction tests were primarily used for identifying new cases, but their high cost and delayed results, combined with laboratory capacity and supply chain limitations, led to the adoption of antigen-detection rapid diagnostic tests (Ag-RDTs) at point-of-care locations. On August 11, 2022, the US Food and Drug Administration (FDA) issued a safety communication recommending serial Ag-RDT testing to reduce the risk of false-negative results from singular testing and combat disease spread, followed by a revision letter on November 1, 2022, to manufacturers to update their authorized product labeling.

Objective: This study evaluates adherence to serial testing using data from the Centers for Disease Control and Prevention (CDC) Increasing Community Access to Testing, Treatment, and Response (ICATT) program.

Methods: This national, retrospective cohort study analyzed ICATT program line-level testing records from August 12, 2022, to January 14, 2025. Patients were included if they had at least 1 Ag-RDT. Adherence was determined based on whether patients followed FDA guidance after a negative test. Multivariable logistic regression models were used to identify demographic features associated with adherence, with an expanded model to identify the impact of additional clinical variables. Rates of testing positive after an initial negative Ag-RDT were also calculated for ICATT pharmacy testing.

Results: There were 2,189,464 patients included in the second-test adherence analysis, and 6813 included in the third-test adherence analysis. Less than 1% (20,290/2,189,464) of patients were adherent to the instructions to use a second Ag-RDT, and 98.2% (2,149,292/2,189,464) had no second test documented in ICATT and therefore were classified as nonadherent in this analysis. Second-test adherence was higher among Asian (adjusted odds ratio [AOR] 1.28, 95% CI 1.21-1.36) and Black (AOR 1.17, 95% CI 1.12-1.23) patients and lower among children (AOR 0.48, 95% CI 0.45-0.51). In the expanded model, recent COVID-19 contact increased adherence odds (AOR 1.87, 95% CI 1.64-2.13), whereas recent COVID-19 infection decreased them (AOR 0.66, 95% CI 0.56-0.76). Among patients with symptoms initially testing negative with an Ag-RDT, 24.9%

XSL•FO
RenderX

(7310/29,315) tested positive on their second test, and 57.4% (4195/7310) of these episodes were adherent. For patients without symptoms, 4.2% (60/1435) tested positive on their third test after testing negative on their first 2 tests, with 60% (36/60) adherence.

Conclusions: This study is the first to evaluate adherence to the FDA instructions to perform serial Ag-RDTs. Overall, adherence was low among patients seeking testing at community sites. Recent COVID-19 exposure was associated with the highest odds of serial testing adherence. These insights can inform targeted public health strategies to improve adherence and reduce the risk of false-negative results from Ag-RDTs in future pandemics.

(JMIR Public Health Surveill 2026;12:e92002) doi:10.2196/92002

KEYWORDS

antigen testing; at-home testing; Centers for Disease Control and Prevention; community testing; COVID-19; diagnostic testing strategy; false-negative; public health surveillance; real-world evidence; repeat testing; risk communication; serial testing; testing adherence

Introduction

Throughout the COVID-19 pandemic, regular population testing for SARS-CoV-2 was a cornerstone of disease monitoring and management. Early in the pandemic, nucleic acid amplification tests (NAATs), which include reverse transcriptase–polymerase chain reaction (RT-PCR) tests, were relied upon as the sole way to identify new cases because they were the first diagnostic tests developed and authorized by the US Food and Drug Administration (FDA). While highly sensitive, RT-PCR tests were expensive, and results may not be available for several days due to shipping delays, staffing shortages, supply chain shortages, and reporting backlogs [1,2]. The emergence of antigen-detection rapid diagnostic tests (Ag-RDTs) improved the accessibility and timeliness of SARS-CoV-2 testing, and these tests were less sensitive than their NAAT counterparts [3].

On August 11, 2022, the US FDA issued a new safety communication for serial Ag-RDT testing for SARS-CoV-2 to reduce the risk of false-negative results for COVID-19 [4-6]. The safety communication instructed individuals with COVID-19–like symptoms who received negative test results to test 48 hours after the first test. Asymptomatic individuals with a negative result were instructed to test 48 hours after the first test and then again after 48 hours if the first repeat test was also negative. Confirmatory “laboratory molecular-based” testing, including RT-PCR, was also encouraged, when possible, to further reduce the risk of false-negative Ag-RDT results. For both symptomatic and asymptomatic individuals, a positive result in any repeat test was considered indicative of a case of COVID-19. A comprehensive study of more than 7000 participants was used to establish these instructions and determined that serial Ag-RDT testing improved sensitivity for both Delta and Omicron SARS-CoV-2 variants [7,8].

The degree to which the US population followed the FDA’s instructions has not been studied, primarily because most Ag-RDTs were performed at home, making studies of Ag-RDT testing more difficult to conduct. While it was possible for patients to voluntarily report at-home test results for 4 Ag-RDT manufacturers, that reporting was only a fraction of the total testing volume, and most results were self-interpreted and self-reported [9]. This was partially rectified by the MakeMyTestCount program launched by the US National Institutes of Health’s Rapid Acceleration of Diagnostics Tech program in December 2022 [10]. However, there have been no studies to date that have evaluated adherence to serial Ag-RDT testing within that program. Studies that have evaluated serial test adherence have either been outside the United States, within a closed cohort, or have been small, limiting understanding of how the specific FDA instructions generalize to a broader US population [11-15]. Understanding whether COVID-19 testing among the public adhered to the authorized instructions for use of antigen tests would provide important information for public health officials and epidemiologists to optimize diagnostic testing for future respiratory outbreaks.

The US Centers for Disease Control and Prevention (CDC) Increasing Community Access to Testing, Treatment, and Response (ICATT) program began in April 2020 to provide no-cost SARS-CoV-2 testing through a public-private partnership with participating pharmacy vendors, including CVS, Rite Aid, Walgreens, and eTrueNorth, among others [16]. Since its creation, the ICATT program has collected data on 61 million SARS-CoV-2 tests at more than 19,000 locations in all 50 states, Puerto Rico, and Washington DC. Of the 61 million tests, 49.4 million were categorized as “ICATT paid” or “ICATT conducted,” and 12 million were categorized as “non-ICATT test” records performed at ICATT locations. Thus, the ICATT dataset captures testing that occurred at participating vendor sites and does not capture at-home testing or testing performed at non-ICATT sites. In July 2022, ICATT program testing vendors began assigning unique patient identification numbers to individual testing records, allowing for a large-scale, longitudinal study of adherence to FDA serial testing instructions. In addition, the ICATT dataset included symptom status, demographic information, and clinical features for each testing record to study factors that affect adherence. This study aims to identify patient-level demographic and clinical features associated with adherence to FDA serial Ag-RDT testing guidelines within a national retrospective cohort.

Methods

Data and Key Variables

ICATT-funded line-level testing records were used as the primary source of data (49.4 million), in addition to Medicare, Medicaid, private insurers, Health Resources and Services Administration (HRSA), and cash-paid pharmacy test records (11.9 million) at ICATT locations. These records reflect on-site

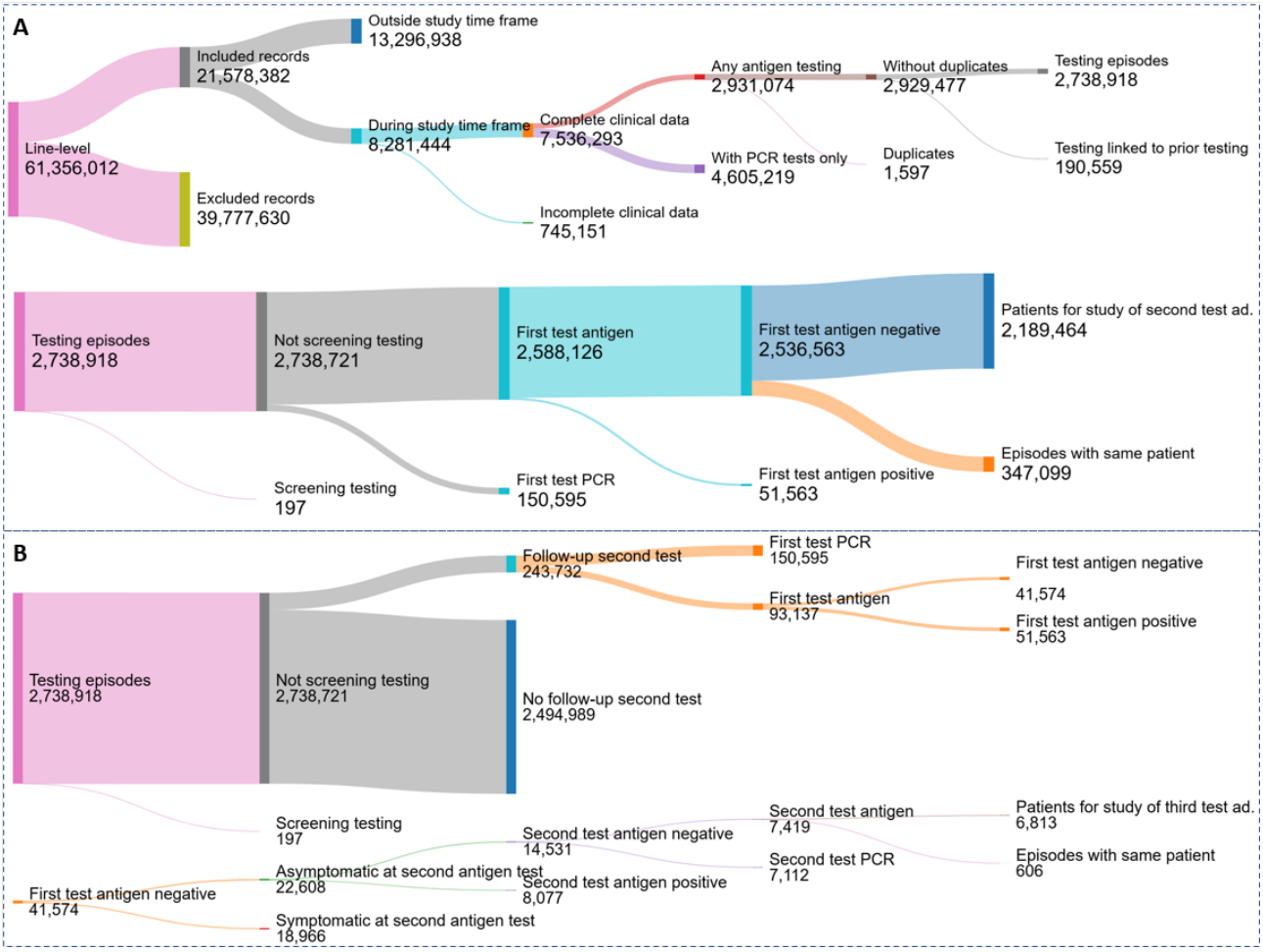
XSL•FO
RenderX

testing at participating ICATT vendor locations and therefore do not include at-home tests or testing performed at non-ICATT sites. ICATT testing records were provided by ICATT testing sites to CDC as part of CDC's One CDC Data Platform (1CDP). Individual testing records were combined into testing events (testing records from the same patient and the same vendor on the same day). Testing events that occurred within 10 days of each other were collapsed into testing episodes via daisy-chaining to link distinct testing events for each patient. Adherence with serial testing was then determined for each testing episode. Finally, summaries of adherence were created for each patient, demographic and clinical features were joined to patient records, and inclusion/exclusion criteria were applied.

There were 745,151 testing records excluded because the symptom status was unreported, and it was used to determine whether a patient was required to take 2 or 3 Ag-RDTs according to the authorized instructions for use. Patients were included if they had at least 1 Ag-RDT between August 12, 2022, and January 14, 2025, at a CVS, Walgreens, or eTrueNorth facility. If the duration of a patient's testing episode exceeded 30 days and the number of tests per day exceeded 0.2, the episode was excluded from the study (197 episodes), as this was considered to be indicative of regular screening testing and not diagnostic testing for acute illness. Since the primary outcome of interest was adherence with FDA serial Ag-RDT testing instructions after a negative first test, patients were also excluded if their first test was a NAAT/RT-PCR test (150,595 episodes) or if their first test result was positive (51,563 episodes). In addition, 347,099 testing episodes were merged with other episodes with the same patient ID, and averages were calculated for each patient's numeric or binary variables, leaving 2,189,464 patients eligible for second-test adherence analysis. The number of tests included and excluded based on different criteria is shown in Figure 1.

In an expanded model, available self-reported clinical variables, including recent contact with a patient infected with COVID-19, vaccination status, COVID-19 infection within the past 90 days, and symptom status at the time of testing, were examined alongside demographic features including age, sex, race, ethnicity, insurance status, test vendor (eg, CVS or Walgreens), and location (collapsed into US Census regions). To account for multiple testing episodes by the same patient, demographic variables for each testing episode were assigned based on the demographic information available for the first test of the testing episode. Clinical variables were represented as the rate of episodes in which that clinical variable occurred (eg, the proportion of episodes where the patient was symptomatic).

Figure 1. Tests-to-patients flow diagram for second-test and third-test adherence analyses. (A) Flow for the second-test adherence analysis; (B) Flow for the third-test adherence analysis. The term polymerase chain reaction (PCR) is used in the figure to denote any nucleic acid amplification test (NAAT).



XSL•FO
RenderX

Patient Summary Adherence Determination

Adherence was defined at the testing-episode level. A testing episode included all tests for the same patient that occurred within 10 days of one another. Episodes were excluded if they lasted more than 30 days and reflected frequent screening (>0.2 tests per day), began with a NAAT/RT-PCR test, or began with a positive Ag-RDT result. A series of complex branching logic (Figure 2) was then applied to the episodes to determine adherence with serial testing.

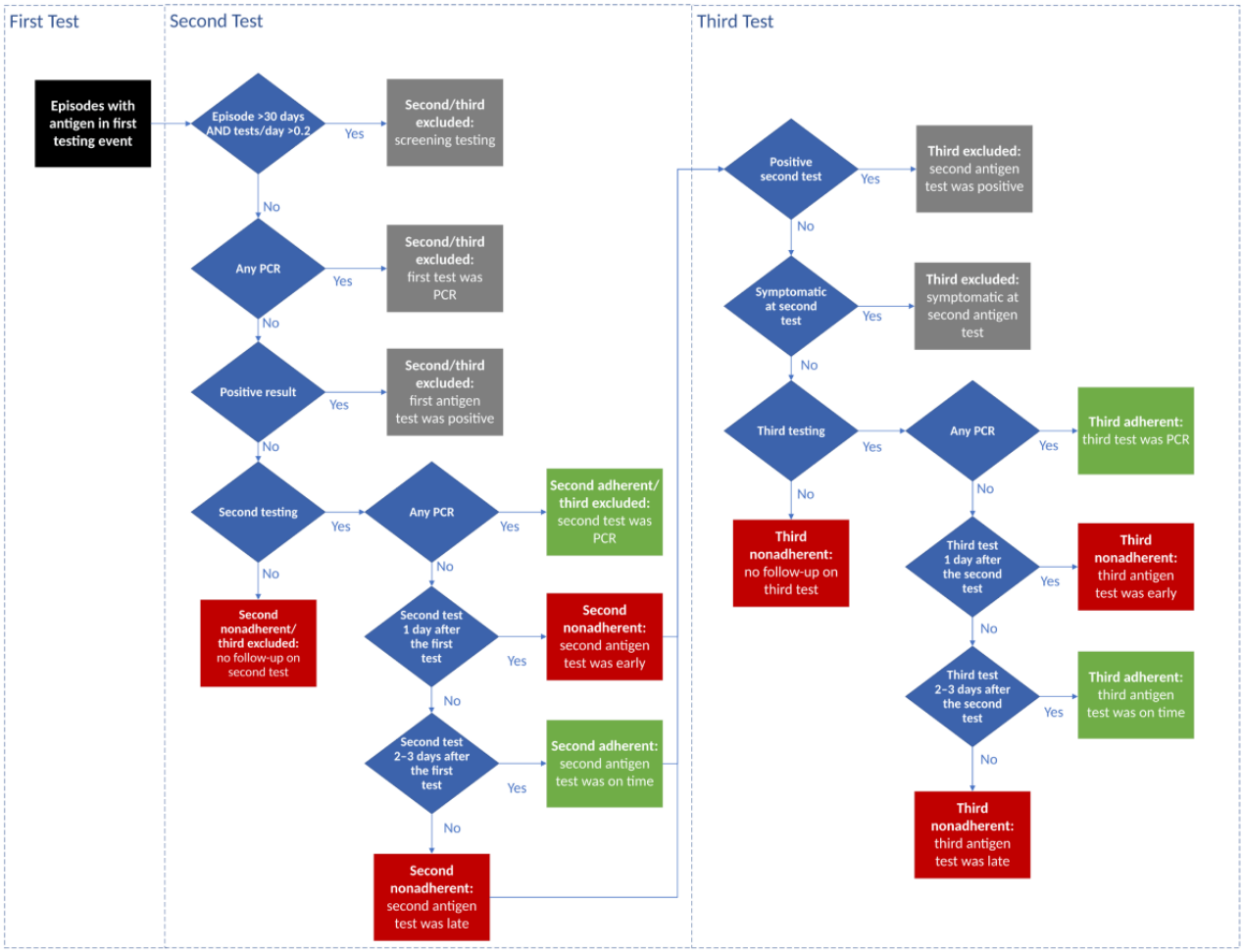
For second-test adherence, the first eligible test had to be a negative Ag-RDT. An episode was classified as second-test adherent if the next documented ICATT test was either (1) a NAAT/RT-PCR test or (2) a second Ag-RDT performed 2-3 days after the first negative Ag-RDT. An episode was classified as second-test nonadherent if no follow-up test was documented in ICATT, if the second Ag-RDT was performed 1 day after the first test, or if it was performed 4 or more days later.

For third-test adherence, only episodes in which the patient was asymptomatic and the second Ag-RDT was negative were eligible. These episodes were classified as third-test adherent if the next documented ICATT test was either (1) a NAAT/RT-PCR test or (2) a third Ag-RDT performed 2-3 days after the second negative Ag-RDT. They were classified as third-test nonadherent if no third test was documented in ICATT, if the third Ag-RDT was performed 1 day after the second test, or if it was performed 4 or more days later. Episodes with patients with symptoms or a positive second test were excluded from the third-test analysis because a third test was not recommended under FDA instructions [4]. For example, a patient without symptoms with negative Ag-RDTs on day 0, day 2, and day 4 would be classified as adherent for both the second and third tests; if the second test occurred on day 1 or day 4, or if no follow-up test was documented in ICATT, the episode would be classified as nonadherent.

Because some patients contributed more than 1 eligible testing episode, second-test adherence and third-test adherence were summarized at the patient level as the proportion of that patient's eligible episodes that were adherent. This approach was used to avoid overweighting patients with repeated episodes and to reduce bias that could arise from treating correlated episodes from the same individual as independent observations. Although this aggregation may smooth within-patient variability across episodes, it allowed each patient to contribute a single outcome value to the regression analyses. For example, a patient with 2 adherent testing episodes and 1 nonadherent testing episode was considered 66.67% adherent.

XSL•FO
RenderX

Figure 2. Serial testing adherence logic flowchart. The term polymerase chain reaction (PCR) is used in the figure to denote any nucleic acid amplification test (NAAT). Testing results were collapsed into episodes by grouping patient tests that occurred within 10 days of each other. Testing episodes exceeding both 30 days and 0.2 tests per day were excluded. After the initial test, testing was considered second-test adherent or third-test adherent based on the FDA safety communication [4]. Blue diamonds represent decision points, gray boxes indicate excluded testing episodes, red boxes indicate nonadherent episodes, and green boxes indicate adherent episodes.



Statistical Modeling

Two separate generalized linear models with a binomial family and logit link function were fit to model the second- and third-test Ag-RDT serial testing adherence. For each model, the outcome variable was the patient-level proportion of eligible testing episodes that were adherent. A patient-level analytic approach was selected because some individuals contributed multiple testing episodes; analyzing episodes as independent observations could disproportionately weight frequent testers and underestimate within-patient correlation. All demographic variables and symptom status were included in the model as candidate predictors. To understand whether the time since the initial safety communication affected adherence, we also included calendar quarters since the initial announcement in the model. In an expanded model, the remaining 3 clinical variables were included in the model; these were included separately due to the high degree of missingness to avoid skewing results in the main analysis. An α of .05 was used to determine statistical significance. Bonferroni corrections were employed to compensate for family-wise error associated with 35 comparative odds ratios.

Rate of Testing Positive After Initially Testing Negative

To assess the actual impact of adherence with serial testing, 2 rates (1 for patients with symptoms and 1 for patients without symptoms) were calculated to determine the percentage of COVID-19 cases that were not detected on the first Ag-RDT but were detected subsequently by serial testing.

Serial Testing of Patients With Symptoms

To estimate the rate of episodes whereby patients with symptoms and initially negative Ag-RDTs subsequently tested positive during serial testing, testing episodes with a first Ag-RDT in a patient with symptoms that were followed by any second test were compiled. Then, the fraction of the second tests that were positive and negative was calculated, and the fraction of each group (positive and negative second tests) that was adherent to the FDA serial testing instructions was calculated.

Serial Testing of Patients Without Symptoms

To estimate the rate of episodes whereby patients with initially negative Ag-RDTs subsequently tested positive during serial testing, testing episodes with patients without symptoms who had negative first and second Ag-RDTs, and a third test were compiled. Then, the fraction of the third tests that were positive

XSL•FO
RenderX

was calculated, and the fraction of each group (positive and negative third tests) that was adherent to the FDA serial testing instructions was calculated.

Software

We used R (version 4.4.1, R Foundation for Statistical Computing, Vienna, Austria) for all statistical analyses.

Ethical Considerations

This study is not human subjects research because it only used data from public health surveillance activities [17]. Therefore, this study was exempt from requiring the Johns Hopkins University's Institutional Review Board approval. This activity was reviewed by CDC, deemed not research, and was conducted consistent with applicable federal law and CDC policy [18].

Results

Patient Characteristics

Of the 61,356,012 ICATT testing records, a total of 2,929,477 unique records representing 2,189,464 patients with Ag-RDTs between August 12, 2022, and January 14, 2025, were included in the analysis of second-test adherence (Figure 1). The majority of records were excluded for falling outside the study window (n=13,296,938), followed by missing clinical data on self-reported symptom status (n=745,151), and only having NAAT/RT-PCR tests (n=4,605,219). Included records were then collapsed into 2,738,918 "testing episodes." There were an additional 197 episodes removed since they were assumed to be screening testing because the total duration of the testing episode exceeded 30 days and the number of tests per day within the episode exceeded 0.2. In addition, testing episodes were excluded if the first test was a NAAT/RT-PCR test (n=150,595) or a positive antigen test (n=51,563). Finally, episodes were summarized by patient ID, with 347,099 episodes associated with another testing episode with the same patient ID, leaving 2,189,464 patients for analysis of second-test adherence. To summarize episodes by patient ID, rates were calculated for the fraction of testing episodes in which the patient was symptomatic and adherent for the second test or third test. In addition, the first testing record in the line-level dataset with a given patient ID was used as the source of all demographic information.

Of the patients included in the final analytic sample of second-test adherence (Table 1), the majority were adults (1,676,033/2,189,464, 76.5%), non-Hispanic or Latino (1,583,019/2,189,464, 72.3%), and tested at a CVS location (2,153,391/2,189,464, 98.4%; Table 1). Approximately 60% (1,312,464/2,189,464) of patients were women, and 14% (296,851/2,189,464) were Black or African American. Geographically, the majority of patients were from the South (1,251,495/2,189,464, 57.2%), followed by the Midwest (441,296/2,189,464, 20.2%) and the Northeast (297,278/2,189,464, 13.6%). Patients with private insurance comprised 70.5% (1,543,902/2,189,464) of the patient cohort. Uninsured patients represented 12.1% (264,071/2,189,464), Medicaid patients represented 7.9% (173,486/2,189,464), and Medicare patients represented 8% (176,015/2,189,464).

The analysis of third-test adherence followed the same process as second-test adherence for compiling testing episodes (Figure 1), but additional criteria were applied to exclude episodes inappropriate for analysis of third-test adherence. These exclusion criteria included having no follow-up second test (n=2,494,989), being symptomatic at the second antigen test (n=18,966), having a positive second antigen test (n=8077), and having a second NAAT/RT-PCR test (n=7112). Finally, episodes were summarized by patient ID, with 606 episodes associated with another testing episode with the same patient ID, leaving 6813 patients for analysis of third-test adherence.

Of the 6813 patients included in the final analytic sample of third-test adherence (Table 2), the majority were adults (5237/6813, 76.9%), non-Hispanic or Latino (4465/6813, 65.5%), and tested at a CVS location (6664/6813, 97.8%; Table 2). Approximately 60% (3939/6813) of participants were women, and 23% (1560/6813) were Black or African American. Geographically, the largest proportion of patients was from the South (3202/6813, 47%), followed by the Northeast (1570/6813, 23%) and the Midwest (1116/6813, 16.4%). Patients with private insurance comprised 56% (3812/6813) of the patient cohort, uninsured patients represented 18.6% (1265/6813), Medicaid patients represented 12.4% (842/6813), and Medicare patients represented 12.4% (846/6813).

XSL•FO
RenderX

Table 1. A summary of clinical and demographic data for patients included in the second-test analysis (N=2,189,464).

Characteristic	Value
Symptomatic rate (%), mean (SD)	87.3 (32.8)
Age group (years), n (%)	
Adult (18-64)	1,676,033 (76.5)
Child (<18)	281,575 (12.9)
Older adult (>65)	231,856 (10.6)
Race, n (%)	
American Indian or Alaska Native	8896 (0.4)
Asian	139,251 (6.4)
Black or African American	296,851 (13.6)
Native Hawaiian or Other Pacific Islander	6837 (0.3)
Not reported	397,562 (18.2)
Other	117,439 (5.4)
White	1,222,628 (55.8)
Ethnicity, n (%)	
Hispanic/Latino	358,182 (16.4)
Not Hispanic/Latino	1,583,019 (72.3)
Not reported	248,263 (11.3)
Region, n (%)	
Midwest	441,296 (20.2)
Northeast	297,278 (13.6)
Other/unspecified	917 (0.0)
South	1,251,495 (57.2)
West	198,478 (9.1)
Sex, n (%)	
Female	1,312,464 (59.9)
Male	869,172 (39.7)
Undetermined	7828 (0.4)
Insurance Status, n (%)	
Cash	29,013 (1.3)
HRSA[a]	2961 (0.1)
Medicaid	173,486 (7.9)
Medicare	176,015 (8)
Not Reported	16 (0.0)
Private insurance	1,543,902 (70.5)
Uninsured	264,071 (12.1)
Site Contractor, n (%)	
CVS	2,153,391 (98.4)
eTrueNorth	494 (0.0)
Walgreens	35,579 (1.6)
Year quarter, n (%)	
2022Q3	223,654 (10.2)
2022Q4	363,206 (16.6)

XSL•FO
RenderX

Characteristic	Value
2023Q1	360,292 (16.5)
2023Q2	195,041 (8.9)
2023Q3	231,471 (10.6)
2023Q4	246,923 (11.3)
2024Q1	219,990 (10.0)
2024Q2	97,829 (4.5)
2024Q3	134,493 (6.1)
2024Q4	99,484 (4.5)
2025Q1	17,081 (0.8)

[a]HRSA: Health Resources and Services Administration.

XSL•FO
RenderX

Table 2. A summary of clinical and demographic data for patients included in the third-test analysis (N=6813).

Characteristic	Value
Symptomatic rate (%), mean (SD)	87.3 (32.8)
Age group (years), n (%)	
Adult (18-64)	5237 (76.9)
Child (<18)	452 (6.6)
Older adult (>65)	1124 (16.5)
Race, n (%)	
White	2689 (39.5)
American Indian or Alaska Native	41 (0.6)
Asian	477 (7)
Black or African American	1560 (22.9)
Native Hawaiian or Other Pacific Islander	41 (0.6)
Not reported	1675 (24.6)
Other	330 (4.8)
Ethnicity, n (%)	
Not Hispanic/Latino	4465 (65.5)
Hispanic/Latino	1480 (21.7)
Not reported	868 (12.7)
Region, n (%)	
Northeast	1570 (23)
Midwest	1116 (16.4)
Other/Unspecified	1 (0.0)
South	3202 (47)
West	924 (13.6)
Sex, n (%)	
Male	2844 (41.7)
Female	3939 (57.8)
Undetermined	30 (0.4)
Insurance Status, n (%)	
Private insurance	3812 (56)
Cash	33 (0.5)
HRSA[a]	15 (0.2)
Medicaid	842 (12.4)
Medicare	846 (12.4)
Uninsured	1265 (18.6)
Site Contractor, n (%)	
CVS	6664 (97.8)
eTrueNorth	6 (0.1)
Walgreens	143 (2.1)
Year quarter, n (%)	
2022Q3	1906 (28)
2022Q4	1738 (25.5)
2023Q1	1694 (24.9)

XSL•FO
RenderX

Characteristic	Value
2023Q2	473 (6.9)
2023Q3	419 (6.2)
2023Q4	212 (3.1)
2024Q1	149 (2.2)
2024Q2	62 (0.9)
2024Q3	126 (1.8)
2024Q4	31 (0.5)
2025Q1	3 (0.0)

[a]HRSA: Health Resources and Services Administration.

Patient and Testing Episode Adherence

At the patient level, second-test adherence was rare. Less than 1% (20,290/2,189,464) of patients included in the second-test adherence analysis were adherent to the instructions for a second test after an initial negative Ag-RDT, and 98.2% (2,149,292/2,189,464) had no second test documented in ICATT and therefore were classified as nonadherent in this analysis. At the testing-episode level, of the 2,536,563 testing episodes eligible for second-test adherence analysis, over 98% (n=2,494,989) were not adherent for either the second or third test (Table 3). The reason for lack of adherence was almost entirely a lack of documented follow-on testing in ICATT (2,494,989/2,536,563, 98.4% for the second test and 5894/7419, 79.4% for the third test), which may reflect loss to follow-up or follow-up testing done outside the program in addition to true nonadherence. Among testing episodes with a second test and only a negative Ag-RDT (n=41,574), the average time between tests was 4.15 (SD 2.59) days. Roughly half (20,481/41,574, 49.3%) of those testing episodes were adherent, either with a second Ag-RDT performed 2-3 days after the first negative Ag-RDT (13,369/41,574, 32.2%) or with a NAAT/RT-PCR test (7112/41,574, 17.1%). However, of the remaining 50.7% of included tests, 40.2% (16,700/41,574) received the follow-up Ag-RDT after the required window, while 10.6% (4393/41,574) received the follow-up Ag-RDT too soon. Patients who received a second Ag-RDT waited an average of 4.15 (SD 2.59) days after their first Ag-RDT.

Among the 7419 testing episodes with a recommended third test per FDA instructions, a total of 20.6% (1525/7419) had a third test. Of those, 75.9% (1157/1525) were nonadherent due to late Ag-RDT testing, 4.5% (69/1525) were nonadherent due to early antigen testing, 14% (214/1525) were adherent from on-time antigen testing, and 5.6% (85/1525) were adherent from NAAT/RT-PCR testing at the third testing encounter. Patients who received a third Ag-RDT waited an average of 5.84 (SD 2.25) days after their second Ag-RDT.

Similar results (Table 4) were found for serial testing adherence of testing episodes used for the expanded model that restricted the inclusion criteria to testing records with complete reporting for the following clinical variables: symptom status, recent COVID-19 contact, having had COVID-19 within 90 days, and having received a vaccine.

XSL•FO
RenderX

Table 3. A summary of testing episode adherence for the second and third test analyses. Episodes with no documented follow-up test in Increasing Community Access to Testing, Treatment, and Response (ICATT) were classified as nonadherent for analytic purposes; this category may include loss to follow-up or follow-up testing outside ICATT.

Characteristic	Value
Second-test adherence, n/N (%)	
Adherent: second test antigen on-time	13,369/2,536,563 (0.5)
Adherent: second test PCR[a]	7112/2,536,563 (0.3)
Nonadherent: no follow-up second test	2,494,989/2,536,563 (98.4)
Nonadherent: second test antigen early	4393/2,536,563 (0.2)
Nonadherent: second test antigen late	16,700/2,536,563 (0.7)
Third-test adherence, n/N (%)	
Adherent: third test antigen on-time	214/7419 (2.9)
Adherent: third test PCR	85/7419 (1.1)
Nonadherent: no follow-up third test	5894/7419 (79.4)
Nonadherent: third test antigen early	69/7419 (0.9)
Nonadherent: third test antigen late	1157/7419 (15.6)
Timing between tests, mean (SD)	
Days between test 1 and 2	4.15 (2.59)
Days between test 2 and 3	5.84 (2.25)

[a]PCR: polymerase chain reaction.

Table 4. Serial testing adherence of testing episodes for the expanded model, which required clinical variables in addition to symptom status to be reported. Episodes with no documented follow-up test in Increasing Community Access to Testing, Treatment, and Response (ICATT) were classified as nonadherent for analytic purposes; this category may include loss to follow-up or follow-up testing outside ICATT.

Characteristic	Value
Second-test adherence, n/N (%)	
Adherent: second antigen test, on time	572/140,461 (0.4)
Adherent: second PCR[a] test	547/140,461 (0.4)
Nonadherent: no follow-up second test	138,330/140,461 (98.5)
Nonadherent: second antigen test, early	252/140,461 (0.2)
Nonadherent: second antigen test, late	760/140,461 (0.5)
Third-test adherence, n/N (%)	
Adherent: third antigen test, on time	8/450 (1.8)
Adherent: third PCR test	3/450 (0.7)
Nonadherent: no follow-up third test	343/450 (76.2)
Nonadherent: third antigen test, early	6/450 (1.3)
Nonadherent: third antigen test, late	90/450 (20)
Time between tests, mean (SD)	
Days between tests 1 and 2	3.85 (2.26)
Days between tests 2 and 3	3.53 (2.26)

[a]PCR: polymerase chain reaction.

Second-Test Adherence: Demographic Variables

Multiple clinical and patient demographics were significantly associated with second-test adherence (Table 5). Patients with symptoms had 47% lower odds (adjusted odds ratio [AOR] 0.53, 95% CI 0.51-0.55; $P<.001$) of adherence compared with those who were without symptoms. Both children (AOR 0.48, 95% CI 0.45-0.51; $P<.001$) and adults aged 65 years and older (AOR 0.83, 95% CI 0.77-0.88; $P<.001$) had lower odds of adherence compared with adults. In contrast, both Asian (AOR

XSL•FO
RenderX

1.28, 95% CI 1.21-1.36; *P*<.001) and Black (AOR 1.17, 95% CI 1.12-1.23; *P*<.001) individuals had higher odds of adherence compared with White individuals. Hispanic or Latino individuals had 9% higher odds of adherence (AOR 1.09, 95% CI 1.03-1.16; *P*=.003) compared with non-Hispanic individuals; however, this association did not remain statistically significant after correction for multiple comparisons. All geographic regions had lower odds of adherence than the Northeast, with the largest difference observed in the South (AOR 0.82, 95% CI 0.79-0.86; *P*<.001). Women had 8% higher odds of adherence (AOR 1.08, 95% CI 1.05-1.11; *P*<.001) compared with men. Both patients with Medicare (AOR 1.11, 95% CI 1.03-1.20; *P*=.005) and patients with Medicaid (AOR 1.14, 95% CI 1.07-1.20; *P*<.001) had higher odds of adherence than those with private insurance; however, the association for Medicare did not remain statistically significant after correction for multiple comparisons. Estimates for the site contractor eTrueNorth (Table 5) are based on small sample sizes (Table 1), as reflected by wider CIs. Adherence significantly attenuated over time since guideline implementation, ranging from 17% lower odds in the quarter immediately following guideline implementation (AOR 0.83, 95% CI 0.79-0.87; *P*<.001) to 92% lower odds of adherence more than 2 years later (AOR 0.08, 95% CI 0.05-0.13; *P*<.001).

XSL•FO
RenderX

Table 5. Fractional logit regression model of second-test adherence.

Characteristic	OR[a] (95% CI)	*P* value
Symptomatic rate	0.53 (0.51-0.55)	<.001
Age group (years; reference: adult, 18-64)		
Child (<18)	0.48 (0.45-0.51)	<.001
Older adult (>65)	0.83 (0.77-0.88)	<.001
Race (reference: White)		
American Indian or Alaska Native	1.07 (0.84-1.33)	.57
Asian	1.28 (1.21-1.36)	<.001
Black or African American	1.17 (1.12-1.23)	<.001
Native Hawaiian or Other Pacific Islander	1.21 (0.94-1.53)	.13
Not reported	1.16 (1.10-1.24)	<.001
Other	1.00 (0.92-1.08)	.93
Ethnicity (reference: not Hispanic/Latino)		
Hispanic/Latino	1.09 (1.03-1.16)	.003
Not reported	0.83 (0.78-0.88)	<.001
Region (reference: Northeast)		
Midwest	0.83 (0.79-0.88)	< .001
Other/unspecified	0.18 (0.03-0.55)	.01
South	0.82 (0.79-0.86)	<.001
West	0.93 (0.88-0.99)	.02
Sex (reference: male)		
Female	1.08 (1.05-1.11)	<.001
Undetermined	1.25 (0.99-1.55)	.049
Insurance Status (reference: private insurance)		
Cash	1.16 (0.98-1.37)	.08
HRSA[b]	1.23 (0.86-1.70)	.24
Medicaid	1.14 (1.07-1.20)	<.001
Medicare	1.11 (1.03-1.20)	.005
Not Reported	1.67 (0.09-9.33)	.63
Uninsured	1.01 (0.96-1.05)	.74
Site Contractor (reference: CVS)		
eTrueNorth	4.97 (2.90-7.91)	<.001
Walgreens	0.87 (0.76-1.00)	.06
Year quarter (reference: 2022Q3)		
2022Q4	0.83 (0.79-0.87)	<.001
2023Q1	0.78 (0.75-0.82)	<.001
2023Q2	0.51 (0.48-0.54)	<.001
2023Q3	0.51 (0.48-0.54)	<.001
2023Q4	0.34 (0.31-0.36)	<.001
2024Q1	0.27 (0.25-0.29)	<.001
2024Q2	0.23 (0.20-0.26)	<.001
2024Q3	0.28 (0.25-0.31)	<.001

XSL•FO
RenderX

Characteristic	OR[a] (95% CI)	*P* value
2024Q4	0.15 (0.12-0.17)	<.001
2025Q1	0.08 (0.05-0.13)	<.001

[a]OR: odds ratio.

[b]HRSA: Health Resources and Services Administration.

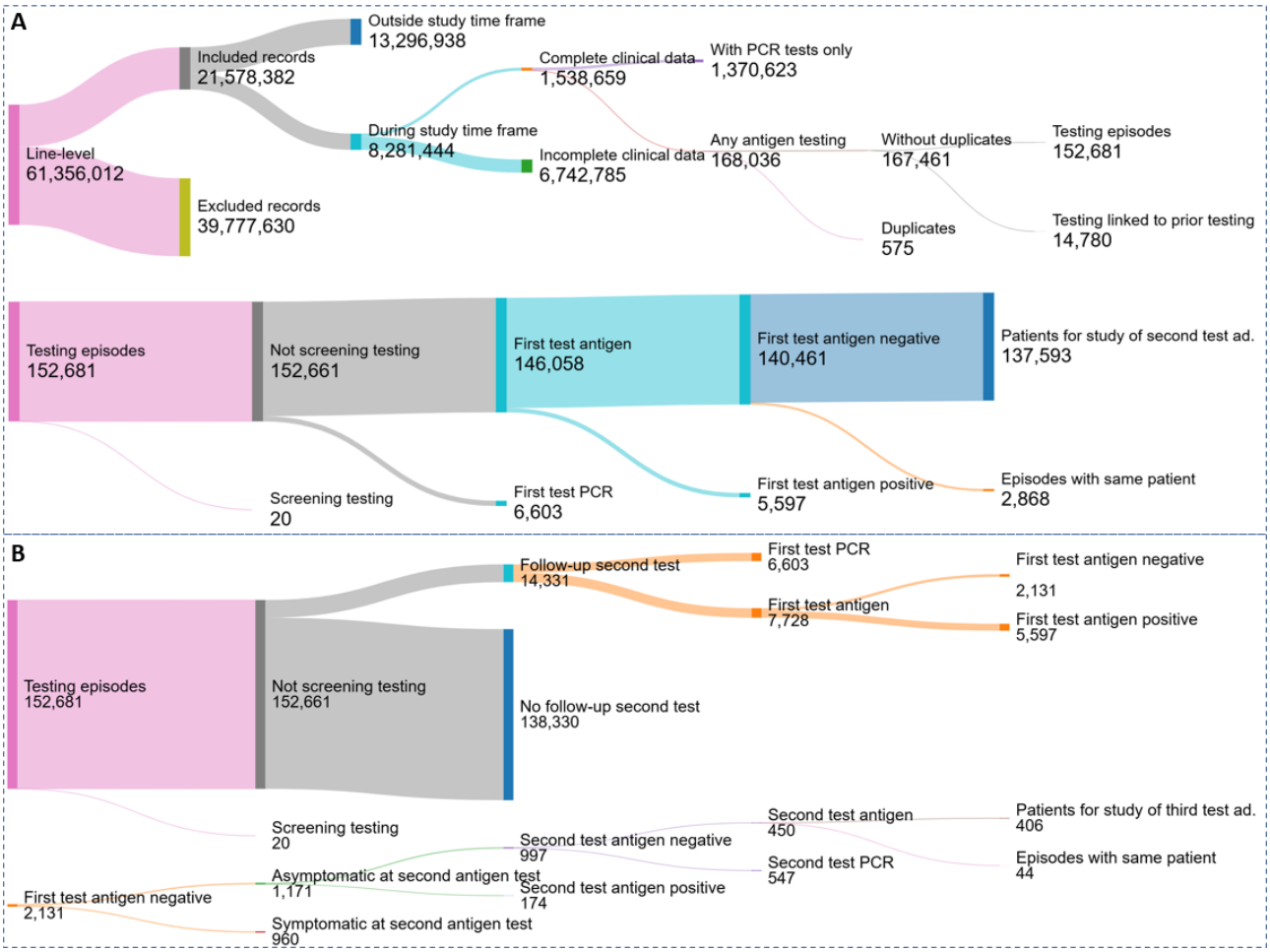
Second-Test Adherence Expanded Model: Demographic and Clinical Variables

There were 137,593 patients (Figure 3) included in the expanded model with complete demographic and clinical information (Table 6), representing a substantially smaller complete-case subset of the primary analytic sample. After reducing the model to this sample and adding in the 3 clinical variables (recent COVID-19 contact, COVID-19 within 90 days, and vaccination status), many of the demographic variables (sex, insurance status, and symptom status) were no longer statistically significant (Table 7). However, the newly added clinical variables were significantly associated with the outcome. For example, testing episodes among patients with a recent COVID-19 contact had 87% higher odds (AOR 1.87, 95% CI 1.64-2.13; *P*<.001) of adherence compared with patients without recent COVID-19 contact. In contrast, testing episodes among patients who had COVID-19 within the past 90 days had 34% lower odds (AOR 0.66, 95% CI 0.56-0.76; *P*<.001) of adherence. Vaccination status (AOR 1.07, 95% CI 0.92-1.26; *P*=.40) was not significantly associated with adherence.

Among demographic variables, children had 46% lower odds of adherence than adults (AOR 0.54, 95% CI 0.40-0.72; *P*<.001), while older adult patients were no longer statistically different from adults (AOR 0.89, 95% CI 0.68-1.16; *P*=.40). Asian individuals had 33% higher odds of adherence (AOR 1.33, 95% CI 1.05-1.67; *P*=.02) compared with White individuals; however, this association did not remain statistically significant after correction for multiple comparisons, and all other racial categories were not statistically significant. Testing episodes in the quarter immediately following the issuance of the FDA recommendations showed the opposite direction of association, with 20% higher odds of adherence (AOR 1.20, 95% CI 1.04-1.38; *P*=.01) compared with the quarter of policy implementation; however, this association did not remain statistically significant after correction for multiple comparisons.

XSL•FO
RenderX

Figure 3. Patient flow diagram for the second-test and third-test adherence analyses (expanded model). (A) Flow for the second-test adherence analysis; (B) Flow for the third-test adherence analysis. The term polymerase chain reaction (PCR) is used in the figure to denote any nucleic acid amplification test (NAAT).



XSL•FO
RenderX

Table 6. Patient demographic and clinical characteristics for the expanded model of second-test adherence study (N=137,593).

Characteristic	Value
Clinical characteristics (%), mean (SD)	
Symptomatic rate	82.0 (38.3)
Recent COVID-19 contact rate	50.0 (49.8)
Had COVID-19 within 90 days rate	55.0 (49.6)
Has received vaccine rate	77.0 (42.3)
Age group (years), n (%)	
Adult (18-64)	112,601 (81.8)
Child (<18)	10,204 (7.4)
Older Adult (>65)	14,788 (10.7)
Race, n (%)	
White	69,085 (50.2)
American Indian or Alaska Native	798 (0.6)
Asian	9631 (7.0)
Black or African American	26,188 (19.0)
Native Hawaiian or Other Pacific Islander	502 (0.4)
Not reported	26,343 (19.1)
Other	5046 (3.7)
Ethnicity, n (%)	
Not Hispanic/Latino	97,082 (70.6)
Hispanic/Latino	27,312 (19.8)
Not reported	13,199 (9.6)
Region, n (%)	
Northeast	19,386 (14.1)
Midwest	26,840 (19.5)
Other/unspecified	146 (0.1)
South	74,331 (54.0)
West	16,890 (12.3)
Sex, n (%)	
Male	55,696 (40.5)
Female	81,242 (59.0)
Undetermined	655 (0.5)
Insurance status, n (%)	
Private insurance	82,835 (60.2)
Cash	2667 (1.9)
HRSA[a]	220 (0.2)
Medicaid	11,377 (8.3)
Medicare	10,071 (7.3)
Uninsured	30,423 (22.1)
Site Contractor, n (%)	
CVS	119,592 (86.9)
Walgreens	18,001 (13.1)
Year quarter, n (%)	

XSL•FO
RenderX

Characteristic	Value
2022Q3	46,086 (33.5)
2022Q4	34,580 (25.1)
2023Q1	20,164 (14.7)
2023Q2	4614 (3.4)
2023Q3	7776 (5.7)
2023Q4	5027 (3.7)
2024Q1	3930 (2.9)
2024Q2	944 (0.7)
2024Q3	8224 (6.0)
2024Q4	5326 (3.9)
2025Q1	922 (0.7)

[a]HRSA: Health Resources and Services Administration.

XSL•FO
RenderX

Table 7. Fractional logit regression for the expanded model of second-test adherence.

Characteristic	OR[a] (95% CI)	*P* value
Clinical characteristics		
Symptomatic rate	0.91 (0.78-1.05)	.20
Recent COVID-19 contact rate	1.87 (1.64-2.13)	<.001
Had COVID-19 within 90 days rate	0.66 (0.56-0.76)	<.001
Has received vaccine rate	1.07 (0.92-1.26)	.40
Age group (years; reference: adult, 18-64)		
Child (<18)	0.54 (0.40-0.72)	<.001
Older Adult (>65)	0.89 (0.68-1.16)	.40
Race (reference: White)		
American Indian or Alaska Native	0.53 (0.13-1.38)	.27
Asian	1.33 (1.05-1.67)	.02
Black or African American	1.03 (0.86-1.22)	.77
Native Hawaiian or Other Pacific Islander	0.57 (0.09-1.79)	.43
Not reported	0.85 (0.67-1.09)	.21
Other	1.07 (0.76-1.47)	.69
Ethnicity (reference: not Hispanic/Latino)		
Hispanic/Latino	1.16 (0.92-1.44)	.20
Not reported	0.78 (0.59-1.02)	.08
Region (reference: Northeast)		
Midwest	0.71 (0.58-0.86)	<.001
Other/unspecified	1.00 (0.10-4.44)	.88
South	0.70 (0.60-0.83)	<.001
West	0.57 (0.44-0.73)	<.001
Sex (reference: male)		
Female	1.02 (0.90-1.16)	.72
Undetermined	1.06 (0.35-2.48)	.90
Insurance Status (reference: private insurance)		
Cash	0.00 (0.00-0.00)	.94
HRSA[b]	2.78 (1.09-5.77)	.02
Medicaid	1.07 (0.86-1.32)	.56
Medicare	0.92 (0.67-1.25)	.60
Uninsured	0.89 (0.74-1.06)	.18
Site contractor (reference: CVS)		
Walgreens	0.96 (0.72-1.25)	.74
Year Quarter (reference: 2022Q3)		
2022Q4	1.20 (1.04-1.38)	.01
2023Q1	1.05 (0.86-1.29)	.62
2023Q2	0.63 (0.39-0.96)	.04
2023Q3	0.11 (0.04-0.22)	<.001
2023Q4	0.03 (0.00-0.14)	<.001
2024Q1	0.08 (0.01-0.26)	<.001

XSL•FO
RenderX

Characteristic	OR[a] (95% CI)	*P* value
2024Q2	0.20 (0.01-0.89)	.11
2024Q3	0.15 (0.07-0.29)	<.001
2024Q4	0.38 (0.21-0.67)	.001
2025Q1	0.00 (0.00-0.00)	.97

[a]OR: odds ratio.

[b]HRSA: Health Resources and Services Administration.

Third-Test Adherence: Demographic Variables

Some similar trends among demographic variables were observed for third-test adherence (Table 8). Patients with symptoms had 35% lower odds of adherence compared with those without symptoms (AOR 0.65, 95% CI 0.45-0.91; *P*=.02), and children had 50% lower odds of adherence than adults (AOR 0.50, 95% CI 0.24-0.94; *P*=.05); however, these associations did not remain statistically significant after correction for multiple comparisons. Asian individuals had 53% lower odds of adherence than White individuals (AOR 0.47, 95% CI 0.24-0.82; *P*=.01); however, this association did not remain statistically significant after correction for multiple comparisons, and all other racial categories were not statistically significant. Among geographic regions, the only significant difference was observed in the West, which had 47% higher odds of adherence than the Northeast (AOR 1.47, 95% CI 1.00-2.15; *P*=.048); however, this association did not remain statistically significant after correction for multiple comparisons. Ethnicity, sex, insurance status, site contractor, and temporality were not statistically significant. Similar to second-test adherence, the odds of third-test adherence also declined over time.

XSL•FO
RenderX

Table 8. Fractional logit regression model of third-test adherence.

Characteristic	OR[a] (95% CI)	*P* value
Symptomatic rate	0.65 (0.45-0.91)	.02
Age group (years; reference: adult, 18-64)		
Child (<18)	0.50 (0.24-0.94)	.05
Older adult (>65)	0.94 (0.60-1.44)	.79
Race (reference: White)		
American Indian or Alaska Native	0.00 (0.00-247,000)	.98
Asian	0.47 (0.24-0.82)	.01
Black or African American	0.75 (0.52-1.05)	.20
Native Hawaiian or Other Pacific Islander	0.00 (0.00-373,000)	.98
Not reported	0.74 (0.47-1.16)	.18
Other	1.22 (0.71-2.00)	.46
Ethnicity (reference: not Hispanic/Latino)		
Hispanic/Latino	0.99 (0.62-1.53)	.96
Not reported	1.59 (1.03-2.41)	.03
Region (reference: Northeast)		
Midwest	1.06 (0.72-1.54)	.78
Other/unspecified	7.09E-07 (N/A[b]-infinity)	>.99
South	0.79 (0.57-1.10)	.16
West	1.47 (1.00-2.15)	.048
Sex (reference: male)		
Female	1.09 (0.85-1.40)	.50
Undetermined	0.00 (0.00-1,410,000,000)	.98
Insurance status (reference: private insurance)		
Cash	0.00 (0.00-95,800,000)	.98
HRSA[c]	5.60E-07 (N/A-infinity)	.99
Medicaid	1.17 (0.79-1.70)	.41
Medicare	0.78 (0.46-1.31)	.36
Uninsured	1.17 (0.84-1.62)	.34
Site contractor (reference: CVS)		
eTrueNorth	1.41E-06 (N/A-infinity)	.99
Walgreens	0.66 (0.22-1.53)	.38
Year Quarter (ref: 2022Q3)		
2022Q4	0.87 (0.63-1.19)	.38
2023Q1	0.77 (0.55-1.07)	.13
2023Q2	0.95 (0.57-1.52)	.84
2023Q3	0.75 (0.40-1.30)	.33
2023Q4	0.64 (0.24-1.36)	.29
2024Q1	0.31 (0.05-1.01)	.11
2024Q2	0.00 (0.00-106)	.98
2024Q3	0.39 (0.06-1.27)	.20
2024Q4	0.85 (0.05-4.15)	.88

XSL•FO
RenderX

Characteristic	OR[a] (95% CI)	*P* value
2025Q1	5.87E-07 (N/A-infinity)	>.99

[a]OR: odds ratio.

[b]N/A: not applicable.

[c]HRSA: Health Resources and Services Administration.

Third-Test Adherence Expanded Model: Demographic and Clinical Variables

There were 406 patients included in the expanded model with complete demographic and clinical variables (Table 9), reflecting a markedly reduced complete-case sample. After adding additional clinical variables into the model and restricting the analysis to this sample, the majority of demographic variables were no longer statistically significant (Table 10). The only significant relationship was observed for recent COVID-19 contacts, as those individuals had 1643% higher odds of adherence (AOR 17.43, 95% CI 2.14-546.41; *P*=.03); however, the wide CI around this estimate indicates limited precision, and the association did not remain statistically significant after correction for multiple comparisons.

XSL•FO
RenderX

Table 9. Patient demographic and clinical characteristics for the expanded model of the third-test adherence study (N=406).

Characteristic	Value
Clinical characteristics (%), mean (SD)	
Symptomatic rate	22.0 (41.4)
Recent COVID-19 contact rate	57.0 (49.2)
Had COVID-19 within 90 days rate	48.0 (49.5)
Has received vaccine rate	62.0 (48.5)
Age group (years), n (%)	
Adult (18-64)	327 (80.5)
Child (<18)	12 (3)
Older adult (>65)	67 (16.5)
Race, n (%)	
White	196 (48.3)
American Indian or Alaska Native	0 (0)
Asian	25 (6.2)
Black or African American	99 (24.4)
Native Hawaiian or Other Pacific Islander	2 (0.5)
Not reported	71 (17.5)
Other	13 (3.2)
Ethnicity, n (%)	
Not Hispanic/Latino	314 (77.3)
Hispanic/Latino	61 (15)
Not reported	31 (7.6)
Region, n (%)	
Northeast	122 (30)
Midwest	69 (17)
South	176 (43.3)
West	39 (9.6)
Sex, n (%)	
Male	145 (35.7)
Female	257 (63.3)
Undetermined	4 (1)
Insurance status, n (%)	
Private insurance	257 (63.3)
Cash	1 (0.2)
HRSA[a]	0 (0)
Medicaid	54 (13.3)
Medicare	40 (9.9)
Uninsured	54 (13.3)
Site contractor, n (%)	
CVS	365 (89.9)
Walgreens	41 (10.1)
Year Quarter, n (%)	
2022Q3	209 (51.5)

XSL•FO
RenderX

Characteristic	Value
2022Q4	123 (30.3)
2023Q1	57 (14.0)
2023Q2	6 (1.5)
2023Q3	5 (1.2)
2023Q4	2 (0.5)
2024Q1	1 (0.2)
2024Q2	1 (0.2)
2024Q3	1 (0.2)
2024Q4	1 (0.2)
2025Q1	0 (0)

[a]HRSA: Health Resources and Services Administration.

XSL•FO
RenderX

Table 10. Fractional logit regression for the expanded model of third-test adherence. Note: some estimates are unstable due to sparse cells in the expanded third-test model.

Characteristic	OR[a] (95% CI)	*P* value
Clinical characteristics		
Symptomatic rate	2.40 (0.39-13.56)	.32
Recent COVID-19 contact rate	17.43 (2.14-546.41)	.03
Had COVID-19 within 90 days rate	0.85 (0.14-4.80)	.85
Has received vaccine rate	0.30 (0.05-1.59)	.16
Age group (years; reference: adult, 18-64)		
Child (<18)	0 (N/A[b]-infinity)	>.99
Older Adult (65+)	2.00 (0.08-20.61)	.59
Race (reference: White)		
Asian	2.49 (0.09-35.39)	.52
Black or African American	0.82 (0.10-4.73)	.83
Native Hawaiian or Other Pacific Islander	0 (N/A-infinity)	>.99
Not reported	0.09 (0.00-2.15)	.16
Other	9.62E-09 (0-infinity)	>.99
Ethnicity (reference: not Hispanic/Latino)		
Hispanic/Latino	1.80 (0.06-33.80)	.70
Not reported	20.61 (0.65-714.67)	.08
Region (reference: Northeast)		
Midwest	0 (N/A-infinity)	>.99
South	0.68 (0.12-3.94)	.66
West	1.20 (0.09-13.29)	.88
Sex (reference: male)		
Female	0.99 (0.19-6.11)	.99
Undetermined	0.38 (0-infinity)	>.99
Insurance status (reference: private insurance)		
Cash	0 (N/A-infinity)	>.99
Medicaid	0.32 (0.01-3.14)	.39
Medicare	0 (N/A-infinity)	>.99
Uninsured	0 (0-infinity)	>.99
Site contractor (reference: CVS)		
Walgreens	0.92 (0.08-7.14)	.94
Year Quarter (ref: 2022Q3)		
2022Q4	0.21 (0.01-1.46)	.18
2023Q1	0.38 (0.01-4.15)	.47
2023Q2	6.14 (0.20-114.87)	.23
2023Q3	0 (0-infinity)	>.99
2023Q4	2.34 (N/A-infinity)	>.99
2024Q1	0.64 (0-infinity)	>.99
2024Q2	7.98 (N/A-infinity)	>.99
2024Q3	N/A (N/A-N/A)	N/A
2024Q4	0 (N/A-infinity)	>.99

XSL•FO
RenderX

[a]OR: odds ratio.

[b]N/A: not applicable.

Analysis of the Rate of Testing Positive After Initially Testing Negative

Serial Testing of Patients With Symptoms

Among 2,025,975 episodes with a negative first Ag-RDT, 1,742,160 (86%) were symptomatic at the first test. Of the episodes with patients with symptoms, a total of 29,315 (1.7%) had a second test, and 7310 (24.9%) of those receiving a second Ag-RDT were positive. Among this group, a total of 4195 (57.4%) of the episodes were adherent with serial testing (eg, an Ag-RDT taken 2-3 days after the first test or an RT-PCR test). Of the 22,005 (75.1%) episodes with a negative second Ag-RDT, a total of 10,818 (49.2%) were adherent with serial testing. Therefore, 24.9% (7310/29,315) of episodes among patients with symptoms tested negative on the first antigen test (95% CI 24.4%-25.4%) and positive on the second test when performing serial testing with 2 antigen tests.

Serial Testing of Patients Without Symptoms

Among 2,025,975 episodes with a negative first Ag-RDT, a total of 7792 (0.39%) were asymptomatic at the first and second tests, and 5957 (0.29%) were also negative at the second Ag-RDT. Among the 5957 episodes among patients without symptoms and with 2 negative Ag-RDTs, a total of 1435 (24.1%) had a third test, of whom 60 (4.2%) were positive, and 1375 (95.8%) were negative for the third Ag-RDT. Of the 60 episodes with a positive third test, 36 (60%) were adherent with serial testing. Of the 1375 episodes with a negative third test, 246 (17.9%) were adherent with serial testing. Therefore, the rate of testing positive through serial testing for this population was 4.2% (60/1435, 95% CI 3.2%-5.2%). In addition, serial testing was also effective for 24 patients without symptoms who were positive on the third test but nonadherent, indicating that serial testing can still be effective in situations where patients do not strictly adhere to testing timelines.

Discussion

Overview

This study is the first to evaluate serial Ag-RDT testing adherence following the FDA's safety communication issued in August 2022. Among 2,189,464 patients who received at least 1 Ag-RDT through the CDC ICATT program, less than 2% also received a follow-up test captured by the program. Accordingly, the large group classified as nonadherent because no follow-up test was observed likely includes a mixture of true nonadherence, loss to follow-up, and follow-up testing done outside ICATT. Multiple demographic variables (including age, race, and region) were significantly associated with adherence, although some variables (eg, insurance, sex, or timeliness relative to the issuance of the recommendations) were only significant for second-test adherence. In the expanded model that included 3 clinical variables, however, the majority of demographic variables were no longer statistically significant. Instead, interaction with a recent COVID-19 contact significantly increased the odds of second- and third-test adherence, while patients who had COVID-19 themselves within the past 90 days were less likely to be adherent. Taken together, these results suggest that clinical variables and temporal nearness to policy issuance and likelihood of exposure (due to a substantial reduction in cases since policy issuance [19,20]) are more closely related to adherence than nonclinical demographic variables.

Within this study, there were only a few demographic differences associated with serial testing adherence, and these distinctions differed from those identified in other studies. In university environments, women were more likely than men to adhere to serial testing policies [11,14], while Asian and Black individuals were less likely to adhere [11]. This partially aligned with the results presented here; while women were also found to be more adherent than men for their second test, so were Asian and Black individuals (although both became statistically insignificant after adjusting for additional clinical variables). These differences may be attributable to the setting of the study (eg, a university setting in a relatively closed environment vs a broader geographic sample) or the time frame in which the studies occurred and the corresponding availability of resources. In particular, over-the-counter Ag-RDT tests were not approved by the FDA until December 2020, which was after the Wisconsin university study occurred [11,21]. As a result, patients in that study may have been more likely to adhere to serial testing because they did not have accessible over-the-counter tests.

In contrast, patients in this study may appear to be nonadherent with serial testing when, in actuality, they took a second Ag-RDT that was not captured by the ICATT program. As a result, the availability of at-home Ag-RDTs is a notable confounding variable. However, it is unlikely that all patients who did not receive a follow-up Ag-RDT within the ICATT data took an at-home Ag-RDT. A previous cross-sectional online survey found that approximately 20% of participants had engaged in at-home COVID-19 testing by March 2022, especially those who had either a COVID-19 exposure or COVID-19-like symptoms [22]. Even if a similar proportion of testing episodes without a second Ag-RDT in the ICATT data had an Ag-RDT, that would still leave over 1,995,991 testing episodes without a follow-up test. This may still be an underestimation because a significant portion of test takers were uninsured, and at-home testing rates are lower among these individuals [16,22].

In the expanded model that included only persons with complete clinical data, the majority of the demographic variables were no longer statistically significant (especially within the second-test adherence model). This suggests that a patient's recent COVID-19 environment is more of a driver to testing than various demographic variables. Furthermore, the conflicting directionality of recent COVID-19 contact (which was associated with higher adherence) and recent COVID-19 status (which was associated with lower adherence) could be attributable to the rationale behind why an individual sought ICATT testing. For example, a patient who recently had

XSL•FO
RenderX

COVID-19 may seek out an ICATT test to confirm that they no longer have the disease, as opposed to testing for a new instance of the disease. In contrast, patients with a recent COVID-19 contact may be more motivated to test because of concern that they may actively have COVID-19. These expanded-model findings should be interpreted cautiously because they were based on complete-case analyses using substantially smaller samples than the primary models. If the completeness of the clinical variables was nonrandom, the expanded-model estimates may be subject to selection bias. The smaller analytic samples, particularly for third-test adherence, also reduced precision and statistical power and may partly explain why some associations were attenuated or no longer statistically significant.

Interestingly, symptom status and vaccination status were not significant after adjusting for COVID-19 contact status, suggesting that actual contact with another patient may be more motivating than a person's own symptoms or vaccination history. This aligns with findings from a scoping review that found that a person's perception of having COVID-19 and motivation for testing are heavily influenced by their social environment [23]. This may also explain why adherence declined over time relative to the announcement of a new safety communication, as motivation to comply may be higher directly after the initial announcement occurs. In addition, independent pharmacies (eTrueNorth sites) showed significantly higher odds of test adherence than chain pharmacies. Similar results have been found in terms of medication adherence by independent pharmacy users [24]. Because all ICATT vendors were required to follow FDA regulations and manufacturers' instructions for use, this difference is unlikely due to procedural variations and warrants further investigation. Overall, these results suggest the need for public health officials to make frequent reminders about changes in instructions to stress their importance and reinforce positive behaviors, although additional work is needed to understand the most effective messaging strategy and how over-the-counter testing may have played a role in serial testing adherence. Although we did not directly measure exposure to public health messaging or pharmacist (testing-site) counseling, the low adherence observed here suggests a need for future work on how serial-testing guidance is communicated to patients and how access to convenient follow-up testing can be improved during future pandemics. Although several ORs were statistically significant, many were close to 1 and therefore likely represent modest practical differences at the individual level. The more important public health finding is the uniformly low uptake of serial testing across the cohort, suggesting broad barriers to adherence rather than large disparities concentrated in a few subgroups. Taken together, these findings suggest that low adherence likely reflects a combination of behavioral factors, including perceived exposure risk and declining motivation over time, and structural factors, including the availability of at-home testing, differences in access, and possible variation in counseling across testing sites.

The analysis on the rate of testing positive through serial testing indicated that serial testing can reduce the risk of false-negative Ag-RDT results, aligning with the FDA safety communication. For patients with symptoms, serial testing (eg, a follow-up Ag-RDT or NAAT/RT-PCR test after the initial test) identified 24.9% of initially negative cases as positive upon retesting, underscoring the importance of multiple tests to catch infections missed by earlier tests. Adherence with serial testing was higher among those who tested positive on the second test (57.4%) compared with those who remained negative (49.2%). Of the 60 episodes with a positive third test, 36 (60%) were adherent with serial testing. Of the 1375 episodes with a negative third test, 246 (17.9%) were adherent with serial testing. For patients without symptoms, the rate of testing positive via serial testing was lower at 4.2%. This lower rate of testing positive for patients without symptoms is logical, as patients without symptoms are less likely to have COVID-19.

Limitations

This study's strength is its large sample size of over 2 million patients over a multiyear time frame. However, there are several limitations that are important to note. Testing data are limited to tests that occurred at vendor sites and do not include at-home tests or tests performed outside of ICATT, creating the potential for misclassification bias in the adherence outcome. As a result, some patients classified as nonadherent may in fact have completed follow-up testing outside the observed pharmacy network. Similarly, it is assumed that the first ICATT test is the start of an individual's testing episode, when a patient may be using ICATT as a follow-up for a test they took at home or at a non-ICATT facility. In addition, observation of repeat testing was limited within the pharmacy chain because patient IDs were generated separately by vendors and were not used across vendors, so linking testing records for the same patient between different pharmacy chains was not possible. Finally, vendors performed testing according to the authorized instructions for use and counseled patients on the serial testing protocol; however, there was no measure of vendor compliance or patient compliance, and there was no measure of whether patients knew the importance of serial testing. Future work could seek to validate the findings presented here by measuring at-home testing rates or examining testing in other health care settings among demographic groups targeted by the ICATT program, as well as measuring knowledge of the serial testing safety communication.

Furthermore, ICATT is a nonprobability sample that targets underresourced communities; as a result, the findings presented here may not be representative of the broader US population. In addition, the study's cut-off date extends well beyond the issued safety communication date, and it is possible that the likelihood of adherence due to the safety communication was lower for dates further out from the announcement. Third, a gender variable closely linked to sex was used to determine a patient's sex because sex was not collected by the ICATT vendors. Finally, there was a high degree of missingness among self-reported clinical variables. Consequently, the expanded models were restricted to complete cases, and if the completeness of the clinical variables was nonrandom, those estimates may be affected by selection bias. The markedly smaller complete-case samples—especially for the third-test expanded model—also reduced precision and statistical power.

XSL•FO
RenderX

Conclusion

This study is the first to evaluate adherence with SARS-CoV-2 serial Ag-RDT testing following the issuance of an FDA safety communication. Our results highlight the value of public-private programs such as ICATT, which not only directly delivered testing services to uninsured individuals during the COVID-19 pandemic but also provided a detailed source of data to evaluate the effectiveness of government safety communications, such as serial testing instructions for Ag-RDTs, as well as the real-world performance of medical countermeasures such as diagnostic testing. The CDC ICATT program can serve as a model for public health programs and data infrastructure that can improve the readiness, response, and monitoring of the US public health authorities to future biological threats.

Using a large national dataset, this study demonstrates that nearly all patients did not have follow-up testing captured in the dataset according to the FDA's test instructions. It is possible that this result may be attributable to the increased availability of at-home tests over time, and it justifies the further development of data systems that can track individual patient testing across modalities for additional study. Despite low adherence to testing and noncompliant timing of testing, serial testing successfully identified infections missed by earlier tests at significant rates, indicating the public health benefit of serial testing.

With the availability of at-home testing, a significant proportion of second- and third-test patients likely used at-home tests rather than going back to the pharmacy, reducing the rates of detection through serial testing compared with those observed in this study. There was significant variability in second-test and third-test adherence rates among distinct demographic groups. Individuals with symptoms and children had lower odds of both second- and third-test adherence; Asian individuals had higher odds of second-test adherence but lower odds of third-test adherence; and individuals in the West had lower odds of second-test adherence but higher odds of third-test adherence. When clinical factors were incorporated into the model, adherence was significantly higher among patients with recent COVID-19 contact, but lower among those who had COVID-19 within the past 90 days. The rate of testing positive through serial testing was 24.9% for patients with symptoms on the second test and 4.2% for patients without symptoms on the third test. Finally, adherence was greater at community pharmacy sites compared with chain pharmacies, and second-test adherence for all test takers declined steadily over time. Altogether, these findings emphasize the need for additional targeted public health strategies to improve adherence, particularly among groups, sites, and later time periods with lower rates, to improve overall disease detection and management in future outbreaks.

Acknowledgments

We appreciate the millions of patients and thousands of health care providers who pursued and conducted testing during the pandemic. The authors would also like to thank the following people for their assistance in reviewing the manuscript: Dr Matthew Bevan, Dr Lucy Carruth, Dr Zahra Chaudhry, Mr Jim Hess, Dr Travis Lim, Mr Edward McDonald, and Mr Michael Obringer from the Johns Hopkins University Applied Physics Laboratory. Finally, we would like to acknowledge the initial vision for the Community-Based Testing program by Admiral Brett Giroir, Admiral Erica Schwartz, and Lieutenant Colonel Tiffany Bilderback.

The authors declare the use of generative AI (GenAI) in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to GenAI tools under full human supervision: text generation, refinement of the layout of Figure 2, proofreading and editing, and summarizing text. The GenAI tools used were Gemini 2.0, GPT 5.2, GPT 5.4 Pro, and GPT 5.5 Pro. Responsibility for the final manuscript lies entirely with the authors. GenAI tools are not listed as authors and do not bear responsibility for the final outcomes. Declaration submitted by: collective responsibility.

Funding

This study was funded by the US Centers for Disease Control and Prevention (CDC) Increasing Community Access to Testing, Treatment, and Response (ICATT) program.

Data Availability

The datasets analyzed during this study are available in the One CDC Data Platform (1CDP). Due to the sensitivity of the public health data and privacy regulations, the datasets are not publicly available but may be requested from the CDC. Anonymized line-level data are available upon request to the corresponding author and subject to an approved data use agreement.

Authors' Contributions

MZM, EAC, ZRS, ACG, and JDM were significantly involved in the conception of the study. MZM and CCP wrote the manuscript and conducted all statistical analyses. ACG, MS, GMT, RMS, TCT, MFI, and JDM provided critical intellectual revisions to the content. All authors reviewed and approved the final manuscript.

Conflicts of Interest

ZRS was employed at Walgreens Co. at the date of article publication.

XSL•FO
RenderX

References

1. CAP survey of laboratories confirms COVID-19 testing challenges, supply shortages, and excess capacity. College of American Pathologists. 2020. URL: https://newsroom.cap.org/latest-news/cap-survey-of-laboratories-confirms-covid-19-testing-challenges--supply-shortages--and-excess-capaci/s/c8c8c213-d4f3-4dc1-86d0-b0aa311bb86b [accessed 2025-09-24]
2. Huq Ronny FM, Sherpa T, Choesang T, Ahmad S. Looking into the laboratory staffing issues that affected ambulatory care clinical laboratory operations during the COVID-19 pandemic. Lab Med 2023;54(4):e114-e116 [FREE Full text] [doi: 10.1093/labmed/lmac139] [Medline: 36282479]
3. Chu VT, Schwartz NG, Donnelly MAP, Chuey MR, Soto R, Yousaf AR, COVID-19 Household Transmission Team. Comparison of home antigen testing with RT-PCR and viral culture during the course of SARS-CoV-2 infection. JAMA Intern Med 2022;182(7):701-709 [FREE Full text] [doi: 10.1001/jamainternmed.2022.1827] [Medline: 35486394]
4. At-home COVID-19 antigen tests-take steps to reduce your risk of false negative results: FDA safety communication. US Food and Drug Administration. 2022. URL: https://web.archive.org/web/20221110193958/https://www.fda.gov/medical-devices/safety-communications/home-covid-19-antigen-tests-take-steps-reduce-your-risk-false-negative-results-fda-safety [accessed 2025-01-03]
5. Lab advisory: FDA advises repeat testing to reduce false negative results with COVID-19 antigen tests. Centers for Disease Control and Prevention. 2022. URL: https://web.archive.org/web/20250201025611/https://www.cdc.gov/locs/2022/08-12-2022-Lab-Advisory-FDA_Advises_Repeat_Testing_Reduce_False_Negative_Results_COVID-19_Antigen_Tests.html [accessed 2026-06-17]
6. In vitro diagnostics EUAs - antigen diagnostic tests for SARS-CoV-2. US Food and Drug Administration. 2025. URL: https://www.fda.gov/medical-devices/covid-19-emergency-use-authorizations-medical-devices/in-vitro-diagnostics-euas-antigen-diagnostic-tests-sars-cov-2#SerialTesting [accessed 2026-06-17]
7. Soni A, Herbert C, Pretz C, Stamegna P, Filippaios A, Shi Q, et al. Design and implementation of a digital site-less clinical study of serial rapid antigen testing to identify asymptomatic SARS-CoV-2 infection. J Clin Transl Sci 2023;7(1):e120 [FREE Full text] [doi: 10.1017/cts.2023.540] [Medline: 37313378]
8. Soni A, Herbert C, Lin H, Yan Y, Pretz C, Stamegna P, et al. Performance of rapid antigen tests to detect symptomatic and asymptomatic SARS-CoV-2 infection : a prospective cohort study. Ann Intern Med 2023;176(7):975-982 [FREE Full text] [doi: 10.7326/M23-0385] [Medline: 37399548]
9. Ritchey MD, Rosenblum HG, Del Guercio K, Humbard M, Santos S, Hall J, et al. COVID-19 self-test data: challenges and opportunities - United States, October 31, 2021-June 11, 2022. MMWR Morb Mortal Wkly Rep 2022;71(32):1005-1010 [FREE Full text] [doi: 10.15585/mmwr.mm7132a1] [Medline: 35951486]
10. MakeMyTestCount. 2024. URL: https://www.makemytestcount.net/ [accessed 2026-06-17]
11. Bigouette JP, Ford L, Pray I, Langolf K, Kahrs J, Zochert T, et al. Application of a serial antigen-based testing strategy for severe acute respiratory syndrome coronavirus 2 and student adherence in a university setting: Wisconsin, October-November 2020. Open Forum Infect Dis 2021;8(10):ofab472 [FREE Full text] [doi: 10.1093/ofid/ofab472] [Medline: 34692890]
12. Prabhu RS, Dhakal R, Hicks AS, McBride J, Patrick AL, Corso CD, et al. Implementation, adherence, and results of systematic SARS-CoV-2 testing for asymptomatic patients treated at a tertiary care regional radiation oncology network. Radiat Oncol 2021;16(1):28 [FREE Full text] [doi: 10.1186/s13014-021-01760-2] [Medline: 33541359]
13. Williams LR, Emary KRW, Phillips DJ, Hay J, Larwood JPJ, Ramasamy MN, et al. Implementation and adherence to regular asymptomatic testing in a COVID-19 vaccine trial. Vaccine 2024;42(21):126167 [FREE Full text] [doi: 10.1016/j.vaccine.2024.126167] [Medline: 39060202]
14. Ratnayake A, Hassig S, He H, Lederer A, Kissinger PJ. Predictors of adherence to a COVID-19 serial testing program among university affiliates in the Deep South. J Public Health Manag Pract 2025;31(4):585-590. [doi: 10.1097/PHH.0000000000002124] [Medline: 39964176]
15. Wu S, Archuleta S, Lim SM, Somani J, Quek SC, Fisher D. Serial antigen rapid testing in staff of a large acute hospital. Lancet Infect Dis 2022;22(1):14-15 [FREE Full text] [doi: 10.1016/S1473-3099(21)00723-4] [Medline: 34883064]
16. Tsai TC, Yu E, Godoshian A, Hoelscher M, Iademarco MF, Walke H, et al. The increasing community access to testing, treatment, and response program — A model for future public health emergency response. NEJM Catal Innov Care Deliv 2026;7(5). [doi: 10.1056/cat.25.0149]
17. Public health surveillance activities. Johns Hopkins Medicine. URL: https://www.hopkinsmedicine.org/institutional-review-board/guidelines-policies/guidelines/public-health-surveillance-activities [accessed 2026-06-17]
18. Protection of Human Subjects (45 CFR pt 46); Institutional Review Boards (21 CFR pt 56). Electronic Code of Federal Regulations (eCFR). URL: https://www.ecfr.gov/current/title-45/subtitle-A/subchapter-A/part-46 [accessed 2026-06-25]
19. COVID-19 national wastewater data. Centers for Disease Control and Prevention. 2025. URL: https://www.cdc.gov/nwss/rv/COVID19-national-data.html [accessed 2025-09-24]
20. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis 2020;20(5):533-534 [FREE Full text] [doi: 10.1016/S1473-3099(20)30120-1] [Medline: 32087114]

XSL•FO
RenderX

21. Kim AE, Bennett JC, Luiten K, O'Hanlon JA, Wolf CR, Magedson A, et al. Comparative diagnostic utility of SARS-CoV-2 rapid antigen and molecular testing in a community setting. J Infect Dis 2024;230(2):363-373 [FREE Full text] [doi: 10.1093/infdis/jiae150] [Medline: 38531685]
22. Rader B, Gertz A, Iuliano AD, Gilmer M, Wronski L, Astley CM, et al. Use of at-home COVID-19 tests—United States, August 23, 2021–March 12, 2022. MMWR Morb Mortal Wkly Rep 2022;71(13):489-494 [FREE Full text] [doi: 10.15585/mmwr.mm7113e1] [Medline: 35358168]
23. Bevan I, Stage Baxter M, Stagg HR, Street A. Knowledge, attitudes, and behavior related to COVID-19 testing: a rapid scoping review. Diagnostics (Basel) 2021;11(9):1685 [FREE Full text] [doi: 10.3390/diagnostics11091685] [Medline: 34574026]
24. Nind J, Smith A, Scahill S, Marra CA. A systematic review of independent and chain pharmacies effects on medication adherence. Pharmacy (Basel) 2022;10(5):124 [FREE Full text] [doi: 10.3390/pharmacy10050124] [Medline: 36287445]

Abbreviations

Ag-RDT: antigen-detection rapid diagnostic test
AOR: adjusted odds ratio
CDC: Centers for Disease Control and Prevention
FDA: Food and Drug Administration
HRSA: Health Resources and Services Administration
ICATT: Increasing Community Access to Testing, Treatment, and Response
NAAT: nucleic acid amplification test
RT-PCR: reverse transcriptase–polymerase chain reaction

Edited by A Mavragani, T Sanchez; submitted 09.Feb.2026; peer-reviewed by N Hafer, I Rinaldi; comments to author 08.Apr.2026; revised version received 27.Apr.2026; accepted 30.Apr.2026; published 15.Jul.2026.

Please cite as:
McGarity MZ, Pollack CC, Currier EA, Godoshian AC, Smith ZR, Siddall M, Tavolacci GM, Salerno RM, Tsai TC, Iademarco MF, Miller JD
Characteristics Influencing Adherence to SARS-CoV-2 Serial Antigen Testing Following a Federal Safety Communication: Cohort Study
JMIR Public Health Surveill 2026;12:e92002
URL: https://publichealth.jmir.org/2026/1/e92002
doi:10.2196/92002
PMID:42456150

©Micajah Z McGarity, Catherine C Pollack, Elizabeth A Currier, Andrew C Godoshian, Zachary R Smith, Marcus Siddall, Gregg M Tavolacci, Reynolds M Salerno, Thomas C Tsai, Michael F Iademarco, Joseph D Miller. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 15.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Hand, Foot, and Mouth Disease Surveillance in Anhui, After Introduction of the Enterovirus 71 Vaccine and During the COVID-19 Pandemic: Ecological Interrupted Time Series Study

Kangguo Li[1*], BD; Jiadong Wu[1,2*], MD; Wanwan Ma[3*], MD; Tao Chen[1], BD; Yunzhi Zenghuang[1], BD; Jia Rui[4], PhD; Zeyu Zhao[5], PhD; Jiabing Wu[3], MD; Tianmu Chen[1], PhD

[1]State Key Laboratory of Molecular Vaccinology and Molecular Diagnostics, School of Public Health, Xiamen University, Xiamen, Fujian, China
[2]Fuzhou Center for Disease Control and Prevention, Fuzhou, Fujian, China
[3]Department of Acute Infectious Disease Prevention and Control, Anhui Provincial Center for Disease Control and Prevention, Hefei, Anhui, China
[4]Department of Epidemiology and Health Statistics, Xiangya School of Public Health, Central South University, Changsha, Hunan, China
[5]WorldPop, School of Geography and Environmental Science, University of Southampton, Southampton, England, United Kingdom
[*]these authors contributed equally

Corresponding Author:
Tianmu Chen, PhD
State Key Laboratory of Molecular Vaccinology and Molecular Diagnostics, School of Public Health
Xiamen University
4221-117, XiangAn South Road
Xiamen, Fujian
China
Phone: 86 13661934715
Email: chentianmu@xmu.edu.cn

Abstract

Background: Routine surveillance is important for understanding hand, foot, and mouth disease (HFMD) after monovalent enterovirus 71 (EV71) vaccine introduction, but interpretation is difficult when etiologic typing is sparse and the COVID-19 pandemic overlaps with the postvaccine era.

Objective: This study described surveillance-derived changes in HFMD burden and typed serotype composition in Anhui Province, China, after EV71 vaccine introduction and during the COVID-19 pandemic.

Methods: We conducted a retrospective ecological surveillance study using HFMD records from January 2008 to December 2023. The primary analysis used negative-binomial interrupted time series (ITS) counterfactual models for all reported HFMD cases and typed EV71 monthly counts. Vaccine-era models were trained using data from January 2013 through May 2016; the previous specification covering January 2010 through May 2016 was retained as a sensitivity analysis. COVID-19–period models were trained using data from January 2008 through December 2019. Secondary analyses assessed spatial, temporal, and age patterns. Vaccine coverage histories, individual vaccination records, specimen submission metadata, and laboratory platform-change metadata were unavailable; therefore, models estimated calendar-based surveillance deviations rather than vaccine effectiveness.

Results: Anhui reported 1,292,579 HFMD cases from 2008 to 2023, including 47,990 (3.7%) typed cases. EV71 declined from 45.8% (8479/18,534) of typed cases before vaccine introduction to 3.4% (578/16,930) during COVID-19, while other enteroviruses increased from 29.6% (5487/18,534) to 81.8% (13,848/16,930). There were 123 deaths from 2008 to 2023, corresponding to 9.52 deaths per 100,000 reported cases; no deaths were recorded from 2020 to 2023. In primary vaccine-era ITS models, typed EV71 counts remained below the predicted trajectory in 2017 to 2019: 470 observed vs 659 predicted cases (95% prediction interval [PI] 467-933) in 2017, 181 vs 531 (95% PI 379-798) in 2018, and 63 vs 427 (95% PI 305-662) in 2019. All reported HFMD cases were not uniformly below the vaccine-era counterfactual. During the COVID-19 pandemic, all reported HFMD cases remained below prediction through 2023, with the largest relative difference in 2022 (28,385 observed vs 175,601 predicted; 95% PI 111,465-268,202). The proportion of reported children aged ≥4 years increased from 19% (50,503/266,132) during 2017 to 2019 to 29.1% (44,133/151,795) during 2021 to 2023 (odds ratio 1.75, 95% CI 1.72-1.78; $P<.001$).

Conclusions: Typed EV71 surveillance signals declined after EV71 vaccine introduction, and overall HFMD burden was strongly disrupted during the COVID-19 pandemic. However, sparse typing, unavailable vaccination and testing policy metadata,

XSL•FO
RenderX

and the aggregated other enteroviruses category limit causal interpretation. These findings should be read as surveillance signals, not as direct estimates of population-level serotype replacement or individual-level vaccine protection.

(JMIR Public Health Surveill 2026;12:e99663) doi:10.2196/99663

KEYWORDS

hand, foot, and mouth disease; enterovirus A71; enterovirus 71 vaccine; coxsackievirus A16; disease surveillance; interrupted time series analysis; COVID-19; spatiotemporal analysis; China; public health surveillance

Introduction

Hand, foot, and mouth disease (HFMD) remains an important childhood infection in China and is caused by multiple enteroviruses [1,2]. Enterovirus 71 (EV71) has long been of particular concern because it is disproportionately associated with severe neurological complications and death [3-5]. Since inactivated monovalent EV71 vaccines were introduced in China in 2016, postlicensure surveillance has been used to examine whether typed HFMD patterns have shifted away from EV71 and toward non-EV71 enteroviruses [6-9].

Routine surveillance is essential for detecting population-level signals, but it is not equivalent to an individually linked vaccine effectiveness study [10]. HFMD surveillance systems do not type every reported case, and the typed subset may be affected by case selection, disease severity, reagent availability, local laboratory capacity, and changes in diagnostic platforms [11,12]. These constraints are especially important when broad categories such as “other enteroviruses” are used because changes in testing breadth may alter which infections become visible to the surveillance system [12-16].

Interpretation became more difficult after January 2020, when the COVID-19 pandemic altered health care use, child contact patterns, and disease reporting [17-23]. Intensive nonpharmaceutical interventions (NPIs) coincided with sharp short-term reductions in HFMD in multiple settings [17-20], while the later rebound has been interpreted as compatible with renewed transmission and an expanded susceptible pool [21-23]. In practice, EV71 vaccine introduction and the COVID-19 period are temporally contiguous disruptions in the same surveillance series. For an ecological dataset, this overlap limits the clean separation of vaccine-era trend shifts from pandemic-era perturbation.

Many province-level HFMD reports have described declining EV71 detection and rising non-EV71 detection after vaccine introduction [6,19,24-26]. What is often less explicit is the inferential boundary: with sparse typing and incomplete testing metadata, typed serotype composition is a surveillance signal rather than a direct estimate of population serotype structure [27]. Long-term provincial surveillance remains valuable because it can show how routine reporting, etiologic typing, seasonality, spatial heterogeneity, severe outcomes, and pandemic-era rebound evolved in the same public health system.

Using provincial HFMD surveillance data from Anhui Province from 2008 to 2023, we aimed to describe how reported burden and typed serotype composition changed after EV71 vaccine introduction and during the COVID-19 period. We prioritized ecological interrupted time series (ITS) counterfactual analyses of all reported HFMD cases and typed EV71 cases, while treating spatial analyses, county-level typed-dominance summaries, seasonality assessment, and joinpoint regression as secondary descriptive analyses. Because individual vaccination records and province-wide vaccine coverage histories were unavailable, the study was framed as an ecological surveillance analysis rather than a causal assessment of vaccine effectiveness.

Methods

Study Design and Data Source

We conducted a retrospective ecological surveillance study using deidentified HFMD records provided by the Anhui Provincial Center for Disease Control and Prevention. The analytic window was January 1, 2008, through December 31, 2023. Anhui Province comprises 16 prefecture-level cities and 104 county-level administrative units. Because different surveillance outputs were available at different spatial resolutions, analyses were conducted at mixed levels by design: prefecture-level city units were used for incidence mapping and spatiotemporal cluster detection; county-level units were used for typed-dominance mapping; and province-level aggregation was used for seasonality, joinpoint, and ITS analyses.

The surveillance extract included date of symptom onset, administrative unit, age, sex, clinical severity, and etiologic typing result when available. Population denominators were obtained from the Anhui Statistical Yearbook [28]. HFMD cases were reported through the statutory notifiable disease surveillance system. Clinical diagnosis followed the national HFMD diagnostic standard, which requires compatible symptoms, such as fever with rash or vesicles affecting the hands, feet, mouth, or buttocks; cases without rash are not clinically diagnosed as HFMD [29]. We followed the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) reporting principles, where applicable, to an aggregated surveillance analysis [30].

Surveillance and Testing Context

Under the national diagnostic standard, severe HFMD includes central nervous system involvement, such as poor mental status, somnolence, startle, delirium, headache, vomiting, limb tremor, myoclonus, nystagmus, ataxia, weakness, acute flaccid paralysis, convulsions, or meningeal signs; critical disease includes severe central nervous system injury, respiratory dysfunction, or circulatory dysfunction [29]. In this analysis, severe cases were records marked as severe in the surveillance extract. The extract contained a death-date field but did not contain follow-up information on neurological sequelae.

XSL•FO
RenderX

Several surveillance features define the boundary of interpretation. June 1, 2016, was used as the EV71 vaccine–era boundary, and January 1, 2020, to December 31, 2023, was treated as the COVID-19 period. Etiologic typing was incomplete and time varying: 47,990 (3.7%) of 1,292,579 reported cases had a recorded etiologic result, with annual typing fractions ranging from 0.3% (76/26,359) in 2008 to 8.4% (4429/52,712) in 2020. Typed analyses therefore describe surveillance-derived typed composition rather than population serotype structure. The surveillance extracts also lacked individual vaccination records, province-wide vaccine coverage histories, annual specimen submission criteria, and laboratory platform-change records. These gaps precluded direct estimation of vaccine effectiveness, adjustment for testing policy changes, or attribution of changes in the aggregated other enteroviruses category to true population-level serotype replacement.

Outcomes and Definitions

The primary burden end point was the monthly number of reported HFMD cases. The primary typing end point was the monthly number of typed EV71 cases. Typed cases were also categorized as coxsackievirus A16 (CV-A16) or other enteroviruses. In this surveillance extract, other enteroviruses represented typed results recorded as neither EV71 nor CV-A16; the category could not distinguish coxsackievirus A6 (CV-A6), coxsackievirus A10 (CV-A10), or other individual serotypes. Deaths were identified from nonmissing death dates, and case fatality was summarized as deaths per 100,000 reported cases.

For descriptive analyses, we summarized annual reported cases, annual incidence, typing fractions, typed serotype composition, sex distribution, age distribution, severe cases, deaths, and case fatality. Age-shift analyses grouped cases with known age as children aged ≤1 year, aged 2 to 3 years, or aged ≥4 years. For county-level typed-dominance maps, we masked county-year observations with <10 typed cases because dominant-serotype assignments based on small number of typed cases are not epidemiologically stable.

Statistical Analysis

We first described overall burden, typed composition, and age-sex structure over time. Seasonality was summarized with epidemic curves and a wavelet analysis of weekly reported cases. Annual incidence maps at the prefecture-level city scale, spatiotemporal scan statistics, Moran *I*, and joinpoint regression were retained as secondary descriptive analyses. Retrospective spatiotemporal scan statistics used a discrete Poisson model, a maximum spatial cluster size of 15% of the population at risk, a maximum temporal window of 1 year, and monthly temporal aggregation. Cluster-size sensitivity analyses used maximum spatial windows of 10% and 20%. Joinpoint models were used to summarize annual trend inflections for all reported HFMD incidence, sex-specific incidence, age group counts, and typed serotype counts.

The primary analysis used negative-binomial ITS counterfactual models with month-specific seasonal terms. Following recent counterfactual-prediction guidance, we specified the target observed series, intervention dates, training windows, evaluation metrics, and prediction-interval summaries before interpreting postintervention contrasts [31,32]. For vaccine-era primary models, we used a shorter and more stable preintervention window (January 2013 to May 2016) and specified a linear secular trend to avoid overfitting the short training set. The previous January 2010 to May 2016 quadratic specification was retained as a sensitivity analysis to preserve comparability with previous analyses. For COVID-19–period models, we trained models using data from January 2008 to December 2019 with a linear secular trend. To assess whether the COVID-19 counterfactual depended on including vaccine-era months in the baseline, we also fit a prevaccine sensitivity model using January 2008 to December 2015 as the training window.

Model performance was evaluated on the training period by comparing observed monthly counts with model-fitted counts. We reported mean absolute error (MAE) and root mean square error (RMSE) as primary fit diagnostics because they are expressed in the same count scale as the modeled outcomes; mean absolute percentage error (MAPE) was also tabulated but interpreted cautiously because sparse-typed monthly counts can inflate percentage errors. For each fitted model, we generated postintervention counterfactual predictions and estimated 95% prediction intervals (PIs) by simulating negative-binomial monthly counts from the fitted model and aggregating simulated postintervention counts by year. We summarized observed counts, point predictions, PIs, signed count differences, and relative differences by year. Because province-wide monthly denominators for specimen submission and testing were unavailable, we did not include typing fractions as regressors in the primary models. Instead, we retained previously generated typing-fraction sensitivity analyses as context in Multimedia Appendix 1.

To assess potential severity-related differences within the typed subset, we compared EV71 proportions between severe and mild typed records and fitted adjusted multinomial models of typed serotype category with calendar year, sex, age group, and severity as covariates. To examine the postpandemic age shift, we compared the age composition of reported HFMD cases with known age from 2017 to 2019 vs from 2021 to 2023, performed a chi-square test across age group distributions, and fitted a binomial model comparing the odds that a reported case belonged to the ≥4-year age group of children from 2021 to 2023 vs from 2017 to 2019.

All analyses were performed in R (version 4.5.2; R Foundation for Statistical Computing). Spatiotemporal scan statistics used SaTScan (version 10.2.5; Martin Kulldorff and Information Management Services Inc), and joinpoint analyses used the *nih.joinpoint* R package.

Ethical Considerations

This study analyzed deidentified routine HFMD surveillance data provided by the Anhui Provincial Center for Disease Control and Prevention as part of the statutory notifiable infectious disease surveillance system. Permission to use the deidentified surveillance data for retrospective analysis was granted by the Anhui Provincial Center for Disease Control and Prevention. The study involved secondary analysis of anonymized records, without direct participant contact, intervention, recruitment, or compensation. No names,

XSL•FO
RenderX

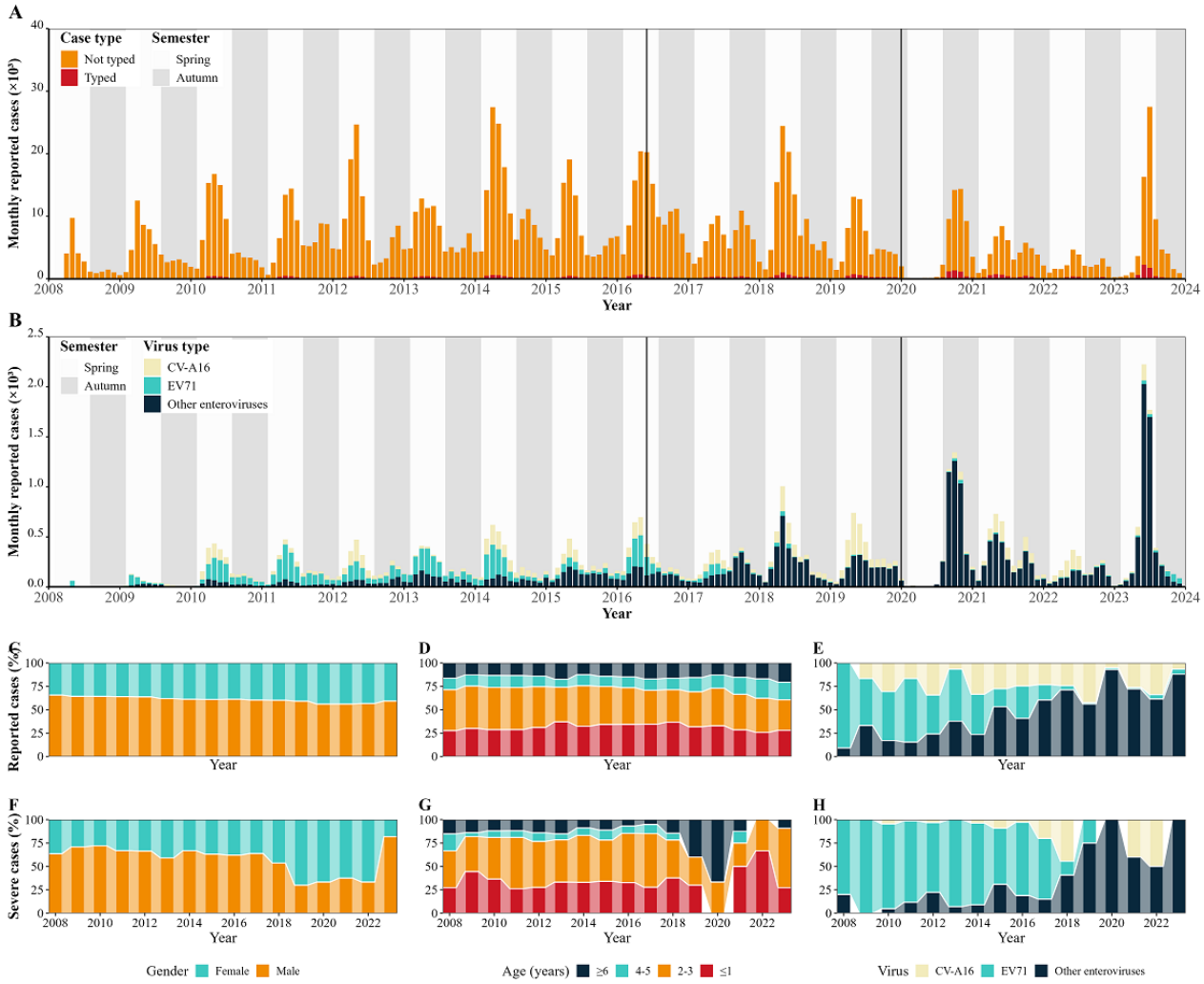
identification numbers, addresses, telephone numbers, or other direct personal identifiers were available to the authors, and all public outputs are reported as aggregated temporal or spatial summaries. Under local requirements for routine public health surveillance data [33], separate ethics committee approval and individual informed consent were not required.

Results

Overview of Reported and Typed HFMD Cases

From January 2008 through December 2023, Anhui reported 1,292,579 HFMD cases, including 47,990 (3.7%) typed cases. Annual incidence peaked in 2014, with 145,528 reported cases (241 per 100,000 population), while the lowest annual incidence occurred in 2022, with 28,385 reported cases (46 per 100,000 population; Table S1 in Multimedia Appendix 1). Monthly burden retained a bimodal seasonal pattern, with a larger spring–early summer peak and a smaller autumn peak (Figure 1A; Figure S1 in Multimedia Appendix 1). Wavelet analysis identified significant 6- and 12-month periodicities in the reported case series (Figure S2 in Multimedia Appendix 1). The single highest monthly count occurred in July 2023, with 27,489 reported cases.

Figure 1. Epidemic curves and case composition of hand, foot, and mouth disease (HFMD) in Anhui Province (2008-2023). (A) Monthly reported HFMD cases, categorized as typed or not typed; (B) monthly typed HFMD cases, displayed as enterovirus 71 (EV71), coxsackievirus A16 (CV-A16), and other enteroviruses; (C, D, and E) annual proportions of cases by sex, age group, and typed serotype, respectively, among all reported cases or typed cases, as appropriate; and (F, G, and H) annual proportions of cases by sex, age group, and typed serotype, respectively, among all severe cases or typed severe cases, as appropriate. Vertical lines indicate June 1, 2016 (EV71 vaccine–era analytic boundary), and January 1, 2020 (start of the COVID-19 period).



HFMD remained concentrated in young children across the full study period. Among the 1,292,579 reported cases, 540,368 (41.8%) occurred among children aged ≤1 year, 517,445 (40%) among children aged 2 to 3 years, and 234,766 (18.2%) among children aged ≥4 years. Male participants accounted for 791,458 (61.2%) reported cases, compared with 501,121 (38.8%) among female participants. Overall, 3064/1,292,579 (0.24%) severe cases were recorded during 2008 to 2023. Severe cases became rare after 2018, with 10/70,759 (0.014%) severe cases in 2019 and 3 to 11 severe cases annually from 2020 to 2023. The death-date field identified 123 deaths during 2008 to 2023, corresponding to 9.52 deaths per 100,000 reported cases; no deaths were recorded during 2020 to 2023 (Figure S5 and Tables S6 and S7 in Multimedia Appendix 1).

XSL•FO
RenderX

Typed Serotype Composition

Among typed cases, the January 2008 to May 2016 period included 18,534 typed detections: EV71 represented 8479 (45.8%), CV-A16 represented 4568 (24.6%), and other enteroviruses represented 5487 (29.6%; Figure 1B; Table S5 in Multimedia Appendix 1). The immediate postrollout prepandemic period (June 2016 to December 2019) included 12,526 typed detections: EV71 accounted for 1114 (8.9%), CV-A16 accounted for 3718 (29.7%), and other enteroviruses accounted for 7694 (61.4%). The COVID-19 period (January 2020 to December 2023) included 16,930 typed detections: other enteroviruses accounted for 13,848 (81.8%), EV71 accounted for 578 (3.4%), and CV-A16 accounted for 2504 (14.8%) (Figure S3 in Multimedia Appendix 1).

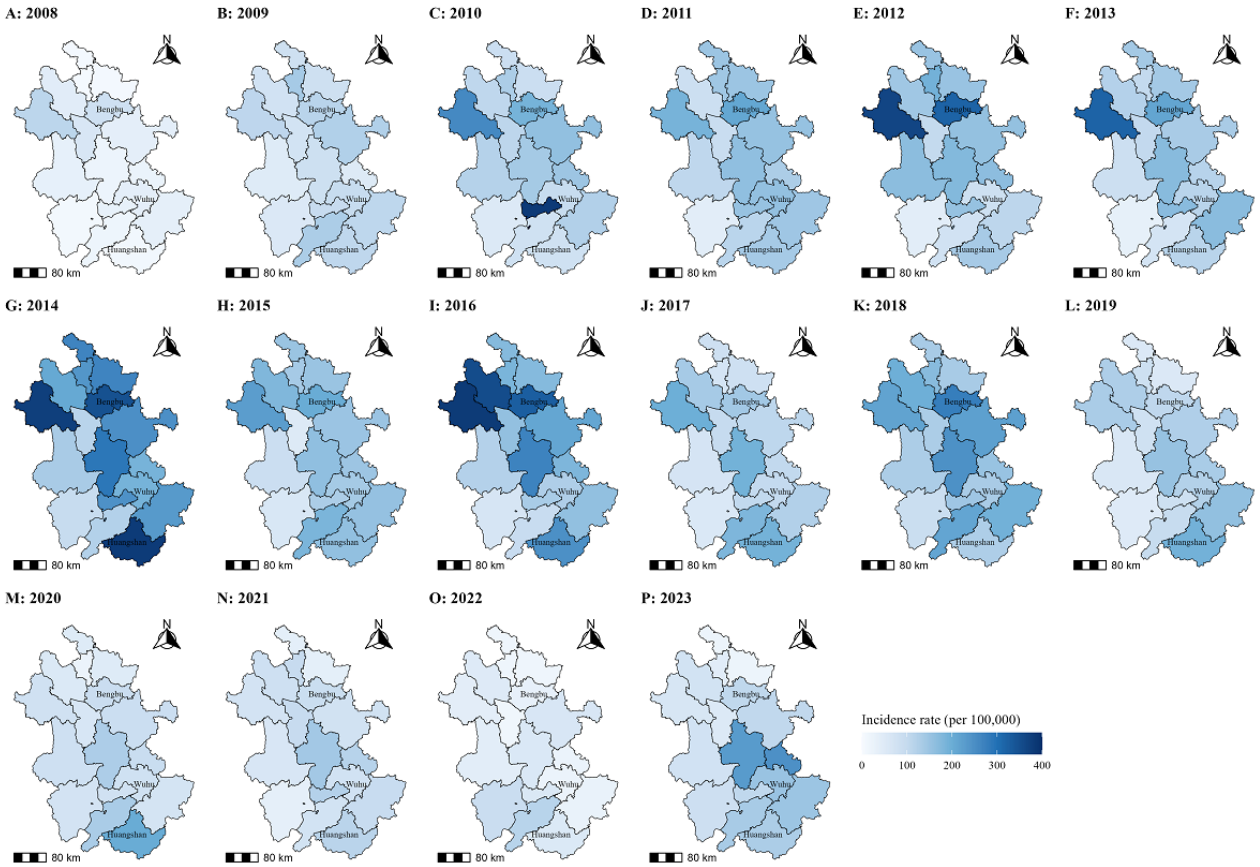
The annual proportion of reported cases with typing increased from 0.3% (76/26,359) of reported cases in 2008 to 5.5% (3919/70,759) in 2019, 8.4% (4429/52,712) in 2020, and 8.4% (5884/70,288) in 2023. In annual typed composition, EV71 accounted for 90.8% (69/76) of typed records in 2008, 34.6% (1266/3664) in 2016, 16.3% (470/2885) in 2017, and 1.6% (63/3919) in 2019. Other enteroviruses accounted for 93.1% (4122/4429) of typed records in 2020 and 88.1% (5186/5884) in 2023 (Table S5 in Multimedia Appendix 1).

In an adjusted multinomial model of typed records, severe cases were more likely than mild cases to be typed as EV71 rather than other enteroviruses (relative risk ratio 7.08, 95% CI 6.01-8.34; Table S5, and Table S8 in Multimedia Appendix 1). Older age groups (2-3 years and ≥4 years) were also more likely than the reference group (≤1 year) to be typed as EV71 or CV-A16 rather than other enteroviruses, while male sex had little association with EV71 classification (relative risk ratio 1.01, 95% CI 0.95-1.07).

Secondary Spatial and Temporal Analyses

Annual incidence maps showed persistent geographic heterogeneity across Anhui (Figure 2; Figure S4 in Multimedia Appendix 1). Retrospective spatiotemporal scan statistics identified 26 cluster windows, of which 25 were significant at $P<.05$ (Table S2 in Multimedia Appendix 1). Wuhu was the most frequently identified cluster center (11 windows), followed by Huangshan, Ma'anshan, and Bengbu (4 windows each). The largest log-likelihood cluster was centered in Huangshan from April to October 2014, with 18,324 observed cases vs 5195 expected cases (relative risk 3.56). The highest relative risk cluster was centered in Ma'anshan from April to June 2012, with 11,585 observed cases vs 2444 expected cases (relative risk 4.77). Annual global Moran I was significant in 2008 (Moran $I=0.43$; $P<.001$), 2016 (Moran $I=0.38$; $P=.002$), and 2021 (Moran $I=0.26$; $P=.02$), but not in most other years (Table S3 in Multimedia Appendix 1).

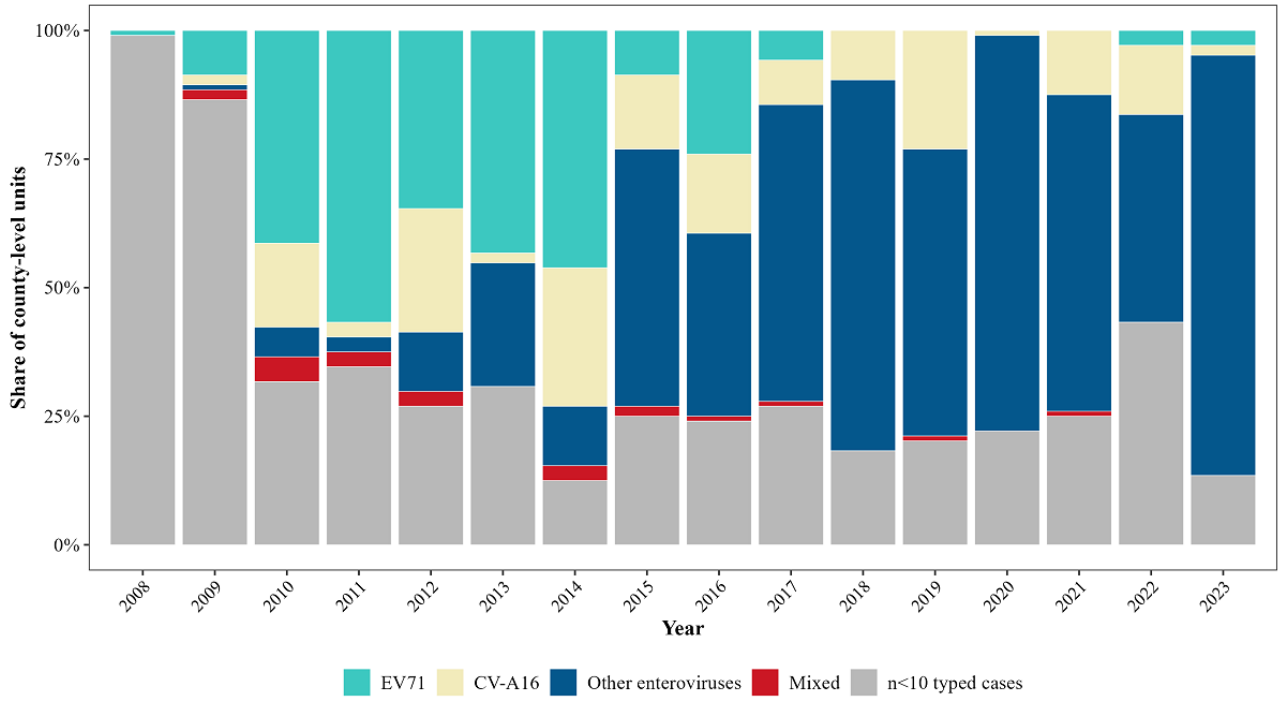
Figure 2. Annual incidence maps of reported hand, foot, and mouth disease (HFMD) in Anhui Province, China (2008-2023). Each panel shows the incidence rate for each prefecture-level city in the corresponding year. Labels for Bengbu, Wuhu, and Huangshan are added as spatial references because these locations are discussed in the text.



XSL•FO
RenderX

County-level typed-dominance summaries showed that apparent EV71-dominant county-years decreased over time, while other enterovirus–dominant county-years became more frequent (Figure 3; Figure S8 in Multimedia Appendix 1). County-year observations with <10 typed cases were masked. Dominant-serotype labels indicate only the largest typed category in a county-year and do not measure the margin between the leading and second-leading typed categories.

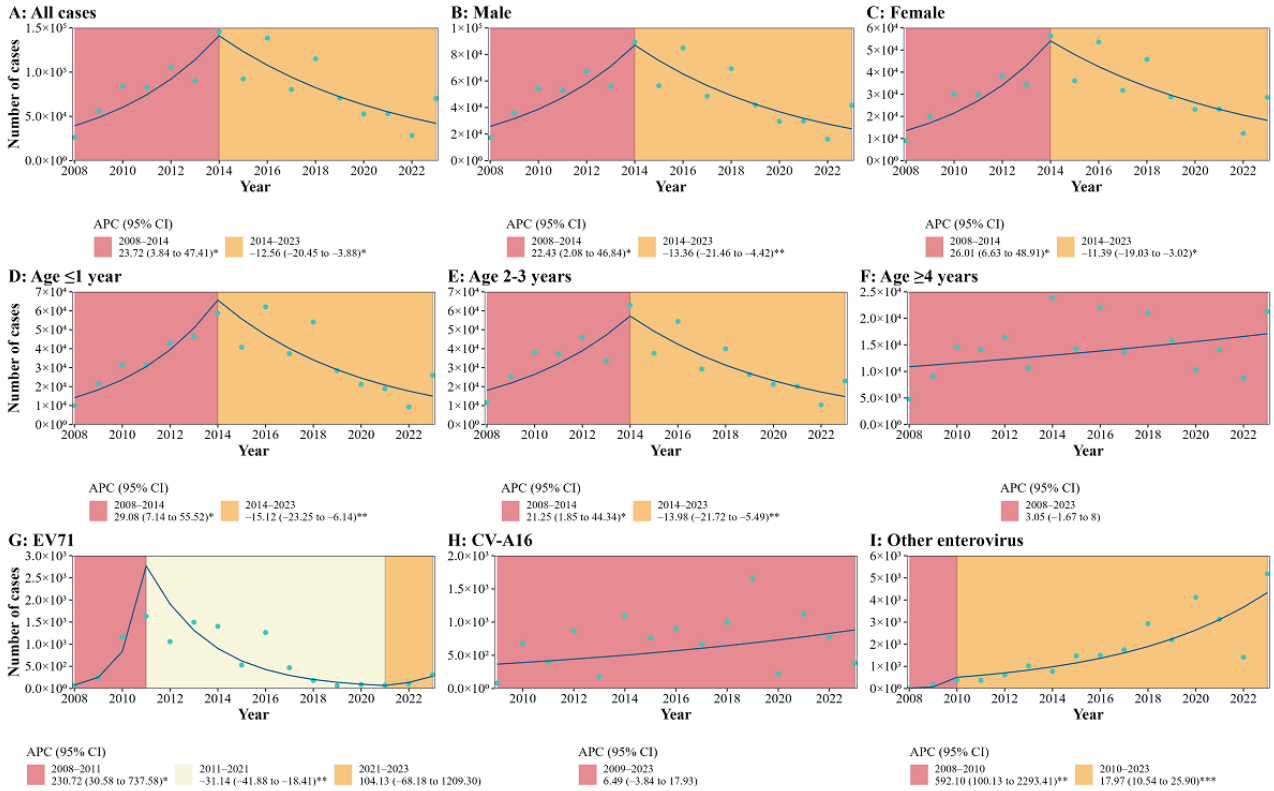
Figure 3. Annual county-level dominance of typed serotype after masking sparse observations in Anhui Province, China (2008-2023). The 100% stacked bar chart shows the share of county-level units classified as enterovirus 71 (EV71)–dominant, coxsackievirus A16 (CV-A16)–dominant, other enterovirus–dominant, mixed, or masked because of <10 typed cases. Counties with <10 typed cases in a given year were not assigned a dominant typed serotype and are shown as masked observations. Dominance indicates the largest typed category in a county-year and should not be interpreted as evidence that one serotype was substantially more frequent than the next most common typed serotype.



Joinpoint regression suggested a province-wide inflection around 2014 for all reported HFMD incidence and a sustained decline thereafter, whereas typed EV71 counts showed a downward long-term pattern after the early years of laboratory expansion (Figure 4). Count-regression sensitivity analyses for annual typed counts showed a negative annual trend for EV71 in the negative-binomial model (annual percentage change −14.4%, 95% CI −22% to −6%), a positive annual trend for other enteroviruses (annual percentage change 19.5%, 95% CI 19.2%-19.9%), and no statistically clear annual trend for CV-A16.

XSL•FO
RenderX

Figure 4. Joinpoint regression of annual hand, foot, and mouth disease (HFMD) trends in Anhui Province, China (2008-2023). (A) Overall incidence, (B) male incidence, (C) female incidence, (D) age ≤1 year, (E) age 2 to 3 years, (F) age ≥4 years, (G) typed enterovirus 71 (EV71) counts, (H) typed coxsackievirus A16 (CV-A16) counts, and (I) typed other enterovirus counts. Dots indicate observed values, and lines indicate fitted segments with 95% CIs. APC: annual percent change. P<0.05, P<0.01, P<0.001.



ITS Analyses of Vaccine-Era Models

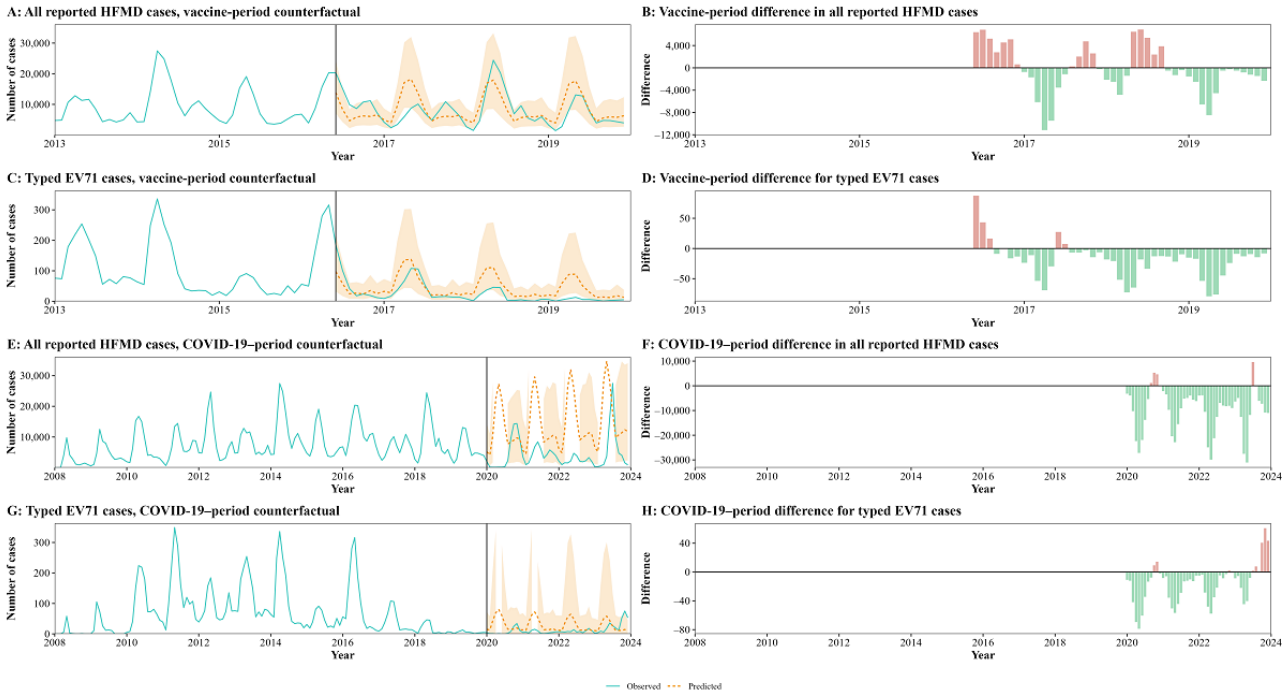
Primary vaccine-era ITS models were trained using data from January 2013 to May 2016 and included a linear secular trend (Figures 5A and 5B). Training-period MAE and RMSE were 2065.5 and 2950.6 monthly cases, respectively, for all reported HFMD cases and 35.5 and 53.4 monthly cases, respectively, for typed EV71 cases. Observed annual typed EV71 counts remained below the predicted trajectory throughout 2017 to 2019 (Figures 5C and 5D; PI summaries are provided in Table S4 in Multimedia Appendix 1): 470 observed vs 659 predicted cases (95% PI 467–933) in 2017, 181 vs 531 (95% PI 379–798) in 2018, and 63 vs 427 (95% PI 305–662) in 2019. These corresponded to annual relative differences of −28.7%, −65.9%, and −85.2%, respectively.

For all reported HFMD cases, observed burden was lower than predicted in 2017 (80,344 observed vs 104,712 predicted; 95% PI 84,388-129,080) and 2019 (70,759 observed vs 101,494 predicted; 95% PI 82,487-132,654), but higher than the point prediction in 2018 (115,029 observed vs 103,091 predicted; 95% PI 82,975-130,561; Figures 5A and 5B). Secondary typed series showed lower-than-predicted CV-A16 counts in 2017 to 2019, while typed other enterovirus counts were below prediction in 2017 and 2019 but above the point prediction in 2018 (Table S3, Table S4 and Figure S7 in Multimedia Appendix 1).

Sensitivity analyses using the previous January 2010 to May 2016 quadratic training window preserved the same direction of post-2016 negative divergence for typed EV71, although effect magnitudes were attenuated compared with the primary model (Table S9 in Multimedia Appendix 1). The longer-window specification yielded 2017 to 2019 typed EV71 counts that were 17%, 49.7%, and 70% below prediction, respectively, compared with 28.7%, 65.9%, and 85.2% in the primary short-window model.

XSL•FO
RenderX

Figure 5. Primary ecological interrupted time series counterfactual analyses for hand, foot, and mouth disease (HFMD) surveillance in Anhui Province, China. The primary models focused on all reported HFMD cases and typed enterovirus 71 (EV71) cases. Panels show observed monthly case counts, predicted counterfactual trajectories, and 95% prediction intervals (PIs), along with postintervention difference plots. (A-D) Vaccine-era models (trained January 2013–May 2016) showing (A, B) all reported HFMD cases and (C, D) typed EV71 cases. (E-H) COVID-19–period models (trained January 2008–December 2019) showing (E, F) all reported HFMD cases and (G, H) typed EV71 cases.



ITS Analyses of COVID-19–Period Models

COVID-19–period ITS models were trained from January 2008 to December 2019. Training-period MAE and RMSE were 2756.2 and 3821.3 monthly cases, respectively, for all reported HFMD and 47.4 and 70.0 monthly cases, respectively, for typed EV71. All reported HFMD cases remained below the pre-2020 counterfactual throughout 2020 to 2023 (Figures 5E and 5F; Table S4 in Multimedia Appendix 1): 52,712 observed vs 149,675 predicted (95% PI 96,603-222,919) in 2020, 53,122 vs 162,121 (95% PI 105,863-243,423) in 2021, 28,385 vs 175,601 (95% PI 111,465-268,202) in 2022, and 70,288 vs 190,201 (95% PI 127,746-284,820) in 2023. The corresponding annual relative differences were −64.8%, −67.2%, −83.8%, and −63%.

Typed EV71 counts were also below predicted values in 2020 to 2022 (Figure 5G and 5H): 87 observed vs 395 predicted in 2020, 72 vs 355 in 2021, and 106 vs 319 in 2022. In 2023, typed EV71 was slightly above the point prediction but remained within the 95% PI (313 observed vs 287 predicted; 95% PI 113-678). Typed CV-A16 and typed other enterovirus counts were below predicted values throughout 2020 to 2023 in the primary COVID-19–period models (Table S4 and Figure S7 in Multimedia Appendix 1).

A sensitivity model trained only on the prevaccine period (January 2008 to December 2015) yielded the same direction of post-2020 suppression for all reported HFMD and the typed serotype series (Table S10 in Multimedia Appendix 1).

Postpandemic Age Shift

The age composition of reported HFMD cases shifted upward after the most restrictive pandemic period (Figure S6 in Multimedia Appendix 1). Among cases with known age, the proportion of children aged ≥4 years increased from 19% (50,503/266,132) in 2017 to 2019 to 29.1% (44,133/151,795) in 2021 to 2023, while the proportion aged ≤1 year decreased from 45% (119,868/266,132) to 35.6% (54,102/151,795). The ≥4-year age group accounted for 26.5% (14,060/53,122) in 2021, 30.9% (8765/28,385) in 2022, and 30.3% (21,308/70,288) in 2023.

The age distribution comparison between 2017 and 2019 and 2021 and 2023 was statistically significant (χ^2_2=6417.0; P<.001). The odds that a reported case belonged to the ≥4-year age group were higher in 2021 to 2023 than in 2017 to 2019 (odds ratio 1.75, 95% CI 1.72-1.78).

Discussion

Principal Results

This study should be read as a long-term surveillance analysis, not as a formal assessment of vaccination-related protection. The main findings were that typed EV71 surveillance signals declined after EV71 vaccine introduction, overall reported HFMD burden did not show a uniform postrollout decrease, and the COVID-19 pandemic created a large disruption in the same ecological time series. The revised analyses also showed that severe and fatal outcomes became rare in the later years, that other enterovirus records became dominant within the typed subset, that spatial clustering persisted in selected years and

XSL•FO
RenderX

locations, and that the postpandemic rebound was accompanied by an older age distribution among reported cases.

The most important interpretive boundary is the typed subset. Only 3.7% (47,990/1,292,579) reported cases had etiologic typing. Individual vaccination records, province-wide vaccine coverage histories, testing policy metadata, and laboratory platform-change records were unavailable. In addition, severe cases were more likely than mild cases to be classified as EV71 rather than other enteroviruses among typed records. These features mean that the typed EV71 series cannot be treated as an unbiased proxy for community serotype incidence.

Interpretation of Typed EV71 and Vaccine-Era Signals

Within those limits, the vaccine-era primary ITS models showed a sustained negative divergence of typed EV71 counts below the counterfactual trajectory during 2017 to 2019. This pattern is directionally consistent with reports from other Chinese provinces after EV71 vaccine introduction [6,24-26]. However, the magnitude of our typed EV71 counterfactual contrast was larger than several published population-level estimates. Guangdong reported a 41.4% reduction in EV71-associated HFMD in 2017 to 2019 with increasing vaccine coverage, and Jiangsu reported an estimated 45.6% reduction with a wide credible interval [24,25]. By comparison, our primary short-window model estimated that typed EV71 counts were 28.7%, 65.9%, and 85.2% below predicted values in 2017, 2018, and 2019, respectively, while the longer-window sensitivity analysis estimated smaller reductions of 17%, 49.7%, and 70%, respectively.

Several mechanisms could make the Anhui-typed EV71 contrast appear larger than a true population-level vaccine effect. Early typing fractions were low, the typed subset may have overrepresented severe cases, severe EV71-associated disease may have been preferentially depleted after vaccine rollout, and changes in specimen submission or laboratory platforms could have altered the case mix entering the typed series. For that reason, the relevant message is not that Anhui had unusually high vaccine impact, but that routine typed surveillance recorded a sustained decline in EV71 signals while the overall reported HFMD burden remained substantial.

The aggregated other enteroviruses category requires similar caution. In this dataset, other enteroviruses were recorded only as typed results that were neither EV71 nor CV-A16. That grouping does not identify which serotypes drove the increase, and it could be sensitive to changes in diagnostic breadth over time [12,15,16]. If multiplex polymerase chain reaction became more common during the later study period, previously underdetected CV-A6, CV-A10, or other enteroviruses could have been more readily classified into the other enteroviruses group [12-16]. We therefore interpret the rise of other enteroviruses as a change in typed surveillance composition rather than definitive evidence of population-level serotype replacement.

COVID-19 Disruption and Postpandemic Age Shift

The COVID-19 period created a distinct ecological disruption. All reported HFMD counts remained far below the pre-2020 counterfactual through 2023, consistent with documented suppression of contact-based infectious diseases during intensive NPIs [17-22]. The same direction held when the COVID-19 counterfactual was anchored only to prevaccine data, reducing concern that the pandemic signal was an artifact of including vaccine-era months in the baseline. At the same time, the postpandemic age composition shifted toward older children. This local age shift is compatible with an immunity-gap or immunity-debt interpretation, but the evidence remains indirect because the surveillance system does not measure susceptibility or prior infection history [23,34-37]. We therefore present this mechanism as a plausible interpretation supported by age-structured surveillance patterns, not as a proven causal explanation. A pre-vaccine meta-analysis from China documented increasing EV71 antibody seroprevalence with age among preschool children [38], providing historical context for age-related immunity rather than direct evidence of a pandemic-related immunity gap. From a policy perspective, the age shift suggests that postpandemic HFMD preparedness should not focus only on infants and very young preschool children. If linked vaccination and coverage data become available, Anhui could evaluate whether unvaccinated older preschool or school-entry children would benefit from catch-up EV71 vaccination; however, such a strategy cannot be recommended from the current ecological data alone, especially because absolute typed EV71 counts remained low and the vaccine does not protect against non-EV71 HFMD [3-8].

Spatial and Seasonal Surveillance Findings

The secondary spatial and seasonal analyses add descriptive context to the ITS findings. The bimodal seasonal pattern and significant 6- and 12-month periodicities indicate that seasonal forcing remained a recurring feature of Anhui HFMD activity across the long surveillance window [34,35]. Spatiotemporal scan statistics repeatedly identified clusters centered in Wuhu, Huangshan, Bengbu, and Ma'anshan, but annual Moran *I* was significant only in selected years. These findings indicate spatial heterogeneity in reported burden, but they do not identify causal drivers such as mobility, school calendars, climate, health care access, socioeconomic conditions, or vaccine uptake [34,35].

County-level typed-dominance maps should be read even more cautiously than incidence maps. The dominance label is categorical, depends only on typed records, and can be unstable when typed counts are small. Masking county-years with <10 typed cases reduces the most extreme instability, but it does not remove selection into testing or measure how close the leading typed category was to the next most common category. These maps are therefore best used to flag areas for enhanced surveillance rather than to infer local replacement of serotypes.

Comparison With Prior Work

Our findings overlap with prior Chinese surveillance studies showing reduced EV71 detection after vaccine introduction and persistent circulation of non-EV71 enteroviruses [6,9,19,24-26]. The added value of the current study is not mechanistic discovery or broad causal generalization beyond Anhui. Instead, it is the explicit framing of what surveillance data can and cannot support when typing is sparse, when the non-EV71 category is aggregated, and when the COVID-19 period interrupts the same time series [10,12,13,27]. This framing aligns more closely with

XSL•FO
RenderX

public health surveillance priorities than with causal inference regarding vaccine impact.

The findings also align with broader reports that COVID-19 pandemic control measures suppressed multiple pediatric and contact-transmitted infections, followed in some settings by rebound or altered age distributions when contact patterns normalized [17-23]. Anhui adds a province-level example in which pandemic-era suppression occurred against the background of an already changing EV71 vaccine–era surveillance series.

Limitations

This study has several limitations. First, the absence of individual vaccination records and province-wide vaccine coverage histories means that we could not account for actual vaccine uptake, estimate vaccine effectiveness, or distinguish vaccine-program effects from other contemporaneous changes [10,24,25]. Second, the surveillance system was passive, so reporting may have been influenced by health care–seeking behavior and pandemic-era service disruption. Third, etiologic typing remained sparse and likely nonrandom, with severe cases more likely to be typed and the association between severity and typing changing over time. We did not have province-wide metadata on specimen submission criteria, the balance between mild and severe cases selected for testing, or annual changes in laboratory platform use [10-13]. Fourth, the aggregated other enteroviruses category may have been influenced by diagnostic drift; wider adoption of multiplex polymerase chain reaction could make CV-A6, CV-A10, or other previously underdetected serotypes more visible without proving population-level niche replacement [12-16,27]. Fifth, because the postvaccine period and the COVID-19 period are contiguous in the same ecological series, no stable interdisruption interval was available for formal statistical separation of their independent effects. Sixth, although the death-date field allowed us to summarize fatal outcomes, the surveillance extract did not include follow-up information on neurological sequelae, so we could not estimate the proportion of survivors with neurological sequelae after severe HFMD. Finally, county-level dominant-serotype maps are sensitive to sparse-typed counts even after masking county-year cells with <10 typed cases.

Public Health Implications

The practical contribution of this analysis is a surveillance blueprint rather than an estimate of vaccine effectiveness. Future HFMD surveillance would be more interpretable if specimen submission rules were standardized across sites and years; if sentinel hospitals routinely used multiplex polymerase chain reaction or sequencing capable of distinguishing CV-A6, CV-A10, and other non-EV71 serotypes; and if surveillance extracts preserved the denominator of submitted and successfully tested specimens [11-16]. Linkage to immunization registries would be needed before the same surveillance system could support vaccine effectiveness analyses [10,24,25]. Wastewater or environmental surveillance could provide an additional low-cost early-warning stream for non-EV71 enteroviruses, especially when clinical typing remains sparse [39,40].

Conclusions

Routine HFMD surveillance in Anhui showed a decline in typed EV71 surveillance signals after EV71 vaccine introduction, a concurrent rise in typed other enterovirus detections, and strong pandemic-era suppression of overall HFMD burden followed by partial rebound and an older-age shift. These findings support extending surveillance priorities beyond EV71 alone. For practice, this means standardizing specimen submission rules, introducing routine multiplex polymerase chain reaction in eligible sentinel hospitals, preserving testing denominators and vaccination histories in surveillance extracts, and considering wastewater surveillance as a lower-cost supplementary early-warning approach. For interpretation, this requires maintaining ecological humility: changes in typed surveillance composition can illuminate temporal patterns but cannot, on their own, establish vaccine-attributable protection or population-level serotype replacement.

Acknowledgments

The authors thank the Anhui Provincial Center for Disease Control and Prevention for data access and surveillance support. During revision, the authors used ChatGPT (version 5.6; OpenAI) to assist with language editing and consistency checks across the manuscript and multimedia appendix. The authors reviewed and verified all generated text, analyses, and interpretations and take full responsibility for the final content.

Funding

This research was supported by the National Science and Technology Major Project (grant 2027ZD01999504), the Major Project of Guangzhou National Laboratory (grant GZNL2024A01004), and the National Key Research and Development Program of China (grant 2024YFE0214800).

Data Availability

Supplementary figures and tables are provided in Multimedia Appendix 1. Aggregated analytic data, analysis scripts, and analytic outputs supporting this study are publicly available in our GitHub repository [41]. The individual-level surveillance records are not publicly available because they contain sensitive statutory notifiable disease surveillance information and are governed by the Anhui Provincial Center for Disease Control and Prevention; requests for additional access would require approval from the

XSL•FO
RenderX

data custodian. Publicly released materials contain aggregated temporal and spatial counts rather than individual-level surveillance records.

Authors' Contributions

KL, Jiadong Wu, and Tianmu Chen conceptualized the study. Jiadong Wu and KL developed the analytic strategy, wrote the analysis code, performed the formal analyses, validated the results, and prepared the figures and tables. WM, Jiabing Wu, Tao Chen and YZ curated the surveillance data, verified variable definitions, and provided context on the surveillance system. JR contributed to epidemiological interpretation and review of the ecological inference framework. ZZ contributed to the interpretation of spatial and population data and reviewed the mapping analyses. Jiabing Wu oversaw surveillance data access, data governance, and interpretation of the Anhui reporting context. Tianmu Chen provided overall supervision, project administration, and funding acquisition. KL, Jiadong Wu, and WM drafted the manuscript. Jiadong Wu, ZZ, Jia Rui, Jiabing Wu, and Tianmu Chen critically revised the manuscript for important intellectual content. All authors approved the final version and agreed to be accountable for the work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary figures and tables for descriptive epidemiology, spatial and temporal analyses, and interrupted time series analyses of hand, foot, and mouth disease in Anhui, China.
[DOCX File , 4099 KB - publichealth_v12i1e99663_app1.docx]

References

1. Xing W, Liao Q, Viboud C, Zhang J, Sun J, Wu JT, et al. Hand, foot, and mouth disease in China, 2008-12: an epidemiological study. Lancet Infect Dis 2014 Apr;14(4):308-318 [FREE Full text] [doi: 10.1016/S1473-3099(13)70342-6] [Medline: 24485991]
2. Hong J, Liu F, Qi H, Tu W, Ward MP, Ren M, et al. Changing epidemiology of hand, foot, and mouth disease in China, 2013-2019: a population-based study. Lancet Reg Health West Pac 2022 Jan 2;20:100370 [FREE Full text] [doi: 10.1016/j.lanwpc.2021.100370] [Medline: 35036978]
3. Zhu F, Xu W, Xia J, Liang Z, Liu Y, Zhang X, et al. Efficacy, safety, and immunogenicity of an enterovirus 71 vaccine in China. N Engl J Med 2014 Feb 27;370(9):818-828. [doi: 10.1056/NEJMoa1304923] [Medline: 24571754]
4. Ooi MH, Wong SC, Lewthwaite P, Cardosa MJ, Solomon T. Clinical features, diagnosis, and management of enterovirus 71. Lancet Neurol 2010 Nov;9(11):1097-1105. [doi: 10.1016/S1474-4422(10)70209-X] [Medline: 20965438]
5. Solomon T, Lewthwaite P, Perera D, Cardosa MJ, McMinn P, Ooi MH. Virology, epidemiology, pathogenesis, and control of enterovirus 71. Lancet Infect Dis 2010 Nov;10(11):778-790. [doi: 10.1016/S1473-3099(10)70194-8] [Medline: 20961813]
6. Meng XD, Tong Y, Wei ZN, Wang L, Mai JY, Wu Y, et al. Epidemical and etiological study on hand, foot and mouth disease following EV-A71 vaccination in Xiangyang, China. Sci Rep 2020 Dec 01;10(1):20909 [FREE Full text] [doi: 10.1038/s41598-020-77768-7] [Medline: 33262488]
7. Li X, Chen S, Chen Y, Han S, Dai B, Li T, et al. Epidemiological and genetic characterizations of hand, foot, and mouth disease and acute respiratory infections due to CV-A6 infection in Henan Province, China between 2021 and 2022. BMC Pediatr 2025 Apr 02;25(1):264 [FREE Full text] [doi: 10.1186/s12887-025-05527-6] [Medline: 40170027]
8. Mao Q, Wang Y, Bian L, Xu M, Liang Z. EV-A71 vaccine licensure: a first step for multivalent enterovirus vaccine to control HFMD and other severe diseases. Emerg Microbes Infect 2016 Jul 20;5(7):e75 [FREE Full text] [doi: 10.1038/emi.2016.73] [Medline: 27436364]
9. Head JR, Collender PA, Lewnard JA, Skaff NK, Li L, Cheng Q, et al. Early evidence of inactivated enterovirus 71 vaccine impact against hand, foot, and mouth disease in a major center of ongoing transmission in China, 2011-2018: a longitudinal surveillance study. Clin Infect Dis 2020 Dec 15;71(12):3088-3095 [FREE Full text] [doi: 10.1093/cid/ciz1188] [Medline: 31879754]
10. Mues KE, Zhou CK, Gerber JE, van Hunsel F, Klein NP, Izurieta HS, et al. A review of methodologic & data considerations for vaccine safety surveillance in the wake of the COVID-19 pandemic. Vaccine 2025 Oct 03;64:127691. [doi: 10.1016/j.vaccine.2025.127691] [Medline: 40913819]
11. Kim BI, Achangwa C, Cho S, Ahn J, Won J, Do H, et al. The hand, foot, and mouth disease sentinel surveillance system in South Korea: retrospective evaluation study. JMIR Public Health Surveill 2024 Jul 23;10:e59446 [FREE Full text] [doi: 10.2196/59446] [Medline: 39045828]
12. Liu F, Ren M, Chen S, Nie T, Cui J, Ran L, et al. Pathogen spectrum of hand, foot, and mouth disease based on laboratory surveillance - China, 2018. China CDC Wkly 2020 Mar 14;2(11):167-171 [FREE Full text] [Medline: 34594617]

XSL•FO
RenderX

13. Li Y, Chang Z, Wu P, Liao Q, Liu F, Zheng Y, et al. Emerging enteroviruses causing hand, foot and mouth disease, China, 2010-2016. Emerg Infect Dis 2018 Oct;24(10):1902-1906 [FREE Full text] [doi: 10.3201/eid2410.171953] [Medline: 30226172]
14. He YQ, Chen L, Xu WB, Yang H, Wang HZ, Zong WP, et al. Emergence, circulation, and spatiotemporal phylogenetic analysis of coxsackievirus a6- and coxsackievirus a10-associated hand, foot, and mouth disease infections from 2008 to 2012 in Shenzhen, China. J Clin Microbiol 2013 Nov;51(11):3560-3566 [FREE Full text] [doi: 10.1128/JCM.01231-13] [Medline: 23966496]
15. Nguyen AT, Ta HT, Dao AT, Vu HM, Ngu ND, Tran DN, et al. Lineage replacement of coxsackievirus A10 in Vietnam associated with increased detection and altered disease severity grading of hand, foot, and mouth disease. Arch Virol 2026 Feb 24;171(3):94. [doi: 10.1007/s00705-026-06526-3] [Medline: 41733718]
16. Wang Z, Wang J, Wang Y, Li Y, Dong Z, Ji J, et al. Coxsackievirus A6 was the predominant pathogen of hand, foot, and mouth disease with frequent recombination and mutations in Shandong Province, 2023. Virus Evol 2025 Dec 18;12(1):veaf105 [FREE Full text] [doi: 10.1093/ve/veaf105] [Medline: 41583791]
17. Li K, Rui J, Song W, Luo L, Zhao Y, Qu H, et al. Temporal shifts in 24 notifiable infectious diseases in China before and during the COVID-19 pandemic. Nat Commun 2024 May 08;15(1):3891 [FREE Full text] [doi: 10.1038/s41467-024-48201-8] [Medline: 38719858]
18. Brett TS, Rohani P. Collateral effects of COVID-19 pandemic control on the US infectious disease landscape. Science 2025 Oct 30;390(6772):510-515. [doi: 10.1126/science.adw4964] [Medline: 41166479]
19. Hu J, Yu L, Peng Z, Pan H, Jin Y, Zeng Z, et al. Dynamic changes of hand, foot, and mouth disease epidemic crossing three periods in Jiangxi Province, China: a long-term epidemiology retrospective study from 2009 to 2022. BMC Infect Dis 2026 Jan 29;26(1):429 [FREE Full text] [doi: 10.1186/s12879-026-12705-z] [Medline: 41612221]
20. Ang HJ, Menegale F, Preziosi G, Pariani E, Migliari M, Pellegrinelli L, et al. Reconstructing the impact of COVID-19 on the immunity gap and transmission of respiratory syncytial virus in Lombardy, Italy. EBioMedicine 2023 Sep;95:104745 [FREE Full text] [doi: 10.1016/j.ebiom.2023.104745] [Medline: 37566927]
21. Niu Y, Luo L, Rui J, Yang S, Deng B, Zhao Z, et al. Control measures during the COVID-19 outbreak reduced the transmission of hand, foot, and mouth disease. J Saf Sci Resil 2021 Jun;2(2):63-68 [FREE Full text] [doi: 10.1016/j.jnlssr.2021.06.002] [Medline: 40477419]
22. Hourani AZ, Abdelsalam A, Sürmeli AD. The impact of COVID-19 non-pharmaceutical interventions on notifiable infectious diseases in Poland: a comprehensive analysis from 2014-2022. BMC Infect Dis 2026 Feb 03;26(1):271 [FREE Full text] [doi: 10.1186/s12879-025-12478-x] [Medline: 41634598]
23. Nygaard U, Holm M, Rabie H, Rytter M. The pattern of childhood infections during and after the COVID-19 pandemic. Lancet Child Adolesc Health 2024 Dec;8(12):910-920. [doi: 10.1016/S2352-4642(24)00236-0] [Medline: 39572124]
24. Tong Y, Fan H, Wang J, Zeng X, Cheng X, Bao C, et al. Population-level impact of the enterovirus A71 vaccination program on hand, foot, and mouth disease: ecological time-series study. JMIR Public Health Surveill 2026 Mar 10;12:e85604 [FREE Full text] [doi: 10.2196/85604] [Medline: 41805655]
25. Xiao J, Zhu Q, Yang F, Zeng S, Zhu Z, Gong D, et al. The impact of enterovirus A71 vaccination program on hand, foot, and mouth disease in Guangdong, China: a longitudinal surveillance study. J Infect 2022 Oct;85(4):428-435. [doi: 10.1016/j.jinf.2022.06.020] [Medline: 35768049]
26. Nie J, Huang T, Sun Y, Peng Z, Dong W, Chen J, et al. Influence of the enterovirus 71 vaccine and the COVID-19 pandemic on hand, foot, and mouth disease in China based on counterfactual models: observational study. JMIR Public Health Surveill 2024 Dec 17;10:e63146 [FREE Full text] [doi: 10.2196/63146] [Medline: 39692430]
27. Pons-Salort M, Grassly NC. Serotype-specific immunity explains the incidence of diseases caused by human enteroviruses. Science 2018 Aug 24;361(6404):800-803 [FREE Full text] [doi: 10.1126/science.aat6777] [Medline: 30139872]
28. Anhui Provincial Bureau of Statistics, Survey Office of the National Bureau of Statistics in Anhui. Anhui Statistical Yearbook 2023. Beijing, China: China Statistics Press; 2023.
29. Diagnosis of hand, foot and mouth disease. National Health Commission of the People's Republic of China. 2018. URL: https://www.nhc.gov.cn/wjw/s9491/201803/fd30701a6bb24564b081349aaf6ff024.shtml [accessed 2026-06-04]
30. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med 2007 Oct 16;4(10):e296 [FREE Full text] [doi: 10.1371/journal.pmed.0040296] [Medline: 17941714]
31. Keogh RH, Van Geloven N. Prediction under interventions: evaluation of counterfactual performance using longitudinal observational data. Epidemiology 2024 May 01;35(3):329-339. [doi: 10.1097/EDE.0000000000001713] [Medline: 38630508]
32. Boyer CB, Dahabreh IJ, Steingrimsson JA. Estimating and evaluating counterfactual prediction models. Stat Med 2025 Oct;44(23-24):e70287. [doi: 10.1002/sim.70287] [Medline: 41055572]
33. Notice on issuing the measures for ethical review of life science and medical research involving human subjects [Article in Chinese]. National Health Commission of the People's Republic of China. 2023. URL: https://www.nhc.gov.cn/qjjys/c100016/202302/6b6e447b3edc4338856c9a652a85f44b.shtml [accessed 2026-09-08]

XSL•FO
RenderX

34. Van Boeckel TP, Takahashi S, Liao Q, Xing W, Lai S, Hsiao V, et al. Hand, foot, and mouth disease in China: critical community size and spatial vaccination strategies. Sci Rep 2016 Apr 29;6:25248 [FREE Full text] [doi: 10.1038/srep25248] [Medline: 27125917]
35. Zhao J, Li X. Determinants of the transmission variation of hand, foot and mouth disease in China. PLoS One 2016 Oct 4;11(10):e0163789 [FREE Full text] [doi: 10.1371/journal.pone.0163789] [Medline: 27701445]
36. Zhao J, Jiang F, Zhong L, Sun J, Ding J. Age patterns and transmission characteristics of hand, foot and mouth disease in China. BMC Infect Dis 2016 Nov 21;16(1):691 [FREE Full text] [doi: 10.1186/s12879-016-2008-y] [Medline: 27871252]
37. Le J, Hong J, Zhao Z, Chen Y, Hu Y, Chang Z, et al. Age-specific transmission for different virus serotypes of hand, foot and mouth disease and the impact of interventions in East China, 2009-2015. Heliyon 2022 Dec 01;8(12):e12042 [FREE Full text] [doi: 10.1016/j.heliyon.2022.e12042] [Medline: 36478843]
38. Yang B, Wu P, Wu JT, Lau EH, Leung GM, Yu H, et al. Seroprevalence of enterovirus 71 antibody among children in China: a systematic review and meta-analysis. Pediatr Infect Dis J 2015 Dec;34(12):1399-1406 [FREE Full text] [doi: 10.1097/INF.0000000000000900] [Medline: 26368058]
39. Shrestha S, Malla B, Haramoto E. Evaluation of the enterovirus serotype monitoring approach for wastewater surveillance of hand foot and mouth disease using secondary epidemiological surveillance data. Sci Total Environ 2025 Mar 15;969:178896 [FREE Full text] [doi: 10.1016/j.scitotenv.2025.178896] [Medline: 40010248]
40. Ozawa H, Yoshida H, Usuku S. Environmental surveillance can dynamically track ecological changes in enteroviruses. Appl Environ Microbiol 2019 Nov 27;85(24):e01604-19 [FREE Full text] [doi: 10.1128/AEM.01604-19] [Medline: 31585989]
41. xmusphlkg/Anhui-HFMD. GitHub. URL: https://github.com/xmusphlkg/Anhui-HFMD [accessed 2026-09-08]

Abbreviations

CV-A6: coxsackievirus A6
CV-A10: coxsackievirus A10
CV-A16: coxsackievirus A16
EV71: enterovirus 71
HFMD: hand, foot, and mouth disease
ITS: interrupted time series
MAE: mean absolute error
MAPE: mean absolute percentage error
NPI: nonpharmaceutical intervention
PI: prediction interval
RMSE: root mean square error
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by F Pabon-Rodriguez; submitted 28.Apr.2026; peer-reviewed by Y Cho, R Chacon, T Yu; comments to author 02.Jun.2026; revised version received 29.Jun.2026; accepted 02.Jul.2026; published 24.Sep.2026.

Please cite as:
Li K, Wu J, Ma W, Chen T, Zenghuang Y, Rui J, Zhao Z, Wu J, Chen T
Hand, Foot, and Mouth Disease Surveillance in Anhui, After Introduction of the Enterovirus 71 Vaccine and During the COVID-19 Pandemic: Ecological Interrupted Time Series Study
JMIR Public Health Surveill 2026;12:e99663
URL: https://publichealth.jmir.org/2026/1/e99663
doi:10.2196/99663
PMID:

©Kangguo Li, Jiadong Wu, Wanwan Ma, Tao Chen, Yunzhi Zenghuang, Jia Rui, Zeyu Zhao, Jiabing Wu, Tianmu Chen. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 24.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Social Determinants of Childhood Vaccination Coverage in the United States Using National Immunization Survey Data From 2010 to 2023: Cross-Sectional Study

Ramya Keerthi Majji[1*], BDS, MS; Yan Zhuang[1], BS, PhD; Felix M Pabon-Rodriguez[1,2*], BS, MS, PhD

[1]Department of Biomedical Engineering and Informatics, Luddy School of Informatics, Computing, and Engineering, Indiana University, Indianapolis, IN, United States

[2]Department of Biostatistics and Health Data Science, School of Medicine, Indiana University, Indianapolis, IN, United States

[*]these authors contributed equally

Corresponding Author:
Felix M Pabon-Rodriguez, BS, MS, PhD
Department of Biostatistics and Health Data Science
School of Medicine
Indiana University
410 W 10th St
Indianapolis, IN, 46202
United States
Phone: 1 3172786135
Email: fpabonrodriguez@gmail.com

Abstract

Background: Vaccination in early childhood is essential to prevent serious infectious diseases and protect community health. In the United States, 7 vaccines (diphtheria, tetanus, and acellular pertussis; polio; measles, mumps, and rubella; *Haemophilus influenzae* type b; hepatitis B; varicella; and pneumococcal conjugate) are recommended by the Centers for Disease Control and Prevention for children aged 19-35 months, along with 3 others (hepatitis A, influenza, and rotavirus), which are recommended for full protection. Despite these guidelines, recent measles outbreaks in the United States have drawn attention to persistent gaps in coverage.

Objective: This cross-sectional study explores immunization patterns among children aged 19-35 months using provider-verified data from the National Immunization Survey-Child between 2010 and 2023.

Methods: Survey-weighted logistic regression models were used to assess associations between vaccination status and social determinants of health, including child age group, maternal education, insurance status, birth order, and region. Separate survey-weighted logistic regression models were fit for each survey year from 2010 to 2023. We considered the complex survey design to calculate vaccination coverage rates and their CIs, accounting for clustering and stratification in the sampling methodology. A composite variable was created to measure full coverage of the 7 recommended vaccines, and within each year, models were estimated separately for each individual vaccine and for the 7-vaccine series composite outcome to identify vaccine-specific barriers to immunization.

Results: Provider-verified response rates declined from 70.8% (17004/24013) in 2010 to 47.7% (18412/38619) in 2023, although the analytic sample size remained stable. Coverage for the Combined 7-Vaccine Series increased from 70.8% (11893/16798) in 2010 to 77.4% (13957/18032) in 2023. Older age was consistently associated with higher odds of complete vaccination (2010: odds ratio [OR] 1.10, 95% CI 1.02-1.19; 2023: OR 1.24, 95% CI 1.13-1.35), whereas lack of insurance and Hispanic origin were associated with lower uptake for selected vaccines. In 2022, uninsured children had 26% lower odds of complete vaccination compared with insured children (OR 0.74, 95% CI 0.65-0.84). Higher income-to-poverty ratio was consistently associated with increased vaccination uptake, including influenza vaccination in 2020 (OR 1.25, 95% CI 1.13-1.39). Regional and language-related disparities persisted across survey years.

Conclusions: Persistent socioeconomic and structural gradients in childhood vaccination coverage highlight the need for equity-focused immunization strategies that address insurance access, language barriers, and geographic disparities.

(JMIR Public Health Surveill 2026;12:e81746) doi:10.2196/81746

XSL•FO
RenderX

KEYWORDS

childhood immunization; national immunization survey; social determinants of health; survey-weighted regression; vaccination coverage

Introduction

Background

Vaccination programs have transformed child health outcomes by lowering the incidence of vaccine-preventable diseases worldwide. These efforts prevent an estimated 2-3 million deaths annually across all age groups, making childhood immunization one of the most effective public health interventions to date [1]. Despite decades of global investment in vaccination infrastructure, gaps remain in immunization coverage, particularly among marginalized communities and for specific vaccines [2]. In the United States, early childhood vaccinations protect against conditions such as measles, pertussis, and polio. Yet, recent resurgence of measles in the United States, along with suboptimal uptake of influenza and hepatitis A vaccines, underscore persistent disparities. These patterns raise concern about whether emerging infectious threats, vaccine hesitancy, and systemic inequities are eroding past gains in population protection [3]. The persistence of undervaccination calls for a deeper understanding of social and structural barriers to vaccine access and trust [4]. Understanding the factors that limit complete immunization is essential for improving equitable vaccine delivery and maintaining high coverage across the population [5].

Prior Work

Social determinants of health (SDoH) are increasingly recognized as critical drivers of health outcomes, including vaccine access and uptake [6,7]. These determinants include socioeconomic status, maternal education, employment, household size, geographic location, access to health care, and cultural or language barriers [8-10], among others. Financial constraints and limited-service availability are common barriers for low-income families, particularly those without private insurance or who reside in underserved areas. County-level data from the United States have revealed pronounced geographic disparities in immunization, often related to income and education levels [11]. In such settings, tools like small area estimation have been applied to identify pockets of low coverage and guide outreach strategies [11]. These disparities often intersect with race, ethnicity, and immigration status, further limiting vaccination access in vulnerable populations [12]. Together, these interrelated structural and socioeconomic factors shape access to pediatric services and influence the likelihood of full immunization [13,14].

In parallel, vaccine hesitancy, defined as a delay in acceptance or refusal despite availability, remains a serious threat to public health and can exacerbate inequities [15]. Misinformation, perceived safety concerns, and social influences have been identified as key factors in vaccine decision-making [2]. Hesitancy tends to cluster within populations and regions, making it harder to maintain herd immunity [16,17]. When combined with structural disadvantages, these behavioral factors further widen gaps in immunization coverage.

Recognizing and addressing these social and structural factors is central to reducing health inequities. Research incorporating SDoH into public health analysis has demonstrated that targeted interventions, including Medicaid expansion, community-based care investment, culturally tailored communication strategies, and patient navigation services, can improve vaccine uptake [18,19]. Community-driven models that consider cultural norms, transportation access, and flexible scheduling have shown promise in overcoming barriers to care and should be prioritized in public health planning [20]. Partnerships between public health officials, policymakers, researchers, and local organizations are essential to operationalize solutions that account for the full social context in which health decisions occur [21,22]. Integrating SDoH into immunization research not only informs policy decisions but also helps develop more tailored interventions that are likely to succeed across diverse populations [14,23]. Furthermore, vaccination efforts contribute directly to broader Sustainable Development Goals of the United Nations by reducing disease burden, supporting educational attainment, and enabling economic participation [6].

Although prior research has documented associations between social determinants and childhood vaccination coverage, many studies rely on pooled multiyear data, limited time windows, or focus on single vaccine outcomes. Less is known about the temporal stability of these associations across independent survey cycles, particularly before and after the COVID-19 pandemic, and about whether disparities differ systematically across specific vaccine types. Addressing these gaps is essential for understanding whether observed inequities represent persistent structural patterns or time-specific disruptions.

Study Objective

This study conducts year-specific survey-weighted analyses of National Immunization Survey-Child (NIS-Child) data from 2010 to 2023 to examine the temporal stability, pandemic-related shifts, and vaccine-specific heterogeneity of associations between social determinants and childhood immunization coverage. Focusing on children aged 19-35 months, we assess how maternal education, insurance status, income-to-poverty ratio, language, and geographic region are associated with individual vaccine uptake and completion of the Combined 7-Vaccine Series. These analyses provide updated evidence on disparities in childhood immunization coverage across prepandemic and postpandemic periods.

Methods

Data

This cross-sectional study uses data from the NIS-Child, which is publicly available through the Centers for Disease Control and Prevention (CDC) [24,25]. The NIS-Child is an annual, nationally representative survey conducted by the CDC using a random digit–dial sampling design of landline and cellular telephone numbers. Households with children aged 19-35

XSL•FO
RenderX

months are identified through telephone screening interviews. With parental consent, vaccination providers are contacted by mail to obtain official immunization records. Final survey weights adjust for sampling probability, nonresponse, and poststratification to US population benchmarks. The survey produces reliable, population-based estimates at the national, state, and selected local levels using a consistent sampling and verification process. The survey is administered by the University of Chicago under the direction of the CDC [26]. This study used all available NIS-Child data meeting inclusion criteria from 2010 to 2023. No formal sample size calculation was performed as this was a secondary analysis of existing surveillance data. The final analytic sample included all children aged 19-35 months with provider-verified immunization records across the 13-year study period.

Variables

To assess childhood immunization trends and disparities, we selected key variables grounded in demographic, socioeconomic, and household contexts. These variables align with SDoHs identified in prior literature as influential factors in vaccine uptake. In addition to child characteristics like age group, sex, and birth order, we considered household and maternal variables, such as maternal education, income-to-poverty ratio, and language of interview, to capture broader influences on immunization behavior.

For every year from 2010 to 2023, we derived new binary coverage variables indicating whether a child received the minimum required doses based on provider-verified data (adequate provider data identifier [PDAT]). These thresholds were based on the CDC's standard immunization schedule. The thresholds recommended by the CDC for children aged 19-35 months are diphtheria, tetanus, and acellular pertussis (DTaP) ≥4 doses, polio ≥3 doses, measles, mumps, and rubella (MMR) ≥1 dose, *Haemophilus influenzae* type b (Hib) ≥3 doses, hepatitis B ≥3 doses, varicella ≥1 dose, and pneumococcal conjugate (PCV) ≥4 doses. These dose thresholds were based on CDC recommendations for children aged 19-35 months and were applied consistently across all survey years (2010-2023) to ensure comparability over time.

The outcome of interest throughout our analysis was whether the child was up to date on individual vaccines. All other variables are listed in Table S1 in Multimedia Appendix 1, which were treated as covariates for understanding patterns in coverage across years. In addition to individual vaccine indicators, we created a composite variable denoted as the Combined 7-Vaccine Series, capturing whether a child was fully vaccinated with all 7 core vaccines (DTaP, polio, MMR, Hib, hepatitis B, varicella, and PCV) as defined by the CDC for national coverage monitoring. Although school entry vaccine requirements vary by state and year, this composite reflects the standardized CDC core series definition rather than uniform state mandates. Additional vaccines (hepatitis A, influenza, and rotavirus) were analyzed separately as individual outcomes. This composite variable served as a unified measure of complete immunization and played a key role in year-over-year comparisons.

Observations were restricted to children with provider-verified immunization records (using PDAT=1 as a filter), consistent with the study eligibility criteria. No additional manual row deletion was performed. Across study years, vaccine coverage variables exhibited no missing values, and most covariates demonstrated <6% missingness. Survey-weighted logistic regression models were estimated using complete-case analysis inherent to the modeling procedure, whereby observations with missing covariate values were excluded automatically at the model estimation stage.

One variable, representing the participation in the Supplemental Nutrition Program for Women, Infants, and Children (WIC) and denoted as CWIC_02, however, exhibited a markedly higher proportion of missing responses, ranging approximately from 50% to 60% across years. This pattern reflects survey flow structure and item-level nonresponse within the NIS-Child rather than data processing decisions. Because WIC participation is captured through multiple related indicators within the survey, including an alternative measure (CWIC_01) with lower missingness and comparable conceptual relevance, sensitivity analyses were conducted to evaluate whether associations were robust to alternative WIC specifications and resulting differences in analytic sample size.

Survey Design

We incorporated the core survey design variables required to ensure valid, nationally representative estimates across all years of analysis. To maintain data integrity, only records with provider-verified data were included, reflecting the study's eligibility criteria: children aged 19-35 months with completed NIS-Child interviews and verified immunization records from health care providers. Each year's dataset was structured using a consistent survey design framework that aligned with the NIS-Child's complex sampling methodology. The unique child identification was used as the primary sampling unit to account for clustering within the sample, ensuring that variance estimates reflected the survey's multistage sampling design. A stratification variable was used to optimize precision and reduce sampling error in national estimates. Sampling weights for the children were applied to each observation to adjust for unequal probabilities of selection, survey nonresponse, and to align the weighted data with US population totals.

Bias

Several measures were implemented to minimize potential bias. Provider verification of immunization records reduced recall and reporting bias compared to parental self-report. Survey weights accounted for differential selection probabilities and nonresponse bias. It is important to mention that the restriction to provider-verified data may have introduced selection bias toward children with better health care access.

Analysis

Statistical analyses were conducted using survey-weighted methods to account for the complex sampling design of NIS-Child. Complete-case analysis was used within each survey-weighted model, excluding records with missing survey weights or covariate values at the estimation stage. Out of all 15 independent variables, the income-to-poverty ratio was the

XSL•FO
RenderX

only variable analyzed as continuous, while the remaining 14 predictors were treated as categorical variables with predefined NIS-Child groupings. No interaction terms were included in the primary models. However, subgroup analyses and sensitivity analyses were conducted to evaluate the robustness of findings with the inclusion of interaction effects. Subgroup analyses examining associations across key demographic strata are presented in Multimedia Appendix 1. To examine annual trends in childhood immunization and assess how they relate to demographic and social factors, we used survey-weighted logistic regression models. Two sets of analyses were conducted. The first focused on the Combined 7-Vaccine Series, which is the binary variable created to reflect whether a child received all 7 core vaccines recommended by the CDC. The second set of models was run for each individual vaccine, allowing us to explore differences in coverage patterns across specific immunizations. We applied the same modeling process across all years from 2010 to 2023. This ensured consistent treatment of variables, uniform survey design application, and systematic evaluation of associations across time. All models used the appropriate sampling weights, clustering, and stratification variables provided by the NIS-Child to produce estimates that represent the US child population.

To evaluate the stability of associations related to WIC participation, 3 model specifications were examined. A base model excluded WIC participation entirely. A second model included CWIC_01, an indicator of WIC participation with lower missingness. The third model included CWIC_02, denoting current participation. Comparing these specifications allowed assessment of whether associations between vaccination coverage and social determinants were sensitive to WIC measurement and resulting analytic sample size differences.

To ensure methodological transparency, categorical predictors were coded using treatment contrasts in R software (version 2024.12.1+563; R Foundation for Statistical Computing), with the first-factor level serving as the reference category. Variables were harmonized across survey years to maintain consistent category definitions and reference groups. The income-to-poverty ratio was modeled as a continuous predictor to preserve information. Although formal spline-based testing of the linearity assumption was not performed, coefficient stability across years did not suggest substantial deviation from linearity on the logit scale. Formal variance inflation factors were not calculated within the survey-weighted framework; however, inspection of coefficient and SE stability indicated no evidence of meaningful multicollinearity. Statistical significance was evaluated at $\alpha=0.05$ without formal adjustment for multiple comparisons, as each year-specific model addressed a substantively distinct survey population and temporal comparison. Interpretation focused on effect sizes, CIs, and consistency of associations across survey years rather than isolated *P* values.

The survey-weighted logistic regression models estimate the likelihood of a child being up to date on recommended vaccines while accounting for the complex sampling design of the NIS-Child. The mathematical formulation of the survey-weighted logistic regression models is provided in Section S1 in Multimedia Appendix 1.

All data processing and statistical analyses were conducted using R statistical programming [27]. The *survey* package was used to implement complex survey design and fit survey-weighted logistic regression models [28]. From this, it is important to note that variance estimation was obtained from using Taylor linearization via the indicated R package. For each survey year, a year-specific survey design object was created, and associations between vaccination coverage and contextual factors (as listed in Table S1 in Multimedia Appendix 1) were examined. This modeling process was conducted separately for each vaccine outcome variable, including the derived Combined 7-Vaccine Series indicator. To visualize trends across years, we used heat maps showing the direction and significance of associations at the 5% significance level, where green indicates statistically significant effects in the positive direction, red indicates statistically significant effects in the negative direction, and gray indicates nonsignificant results. Additional gradient heat maps were generated to display the magnitude and direction of statistically significant log-odds estimates across model specifications. Separate heat maps were also created for each vaccine. Selected figures are included in the main text of this manuscript, and additional figures are present in Multimedia Appendix 1.

Ethical Considerations

The Indiana University Human Research Protection Program staff determined that the analysis done in this study was not human subject research and did not require further institutional review board (IRB) review before conducting the study.

This study used publicly available, deidentified data from the NIS-Child, conducted annually by the CDC. The NIS-Child protocol is reviewed and approved by the CDC IRB, and informed consent is obtained from participating households, with additional parental consent for provider record verification. Because this analysis relied on secondary, deidentified data that is publicly available, IRB review was not required. The dataset contains no personally identifiable information, and no participant compensation was provided as part of this secondary analysis.

Results

Figure 1 shows the trends in the number of children with provider-verified vaccination data (green) versus those with insufficient provider data (red), stratified by age groups (19-23, 24-29, and 30-35 months). A notable shift is observed during the COVID-19 pandemic (highlighted region in red), with a marked increase in missing provider data across all age groups starting in 2020, which coincides with the emergence of COVID-19 in the United States. By 2023, the number of unverified records surpassed verified ones in all age categories, especially among older children (30-35 months), suggesting disruptions in provider follow-up and data collection practices during and after the pandemic. This pattern is further corroborated by Figure 2, which shows a sharp rise in insufficient provider data following the pandemic across all age groups, particularly for children aged 30-35 months. Vaccination reporting by age group, as seen in Figure 2, shows that children aged 30-35 months consistently had higher rates of up-to-date

XSL•FO
RenderX

vaccination, while children aged 19-23 months lagged. The disparity widened after 2020, possibly reflecting pandemic-related disruptions to scheduled vaccinations among younger children.

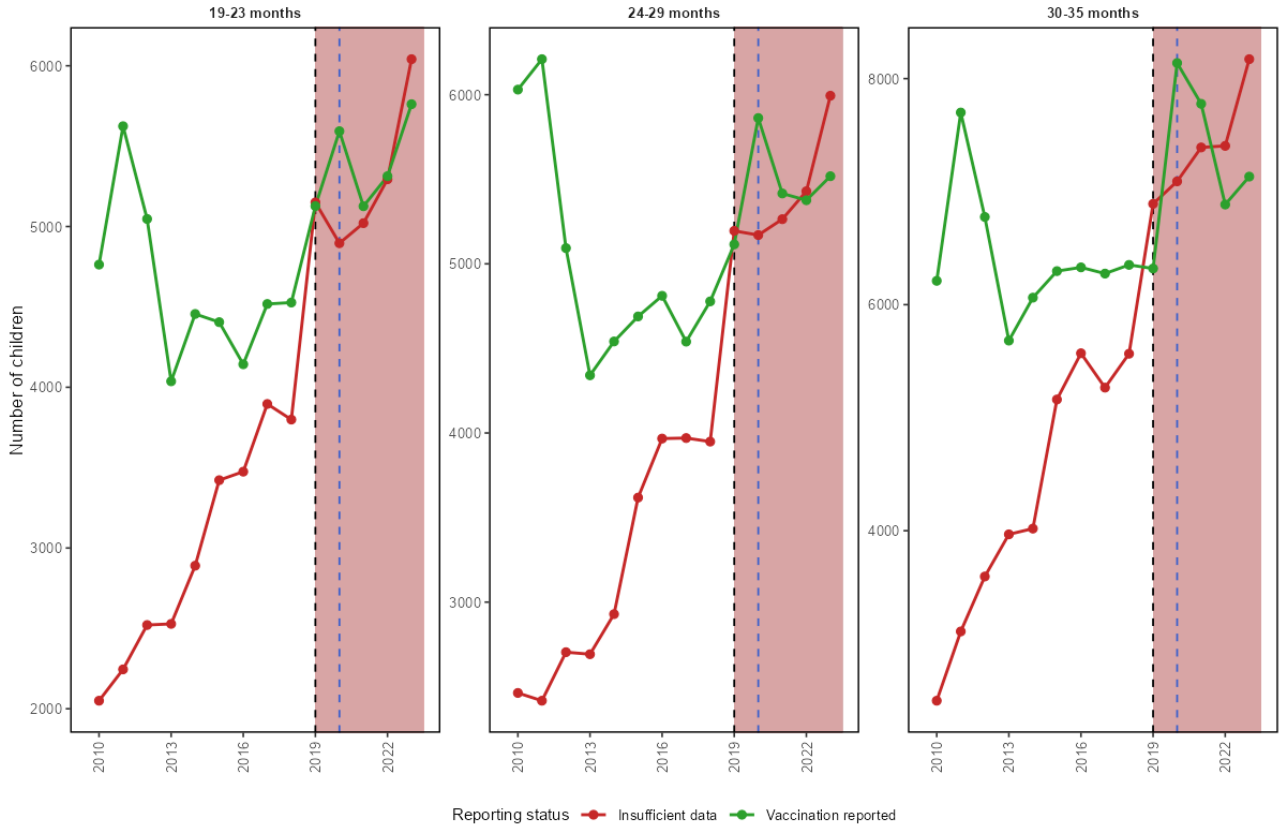
Maternal education levels, a key variable in our models, showed consistent regional differences across time. Figure 3 illustrates that the Northeast and Midwest regions had a higher proportion of college-educated mothers, with an upward trend over time. By contrast, the South had persistently higher proportions of mothers with less than 12 years of education. These regional disparities in maternal education may contribute to the observed geographic differences in vaccination outcomes. Figure 4 presents the mean number of vaccines received per child by region. The Northeast consistently achieved higher average vaccine counts compared to other regions, while the South and West trailed behind.

Figures S5 and S6 in Multimedia Appendix 1 illustrate trends in the language of interview among Hispanic respondents and the weighted proportion of Hispanic to non-Hispanic children included in the NIS-Child survey over time. Figure S5 shows a marked shift from Spanish to English interviews among Hispanic families between 2010 and 2023. Early in the time series, Spanish-language interviews comprised a considerable proportion of the total, suggesting a need for linguistically appropriate outreach. The increasing predominance of English-language responses in later years may reflect demographic transitions or changes in data collection protocols, but could also indicate reduced inclusion of less acculturated families. Figure S6 demonstrates a gradual decline in the number of Hispanic children included in the weighted sample over the years. This trend raises questions about representation and potential undercounting of vulnerable subgroups.

Figure 5 shows year-wise vaccination coverage trends for each individual vaccine. While coverage remained above 90% for core vaccines such as polio, MMR, and PCV, substantial gaps persisted for influenza and hepatitis A. Rotavirus and influenza saw visible declines in 2022 and 2023, diverging from otherwise stable trajectories. Figure 6 reinforces these trends, showing relatively tight 95% CIs for high-coverage vaccines, but wider intervals for hepatitis A and influenza.

To illustrate model-based coverage estimates, a sample of 2022 CIs is shown in Table 1. This includes estimated means and 95% CIs for each vaccine coverage. For example, polio had a mean coverage of 93.5% (95% CI 92.7%-94.3%), while influenza had 72.9% (95% CI 71.6%-74.3%). Complete year-specific coverage estimates for all vaccines are provided in Table S3 in Multimedia Appendix 1.

Figure 1. Provider-verified (green) versus insufficient provider data (red) before and after COVID-19 emergence in the United States. Insufficient provider data refers to completed interviews without adequate provider-verified vaccination records. The dashed line marks 2020; the shaded region indicates the post–COVID-19 period.



XSL•FO
RenderX

Figure 2. Trends in provider-verified data by age group during the study period (2010-2023), highlighting the impact of the COVID-19 pandemic.

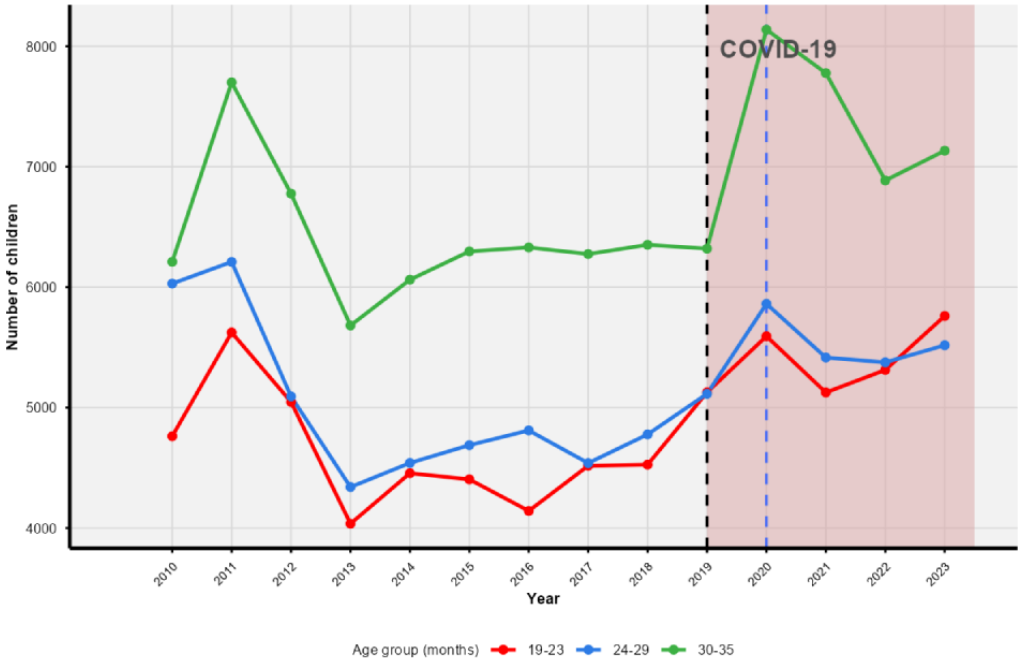
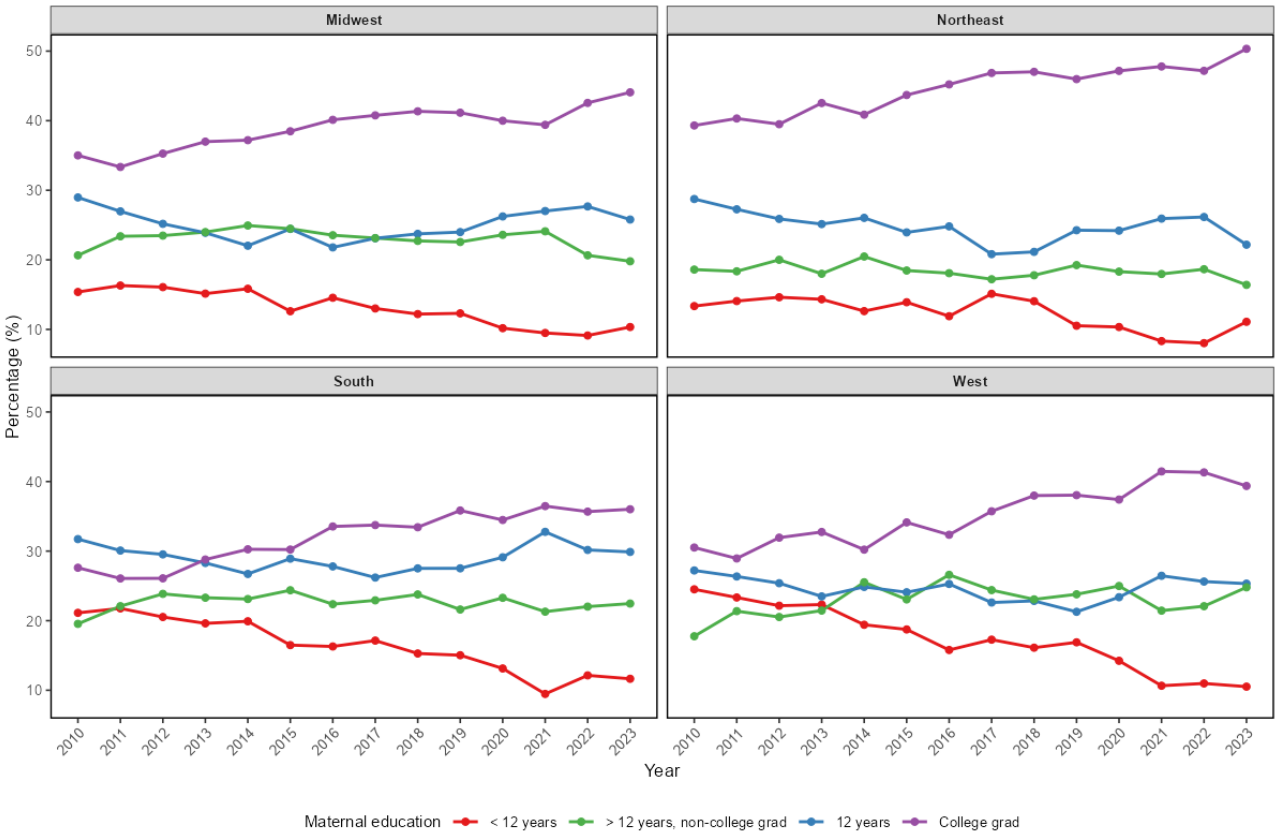


Figure 3. Maternal education trends by regions. grad: graduate.



XSL•FO
RenderX

Figure 4. Mean number of vaccines received by children by regions.

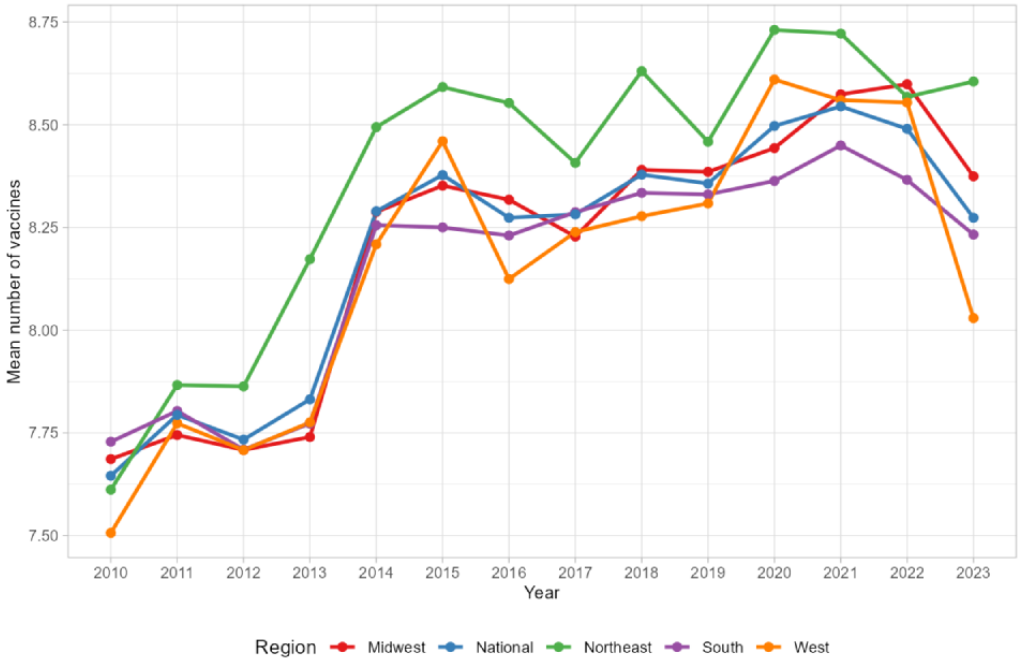
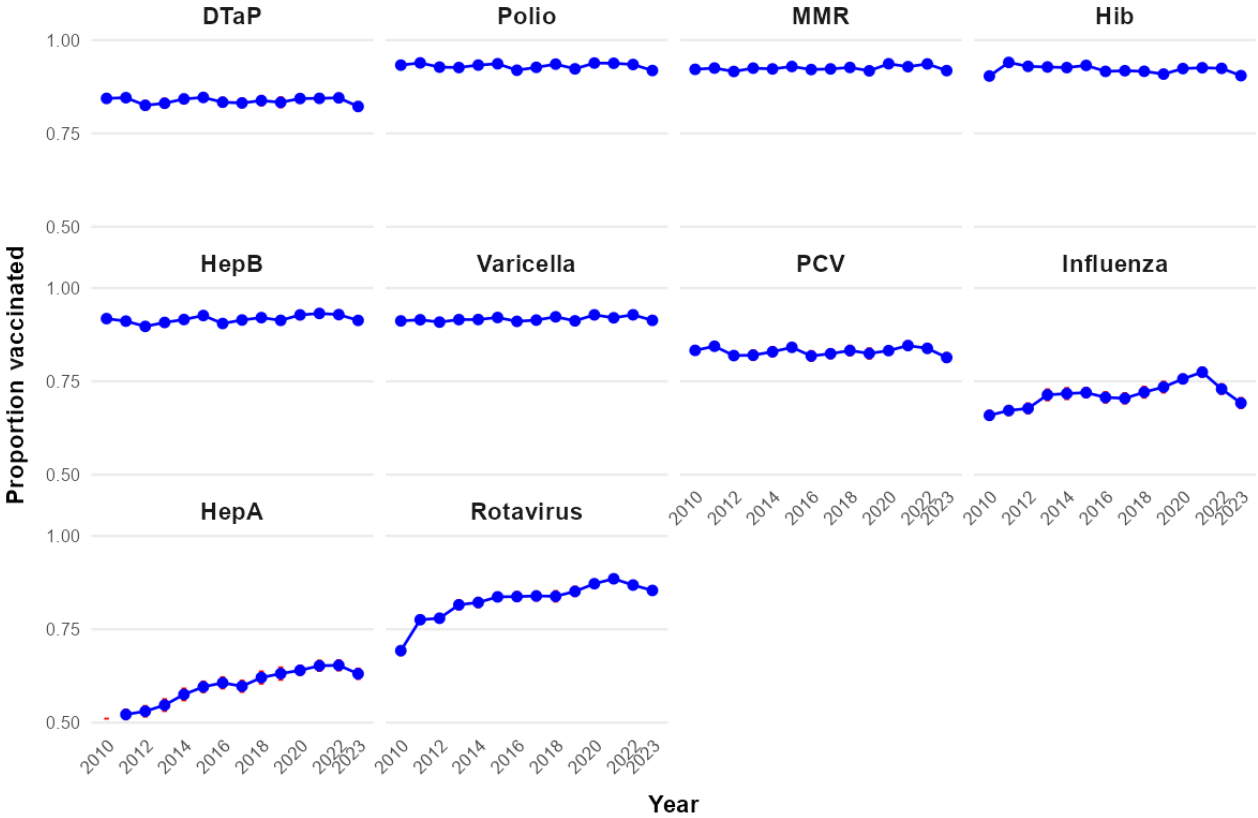


Figure 5. Trends in individual vaccine coverage. Year-wise (2010-2023) coverage rates for diphtheria, tetanus, and acellular pertussis (DTaP); polio; measles, mumps, and rubella (MMR); Haemophilus influenzae type b (Hib); hepatitis B; varicella; pneumococcal conjugate (PCV); influenza; hepatitis A; and rotavirus.



XSL•FO
RenderX

Figure 6. Vaccine coverage with 95% CIs. Combined coverage trend with bootstrapped confidence bands from 2010 to 2023. DTaP: diphtheria, tetanus, and acellular pertussis; HepA: hepatitis A; HepB: hepatitis B; Hib: Haemophilus influenzae type b; MMR: measles, mumps, and rubella; PCV: pneumococcal conjugate.

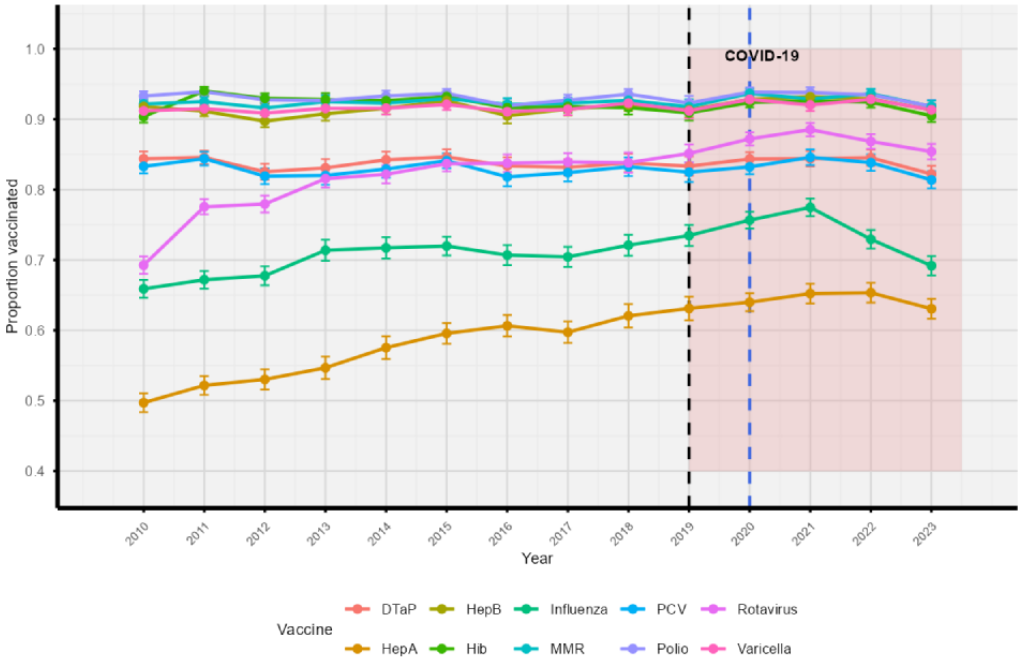


Table 1. Estimated vaccine coverage for example year 2022 with 95% CIs.

Vaccine	Mean (95% CI)
DTaP[a]	0.8453 (0.8340-0.8567)
Polio	0.9347 (0.9268-0.9426)
MMR[b]	0.9361 (0.9291-0.9432)
Hib[c]	0.9243 (0.9164-0.9322)
HepB[d]	0.9288 (0.9213-0.9364)
Varicella	0.9283 (0.9208-0.9359)
PCV[e]	0.8383 (0.8267-0.8501)
Influenza	0.7294 (0.7161-0.7427)
HepA[f]	0.6534 (0.6394-0.6675)
Rotavirus	0.8685 (0.8581-0.8788)

[a]DTaP: diphtheria, tetanus, and acellular pertussis.
[b]MMR: measles, mumps, and rubella.
[c]Hib: *Haemophilus influenzae* type b.
[d]HepB: hepatitis B.
[e]PCV: pneumococcal conjugate.
[f]HepA: hepatitis A.

The primary multivariable model included demographic, socioeconomic, household, and maternal factors as detailed in Table S1 in Multimedia Appendix 1, alongside an indicator reflecting whether the child had ever received benefits through the WIC program. This specification was selected due to lower missingness and broader representation of cumulative program exposure across survey years. Sensitivity analyses were conducted using an alternative indicator reflecting current WIC participation, which exhibited substantial structural missingness and resulted in a marked reduction in analytic sample size. A third model excluded WIC participation entirely while retaining all other social and demographic covariates to evaluate whether inclusion of WIC altered associations among core predictors. Across specifications, the direction and magnitude of non-WIC predictors remained largely consistent, supporting the stability of the primary model.

XSL•FO
RenderX

By 2023, coverage for the Combined 7-Vaccine Series reached 77.4% (14107/18032). In 2010, estimated coverage was 70.8% (SE 0.006), increasing to 71.5% (SE 0.007) by 2012 and continuing to rise thereafter.

In the primary model, older child age was consistently associated with higher odds of being fully vaccinated across the study period (2010: odds ratio [OR] 1.10, 95% CI 1.02-1.19, SE 0.040; *P*=.02; 2023: OR 1.24, 95% CI 1.13-1.36, SE 0.047; *P*<.001). Larger household size was associated with lower odds of vaccination completion in early years (2010: OR 0.89, 95% CI 0.85-0.94, SE 0.027; *P*<.001) and remained a persistent barrier across multiple survey cycles. The language of the interview demonstrated positive associations in selected years. For example, in 2010, children from households interviewed in English had higher odds of complete vaccination (OR 1.50, 95% CI 1.18-1.90, SE 0.123; *P*<.001), and this association reemerged in 2018 (OR 1.50, 95% CI 1.15-1.95, SE 0.407; *P*=.003).

Insurance coverage became increasingly influential in later years. In 2022, insured children had higher odds of complete vaccination compared with uninsured children (OR 0.74, 95% CI 0.65-0.84, SE 0.063; *P*<.001). Maternal education similarly demonstrated positive associations in recent years, indicating a widening gradient by educational attainment. Regional disparities intensified after 2020, with certain geographic regions exhibiting lower odds of complete vaccination (OR 0.74, 95% CI 0.65-0.84, SE 0.063; *P*<.001). Hispanic origin showed intermittent negative associations across survey years but was not uniformly significant throughout the study period. Overall, age-related advantages and household size barriers were the most stable predictors across time.

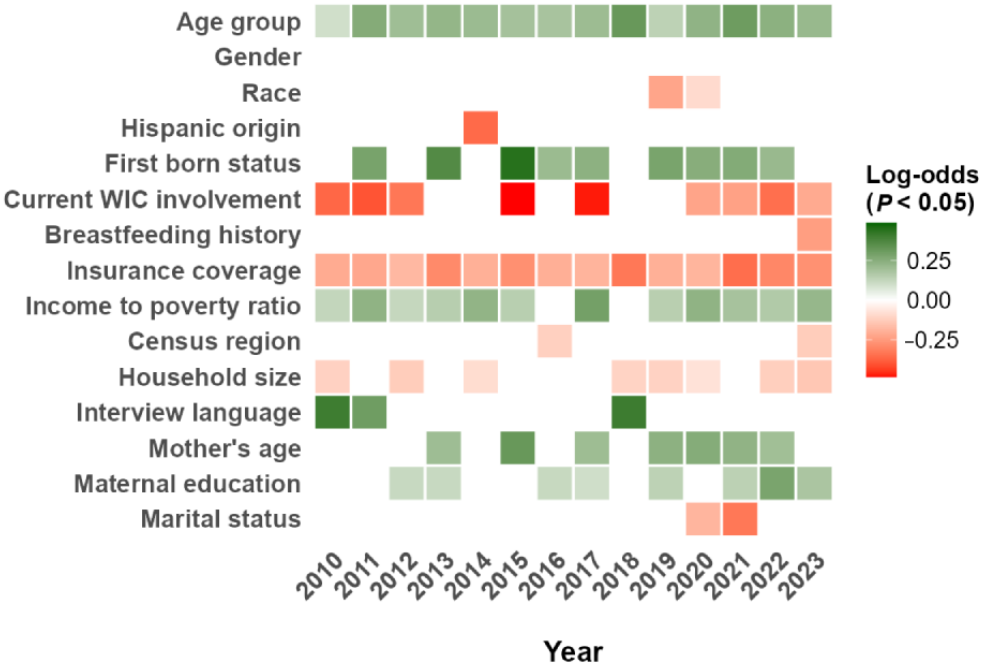
Influenza coverage improved from 65.9% (11470/16798) in 2010 to 77.5% (13408/17232) in 2022 before plateauing. In 2020, marital status (OR 0.75, 95% CI 0.63-0.90, SE 0.087; *P*=.001) and income-to-poverty ratio (OR 1.25, 95% CI 1.13-1.39, SE 0.054; *P*<.001) were significantly associated with influenza vaccination. Insurance coverage and regional disparities also became more prominent after 2020. In 2010, hepatitis A coverage was 49.7% (SE 0.007). Older age (2010: OR 1.94, 95% CI 1.80-2.09, SE 0.038; *P*<.001), smaller household size (2010: OR 0.90, 95% CI 0.85-0.94, SE 0.027; *P*<.001), Hispanic origin (2010: OR 0.75, 95% CI 0.63-0.89, SE 0.086; *P*=.001), language of interview (2010: OR 1.39, 95% CI 1.13-1.73, SE 0.108; *P*=.002), insurance coverage (2010: OR 0.81, 95% CI 0.80-0.98, SE 0.049; *P*=.001) and WIC benefits (2010: OR 0.69, 95% CI 0.56-0.87, SE 0.107; *P*<.001) were significantly associated with uptake. These disparities persisted in 2022 and 2023, indicating continued structural and access-related inequities. PCV and rotavirus vaccines showed persistent negative associations with increasing household size (2010; rotavirus: OR 0.74, 95% CI 0.65-0.84, SE 0.063; *P*<.001; PCV: OR 0.88, 95% CI 0.82-0.93 SE 0.030; *P*<.001), which are highly significant, suggesting logistical challenges among larger families. Regional disparities and Hispanic origin effects were also evident for these vaccines, particularly after 2020.

For MMR vaccination, the strength of age-related associations declined over time. In 2018, age group was strongly associated with higher odds of vaccination (OR 1.37, 95% CI 1.20-1.57, SE 0.069; *P*<.001), and this association remained significant in 2022 (OR 1.34, 95% CI 1.14-1.57, SE 0.081; *P*<.001). However, by 2023, the association weakened and became nonsignificant (OR 1.12, 95% CI 0.97-1.31, SE 0.076; *P*=.12). Participation in WIC demonstrated notable temporal shifts. In 2013, WIC participation was associated with significantly lower odds of MMR vaccination (OR 0.63, 95% CI 0.44-0.90, SE 0.180; *P*=.01). This inverse association attenuated in subsequent years and was not statistically significant in 2016 (OR 0.64, 95% CI 0.40-1.03, SE 0.241; *P*=.06). During the COVID-19 period, the direction of association shifted toward higher vaccination likelihood among WIC participants, although not reaching statistical significance in 2022 (OR 0.72, 95% CI 0.49-1.05, SE 0.194; *P*=.09) and 2023 (OR 0.83, 95% CI 0.62-1.11, SE 0.151; *P*=.21). These temporal fluctuations suggest that socioeconomic program engagement and age-based eligibility effects on MMR vaccination evolved over the study period, particularly during the pandemic years. Additional heat maps for individual vaccines are presented in Figures S6-S15 in Multimedia Appendix 1, demonstrating statistical significance patterns of SDoH across each vaccine type. These disparities remained evident across the study period when looking at the Combined 7-Vaccine Series, reinforcing the persistence of structural and access-related barriers, as seen in Figure 7. Complete table of effect sizes, CIs, and *P* values is provided in Multimedia Appendix 1.

Sensitivity analyses using the current-participation WIC measure yielded comparable directional effects but wider CIs due to reduced sample size. The base model excluding WIC showed no material changes in non-WIC predictor associations, indicating that WIC inclusion did not distort relationships among core social determinants. Detailed model comparisons and corresponding association patterns are provided in Multimedia Appendix 1, where year-specific estimates across WIC specifications are presented.

XSL•FO
RenderX

Figure 7. Heat map of associations between predictors and the uptake of the Combined 7-Vaccine Series. WIC: Supplemental Nutrition Program for Women, Infants, and Children.



Discussion

Principal Findings

This cross-sectional analysis of provider-verified NIS-Child data from 2010 to 2023 documents persistent associations between childhood vaccination coverage and multiple SDoH across the United States. Although coverage for several core vaccines remained high, consistent socioeconomic, regional, and language-related gradients were observed across survey years. These findings describe associations rather than causal effects and highlight structural patterns that may contribute to inequities in immunization uptake.

Maternal education was consistently associated with higher odds of being up to date on vaccines. This pattern aligns with global evidence identifying maternal education as an important determinant of child health outcomes, including immunization coverage [29]. Higher educational attainment may reflect differences in health literacy, access to reliable information, and the ability to navigate health care systems. Caregivers with more education may be better positioned to understand vaccine schedules, communicate with providers, and respond to reminders, which may facilitate timely immunization.

Household income-to-poverty ratio was similarly associated with vaccination coverage, consistent with prior research documenting economic disparities in immunization [30]. Financial constraints may limit transportation access, appointment flexibility, and continuity of care. Even when vaccines are available at low or no cost, structural barriers such as unstable employment, limited clinic availability, and neighborhood-level resource gaps may shape preventive service use. Insurance status demonstrated strong associations, particularly in later survey years. According to a nationwide analysis, children from higher-income families were more likely to complete the full vaccine schedule on time compared to those in lower-income households. Insurance status was another factor with a strong association [31]. Uninsured children were less likely to be fully vaccinated compared to those with public or private coverage. CDC reports have similarly noted lower immunization rates among uninsured children despite the presence of the Vaccines for Children program [32]. Insurance coverage may therefore reflect broader system engagement and continuity of pediatric care.

Ethnicity-related differences were observed for several vaccines, including hepatitis A, influenza, and rotavirus. Hispanic origin was intermittently associated with lower vaccination odds, though not uniformly across all years. These findings are consistent with the literature attributing disparities in preventive care to structural barriers, including language discordance, immigration-related concerns, and differential provider engagement. The Hispanic paradox, which refers to the phenomenon in which Hispanic individuals experience health outcomes that are unexpectedly favorable relative to their socioeconomic status, has been well documented in outcomes such as mortality and chronic disease. However, these results indicate that this paradox does not extend to preventive care measures like childhood immunization, where disparities in access and use remain prominent [33,34]. Preventive services such as vaccination require active health care system engagement and documentation, which may be more sensitive to structural access barriers than mortality outcomes.

Language of interview was associated with vaccination coverage in selected years, with non–English-speaking households less likely to achieve complete immunization. This supports evidence that language discordance between caregivers and providers may contribute to miscommunication, reduced trust, and missed vaccination opportunities [35]. However, the language of the interview likely reflects broader structural barriers beyond

XSL•FO
RenderX

communication alone and should be interpreted as a proxy marker rather than a purely linguistic effect. Improvements in culturally and linguistically appropriate services may therefore play an important role in reducing disparities.

Regional variation persisted throughout the study period and appeared more pronounced after 2020. Differences across census regions likely reflect variation in health care infrastructure, provider density, rural-urban distribution, and state-level immunization policies. Lower vaccination coverage in certain regions may be related to previously documented geographic disparities in pediatric service access, particularly in rural or underserved areas where limited provider availability and long travel distances create barriers to timely immunization [36]. Regional heterogeneity in vaccination coverage may therefore represent structural differences in health care delivery rather than individual-level preferences alone. While regional analyses provide stable comparisons across years, they may mask meaningful state-level variation; future research should examine state-specific differences to better inform policy decisions.

Household size was consistently associated with lower vaccination completion for several vaccines, including PCV and rotavirus. Larger households may face logistical challenges related to scheduling, transportation, and competing caregiving demands. Age-related differences in coverage likely reflect timing of eligibility and health care encounters rather than behavioral differences.

The COVID-19 period coincided with increased insufficient provider data and widening disparities for selected vaccines, particularly influenza and hepatitis A. Disruptions in routine care and shifts in health care use during this period may have influenced vaccination patterns. Recent CDC reports indicate declines in kindergarten vaccination coverage and rising exemption rates, reinforcing concerns about sustained population-level protection. MMR coverage dropped from 93.1% (17931/19342) to 92.7% (17931/18032), a decrease of 0.4 percentage points), while DTaP coverage fell from 92.7% (14909/17232) to 92.3% (15903/18032), also a 0.4 percentage point decline) between the 2022-2023 and 2023-2024 school years. Simultaneously, exemption rates reached a record high of 3.3%, up from 3.0% the previous year [37]. The findings of this study reinforce the need to address coverage disparities through targeted public health strategies. Although the analytic sample size remained stable and standard survey weighting procedures were applied, estimates from 2020-2023 should be interpreted with caution, given potential shifts in provider participation and survey response dynamics during the pandemic period.

As a cross-sectional analysis, this study documents associations between social determinants and vaccination coverage but cannot establish causal relationships. Longitudinal or quasi-experimental studies are needed to determine whether interventions targeting identified social and structural barriers would improve immunization uptake. Given the number of year-specific models estimated, findings should be interpreted in the context of temporal consistency rather than reliance on isolated statistical significance. Overall, the persistence of disparities across education, income, insurance, language, and region suggests that structural inequities remain central to understanding gaps in childhood vaccination coverage.

Limitations

This study has several limitations. First, the cross-sectional design limits inference to associations and does not establish causality. Second, analyses were restricted to children with provider-verified vaccination records, consistent with NIS-Child guidance; however, if children without verified records differ systematically in health care access or sociodemographic characteristics, disparities may be underestimated. Survey-weighted regression models were estimated using complete-case analysis. Although most covariates had low missingness, bias may occur if the data were not missing at random. WIC participation required additional consideration. While the primary WIC indicator used in the main models had relatively low missingness, an alternative WIC variable exhibited substantial structural missingness due to survey design. Sensitivity analyses indicated that this missingness influenced certain associations, particularly involving insurance status, suggesting that the missingness mechanism was not random. However, associations for other core social determinants remained stable across alternative model specifications, including a model excluding WIC, supporting the robustness of the primary findings.

Changes in provider verification patterns and health care use during 2020 to 2023 may affect the interpretation of year-specific estimates, even though provider-phase weights and a consistent analytic framework were applied across all years. Several relevant factors, including immigration status, health literacy, vaccine confidence, and access barriers, were not captured in the public dataset, raising the possibility of unmeasured confounding. Regional variables were measured at an aggregate level and may be subject to ecological bias. Statistical significance was assessed at 5% without formal adjustment for multiple comparisons, and findings should be interpreted with emphasis on effect sizes and consistency over time. Finally, generalizability is limited to US children aged 0-35 months with provider-verified records and may not extend to other populations or health systems.

Future Work

Future research should build on these findings by using longitudinal and quasi-experimental approaches to better understand how structural and policy changes influence vaccination disparities over time. Qualitative and mixed method studies may help clarify mechanisms not captured in NIS-Child, including health literacy, cultural mistrust, transportation barriers, and health care navigation challenges that may contribute to undervaccination among Hispanic, uninsured, and low-income families [38]. State-level and geospatial analyses are also needed to examine within-region heterogeneity and link disparities to policy environments such as Medicaid eligibility, exemption laws, and public health funding differences across states [39]. Incorporating interaction analyses in future models as part of a formal causal inference model, such as ethnicity by language or insurance by income, may further identify subgroups experiencing compounded barriers and inform more targeted interventions.

XSL•FO
RenderX

Conclusion

This study used nationally representative, provider-verified NIS-Child data from 2010 to 2023 to examine associations between SDoH and early childhood vaccination coverage of CDC recommended vaccines, including completion of a Combined 7-Vaccine Series. Across survey years, vaccination coverage was consistently patterned by maternal education, insurance status, household income, language of interview, and region, with particularly persistent gaps for hepatitis A, influenza, and rotavirus. Although provider verification decreased after 2020, the analytic sample of children with adequate provider data remained stable, and results were generally robust across sensitivity specifications. These findings support the need for equity-focused immunization strategies that address structural barriers to accessing preventive care, including insurance coverage, culturally and linguistically appropriate services, and geographically tailored outreach. Continued surveillance and policy-relevant analyses are needed to monitor whether disparities narrow over time and to guide interventions that improve protection against vaccine-preventable diseases for all children.

Funding

The authors declared no financial support was received for this work.

Data Availability

The data used for this study were sourced from the National Immunization Surveys database, available on the Centers for Disease Control and Prevention (CDC) website. Access to the dataset can be obtained from the CDC's official data repository, ensuring transparency and reproducibility in research findings. The study involved secondary analysis of publicly available, deidentified surveillance data. No new datasets were generated. Statistical code used for data processing and analysis is available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: RKM, YZ, FMPR
Data curation: RKM, FMPR
Formal analysis: RKM
Methodology: RKM, YZ, FMPR
Software: RKM
Visualization: RKM
Supervision: YZ, FMPR
Validation: YZ, FMPR
Project administration: YZ, FMPR
Writing – original draft: RKM
Writing – review & editing: RKM, YZ, FMPR

Conflicts of Interest

Author YZ is the Associate Editor of the *Journal of Medical Internet Research*. Author FP-R is a member of the Editorial Board and an Associate Editor of *JMIR Public Health and Surveillance*. The other author has no conflicts to declare.

Multimedia Appendix 1
Additional tables, figures, and information.
[DOCX File , 1827 KB - publichealth_v12i1e81746_app1.docx]

References

1. Andre FE, Booy R, Bock HL, Clemens J, Datta SK, John TJ, et al. Vaccination greatly reduces disease, disability, death and inequity worldwide. Bull World Health Organ 2008;86(2):140-146 [FREE Full text] [doi: 10.2471/blt.07.040089] [Medline: 18297169]
2. Stokley S, Kempe A, Stockwell MS, Szilagyi PG. Improving pediatric vaccination coverage in the United States. Acad Pediatr 2021;21(4S):S1-S2. [doi: 10.1016/j.acap.2021.03.004] [Medline: 33958085]
3. Dubé E, Vivion M, MacDonald NE. Vaccine hesitancy, vaccine refusal and the anti-vaccine movement: influence, impact and implications. Expert Rev Vaccines 2015;14(1):99-117. [doi: 10.1586/14760584.2015.964212] [Medline: 25373435]
4. Wheelock A, Parand A, Rigole B, Thomson A, Miraldo M, Vincent C, et al. Socio-psychological factors driving adult vaccination: a qualitative study. PLoS One 2014;9(12):e113503 [FREE Full text] [doi: 10.1371/journal.pone.0113503] [Medline: 25490542]
5. Opel DJ, Taylor JA, Mangione-Smith R, Solomon C, Zhao C, Catz S, et al. Validity and reliability of a survey to identify vaccine-hesitant parents. Vaccine 2011;29(38):6598-6605. [doi: 10.1016/j.vaccine.2011.06.115] [Medline: 21763384]

XSL•FO
RenderX

6. Closing the gap in a generation: health equity through action on the social determinants of health. WHO Commission on Social Determinants of Health. 2008. URL: https://iris.who.int/bitstream/handle/10665/43943/9789241563703_eng.pdf [accessed 2025-07-31]
7. Solar O, Irwin A. A conceptual framework for action on the social determinants of health?: social determinants of health discussion paper 2. World Health Organization. 2010. URL: https://www.afro.who.int/sites/default/files/2017-06/SDH_conceptual_framework_for_action.pdf [accessed 2025-07-31]
8. Luman ET, Barker LE, Shaw KM, McCauley MM, Buehler JW, Pickering LK. Timeliness of childhood vaccinations in the United States: days undervaccinated and number of vaccines delayed. JAMA 2005;293(10):1204-1211. [doi: 10.1001/jama.293.10.1204] [Medline: 15755943]
9. Freeman RE, Thaker J, Daley MF, Glanz JM, Newcomer SR. Vaccine timeliness and prevalence of undervaccination patterns in children ages 0-19 months, U.S., National Immunization Survey-Child 2017. Vaccine 2022;40(5):765-773 [FREE Full text] [doi: 10.1016/j.vaccine.2021.12.037] [Medline: 34961632]
10. Larson HJ, Jarrett C, Eckersberger E, Smith DMD, Paterson P. Understanding vaccine hesitancy around vaccines and vaccination from a global perspective: a systematic review of published literature, 2007-2012. Vaccine 2014;32(19):2150-2159. [doi: 10.1016/j.vaccine.2014.01.081] [Medline: 24598724]
11. Hill HA, Chen M, Elam-Evans LD, Yankey D, Singleton JA. Vaccination coverage by age 24 months among children born during 2018-2019 - National Immunization Survey-Child, United States, 2019-2021. MMWR Morb Mortal Wkly Rep 2023;72(2):33-38 [FREE Full text] [doi: 10.15585/mmwr.mm7202a3] [Medline: 36634013]
12. McDonald HI, Tessier E, White JM, Woodruff M, Knowles C, Bates C, et al. Early impact of the coronavirus disease (COVID-19) pandemic and physical distancing measures on routine childhood vaccinations in England, January to April 2020. Euro Surveill 2020;25(19):2000848 [FREE Full text] [doi: 10.2807/1560-7917.ES.2020.25.19.2000848] [Medline: 32431288]
13. Braveman P, Egerter S, Williams DR. The social determinants of health: coming of age. Annu Rev Public Health 2011;32:381-398. [doi: 10.1146/annurev-publhealth-031210-101218] [Medline: 21091195]
14. Phelan JC, Link BG, Tehranifar P. Social conditions as fundamental causes of health inequalities: theory, evidence, and policy implications. J Health Soc Behav 2010;51 Suppl:S28-S40. [doi: 10.1177/0022146510383498] [Medline: 20943581]
15. Shen SC, Dubey V. Addressing vaccine hesitancy: clinical guidance for primary care physicians working with parents. Can Fam Physician 2019;65(3):175-181 [FREE Full text] [Medline: 30867173]
16. Phadke VK, Bednarczyk RA, Salmon DA, Omer SB. Association between vaccine refusal and vaccine-preventable diseases in the United States: a review of measles and pertussis. JAMA 2016;315(11):1149-1158 [FREE Full text] [doi: 10.1001/jama.2016.1353] [Medline: 26978210]
17. Olusanya OA, Bednarczyk RA, Davis RL, Shaban-Nejad A. Addressing parental vaccine hesitancy and other barriers to childhood/adolescent vaccination uptake during the coronavirus (COVID-19) pandemic. Front Immunol 2021;12:663074 [FREE Full text] [doi: 10.3389/fimmu.2021.663074] [Medline: 33815424]
18. Graham H. Social determinants and their unequal distribution: clarifying policy understandings. Milbank Q 2004;82(1):101-124 [FREE Full text] [doi: 10.1111/j.0887-378x.2004.00303.x] [Medline: 15016245]
19. Adler NE, Stewart J. Health disparities across the lifespan: meaning, methods, and mechanisms. Ann N Y Acad Sci 2010;1186:5-23. [doi: 10.1111/j.1749-6632.2009.05337.x] [Medline: 20201865]
20. Evans D, Norrbom C, Schmidt S, Powell R, McReynolds J, Sidibe T. Engaging community-based organizations to address barriers in public health programs: lessons learned from COVID-19 vaccine acceptance programs in diverse rural communities. Health Secur 2023;21(S1):S17-S24 [FREE Full text] [doi: 10.1089/hs.2023.0017] [Medline: 37610883]
21. Minkler M, Blackwell AG, Thompson M, Tamir H. Community-based participatory research: implications for public health funding. Am J Public Health 2003;93(8):1210-1213. [doi: 10.2105/ajph.93.8.1210] [Medline: 12893597]
22. Diez Roux AV. Complex systems thinking and current impasses in health disparities research. Am J Public Health 2011;101(9):1627-1634. [doi: 10.2105/ajph.test.2011.300149]
23. Berkman LF, Kawachi I. Social Epidemiology. New York, NY: Oxford University Press; 2000.
24. National Immunization Survey (NIS): data files. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nis/php/datasets-child/index.html [accessed 2025-07-31]
25. NIS-Child datasets. Centers for Disease Control and Prevention. 2024. URL: https://www.cdc.gov/nis/php/datasets-child/index.html [accessed 2025-07-31]
26. National Immunization Survey (NIS). Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nis/about/index.html [accessed 2025-07-31]
27. RStudio: integrated development environment for R posit software, PBC. Posit team. URL: http://www.posit.co/ [accessed 2025-07-31]
28. Lumley T, Gao P, Schneider. Analysis of complex survey samples. GeeksforGeeks. The Comprehensive R Archive Network (CRAN). URL: https://cran.r-project.org/package=survey [accessed 2025-07-31]
29. Gebreyesus A, Tesfay K. Effect of maternal education on completing childhood vaccination in Ethiopia: systematic review and meta-analysis. Sci Rep 2024;14(1):17453 [FREE Full text] [doi: 10.1038/s41598-024-68182-4] [Medline: 39075119]

XSL•FO
RenderX

30. Nandi A, Shet A. Why vaccines matter: understanding the broader health, economic, and child development benefits of routine vaccination. Hum Vaccin Immunother 2020;16(8):1900-1904 [FREE Full text] [doi: 10.1080/21645515.2019.1708669] [Medline: 31977283]
31. Newcomer SR, Michels SY, Albers AN, Freeman RE, Graham JM, Clarke CL, et al. Vaccination timeliness among US children aged 0-19 months, National Immunization Survey-Child 2011-2021. JAMA Netw Open 2024;7(4):e246440 [FREE Full text] [doi: 10.1001/jamanetworkopen.2024.6440] [Medline: 38607623]
32. Valier MR, Yankey D, Elam-Evans LD. Morbidity and mortality weekly report vital signs: trends and disparities in childhood vaccination coverage by vaccines for children program eligibility-National Immunization Survey-Child, United States, 2012-2022. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/vaccines-for-children/vfc-information-for- [accessed 2026-03-20]
33. Herrera GA, Zhao Z, Klevens RM. Variation in vaccination coverage among children of hispanic ancestry. Am J Prev Med 2001;20(4 Suppl):69-74. [doi: 10.1016/s0749-3797(01)00277-x] [Medline: 11331135]
34. Borrell LN, Markides KS. Will the health status of the changing hispanic population remain 'Paradoxical'? Am J Public Health 2024;114(S6):S431-S435. [doi: 10.2105/AJPH.2024.307697] [Medline: 39083746]
35. Al Shamsi H, Almutairi AG, Al Mashrafi S, Al Kalbani T. Implications of language barriers for healthcare: a systematic review. Oman Med J 2020;35(2):e122 [FREE Full text] [doi: 10.5001/omj.2020.40] [Medline: 32411417]
36. Freeman RE, Leary CS, Graham JM, Albers AN, Wehner BK, Daley MF, et al. Geographic proximity to immunization providers and vaccine series completion among children ages 0-24 months. Vaccine 2023;41(17):2773-2780 [FREE Full text] [doi: 10.1016/j.vaccine.2023.03.025] [Medline: 36964002]
37. SchoolVaxView Data. Centers for Disease Control and Prevention. 2024. URL: https://www.cdc.gov/schoolvaxview/data/index.html [accessed 2025-07-31]
38. Gowda C, Dempsey AF. The rise (and fall?) of parental vaccine hesitancy. Hum Vaccin Immunother 2013;9(8):1755-1762 [FREE Full text] [doi: 10.4161/hv.25085] [Medline: 23744504]
39. Omer SB, Pan WKY, Halsey NA, Stokley S, Moulton LH, Navar AM, et al. Nonmedical exemptions to school immunization requirements: secular trends and association of state policies with pertussis incidence. JAMA 2006;296(14):1757-1763. [doi: 10.1001/jama.296.14.1757] [Medline: 17032989]

Abbreviations

CDC: Centers for Disease Control and Prevention
CWIC: Supplemental Nutrition Program for Women, Infants, and Children measure
DTaP: diphtheria, tetanus, and acellular pertussis
Hib: Haemophilus influenzae type b
IRB: institutional review board
MMR: measles, mumps, and rubella
NIS-Child: National Immunization Survey-Child
OR: odds ratio
PCV: pneumococcal conjugate
PDAT: adequate provider data identifier
SDoH: social determinants of health
WIC: Supplemental Nutrition Program for Women, Infants, and Children

Edited by A Mavragani, T Sanchez; submitted 02.Aug.2025; peer-reviewed by MS Hossain, SU Khasbage; comments to author 26.Dec.2025; revised version received 25.Feb.2026; accepted 17.Mar.2026; published 09.Apr.2026.

Please cite as:
Majji RK, Zhuang Y, Pabon-Rodriguez FM
Social Determinants of Childhood Vaccination Coverage in the United States Using National Immunization Survey Data From 2010 to 2023: Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e81746
URL: https://publichealth.jmir.org/2026/1/e81746
doi:10.2196/81746
PMID:

©Ramya Keerthi Majji, Yan Zhuang, Felix M Pabon-Rodriguez. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 09.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction

XSL•FO
RenderX

in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Prevalence and Determinants of Blood Exposure Accidents Among Health Care Workers in Kalemie, Democratic Republic of the Congo: Cross-Sectional Study

Joe Kabamba Dibwe[1,2], MD, MPH; Fiston Ilunga Mbayo[3,4], MD, MPH; Papy Mukalay Muvumbu[5], MD, MPH; Gloire Kasongo Lupitshi[3], MPH, MD; Jean Claude Banze Kabwe Ilunga[6], MD, MPH; Didier Sashila Mukebo[7], MD, MPH; Dieudonné Mukendi Mpunga[8], MD, MPH, PhD

[1]Higher Institute of Medical Techniques of Kongolo, Kongolo, Tanganyika, the Democratic Republic of the Congo
[2]Provincial Health Division of Tanganyika, Kalemie, Tanganyika, the Democratic Republic of the Congo
[3]School of Public Health, University of Malemba Nkulu, Malemba Nkulu, Haut Lomami, the Democratic Republic of the Congo
[4]Public Health, Malemba Nkulu General Reference Hospital, Malemba Nkulu, Haut Lomami, the Democratic Republic of the Congo
[5]Higher Institute of Medical Techniques of Manono, Manono, Tanganyika, the Democratic Republic of the Congo
[6]University of Kamina, Kamina, Haut-Lomami, the Democratic Republic of the Congo
[7]National Malaria Control Program Democratic Republic of Congo, Kinshasa, Kinshasa, the Democratic Republic of the Congo
[8]School of Public Health, Medicine, University of Kinshasa, Kinshasa, Kinshasa, the Democratic Republic of the Congo

Corresponding Author:
Fiston Ilunga Mbayo, MD, MPH
School of Public Health
University of Malemba Nkulu
P.O BOX 365
Malemba Nkulu, Haut Lomami
the Democratic Republic of the Congo
Phone: 243 817689233
Email: ilungambayofiston@gmail.com

Abstract

Background: Blood exposure accidents (BEAs) remain a major occupational hazard for health care providers, particularly in low-resource settings. Despite their preventable nature, BEAs continue to occur frequently due to gaps in compliance with standard precautions and limited institutional surveillance.

Objective: This study aimed to determine the prevalence and factors associated with BEAs among health care providers in Kalemie, Tanganyika Province, Democratic Republic of the Congo.

Methods: A cross-sectional study was conducted in March 2024 among 316 health care providers randomly selected from 34 health facilities: public facilities (20/34, 58.8%), private facilities (9/34, 26.5%), and faith-based facilities (5/34, 14.7%). Data were collected using a structured questionnaire adapted from World Health Organization and Centers for Disease Control and Prevention tools. Logistic regression was performed to identify predictors of BEAs, with adjusted odds ratios (AORs), 95% CIs, and *P* values reported.

Results: The lifetime prevalence of BEAs was 65.8% (208/316), and the 12-month prevalence was 60.4% (191/316). The most frequent causes were unexpected patient movement (81/191, 42.4%) and lack of attention (66/191, 34.6%). Needlestick injuries accounted for 59.2% (113/191) of exposures. Only 36.1% (69/191) of incidents were formally reported. Multivariate analysis identified two independent predictors: (1) noncompliance with standard precautions (AOR 2.921, 95% CI 1.503-5.677; *P*=.001) and (2) facility type (working in secondary-level facilities was protective; AOR 0.310, 95% CI 0.144-0.670; *P*=.003).

Conclusions: BEAs were highly prevalent among health care providers in Kalemie. Most incidents were preventable, and noncompliance with standard precautions significantly increased risk. Strengthening infection prevention and control measures, ensuring consistent availability of personal protective equipment, and establishing effective surveillance systems are urgently needed to protect health care workers and reinforce health system resilience.

(JMIR Public Health Surveill 2026;12:e97009) doi:10.2196/97009

XSL•FO
RenderX

KEYWORDS

blood exposure accidents; occupational hazards; health care workers; infection prevention and control; standard precautions; personal protective equipment; Democratic Republic of the Congo; sub-Saharan Africa

Introduction

Occupational exposure to blood and infectious body fluids remains a major public health concern, particularly for health care workers in low-resource settings. Blood exposure accidents (BEAs) are defined as any accidental percutaneous or mucocutaneous contact with blood or biological fluids potentially contaminated with pathogens such as bacteria, viruses, parasites, or fungi [1,2]. These exposures typically occur through needlestick injuries, cuts, or splashes during clinical procedures. Beyond blood, other biological fluids—including cerebrospinal, synovial, amniotic, and genital secretions—may also transmit infections even when not visibly contaminated [3]. According to the World Health Organization (WHO), approximately 3 million of the 35 million health care workers worldwide are exposed to bloodborne pathogens annually [4,5]. Recent global evidence confirms the persistence of this risk: a meta-analysis published in 2022 reported a lifetime prevalence of occupational blood exposure of 56.6% and a 12-month prevalence of 39% among health care workers worldwide [6]. More recent studies highlight organizational and systemic determinants: a longitudinal cohort in France identified irregular schedules and reliance on external staff as predictors of BEAs [7], whereas a systematic review in 2023 found a pooled global prevalence of needlestick injuries of 40.9% among nurses, with higher rates in low- and middle-income countries [8]. In addition, the WHO reported in 2023 that 41% of health workers in Africa experience at least one percutaneous injury annually, underscoring the magnitude of the problem [9]. The risk of transmission varies by pathogen: 10% to 40% for hepatitis B virus, approximately 2% for hepatitis C virus, and approximately 0.3% for HIV following a needlestick injury [10,11]. The WHO estimates that nearly 2 million needlestick injuries occur annually among health care workers, with 40% to 60% of hepatitis B and hepatitis C virus infections in this population being occupationally acquired [12]. Recent studies confirm the persistence of BEAs: prevalence rates have reached 93% in Cameroon [13], 78.9% among nursing students in Morocco [14], and 31% in India [15]. In Ethiopia and Nigeria, lifetime prevalence exceeds 60% [16,17]. In the Democratic Republic of the Congo (DRC), the problem is particularly acute. A study in Lubumbashi reported a prevalence of 73.2% among health care workers [18,19]. Noncompliance with standard precautions was identified as the strongest predictor (adjusted odds ratio [AOR] 2.921, 95% CI 1.503-5.677), confirming that most BEAs could be prevented through simple measures such as consistent use of personal protective equipment (PPE) and avoidance of needle recapping [19]. These findings highlight systemic gaps in infection prevention and control across provinces in the DRC. Multiple factors contribute to BEAs, including insufficient infection prevention training, inadequate PPE, poor compliance with standard precautions, absence of infection control committees, and overcrowded or poorly designed workspaces [20,21]. While high-income countries have implemented surveillance systems and continuous training programs [22], many African countries still face underreporting and poor documentation of BEAs [23]. This study aimed to determine the factors associated with BEAs among health care providers in Kalemie, DRC, hypothesizing that noncompliance with standard precautions significantly increases the occurrence of BEAs.

Methods

Study Design and Setting

We conducted a cross-sectional study in March 2024 among health care providers in Kalemie, Tanganyika Province, eastern DRC. The city comprises 2 health zones (Kalemie and Nyemba) covering 11 health areas and a total of 45 public, private, and faith-based health facilities.

Sample Size and Sampling Technique

The sample size was calculated using the single population proportion formula [24]: $n = [(z\alpha/2)^2 \times p\ (1 - p)]/d^2$, where z=1.96 (for 95% confidence), p=73.2% (estimated BEA prevalence from the aforementioned Lubumbashi study), and d=0.05 (margin of error). After accounting for a 10% nonresponse rate, the final sample size was 335.

A 2-stage sampling approach was applied. Although clustering was considered, no formal adjustment for design effect was performed as the calculated sample size was deemed sufficient to ensure representativeness. The final number of respondents was 316, corresponding to a response rate of 94.3% (316/335); the difference was due to nonresponses.

Recruitment Procedure

From the 45 health facilities listed in the National Health Information System (NHIS), 34 (75.6%) were randomly selected using a computer-generated list, ensuring proportional representation of public, private, and faith-based institutions. Within each selected facility, staff rosters were systematically validated by managers against official NHIS records to guarantee accuracy and completeness. The number of participants per facility was determined proportionally to staff size, and providers were then chosen using a random number generator. This systematic validation of lists and proportional allocation ensured representativeness and minimized selection bias.

Inclusion and Exclusion Criteria

Inclusion criteria were nurses, physicians, midwives, and laboratory technicians with direct contact with blood or body fluids. Exclusion criteria were providers with less than 1 month of professional experience to reduce onboarding variability.

Data Collection Tools and Procedures

Data were collected using a structured questionnaire adapted from validated WHO and Centers for Disease Control and Prevention (CDC) instruments [25-27]. The tool comprised closed-ended items covering sociodemographic characteristics,

XSL•FO
RenderX

knowledge of BEAs, attitudes toward reporting and prevention, compliance with standard precautions, and history of occupational exposure. Prior to implementation, the questionnaire was pilot-tested at Baraka Health Post and Shalom Polyclinic to ensure clarity and contextual relevance. Seven data collectors and 1 supervisor received 2 days of standardized training focused on interview techniques, ethical considerations, and data quality assurance. Reliability testing yielded a Cronbach α coefficient of 0.766, indicating acceptable internal consistency. Data collection was conducted through face-to-face interviews using the KoboCollect (Kobo) software installed on tablets. Each interviewer completed approximately 10 provider interviews per day over a 5-day period. Supervisory checks were performed daily to validate the completeness and accuracy of entries. The use of electronic data capture minimized transcription errors and ensured secure storage of anonymized records.

Variables

The dependent variable was occupational exposure to blood within the previous 12 months, defined as any percutaneous or mucocutaneous contact with potentially infectious biological fluids. Independent variables included the type of health facility, categorized as secondary level (general referral hospitals, referral health centers, and polyclinics) vs primary level (health centers, health posts, and dispensaries). Compliance with standard precautions was dichotomized as “always” vs “not always.” Needle recapping was assessed through self-reported frequency and classified as “never” vs “ever.” Availability and use of PPE were measured on an ordinal scale (“rarely,” “usually,” and “always”). Syringe reuse was recorded as a binary variable (“yes” or “no”). Sociodemographic characteristics were also collected, including age (categorized), sex, marital status, professional category, years of professional experience, and daily working hours.

Statistical Analysis

Data were exported from KoboCollect to Microsoft Excel 2016 and analyzed using SPSS (version 25.0; IBM Corp). Descriptive statistics were computed, including frequencies, percentages, medians, and IQRs. Normality of continuous variables was assessed using the Kolmogorov-Smirnov test [28]. Knowledge scores were derived from 31 items and categorized as optimal (>85%), moderate (70%-85%), or poor (<70%). Bivariate analyses were performed using chi-square tests, and variables with a *P* value of .20 or lower were retained for multivariate modeling. Logistic regression was conducted to identify independent predictors of BEAs, with AORs and 95% CIs reported. Both crude (odds ratio) and adjusted (AOR) estimates were presented, and potential confounding factors were explicitly discussed. Model diagnostics included assessment of multicollinearity using the variance inflation factor (<2), evaluation of model fit using the Hosmer-Lemeshow goodness-of-fit test, and discrimination using the area under the receiver operating characteristic curve [29]. Statistical significance was set at a *P* value below .05.

Ethical Considerations

This study was conducted and reported in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines [30]. Ethics approval was obtained from the ethics committee of the School of Public Health, University of Kinshasa (reference ESP/CE/38/2024). Written informed consent was obtained from all participants prior to their inclusion. Confidentiality was safeguarded by anonymizing the data and storing them securely, accessible only to authorized researchers. No financial compensation was provided to participants for their involvement in the study.

Results

Sociodemographic and Professional Characteristics

A total of 316 health care providers participated (response rate: 316/335, 94.3%). Most (n=256, 81%) worked in secondary-level facilities (general referral hospitals, referral health centers, and polyclinics). Most were female (n=179, 56.6%), married (n=230, 72.8%), and aged 40 years or older (n=128, 40.5%). Nurses represented 78.2% (n=247) of the sample. Median professional experience was 9 (IQR 5-15) years, and median workload was 8 (IQR 7-9) hours per day, as shown in Table 1.

XSL•FO
RenderX

Table 1. Sociodemographic and professional characteristics of health care providers surveyed in Kalemie, Tanganyika Province, Democratic Republic of the Congo, in March 2024 (cross-sectional study; N=316).

Characteristic and category	Values
Type of health facility, n (%)	
General referral hospital, referral health center, or polyclinic	256 (81)
Health center, health post, or dispensary	60 (19)
Age (y), median (IQR)	38 (30-46)
19-29, n (%)	66 (20.9)
30-39, n (%)	122 (38.6)
≥40, n (%)	128 (40.5)
Sex, n (%)	
Male	137 (43.4)
Female	179 (56.6)
Marital status, n (%)	
Single	57 (18)
Married	230 (72.8)
Common-law union	4 (1.3)
Widowed	25 (7.9)
Professional category, n (%)	
Nurse	247 (78.2)
Midwife or birth attendant	18 (5.7)
Laboratory technician	21 (6.6)
Physician	30 (9.5)
Department of assignment, n (%)	
Surgery or operating room	31 (9.8)
Obstetrics and gynecology	61 (19.3)
Laboratory	31 (9.8)
Internal medicine	40 (12.7)
Pediatrics	26 (8.2)
General services	117 (37)
Emergency	10 (3.2)
Experience (y), median (IQR)	9 (5-15)
<5, n (%)	67 (21.2)
≥5, n (%)	249 (78.8)
Workload (h), median (IQR)	8 (7-9)
≤8, n (%)	209 (66.1)
>8, n (%)	107 (33.9)

Prevalence and Circumstances of BEAs

Overall, 65.8% (208/316) reported at least one BEA during their career, and 60.4% (191/316) reported one within the previous 12 months. The leading causes were unexpected patient movement (81/191, 42.4%) and lack of attention (66/191, 34.6%). Most incidents occurred in treatment rooms (64/191, 33.5%), delivery rooms (33/191, 17.3%), and operating rooms (27/191, 14.1%).

Needlestick injuries accounted for 59.2% (113/191) of exposures, followed by splashes on damaged skin (43/191, 22.5%) and on mucous membranes (18/191, 9.4%). Hollow needles (77/191, 40.3%) and defective surgical equipment (49/191, 25.7%) were the objects most frequently responsible, as shown in Table 2. Only 36.1% (69/191) of BEAs were formally reported; among the 122 non-reported cases, the main reason was the perception that the incident was not risky (70/122, 57.4%).

XSL•FO
RenderX

Table 2. Prevalence and circumstances of blood exposure accidents (BEAs) among health care providers in Kalemie, Tanganyika Province, Democratic Republic of the Congo, in March 2024 (N=316).

Characteristic	Participants, n (%)
BEA occurrence during professional career	
No	108 (34.2)
Yes	208 (65.8)
BEA occurrence in the previous 12 mo	
No	125 (39.6)
Yes	191 (60.4)
Cause of BEA (n=191)	
Unexpected patient movement	81 (42.4)
Lack of attention	66 (34.6)
Lack of safety devices	23 (12)
Work context (stress or overload)	10 (5.2)
Lack of practical experience	7 (3.7)
Restricted workspace	4 (2.1)
Location of occurrence (n=191)	
Treatment room	64 (33.5)
Delivery room	33 (17.3)
Operating room	27 (14.1)
Laboratory	21 (11)
Patient room	17 (8.9)
Emergency room	17 (8.9)
Other	8 (4.2)
Vaccination session	4 (2.1)
Mechanism of occurrence (n=191)	
Needlestick	113 (59.2)
Splash on damaged skin	43 (22.5)
Splash on mucous membrane	18 (9.4)
Blade injury	14 (7.3)
Other	3 (1.6)
Object responsible (n=191)	
Hollow needle (syringe)	77 (40.3)
Defective surgical equipment	49 (25.7)
Catheter	29 (15.2)
Suture needle	24 (12.6)
Blade	12 (6.3)
Risky procedure responsible for BEA (n=191)	
Intramuscular injection	62 (32.5)
Suturing	27 (14.1)
Needle insertion or removal	22 (11.5)
Blood sampling	21 (11)
Other	19 (9.9)
Surgical procedure	14 (7.3)

XSL•FO
RenderX

Characteristic	Participants, n (%)
Delivery management	13 (6.8)
Puncture	7 (3.7)
Recapping used needle	6 (3.1)

Knowledge, Attitudes, and Practices

Nearly all participants (309/316, 97.8%) had heard of BEAs, mainly through formal education (125/316, 40.5%) and training (125/316, 40.5%). However, 72.8% (230/316) had never received specific training on standard precautions, and 94.3% (298/316) demonstrated poor knowledge levels.

Attitudes were mixed: 48.4% (153/316) strongly agreed on the importance of reporting BEAs, 72.2% (228/316) emphasized the importance of knowing the patient's serological status, and 73.4% (232/316) expressed willingness to initiate postexposure prophylaxis.

In practice, 76.6% (242/316) claimed to always follow standard precautions, but only 25% (79/316) consistently washed their hands before and after patient contact. Needle recapping was reported by 27.8% (88/316), PPE was always used by 19.6% (62/316), and syringe reuse occurred in 16.5% (52/316) of cases, as shown in Table 3.

XSL•FO
RenderX

Table 3. Knowledge, attitudes, and practices regarding blood exposure accidents (BEAs) among health care providers in Kalemie, Tanganyika Province, Democratic Republic of the Congo, in March 2024 (N=316).

Characteristic	Participants, n (%)
Had information about BEAs	
No	7 (2.2)
Yes	309 (97.8)
Source of information (n=309)	
School or university	125 (40.5)
Radio	5 (1.6)
Colleague	25 (8.1)
Training	125 (40.5)
Official memo	8 (2.5)
Other sources	21 (6.8)
Training on standard precautions	
No	230 (72.8)
Yes	86 (27.2)
Level of knowledge on BEAs	
Low	298 (94.3)
Moderate	16 (5.1)
Optimal	2 (0.6)
Importance of reporting BEAs	
Strongly disagree	26 (8.2)
Slightly agree	29 (9.2)
Agree	106 (33.5)
Strongly agree	153 (48.4)
Do not know	2 (0.6)
Importance of knowing patients' serostatus	
Strongly disagree	10 (3.2)
Slightly agree	9 (2.8)
Agree	67 (21.2)
Strongly agree	228 (72.2)
Do not know	2 (0.6)
Importance of knowing one's own serostatus	
Strongly disagree	6 (1.9)
Slightly agree	10 (3.2)
Agree	81 (25.6)
Strongly agree	218 (69)
Do not know	1 (0.3)
Importance of PPE[a] use for BEA prevention	
Strongly disagree	28 (8.9)
Slightly agree	69 (21.8)
Agree	139 (44)
Strongly agree	80 (25.3)
Postexposure prophylaxis consideration	

XSL•FO
RenderX

Characteristic	Participants, n (%)
No	84 (26.6)
Yes	232 (73.4)
Compliance with standard precautions	
Never	1 (0.3)
Rarely	36 (11.4)
Usually	37 (11.7)
Always	242 (76.6)
Hand hygiene	
Never	14 (4.4)
Rarely	67 (21.2)
Usually	156 (49.4)
Always	79 (25)
Needle recapping	
Never	89 (28.2)
Rarely	49 (15.5)
Usually	90 (28.5)
Always	88 (27.8)
Mouth pipetting	
Never	182 (57.6)
Rarely	7 (2.2)
Usually	11 (3.5)
Not applicable	116 (36.7)
Availability of PPE	
Rarely	99 (31.3)
Usually	137 (43.4)
Always	80 (25.3)
Use of PPE	
Never	1 (0.3)
Rarely	92 (29.1)
Usually	161 (50.9)
Always	62 (19.6)
Reuse of syringes	
No	264 (83.5)
Yes	52 (16.5)

[a]PPE: personal protective equipment.

Factors Associated With BEAs

Bivariate analysis identified 6 factors associated with BEAs, with exact *P* values: facility type (*P*<.001), daily working hours (*P*=.001), compliance with standard precautions (*P*<.001), needle recapping (*P*<.001), PPE availability (*P*=.03), and PPE use (*P*=.003).

Multivariate logistic regression confirmed two independent predictors: (1) noncompliance with standard precautions (AOR 2.921, 95% CI 1.503-5.677; *P*=.001) and (2) facility type (AOR 0.310, 95% CI 0.144-0.670; *P*=.003), as shown in Table 4.

XSL•FO
RenderX

Table 4. Bivariate and multivariate analysis of factors associated with blood exposure accidents (BEAs) among health care providers in Kalemie, Tanganyika Province, Democratic Republic of the Congo, in March 2024 (N=316).

Characteristic	BEA occurrence, n (%)		Bivariate analysis		Multivariate analysis	
	No (n=125)	Yes (n=191)	OR[a] (95% CI)	*P* value	Adjusted OR (95% CI)	*P* value
Type of health facility						
General referral hospital, referral health center, or polyclinic	88 (70.4)	168 (88.0)	3.071 (1.718-5.490)	*<.001* [b]	0.310 (0.144-0.670)	*.003* [b]
Health center, health post, or dispensary	37 (29.6)	23 (12.0)	1 (reference)	__[c]	1 (reference)	—
Sex						
Male	58 (46.4)	79 (41.4)	0.815 (0.517-1.284)	.38	—	—
Female	67 (53.6)	112 (58.6)	1 (reference)	—	—	—
Age group (y)						
19-29	32 (25.6)	34 (17.8)	0.681 (0.374-1.240)	.21	—	—
30-39	43 (34.4)	79 (41.4)	1.178 (0.704-1.969)	.53	—	—
≥40	50 (40)	78 (40.8)	1 (reference)	—	—	—
Professional category						
Midwife	7 (5.6)	11 (5.8)	1.179 (0.327-4.250)	.80	—	—
Nurse	101 (80.8)	146 (76.4)	1.084 (0.440-2.669)	.86	—	—
Physician	8 (6.4)	22 (11.5)	2.062 (0.631-6.739)	.23	—	—
Laboratory technician	9 (7.2)	12 (6.3)	1 (reference)	—	—	—
Department of assignment						
Surgery or operating room	6 (4.8)	25 (13.1)	2.206 (0.663-7.344)	.20	0.458 (0.130-1.613)	.22
Obstetrics and gynecology	20 (16)	41 (21.5)	1.085 (0.412-2.859)	.87	0.521 (0.171-1.589)	.25
Laboratory	15 (12)	16 (8.4)	0.565 (0.193-1.649)	.30	1.735 (0.314-9.593)	.53
Internal medicine	18 (14.4)	22 (11.5)	0.647 (0.233-1.795)	.40	1.352 (0.463-3.944)	.58
General services	55 (44)	61 (31.9)	0.577 (0.238-1.398)	.22	0.988 (0.370-2.641)	.98
Emergency	1 (0.8)	9 (4.7)	4.765 (0.518-43.798)	.17	0.282 (0.030-2.680)	.27
Pediatrics	9 (7.2)	17 (8.9)	1 (reference)	—	1 (reference)	—
Experience (y)						
<5	29 (23.2)	38 (19.9)	0.822 (0.476-1.420)	.48	—	—
≥5	96 (76.8)	153 (80.1)	1 (reference)	—	—	—
Workload (h)						
>8	56 (44.8)	51 (26.7)	2.228 (1.383-3.588)	*.001* [b]	0.627 (0.359-1.093)	.10
≤8	69 (55.2)	140 (73.3)	1 (reference)	—	1 (reference)	—
Training on standard precautions						
No	95 (76)	135 (70.7)	0.761 (0.455-1.274)	.30	—	—
Yes	30 (24)	56 (29.3)	1 (reference)	—	—	—
Compliance with standard precautions						
Not always	14 (11.2)	38 (19.9)	2.915 (1.707-4.980)	*<.001* [d]	2.921 (1.503-5.677)	*.001* [b]
Always	111 (88.8)	153 (80.1)	1 (reference)	—	1 (reference)	—
Needle recapping						
Always, usually, or rarely	75 (60)	39 (20.4)	2.598 (1.573-4.292)	*<.001* [d]	0.884 (0.459-1.703)	.71
Never	50 (40)	152 (79.6)	1 (reference)	—	1 (reference)	—

XSL•FO
RenderX

Characteristic	BEA occurrence, n (%)		Bivariate analysis		Multivariate analysis	
	No (n=125)	Yes (n=191)	OR[a] (95% CI)	*P* value	Adjusted OR (95% CI)	*P* value
Availability of PPE[e]						
Not always	85 (68)	151 (79.1)	0.563 (0.337-0.940)	*.03* [b]	0.760 (0.397-1.458)	.41
Always	40 (32)	40 (20.9)	1 (reference)	—	1 (reference)	—
Use of PPE during procedures						
Not always	39 (31.2)	54 (20.3)	2.362 (1.344-4.152)	*.003* [b]	0.546 (0.263-1.135)	.11
Always	86 (68.8)	137 (71.7)	1 (reference)	—	1 (reference)	—

[a]OR: odds ratio.

[b]$P<.05$.

[c]Not applicable.

[d]$P<.001$.

[e]PPE: personal protective equipment.

Discussion

Summary of Key Findings

This study highlights that BEAs remain a major occupational hazard among health care providers in Kalemie, Tanganyika Province, DRC. The strongest predictor was noncompliance with standard precautions [15], whereas working in secondary-level facilities appeared protective.

Interpretation and Comparison With Previous Studies

Our findings are consistent with those of recent African studies reporting high BEA prevalence: 93% in Cameroon [10], 78.9% in Morocco [11], more than 60% in Ethiopia [13,17], and more than 60% in Nigeria [14]. The Kalemie results reinforce the persistence of occupational risks across sub-Saharan Africa.

The protective effect of secondary-level facilities may be explained by better infrastructure, availability of PPE, and more structured infection prevention programs, as observed in Ethiopia [17] and Benin [18]. Conversely, primary-level facilities often lack resources, increasing exposure risks [20].

The association between noncompliance with standard precautions and BEAs has been consistently documented [27]. In Kalemie, this factor nearly tripled the risk, confirming that most accidents are preventable through adherence to basic infection control measures. The apparent contradiction between bivariate and multivariate analyses regarding facility type likely reflects confounding factors: secondary facilities host more complex procedures and higher patient volumes, increasing crude risk, but once adjusted for PPE availability, training, and compliance, their structured infection control systems reduced overall risk. Similar methodological challenges have been highlighted in recent epidemiological modeling studies [31].

Strengths and Limitations

The methodological rigor of this study, including the use of a validated questionnaire adapted from WHO and CDC tools [22,26], random sampling, and robust statistical diagnostics such as the Hosmer-Lemeshow test and the area under the receiver operating characteristic curve [28,29], supports the representativeness of the findings. Nevertheless, limitations must be acknowledged. Reliance on self-reported data may have introduced social desirability bias, as suggested by the discrepancy between the high proportion of providers claiming to “always” follow standard precautions (242/316, 76.6%) and the finding that noncompliance was the strongest predictor. Recall bias may also have affected lifetime prevalence estimates. Measurement limitations exist because exposure was assessed exclusively through self-report without external validation, and the cross-sectional design precludes causal inference. Finally, generalizability is limited as this study was conducted in Kalemie and may not fully represent health care providers in other provinces of the DRC.

Broader Implications

BEAs remain highly prevalent among health care providers in Kalemie. Most incidents are preventable through adherence to standard precautions, adequate provision of PPE, and establishment of surveillance systems. Strengthening institutional policies and continuous training are essential to reduce occupational risks and safeguard health care workers.

Beyond Kalemie, these findings highlight systemic gaps in infection prevention and control across sub-Saharan Africa. Addressing these gaps requires coordinated action: investment in PPE, integration of infection control committees, and reinforcement of reporting and monitoring systems. At the global level, the persistence of BEAs underscores the need for harmonized occupational health policies and stronger advocacy for health care worker safety.

Conclusions

BEAs remain a major occupational hazard for health care providers in Kalemie, DRC. With nearly two-thirds of providers reporting lifetime exposure and more than half exposed within the past year, the burden is substantial. Most incidents were preventable, and noncompliance with standard precautions emerged as the strongest predictor. Conversely, working in secondary-level facilities appeared protective once confounding factors were accounted for.

XSL•FO
RenderX

These findings underscore the urgent need to strengthen infection prevention and control measures, ensure consistent availability and use of PPE, and establish effective surveillance and reporting systems. Beyond Kalemie, the results highlight systemic gaps across sub-Saharan Africa, calling for coordinated national and international efforts to safeguard health care workers. Protecting those at the frontline is essential not only for their safety but also for the resilience of health systems.

Acknowledgments

The authors thank the Tanganyika Provincial Health Division for administrative support, the managers of the participating health facilities for their collaboration, and the health care providers who generously contributed their time and insights to this study. All authors declared that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee support for the publication of this article. No generative AI tools were used in the design, conduct, analysis, or writing of this study. All content was produced by the authors, and responsibility for the accuracy and integrity of the work rests entirely with them.

Data Availability

The datasets generated and analyzed during the current study are not publicly available due to confidentiality agreements with participating health facilities. However, anonymized data can be obtained from the corresponding author on reasonable request and with approval from the institutional ethics committee.

Funding

The authors declared no financial support was received for this work.

Authors' Contributions

Conceptualization: JKD, FIM
Data curation: JKD
Formal analysis: JKD, PMM
Investigation: JCBKI
Methodology: JKD, FIM, DMM
Resources: DSM
Supervision: JKD
Validation: DSM, DMM
Writing—original draft: JKD
Writing—review and editing: JKD, FIM, PMM, GKL, JCBKI, DSM, DMM
All authors have read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

References

1. Keubou Boukeng LB, Abo Okala ML, Bevela JY, Minkandi CA, Ebogo Etoa C. Accidental exposures to blood among dental health care workers in five referral hospitals in Yaoundé, Cameroon. Tunis Med 2025 Dec 27;103(10):1402-1408. [doi: 10.62438/tunismed.v103i10.5307] [Medline: 41879689]
2. Nouetchognou JS, Ateudjieu J, Jemea B, Mbanya D. Accidental exposures to blood and body fluids among health care workers in a referral hospital of Cameroon. BMC Res Notes 2016 Feb 15;9:94 [FREE Full text] [doi: 10.1186/s13104-016-1923-8] [Medline: 26879949]
3. Senbato FR, Wolde D, Belina M, Kotiso KS, Medhin G, Amogne W, et al. Compliance with infection prevention and control standard precautions and factors associated with noncompliance among healthcare workers working in public hospitals in Addis Ababa, Ethiopia. Antimicrob Resist Infect Control 2024 Mar 13;13(1):32 [FREE Full text] [doi: 10.1186/s13756-024-01381-w] [Medline: 38475931]
4. Ehimen FA, Akpan IS, Osagiede EF, Abah S, Okukpon P, Airefetalor I. A multi-center study of the comparative evaluation of occupational exposure to blood and body fluids among health care workers in Edo central senatorial district, Nigeria. Pan Afr Med J One Health 2020 Jun 12;2:11 [FREE Full text] [doi: 10.11604/pamj-oh.2020.2.11.23520]
5. Adal O, Abebe A, Feleke Y. Occupational exposure to blood and body fluids among nurses in the emergency department and intensive care units of public hospitals in Addis Ababa city: cross-sectional study. Environ Health Insights 2023 Feb 15;17:11786302231157223 [FREE Full text] [doi: 10.1177/11786302231157223] [Medline: 36818389]

XSL•FO
RenderX

6. Al-Faouri I, Okour SH, Alakour NA, Alrabadi N. Knowledge and compliance with standard precautions among registered nurses: a cross-sectional study. Ann Med Surg (Lond) 2021 Jan 29;62:419-424 [FREE Full text] [doi: 10.1016/j.amsu.2021.01.058] [Medline: 33552505]
7. Bun RS, Aït Bouziad K, Daouda OS, Miliani K, Eworo A, Espinasse F, et al. Identifying individual and organizational predictors of accidental exposure to blood (AEB) among hospital healthcare workers: a longitudinal study. Infect Control Hosp Epidemiol 2024 Apr;45(4):491-500 [FREE Full text] [doi: 10.1017/ice.2023.248] [Medline: 38086622]
8. Hosseinipalangi Z, Golmohammadi Z, Ghashghaee A, Ahmadi N, Hosseinifard H, Mejareh ZN, et al. Global, regional and national incidence and causes of needlestick injuries: a systematic review and meta-analysis. East Mediterr Health J 2022 Mar 29;28(3):233-241 [FREE Full text] [doi: 10.26719/emhj.22.031] [Medline: 35394056]
9. Prevention and control of viral hepatitis infection: framework for global action. World Health Organization. 2012. URL: https://iris.who.int/server/api/core/bitstreams/e664a5c2-4f02-43ad-8308-d00a18f3c13d/content [accessed 2026-07-23]
10. Mengistu DA, Dirirsa G, Mati E, Ayele DM, Bayu K, Deriba W, et al. Global occupational exposure to blood and body fluids among healthcare workers: systematic review and meta-analysis. Can J Infect Dis Med Microbiol 2022 Jun 3;2022:5732046 [FREE Full text] [doi: 10.1155/2022/5732046] [Medline: 35692264]
11. Mekonnin T, Tsegaye A, Berihun A, Kassachew H, Sileshi A. Occupational exposure to blood and body fluids among health care workers in Mizan Tepi University Teaching Hospital, Bench Maji Zone, South West Ethiopia. Med Saf Glob Health 2018;7(2) [FREE Full text] [doi: 10.4172/2574-0407.1000146]
12. Shitu S, Adugna G, Abebe H. Occupational exposure to blood/body fluid splash and its predictors among midwives working in public health institutions at Addis Ababa city Ethiopia, 2020. Institution-based cross-sectional study. PLoS One 2021 Jun 18;16(6):e0251815 [FREE Full text] [doi: 10.1371/journal.pone.0251815] [Medline: 34143794]
13. Cheuyem FZ, Mouangue C, Ajong BN, Edzamba MF, Hamadama DC, Achangwa C, et al. Occupational exposure to blood and other body fluids among healthcare workers in Cameroon: a systematic review and meta-analysis. Glob Health Econ Sustain 2025 Apr 18;3(3):185-196 [FREE Full text] [doi: 10.36922/ghes025090016]
14. Ogoina D, Pondei K, Adetunji B, Chima G, Isichei C, Gidado S. Prevalence and determinants of occupational exposures to blood and body fluids among health workers in two tertiary hospitals in Nigeria. Afr J Infect Dis 2014;8(2):50-54 [FREE Full text] [doi: 10.4314/ajid.v8i2.7] [Medline: 25729538]
15. Babidi B, Bakadia BM, Kalenga MP, Kimuni KC, Ndaya KA, Kasongo PC, et al. Evaluation of knowledge, attitudes and practices of health professionals in front of the exposure accidents to blood in two hospital structures of Lubumbashi. Open Access Libr J 2017 Aug 10;04(08):1-165. [doi: 10.4236/oalib.1103823]
16. Abere G, Yenealem DG, Wami SD. Occupational exposure to blood and body fluids among health care workers in Gondar Town, Northwest Ethiopia: a result from cross-sectional study. J Environ Public Health 2020;2020:3640247 [FREE Full text] [doi: 10.1155/2020/3640247] [Medline: 32508935]
17. Yazie TD, Chufa KA, Tebeje MG. Prevalence of needlestick injury among healthcare workers in Ethiopia: a systematic review and meta-analysis. Environ Health Prev Med 2019 Aug 14;24(1):52 [FREE Full text] [doi: 10.1186/s12199-019-0807-7] [Medline: 31409284]
18. Malonga KF, Mbutshu LH. Factors associated with blood exposure accidents in public hospitals in Lubumbashi in DR Congo. Occup Environ Med 2018;75:A18 [FREE Full text]
19. Yenesew MA, Fekadu GA. Occupational exposure to blood and body fluids among health care professionals in Bahir Dar town, Northwest Ethiopia. Saf Health Work 2014 Mar;5(1):17-22 [FREE Full text] [doi: 10.1016/j.shaw.2013.11.003] [Medline: 24932415]
20. Dilie A, Amare D, Gualu T. Occupational exposure to needle stick and sharp injuries and associated factors among health care workers in Awi Zone, Amhara Regional State, Northwest Ethiopia, 2016. J Environ Public Health 2017;2017:2438713 [FREE Full text] [doi: 10.1155/2017/2438713] [Medline: 28855920]
21. Harorani M, Ghaffari K, Jadidi A, Hezave AK, Davodabadi F, Barati N, et al. Adherence to personal protective equipment against infectious diseases among healthcare workers in Arak-Iran. Open Public Health J 2021 Dec 21;14:519-525. [doi: 10.2174/1874944502114010519]
22. Debelu D, Mengistu DA, Tolera ST, Aschalew A, Deriba W. Occupational-related injuries and associated risk factors among healthcare workers working in developing countries: a systematic review. Health Serv Res Manag Epidemiol 2023;10:23333928231192834 [FREE Full text] [doi: 10.1177/23333928231192834] [Medline: 37636487]
23. Global report on infection prevention and control 2024: executive summary. World Health Organization. 2024. URL: https://www.who.int/publications/i/item/B09195 [accessed 2026-07-23]
24. Mengistu DA, Tolera ST, Demmu YM. Worldwide prevalence of occupational exposure to needle stick injury among healthcare workers: a systematic review and meta-analysis. Can J Infect Dis Med Microbiol 2021;2021:9019534 [FREE Full text] [doi: 10.1155/2021/9019534] [Medline: 33564345]
25. Tawiah PA, Baffour-Awuah A, Appiah-Brempong E, Afriyie-Gyawu E. Identifying occupational health hazards among healthcare providers and ancillary staff in Ghana: a scoping review protocol. BMJ Open 2022 Jan 04;12(1):e058048 [FREE Full text] [doi: 10.1136/bmjopen-2021-058048] [Medline: 34983774]
26. Kofman AD, Struble KA, Heneine W, Gayle B, de Perio MA, Okasako-Schmucker DL, et al. 2025 US Public Health Service guidelines for the management of occupational exposures to human immunodeficiency virus and recommendations

XSL•FO
RenderX

for post-exposure prophylaxis in healthcare settings. Infect Control Hosp Epidemiol 2025 Sep;46(9):863-873. [doi: 10.1017/ice.2025.10254] [Medline: 41569270]
27. Seer-Uke NE, Dahiru U, Tyoakaa AA, Zwawua O, Aondoaver SS. Knowledge and practice of standard precautions among health care workers amidst COVID-19 in public secondary health facilities in Makurdi Metropolis, Nigeria. World J Med Case Rep 2022;3(2):33-37. [doi: 10.11648/j.wjmcr.20220302.15]
28. Hosmer DWJ, Lemeshow S, Sturdivant RX. Applied Logistic Regression. Hoboken, NJ: John Wiley & Sons; Mar 2013.
29. Vittinghoff E, Glidden DV, Shiboski SC, McCulloch CE. Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models. New York, NY: Springer; 2012.
30. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol 2008 Apr;61(4):344-349. [doi: 10.1016/j.jclinepi.2007.11.008] [Medline: 18313558]
31. Lim O, Chua WY, Wong A, Ling RR, Chan HC, Quek SC, et al. The environmental impact and sustainability of infection control practices: a systematic scoping review. Antimicrob Resist Infect Control 2024 Dec 24;13(1):156 [FREE Full text] [doi: 10.1186/s13756-024-01507-0] [Medline: 39719628]

Abbreviations

AOR: adjusted odds ratio
BEA: blood exposure accident
CDC: Centers for Disease Control and Prevention
DRC: Democratic Republic of the Congo
NHIS: National Health Information System
PPE: personal protective equipment
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 02.Apr.2026; peer-reviewed by B Narh Lasidji, O Mukuku; comments to author 20.Apr.2026; revised version received 19.Jul.2026; accepted 20.Jul.2026; published 05.Aug.2026.

Please cite as:
Kabamba Dibwe J, Ilunga Mbayo F, Mukalay Muvumbu P, Kasongo Lupitshi G, Banze Kabwe Ilunga JC, Sashila Mukebo D, Mpunga DM
Prevalence and Determinants of Blood Exposure Accidents Among Health Care Workers in Kalemie, Democratic Republic of the Congo: Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e97009
URL: https://publichealth.jmir.org/2026/1/e97009
doi:10.2196/97009
PMID:42555929

©Joe Kabamba Dibwe, Fiston Ilunga Mbayo, Papy Mukalay Muvumbu, Gloire Kasongo Lupitshi, Jean Claude Banze Kabwe Ilunga, Didier Sashila Mukebo, Dieudonné Mukendi Mpunga. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 05.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Linking Women's Empowerment to Mental Health Symptoms: Population-Based Study

Maggie Craig[1], MPH; Hannah Walter[1], MPH; Masum Ali[2], MSc; Shahriar Faruque[3], MD; Sanjib Saha[1], PhD

[1]Department of Clinical Sciences (Malmö), Lund University, BMC, Sölvegatan 19, Lund, Sweden

[2]Nutrition and Health Science, Laney Graduate School, Emory University, Atlanta, GA, United States

[3]Adult Psychiatry, Directorate General of Health, Directorate General of Health, Dhaka, Bangladesh

Corresponding Author:
Sanjib Saha, PhD
Department of Clinical Sciences (Malmö), Lund University, BMC, Sölvegatan 19, Lund, Sweden

Abstract

Background: Mental health disorders represent a growing global challenge, particularly in low- and middle-income countries, where women are disproportionately affected by social and structural inequalities. Although women's empowerment has been increasingly recognized as a strategy to improve physical health, its relationship with mental health remains less well understood.

Objective: This study examined the association between women's empowerment and mental health symptoms among ever-married women in Bangladesh.

Methods: A cross-sectional analysis was conducted using data from the 2022 Bangladesh Demographic and Health Survey. Women's empowerment was measured using the Survey-Based Women's Empowerment (SWPER) Global index across 3 domains: attitudes toward violence, social independence, and decision-making. Associations between empowerment domains and depressive symptoms, anxiety, and any mental health symptoms were assessed using multivariable logistic regression.

Results: Empowerment in the attitudes toward violence domain was significantly associated with mental health outcomes. Women with low empowerment in this domain had higher odds of depression (adjusted odds ratio [aOR] 1.81, 95% CI 1.16-2.81), anxiety (aOR 1.46, 95% CI 1.10-1.94), and any mental health symptoms (aOR 1.63, 95% CI 1.27-2.10). No significant associations were found for the social independence or decision-making domains.

Conclusions: Higher empowerment in the "attitudes toward violence" domain was associated with a lower prevalence of common mental health symptoms, suggesting that addressing gender-based violence norms may improve women's mental health in Bangladesh.

(JMIR Public Health Surveill 2026;12:e85839) doi:10.2196/85839

KEYWORDS

women's empowerment; mental health; Bangladesh; depression; anxiety; Survey-Based Women's Empowerment Global index; SWPER Global index; intimate partner violence; IPV; low- and middle-income countries; LMICs; demographic and health survey

Introduction

Mental health disorders are an increasingly critical global health concern, accounting for an estimated 17.2% of disability-adjusted life years (DALYs) in 2021 [1]. The burden is disproportionately high in low- and middle-income countries (LMICs), where limited health care access, stigma, and economic barriers restrict both prevention and treatment [2]. Women in LMICs experience a disproportionate burden of common mental disorders because of intersecting social, economic, and gender-related disadvantages, including poverty, limited educational and employment opportunities, restricted decision-making autonomy, gender discrimination, social isolation, intimate partner violence (IPV), and inadequate access to mental health services [2-4]. In South Asia, these structural inequities collectively contribute to depression, anxiety, and other common mental disorders among women [4]. Although IPV is a well-established determinant of women's mental health, it represents only 1 pathway through which gender inequality may influence psychological well-being [5].

Women's empowerment has emerged as an important social determinant of health and is recognized as a key component of Sustainable Development Goal (SDG) 5, which seeks to achieve gender equality and empower all women and girls [6]. Women's empowerment is a multidimensional construct encompassing access to resources, decision-making autonomy, freedom of movement, social independence, and the ability to challenge inequitable gender norms [7]. Greater empowerment has been associated with improved reproductive, maternal, and child health outcomes, increased health care use, and enhanced overall well-being [8]. Emerging evidence also suggests that women's empowerment may influence mental health by increasing

XSL•FO
RenderX

autonomy, self-efficacy, and social support while reducing exposure to adverse social and gender-related stressors [9].

Although an increasing number of studies have examined the relationship between women's empowerment and mental health, findings remain inconsistent across countries and empowerment domains. Some studies have reported that greater empowerment is associated with lower risks of depression and anxiety, whereas others have found weak or domain-specific associations [10-12]. These inconsistencies may partly reflect differences in the conceptualization and measurement of empowerment, as well as variation in sociocultural contexts. Consequently, examining empowerment as a multidimensional construct may provide a more comprehensive understanding of its relationship with women's mental health.

In Bangladesh, these structural disadvantages are particularly salient. The country has made substantial progress in women's education and maternal health, yet persistent gender inequalities continue to affect women's social and economic opportunities [13]. Women continue to face barriers related to employment, household decision-making, child marriage, and gender-based violence [13]. Social norms that justify violence against women remain prevalent in many communities, and IPV remains widespread, with 23.2% of women aged 15 to 49 years reporting physical or sexual violence in the past year [14]. These structural disadvantages may adversely affect both women's empowerment and their mental and physical health. Previous studies conducted in Bangladesh have primarily reported that lower autonomy, limited decision-making power, and exposure to gender-based violence are associated with poorer mental health outcomes [12,15-18]. However, these studies generally focused on specific dimensions of empowerment or selected populations, limiting our understanding of how multidimensional empowerment influences mental health at the national level.

Various indices have been used to measure empowerment, including the Women's Empowerment in Agriculture Index [19], the Gender Development Index [20], and the Gender Inequality Index [21]. Although useful at national and regional levels, these tools offer limited applicability for subnational or subgroup analyses. To address this gap, the Survey-Based Women's Empowerment (SWPER) Global index was developed in 2017 using Demographic and Health Survey (DHS) data from 34 African countries [22]. The index identifies 3 domains of empowerment—attitudes toward violence, social independence, and decision-making—and has since been adapted as the SWPER Global index for use across world regions [22,23]. The SWPER Global index provides a standardized multidimensional measure of women's empowerment across these 3 domains [23]. Although the relationship between IPV victimization and women's mental health has been extensively documented, considerably less is known about whether women's attitudes toward the acceptability of violence—measured as an empowerment domain rather than as exposure to IPV—are independently associated with mental health [24,25]. Understanding this relationship may provide new insights into how gender norms and women's empowerment influence psychological well-being beyond direct experiences of violence.

Evidence on women's empowerment and mental health in South Asia remains limited. A recent study in Nepal, using the SWPER Global index, found that social independence was protective against depression and anxiety, whereas decision-making was associated with increased risk [26]. Previous studies across various contexts have reported mixed effects; some have identified empowerment—particularly autonomy and agency—as protective against mental health symptoms [19,27-29], while others have reported detrimental effects linked to specific domains such as decision-making [11,26]. Such findings highlight the complexity and context-specific nature of the relationship between empowerment and mental health.

Despite growing evidence linking women's empowerment and mental health, important knowledge gaps remain. Few nationally representative studies have examined multidimensional women's empowerment in relation to mental health in Bangladesh, and the relative contributions of different empowerment domains remain unclear, particularly the role of attitudes toward the acceptability of violence. Given persistent gender inequalities and limited mental health resources in Bangladesh, understanding these relationships is essential for informing gender-responsive public health interventions. Therefore, this study examined the associations between the 3 SWPER Global empowerment domains and symptoms of depression, anxiety, and overall mental health among ever-married women in Bangladesh. We hypothesize that lower levels of empowerment across the attitudes toward violence, social independence, and decision-making domains would be associated with greater odds of adverse mental health outcomes.

Methods

Data Source

This study analyzed secondary data from the women's individual recode (IR) file of the nationally representative 2022 Bangladesh Demographic and Health Survey (BDHS). The BDHS was conducted in urban and rural areas across all 8 administrative divisions of Bangladesh between June and December 2022. It used a 2-stage stratified cluster sampling design based on the 2011 national census. In the first stage, 675 enumeration areas were selected with probability proportional to size. In the second stage, an average of 45 households per cluster were selected systematically, yielding a sample of 30,330 households. Ever-married women aged 15 to 49 years who were usual residents of, or who stayed in, selected households on the night before the survey were eligible for the interview. Thirty of the 45 selected households in each cluster were randomly allocated to the long women's questionnaire, which included the Patient Health Questionnaire–9 (PHQ-9) and Generalized Anxiety Disorder–7 (GAD-7) mental health modules. For this analysis, only women with complete data for SWPER Global empowerment indicators and mental health measures were included (N=17,027). Details of the BDHS methodology are available elsewhere [30].

Exposure Variable

Women's empowerment was the exposure variable, assessed using the SWPER Global index, which was constructed from 14 standardized DHS items available in the women's IR file

XSL•FO
RenderX

for married women. The index includes 3 domains: attitudes toward violence, social independence, and decision-making. The attitudes toward violence domain reflects women's acceptance or rejection of justifications for wife beating; the social independence domain captures women's education, age at marriage or first birth, age difference with the partner, and media exposure; and the decision-making domain reflects participation in household decisions. Domain scores were generated using principal component analysis and then categorized into tertiles (low, medium, and high) following established SWPER Global methodology [22,23].

Outcome Variables

This study evaluated 3 mental health outcomes: depression, anxiety, and any mental health symptoms. Anxiety was assessed using the GAD-7 scale [31]. Participants reported the frequency of their symptoms on a scale from 0 ("never") to 3 ("always") over the preceding 2 weeks, with responses to the 7 questions aggregated to yield a total score between 0 and 21. Respondents were categorized as exhibiting anxiety symptoms if they achieved a GAD-7 score of ≥6. This cutoff was used for the present population-based analysis as a symptom-screening definition rather than a clinical diagnosis. Depression was assessed using the PHQ-9 [32]. Participants reported the occurrence of their symptoms on a scale from 0 ("never") to 3 ("always"), with responses to the 9 questions aggregated to yield a cumulative score between 0 and 27. Respondents were categorized as exhibiting depressive symptoms if they achieved a PHQ-9 score of ≥10. Respondents were classified as exhibiting "any mental health symptoms" if they reported experiencing ≥1 symptom of anxiety and/or depression. This study categorized mental health symptoms according to methodologies established in prior research [26]. Although the PHQ-9 and GAD-7 are screening rather than diagnostic instruments, Bangladeshi validation studies have supported their reliability and validity in local samples, including evidence for unidimensionality and acceptable internal consistency [33].

Covariates

This study included various sociodemographic covariates such as age, place of residence, division, highest level of education, number of children, wealth index, religion, the respondent's employment status, the partner's employment status, and the respondent's BMI. These variables were selected a priori as potential confounders because they may be associated with both women's empowerment and mental health. Religion was included to account for cultural and normative differences that may shape gender roles and attitudes toward women's autonomy; partner employment status was included as a marker of household economic stability and financial stress; and BMI was included as an indicator of women's physical and nutritional status, which may be linked to both psychological well-being and social disadvantage. Age was classified into 3 groups: 15 to 29 years, 30 to 39 years, and 40 to 49 years. Residence was classified as urban or rural according to country-specific criteria [30]. Bangladesh comprises 8 divisions: Barishal, Chattogram, Dhaka, Khulna, Mymensingh, Rajshahi, Rangpur, and Sylhet. Education level was classified according to the highest level of schooling completed: no education, primary, secondary, or higher education. Children were classified into 3 categories: 0 or 1 child, 2 children, and ≥3 children, according to population distribution. The wealth index was developed through principal component analysis, a statistical technique frequently used to create composite indicators of economic status derived from household assets and characteristics. The population was subsequently categorized into quintiles, from the poorest to the richest [30]. The respondent's reported religion was categorized as either "Islam" or "Other." The employment status of both the respondent and their partner was categorized as binary variables: employed or unemployed. BMI was classified according to World Health Organization (WHO) standards: underweight (<18.5 kg/m^2), normal weight (18.5 kg/m^2-24.9 kg/m^2), overweight (25.0 kg/m^2-29.9 kg/m^2), and obese (≥30.0 kg/m^2). This study categorized BMI into 2 groups: "normal" (18.5 kg/m^2-24.9 kg/m^2) and "underweight/overweight" (<18.5 kg/m^2 or ≥25.0 kg/m^2) to identify individuals outside the conventional "normal" range.

Statistical Analysis

Descriptive statistics were used to summarize sample characteristics, with proportions estimated for categorical variables and means and SDs reported for continuous variables. The prevalence of anxiety, depression, and any mental health symptoms was calculated by sociodemographic characteristics and SWPER Global domains. Bivariate associations were evaluated using chi-square tests. Associations between empowerment levels and mental health outcomes were examined with bivariate and multivariate logistic regression; covariates with $P<.25$ in bivariate analyses were included in the multivariate models. Collinearity among covariates was assessed using variance inflation factors. Adjusted odds ratios (aORs) with 95% CIs were reported. Statistical significance was determined using an α level of .05, with all tests being 2-sided. Model fit was assessed using the Archer-Lemeshow *F*-adjusted goodness-of-fit test for survey-weighted logistic regression; a nonsignificant result indicated no evidence of lack of fit. All analyses were conducted in Stata (version 18.5; StataCorp) using survey commands (*svy*) to apply the BDHS sampling weights, strata, and primary sampling unit, thereby producing nationally representative estimates and valid SEs.

Sensitivity Analyses

We re-estimated all associations using survey-weighted Poisson regression with a log link to obtain adjusted prevalence ratios (aPRs) and 95% CIs. These models included the same exposures and covariates as the primary logistic regression models and accounted for the BDHS sampling weights, stratification, and clustering. This analysis was undertaken because odds ratios (ORs) can overstate the magnitude of association when outcomes are common in cross-sectional studies [34].

Ethical Considerations

The 2022 BDHS protocol was reviewed and approved by the Institutional Review Board of Inner City Fund (ICF) International and the Bangladesh Medical Research Council (BMRC). The publicly available BDHS documentation does not report the corresponding approval numbers. Before

XSL•FO
RenderX

participation, all respondents provided informed consent to take part in the survey. The present study was a secondary analysis of deidentified BDHS data obtained through the DHS program. The analytic dataset contained no direct personal identifiers, and the research team had no access to information that could identify individual participants. Therefore, no additional participant contact, recruitment, or consent was required for this analysis. No compensation was provided to participants for this secondary analysis. The manuscript and supplementary materials contain no identifiable participant information, including names, photographs, addresses, or other potentially identifying details. Detailed ethical aspects and consent forms are available on the BDHS website and in the report.

Results

This study included a total of 17,027 respondents, representing a population of ever-married women aged 15 to 49 years who responded to the questions required for calculating the SWPER Global index and assessing mental health in the 2022 BDHS. Table 1 presents a summary of the sociodemographic characteristics of the respondents included in the study.

XSL•FO
RenderX

Table . Sociodemographic characteristics of included 2022 Bangladesh Demographic and Health Survey respondents (N=17,027).

Sociodemographic characteristics	Participants, n (%)
Age (y), mean (SD)	32.82 (8.40)
Age group (y), n (%)	
15-29	6492 (38.1)
30-39	6286 (36.9)
40-49	4249 (25)
Residence type, n (%)	
Urban	5920 (34.8)
Rural	11,107 (65.2)
Division, n (%)	
Barishal	1819 (10.7)
Chattogram	2545 (14.9)
Dhaka	2560 (15)
Khulna	2229 (13.1)
Mymensingh	1837 (10.8)
Rajshahi	2172 (12.8)
Rangpur	2082 (12.2)
Sylhet	1783 (10.5)
Highest education level, n (%)	
No education	2354 (13.8)
Primary	4684 (27.5)
Secondary	7677 (45.1)
Higher	2312 (13.6)
Number of children, mean (SD)	2.25 (1.13)
Number of children grouped, n (%)	
0-1	4539 (26.6)
2	6707 (39.4)
≥3	5781 (34)
Household wealth index, n (%)	
Poorest	3089 (18.1)
Poorer	3361 (19.7)
Middle	3384 (19.9)
Richer	3518 (20.7)
Richest	3675 (21.6)
Religion, n (%)	
Islam	15,229 (89.4)
Other	1798 (10.6)
Respondent working, n (%)	
Yes	5402 (31.7)
No	11,625 (68.3)
Husband or partner working, n (%)	
Yes	16,505 (97)
No	506 (3)

XSL•FO
RenderX

Sociodemographic characteristics	Participants, n (%)
BMI of respondents (kg/m^2),[a] mean (SD)	23.97 (4.31)
BMI group[a], n (%)	
Normal (18.5 kg/m^2-24.9 kg/m^2)	4462 (52.5)
Abnormal (<18.5 kg/m^2 or ≥25.0 kg/m^2)	4029 (47.5)

[a]Some sociodemographic characteristics were not reported by all respondents; therefore, the total for these variables may not be equal to the sample size (N=17,027).

Mental health outcomes were stratified by sociodemographic characteristics and summarized in Table 2. Across the sample, 4.9% of respondents were classified as having depression, 19.5% as having anxiety, and 31.9% as having at least 1 mental health symptom in the 2022 BDHS. Education was inversely associated with the prevalence of anxiety, depression, and any mental health symptoms, with lower levels of education associated with a higher prevalence of mental health symptoms. Similarly, the household wealth index was inversely associated with mental health outcomes, with poorer individuals more likely to experience depression, anxiety, and any mental health symptoms. Being underweight or overweight was associated with a higher prevalence of anxiety, depression, and any mental health symptoms (Table 2).

Table . Unweighted frequencies and weighted prevalence of depression, anxiety, and any mental health symptoms (MHSs) by women's empowerment domain.[a]

Empowerment domains or categories	Depression (n=830), n (weighted %)	*P* value	Anxiety (n=3269), n (weighted %)	*P* value	Any MHSs (n=5417), n (weighted %)	*P* value
SWPER[b]—attitudes toward violence		.001		<.001		<.001
Low	48 (7.6)		174 (27.1)		262 (41.7)	
Medium	107 (5.8)		421 (23.5)		695 (39.9)	
High	675 (4.4)		2674 (18.2)		4460 (30)	
SWPER—social independence		.01		<.001		<.001
Low	425 (5.3)		1712 (21.3)		2646 (33)	
Medium	291 (4.3)		1120 (17.5)		1973 (30.8)	
High	114 (3.8)		437 (16.1)		798 (28.4)	
SWPER—decision-making		.60		.048		.007
Low	96 (4.4)		383 (16.9)		653 (29.4)	
Medium	194 (4.5)		834 (18.8)		1518 (33.8)	
High	540 (4.9)		2052 (19.7)		3246 (31)	

[a]Frequencies are unweighted numbers of women with depression, anxiety, or at least 1 mental health symptom. Percentages are weighted estimates accounting for the 2022 Bangladesh Demographic and Health Survey sampling weights, stratification, and clustering. *P* values were calculated using design-based Pearson chi-square tests.

[b]SWPER: Survey-Based Women's Empowerment.

The distribution of respondents' empowerment levels varied across the 3 SWPER Global domains. Most women had high empowerment in the "attitudes toward violence" (86.2%) and decision-making (60.7%) domains, whereas nearly half had low empowerment in the social independence domain (46.3%). Overall, higher empowerment across the 3 domains was more common among women who lived in urban areas, had higher educational attainment, and belonged to wealthier households. Detailed distributions of empowerment levels according to sociodemographic characteristics are presented in Tables S1A to S1C in Multimedia Appendix 1.

Unadjusted ORs for the associations between women's empowerment domains and mental health outcomes are summarized in Table 3. Figure 1 presents graphical representations of aORs after adjustment for covariates including age, residence type, division, education level, number of children, wealth index, religion, work status, and BMI. Results indicated a significant association between mental health outcomes and empowerment in the "attitudes toward violence" domain. Women with low empowerment in this domain had higher odds of depression (aOR 1.81, 95% CI 1.16-2.81), anxiety (aOR 1.46, 95% CI 1.10-1.94), and any mental health symptoms (aOR 1.63, 95% CI 1.27-2.10). Medium empowerment in this domain was also associated with higher odds of anxiety (aOR 1.26, 95% CI 1.04-1.53) and any mental health symptoms (aOR 1.56, 95% CI 1.32-1.85). No significant association was observed between empowerment in the social independence or decision-making domains and mental health

XSL•FO
RenderX

outcomes in this study. The Archer-Lemeshow *F*-adjusted goodness-of-fit test indicated no evidence of lack of fit for all final models (*P*>.05), except for the model examining depression in relation to the "Attitude to Violence" domain (*P*=.03).

Table . Unadjusted associations of SWPER (Survey-Based Women's Empowerment) Global index domains and mental health symptoms (MHSs) among women in Bangladesh[a].

Empowerment domains or categories	Depression, OR[b] (95% CI)	Anxiety, OR (95% CI)	Any MHSs, OR (95% CI)
SWPER—attitudes toward violence			
Low	2.05 (1.46-2.87)[c]	1.78 (1.44-2.22)[c]	1.72 (1.41-2.11)[c]
Medium	1.36 (1.08-1.71)[d]	1.41 (1.22-1.64)[c]	1.58 (1.40-1.80)[c]
High	1.00	1.00	1.00
SWPER—social independence			
Low	1.39 (1.07-1.79)[d]	1.40 (1.23-1.60)[c]	1.24 (1.11-1.39)[c]
Medium	1.13 (0.87-1.45)	1.09 (0.95-1.25)	1.12 (1.00-1.25)[e]
High	1.00	1.00	1.00
SWPER—decision-making			
Low	0.89 (0.69-1.19)	0.83 (0.71-0.98)[e]	0.93 (0.81-1.06)
Medium	0.92 (0.74-1.14)	0.94 (0.84-1.05)	1.14 (1.04-1.25)[d]
High	1.00	1.00	1.00

[a]Statistical significance was determined using 2-sided tests with an α of .05. Empowerment levels were calculated using the SWPER Global index and categorized using the original studies' methodologies [22,23].

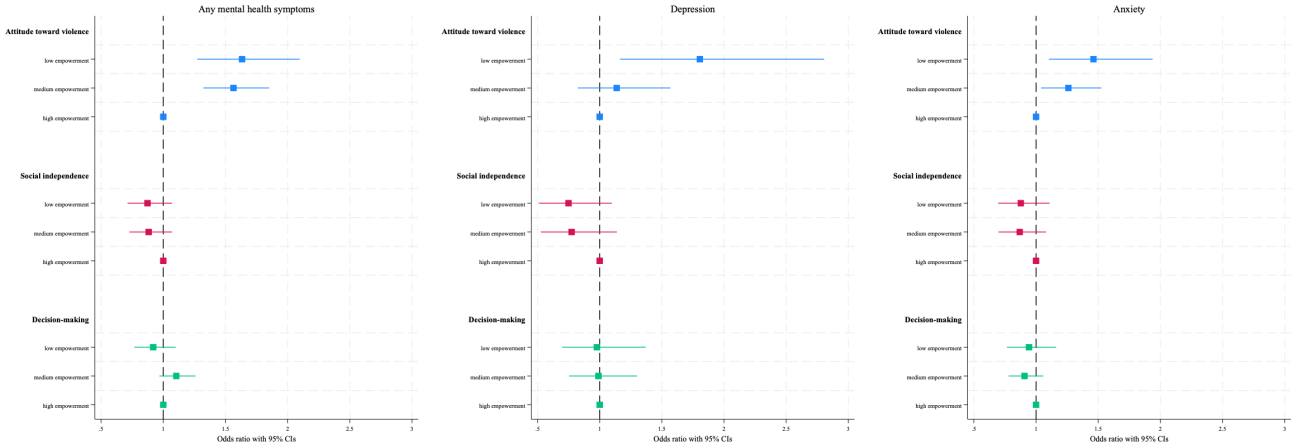
[b]OR: odds ratio; unadjusted ORs are based on bivariate logistic regression models.

[c]*P*<.001.

[d]*P*<.01.

[e]*P*<.05.

Figure 1. Adjusted odds ratios presenting the associations between women's empowerment domains and mental health symptom prevalence. Odds ratios were adjusted for age, residence type, division, education level, number of children, religion, employment status (respondent and partner), and BMI.



Sensitivity analyses using survey-weighted Poisson regression produced aPRs that were consistent in direction and statistical significance with the primary logistic regression findings (Tables S2A-S2C in Multimedia Appendix 1).

Discussion

Principal Findings

This study found a significant association between women's empowerment in the "attitudes toward violence" domain and mental health outcomes among Bangladeshi women, whereas no significant association was observed for the social independence or decision-making domains. These findings indicate that the relationship between empowerment and mental

XSL•FO
RenderX

health is multidimensional and may vary according to the specific dimension of empowerment being assessed.

The present findings are broadly consistent with evidence from neighboring South Asian countries, although the patterns are not uniform across settings. A recent nationally representative study from Nepal reported that greater social independence was associated with lower odds of anxiety and depression, whereas higher decision-making empowerment was associated with increased odds of these symptoms [26]. Similarly, prospective evidence from rural India suggests that the relationship between women's agency and mental distress is complex and context-dependent rather than uniformly protective [11]. In Pakistan, population-based research has shown that IPV is strongly associated with poorer mental health among married women, highlighting the broader regional importance of violence-related gender norms for women's psychological well-being [35].

Attitudes Toward Violence

Women with lower empowerment in the "attitudes toward violence" domain—those less likely to reject justifications for IPV—had significantly higher odds of depression, anxiety, and any mental health symptoms, even after adjusting for confounders. Although the observed aORs were modest in magnitude, they may still be important at the population level because even small shifts in common mental health symptoms can translate into substantial public health impacts in a large population. The findings therefore suggest that attitudes toward violence may be a meaningful but not the sole determinant of women's mental health in Bangladesh. This finding is consistent with broader literature showing that IPV and tolerance of violence are closely linked to women's psychological distress [36,37]. In Bangladesh, previous studies have shown that IPV is associated with depression and that women exposed to violence may face substantial barriers to seeking mental health support [38-40]. Studies in South Asia have also demonstrated that women experiencing IPV report poorer psychological well-being, while attitudinal rejection of IPV can promote resilience and access to social support and services [11,26,35].

Less attention has been paid to the role of women's attitudes toward IPV, as distinct from victimization itself, in shaping mental health outcomes. Recent evidence from Nepal found no significant association between the SWPER Global "attitudes toward violence" domain and mental health symptom prevalence [12], in contrast to the protective effects found in this Bangladeshi sample. The observed benefits in Bangladesh may reflect context-specific pathways. One possible explanation is that women who reject IPV may also differ in other unmeasured ways that are linked to better mental health, although the present cross-sectional data cannot establish such pathways. Given that 23.2% of women in Bangladesh have experienced IPV [8], attitudinal rejection may be a critical focus for both empowerment and mental health interventions.

Taken together, these findings suggest that attitudes toward IPV may capture an important dimension of internalized gender norms that is especially relevant to mental health in Bangladesh. Women who reject the acceptability of IPV may be more likely to recognize abuse, seek support, and resist harmful relationship dynamics, whereas acceptance of IPV may reflect deeper social normalization of violence and reduced psychosocial agency. In a setting where IPV remains prevalent, this domain may therefore be particularly salient for understanding women's mental health vulnerability.

Social Independence

The association between the "social independence" domain and mental health symptoms was not statistically significant after adjustment for confounders, although the unadjusted analyses suggested a modest protective pattern. The absence of a statistically significant association for social independence does not necessarily indicate that no relationship exists. Rather, it may suggest that any effect is small, context-dependent, or not fully captured by the current cross-sectional measures.

This mixed pattern aligns with existing literature reporting both protective and adverse effects of social independence–related empowerment on women's mental health. In Nepal, higher social independence was associated with lower odds of both anxiety and depression [26]. In Bangladesh, related research among female garment workers suggests that empowerment-related resource mobilization may contribute to lower depressive symptoms and better self-esteem [27]. At the same time, evidence from India indicated that women who were more comfortable speaking up in public reported greater mental distress, possibly because of social backlash and increased visibility within conservative communities [11].

These divergent findings underscore the context-dependent and multidimensional nature of social independence as an empowerment construct. Although greater social independence may be associated with improved access to resources, mobility, and support, it may also coincide with increased scrutiny, interpersonal conflict, or other social pressures in more conservative settings. Because this study is cross-sectional, these patterns should be interpreted cautiously and should not be taken to indicate causal pathways. It is also possible that more socially independent women are more likely to recognize or report psychological symptoms, which could influence observed associations. Further longitudinal and qualitative research is needed to clarify how social independence relates to women's mental health across different contexts.

Decision-Making

This study found no statistically significant association between women's empowerment in the "decision-making" domain and mental health symptoms. The absence of a statistically significant association for decision-making does not necessarily imply that there is no relationship. Instead, it may reflect a small, context-dependent effect or a pattern that is not well captured by the current cross-sectional design and measurement approach. However, this finding is consistent with the broader literature showing that household decision-making does not have a uniform relationship with women's mental health. For example, higher decision-making empowerment has been associated with increased odds of anxiety and depression in Nepal [26], whereas studies from Mozambique [41] and other sub-Saharan African settings [28,42] have reported protective associations. Evidence from India further suggests that greater decision-making

XSL•FO
RenderX

autonomy may be accompanied by psychological strain when women assume nontraditional roles in patriarchal settings where such changes are not fully supported by families or communities [11,26].

The lack of a significant relationship in our study should be interpreted cautiously. Because this study is cross-sectional, it does not allow causal inference. One possible explanation is that decision-making autonomy may be associated with both potential benefits and potential burdens: although it may increase women's control over household matters, it may also coincide with greater responsibility, interpersonal tension, or social and familial resistance. These findings suggest that the relationship between household decision-making and mental health may be context-dependent, and that instrumental forms of empowerment do not necessarily correspond to better psychological well-being in all settings.

Mental Health and Empowerment Patterns Across Sociodemographic Groups

In our sample, depression affected a smaller proportion of respondents (4.9%) than that reported in both the 2019 National Mental Health Survey in Bangladesh (7.9%) [43] and the 2021 Global Burden of Disease (GBD) estimate for women aged 15 to 49 years (8.3%). In contrast, anxiety prevalence (19.5% in this study) was notably higher than the estimates from both sources (5.4% and 5.5%, respectively) [43].

Several methodological differences may help explain these discrepancies. The National Mental Health Survey used stricter diagnostic procedures, including assessment by trained research psychiatrists, whereas the present study relied on screening instruments included in the BDHS. The GBD estimates may also reflect older underlying data, and differences in instruments, symptom thresholds, and case definitions across studies may further affect prevalence estimates.

Policy Implications

The present findings have several practical implications for mental health and gender equity programming in Bangladesh. Because the attitudes toward violence domain was the only empowerment dimension consistently associated with depression, anxiety, and overall mental health symptoms, interventions that aim to reduce the social acceptability of IPV may have benefits beyond violence prevention alone. Mental health providers, community health workers, and frontline practitioners could integrate brief screening for violence-related norms, psychological distress, and support needs into routine services for women, particularly in primary care and reproductive health settings. Public health programs may also benefit from combining mental health promotion with community-based efforts to challenge harmful gender norms, strengthen referral pathways, and improve access to psychosocial support for women at risk of both violence and poor mental health.

These findings also suggest several directions for future research and policy. Longitudinal studies are needed to better understand the causal pathways linking different empowerment domains and mental health over time. Further qualitative and mixed methods research could help clarify why attitudes toward violence appear more strongly related to mental health than social independence or decision-making in this setting. At the policy level, integrating women's empowerment indicators into mental health surveillance and embedding gender-transformative approaches within national mental health and violence prevention strategies may help translate these findings into more effective and contextually appropriate interventions.

Strengths and Limitations

This study has several limitations. First, all data were self-reported, raising the possibility of reporting bias—especially on sensitive topics such as mental health and women's empowerment—due to stigma or gender norms in Bangladesh. Highly empowered women may also be more likely to recognize and disclose mental health symptoms, potentially inflating observed associations. The BDHS used screening instruments rather than clinical diagnostic tools, which may underestimate the true burden of mental health symptoms and limit the scope of assessed conditions. Although the SWPER Global index is a validated tool, it may not fully capture the cultural and contextual nuances of empowerment. Despite adjusting for key sociodemographic factors, unmeasured confounders—including childhood trauma, extended family dynamics, or access to mental health services—may influence results. We included religion, partner employment status, and BMI as contextual covariates to better account for the cultural environment, household economic conditions, and women's physical health, all of which may plausibly influence both empowerment and mental health. Although these were not the primary variables of interest, their inclusion was intended to reduce residual confounding and improve interpretation of the observed associations.

The main strengths of this study include its use of a large, nationally representative BDHS sample, which enhances the generalizability of the findings to Bangladeshi women. The use of the SWPER Global index also allowed separate examination of 3 distinct dimensions of empowerment—attitudes toward violence, social independence, and decision-making—rather than treating empowerment as a single construct, which improves the specificity of the findings. In addition, the study used standardized screening measures for anxiety and depression and adjusted for several relevant sociodemographic and contextual covariates, which strengthens internal validity and supports a more nuanced interpretation of the observed associations. Finally, this research is among the first to systematically explore the association between women's empowerment and mental health symptom prevalence in Bangladesh, addressing an important gap in the literature.

Conclusions

In this nationally representative study of Bangladeshi women, only the SWPER Global domain of attitudes toward violence was independently associated with lower prevalence of mental health symptoms after adjusting for sociodemographic factors. Women who rejected justifications for IPV reported lower prevalence of depression, anxiety, and overall mental health symptoms. The lack of association for the social independence and decision-making domains, coupled with inconsistencies across the literature, underscores the need for contextually

XSL•FO
RenderX

tailored interventions. These findings highlight the importance of integrating mental health support and targeting attitudes toward the acceptance of IPV and harmful gender norms within women's empowerment initiatives in Bangladesh. In practice, these results support the inclusion of violence-norm screening, community awareness activities, and culturally responsive psychosocial support within women's health and primary care services. Future research should further investigate the mechanisms linking empowerment domains and mental health to guide more targeted interventions for Bangladeshi women.

Acknowledgments

The authors acknowledge that this study would not have been possible without the public availability of 2022 Bangladesh Demographic and Health Survey data provided by the Demographic and Health Survey Program. The authors declare the use of generative AI (GAI) in the research and writing process. According to the Generative Artificial Intelligence Delegation Taxonomy (2025), the following tasks were delegated to GAI tools under full human supervision: data analysis, proofreading, and editing. The GAI tool used was Perplexity (Perplexity AI Inc). Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Funding

The authors declare that no financial support was received for this work.

Data Availability

The data used in this study are publicly available from the Demographic and Health Surveys (DHS) program. The 2022 Bangladesh Demographic and Health Survey dataset can be accessed upon request through the DHS program website [44].

Authors' Contributions

Conceptualization: MC, MA, SF, SS
Data curation: MA
Formal analysis: MC
Investigation: MC, SS
Methodology: MC, HW, SS
Project management: MC
Resources: SS
Software: MC, SS
Supervision: SS
Validation: MC, HW, SF, SS
Visualization: MC, HW, SS
Writing—original draft: MC
Writing—review and editing: MC, HW, MA, SF, SS

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional analyses.
[DOCX File, 64 KB - publichealth_v12i1e85839_app1.docx]

References

1. Mental health. Institute for Health Metrics and Evaluation. URL: https://www.healthdata.org/research-analysis/health-topics/mental-health-our-approach [accessed 2026-08-24]
2. Freeman M. Investing for population mental health in low and middle income countries-where and why? Int J Ment Health Syst 2022 Aug 11;16(1):38. [doi: 10.1186/s13033-022-00547-6] [Medline: 35953845]
3. Kirkbride JB, Anglin DM, Colman I, et al. The social determinants of mental health and disorder: evidence, prevention and recommendations. World Psychiatry 2024 Feb;23(1):58-90. [doi: 10.1002/wps.21160] [Medline: 38214615]
4. Lund C, Breen A, Flisher AJ, et al. Poverty and common mental disorders in low and middle income countries: a systematic review. Soc Sci Med 2010 Aug;71(3):517-528. [doi: 10.1016/j.socscimed.2010.04.027] [Medline: 20621748]
5. Gunarathne L, Bhowmik J, Apputhurai P, Nedeljkovic M. Factors and consequences associated with intimate partner violence against women in low- and middle-income countries: a systematic review. PLoS One 2023;18(11):e0293295. [doi: 10.1371/journal.pone.0293295] [Medline: 37939106]

XSL•FO
RenderX

6. Odera JA, Mulusa J. SDGs, gender equality and women's empowerment: what prospects for delivery? In: Kaltenborn M, Krajewski M, Kuhn H, editors. Sustainable Development Goals and Human Rights: Springer; 2020:95-118. [doi: 10.1007/978-3-030-30469-0_6]
7. Wiafe PA, Appiah-Boateng S, Tuffour KA, Asante I. A multidimensional approach to women's empowerment in South Africa: understanding the driving force. Cogent Econ Finance 2026;14(1):2676343. [doi: 10.1080/23322039.2026.2676343]
8. Bliznashka L, Udo IE, Sudfeld CR, Fawzi WW, Yousafzai AK. Associations between women's empowerment and child development, growth, and nurturing care practices in sub-Saharan Africa: a cross-sectional analysis of demographic and health survey data. PLoS Med 2021 Sep;18(9):e1003781. [doi: 10.1371/journal.pmed.1003781] [Medline: 34529666]
9. Jamil A. Beyond the physical: understanding the emotional, psychological, and social factors affecting women's health today. Cureus 2025 Apr;17(4):e81874. [doi: 10.7759/cureus.81874] [Medline: 40201044]
10. Rafi MA, Anika US, Hasan MT, Hossain MG. Association between women's empowerment and mental health help-seeking behaviour in Bangladesh: findings from a nationally representative survey. BMJ Open 2025 Sep 14;15(9):e099770. [doi: 10.1136/bmjopen-2025-099770] [Medline: 40953863]
11. Richardson RA, Harper S, Bates LM, Nandi A. The effect of agency on women's mental distress: a prospective cohort study from rural Rajasthan, India. Soc Sci Med 2019 Jul;233:47-56. [doi: 10.1016/j.socscimed.2019.05.052] [Medline: 31176057]
12. Eom YJ, Jang S, Hasman A, Subramanian SV, Kim R. Exploring the link between women's empowerment, depression, and anxiety in Bangladesh: evidence from a nationally representative cross-sectional study. J Glob Health 2026 May 29;16:04189. [doi: 10.7189/jogh.16.04189] [Medline: 42212582]
13. UN Women Bangladesh newsletter: July-December 2025. UN Women. 2025. URL: https://asiapacific.unwomen.org/en/digital-library/publications/2026/01/un-women-bangladesh-newsletter-july-dec-2025 [accessed 2026-08-24]
14. Bangladesh. UN Women. URL: https://data.unwomen.org/country/bangladesh [accessed 2026-08-24]
15. Hoq MN, Ara F, Ahmad M, Hossain MM, Sultana MM. The rural-urban disparity in the women's justification of intimate partner violence and Awareness of Domestic Violence Act 2010 in Bangladesh. World Dev Sustain 2026 Jun;8:100306. [doi: 10.1016/j.wds.2026.100306]
16. Halim SF, Akter O, Islam MK. Link between gender inequality and women's satisfaction with maternal healthcare services: a cross-sectional study in Khulna, Bangladesh. BMJ Open 2024 Nov 7;14(11):e084404. [doi: 10.1136/bmjopen-2024-084404] [Medline: 39515870]
17. Ahmed T, Luthfa ST, Chaahat AH, Kanchon MR, Sharif AB. Determinants of early mental health help-seeking among women in Bangladesh: a nationally representative bootstrapped regression analysis. PLOS Ment Health 2025;2(9):e0000420. [doi: 10.1371/journal.pmen.0000420] [Medline: 41662058]
18. Islam MZ, Sharf-Ul-Alam SM, Rahman MF, et al. Mental health status and associated factors among rural adult women: findings of a community-based cross-sectional study in Bangladesh. BMJ Open 2026 Mar 24;16(3):e109974. [doi: 10.1136/bmjopen-2025-109974] [Medline: 41876157]
19. Leight J, Pedehombga A, Ganaba R, Gelli A. Women's empowerment, maternal depression, and stress: evidence from rural Burkina Faso. SSM Ment Health 2022 Dec;2:100160. [doi: 10.1016/j.ssmmh.2022.100160] [Medline: 36688233]
20. Gender Development Index (GDI) & Gender Inequality Index (GII). UNESCO. 2020. URL: https://www.unesco.org/en/world-media-trends/gender-development-index-gdi-gender-inequality-index-gii [accessed 2026-08-24]
21. Yu S. Uncovering the hidden impacts of inequality on mental health: a global study. Transl Psychiatry 2018 May 18;8(1):98. [doi: 10.1038/s41398-018-0148-0] [Medline: 29777100]
22. Ewerling F, Lynch JW, Victora CG, van Eerdewijk A, Tyszler M, Barros AJ. The SWPER index for women's empowerment in Africa: development and validation of an index based on survey data. Lancet Glob Health 2017 Sep;5(9):e916-e923. [doi: 10.1016/S2214-109X(17)30292-9] [Medline: 28755895]
23. Ewerling F, Raj A, Victora CG, Hellwig F, Coll CV, Barros AJ. SWPER Global: a survey-based women's empowerment index expanded from Africa to all low- and middle-income countries. J Glob Health 2020 Dec;10(2):020343. [doi: 10.7189/jogh.10.020434] [Medline: 33274055]
24. Sardinha L, Nájera Catalán HE. Attitudes towards domestic violence in 49 low- and middle-income countries: a gendered analysis of prevalence and country-level correlates. PLoS One 2018;13(10):e0206101. [doi: 10.1371/journal.pone.0206101] [Medline: 30379891]
25. Ibala RM, Seff I, Stark L. Attitudinal acceptance of intimate partner violence and mental health outcomes for female survivors in sub-Saharan Africa. Int J Environ Res Public Health 2021 May 12;18(10):5099. [doi: 10.3390/ijerph18105099] [Medline: 34065818]
26. Shawon MS, Hossain FB, Ahmed R, Poly IJ, Hasan M, Rahman MR. Role of women empowerment on mental health problems and care-seeking behavior among married women in Nepal: secondary analysis of nationally representative data. Arch Womens Ment Health 2024 Aug;27(4):527-536. [doi: 10.1007/s00737-024-01433-5] [Medline: 38315185]
27. Parvin K, Mamun MA, Gibbs A, Jewkes R, Naved RT. The pathways between female garment workers' experience of violence and development of depressive symptoms. PLoS One 2018;13(11):e0207485. [doi: 10.1371/journal.pone.0207485] [Medline: 30440031]

XSL•FO
RenderX

28. Scott S, Arrieta A, Kumar N, Menon P, Quisumbing A. Multidimensional predictors of common mental disorders among Indian mothers of 6- to 24-month-old children living in disadvantaged rural villages with women's self-help groups: a cross-sectional analysis. PLoS One 2020;15(6):e0233418. [doi: 10.1371/journal.pone.0233418] [Medline: 32574218]
29. Yoshikawa K, Kamiya Y. Married couples' perceptions on women's autonomy and subjective well-being in Lao PDR. Int J Soc Econ 2019 Jun;46(6):792-804. [doi: 10.1108/IJSE-09-2018-0450]
30. Bangladesh demographic and health survey 2022: final report. : National Institute of Population Research and Training (NIPORT) and ICF; 2024 URL: https://dhsprogram.com/pubs/pdf/FR386/FR386.pdf [accessed 2026-08-24]
31. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
32. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613. [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
33. Redwan A, Karim MR, Royle RK, Chowdhury AR. Validation of Bangla generalized anxiety disorder 7 scale among general population. Bangl J Psychiatry 2020;34(1):3-8. [doi: 10.3329/bjpsy.v34i1.71969]
34. Barros AJ, Hirakata VN. Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. BMC Med Res Methodol 2003 Oct 20;3:21. [doi: 10.1186/1471-2288-3-21] [Medline: 14567763]
35. Ali TS, Mogren I, Krantz G. Intimate partner violence and mental health effects: a population-based study among married women in Karachi, Pakistan. Int J Behav Med 2013 Mar;20(1):131-139. [doi: 10.1007/s12529-011-9201-6] [Medline: 22037921]
36. White SJ, Sin J, Sweeney A, et al. Global prevalence and mental health outcomes of intimate partner violence among women: a systematic review and meta-analysis. Trauma Violence Abuse 2024 Jan;25(1):494-511. [doi: 10.1177/15248380231155529] [Medline: 36825800]
37. Li MY, Wang HX, Yeh ST. Examining mental health outcomes of intimate partner violence among female survivors in Taiwan: a population-based study. Womens Health Rep (New Rochelle) 2025;6(1):384-392. [doi: 10.1089/whr.2024.0203] [Medline: 40308357]
38. Esie P, Osypuk TL, Schuler SR, Bates LM. Intimate partner violence and depression in rural Bangladesh: accounting for violence severity in a high prevalence setting. SSM Popul Health 2019;7:100368. [doi: 10.1016/j.ssmph.2019.100368] [Medline: 30766911]
39. Koly KN, Saba J, Mallick T, Rashid F, Watson J, Neves BB. Exploring the pattern of mental health support-seeking behaviour and related barriers among women experiencing intimate partner violence in urban slums of Bangladesh: perspectives from multiple level stakeholders. PLOS Glob Public Health 2025;5(5):e0004568. [doi: 10.1371/journal.pgph.0004568] [Medline: 40343953]
40. Bhowmik J, Biswas RK. Married women's attitude toward intimate partner violence is influenced by exposure to media: a population-based spatial study in Bangladesh. Int J Environ Res Public Health 2022 Mar 15;19(6):3447. [doi: 10.3390/ijerph19063447] [Medline: 35329134]
41. Antabe R, Antabe G, Sano Y, Pienaah CK. Women's household decision-making autonomy and mental health outcomes in Mozambique. Glob Ment Health (Camb) 2025;12:e40. [doi: 10.1017/gmh.2025.29] [Medline: 40303963]
42. Shooshtari S, Abedi MR, Bahrami M, Samouei R. Empowerment of women and mental health improvement with a preventive approach. J Educ Health Promot 2018;7:31. [doi: 10.4103/jehp.jehp_72_17] [Medline: 29629392]
43. Bangladesh national mental health survey 2018-2019. : National Institute of Mental Health (NIMH); 2019 URL: https://ghdx.healthdata.org/record/bangladesh-national-mental-health-survey-2018-2019 [accessed 2026-09-01]
44. Bangladesh: standard DHS, 2022. : The DHS Program; 2022 URL: https://dhsprogram.com/methodology/survey/survey-display-584.cfm [accessed 2026-08-24]

Abbreviations

aOR: adjusted odds ratio
aPR: adjusted prevalence ratio
BDHS: Bangladesh Demographic and Health Survey
BMRC: Bangladesh Medical Research Council
DALY: disability-adjusted life year
DHS: Demographic and Health Survey
GAD-7: Generalized Anxiety Disorder–7
GBD: Global Burden of Disease
ICF: Inner City Fund
IPV: intimate partner violence
IR: individual recode
LMIC: low- and middle-income country
OR: odds ratio

XSL•FO
RenderX

PHQ-9: Patient Health Questionnaire–9
SDG: Sustainable Development Goal
SWPER: Survey-Based Women's Empowerment
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 14.Oct.2025; peer-reviewed by MA Hossain, M Hirschi, M Schouler-Ocak; revised version received 12.Aug.2026; accepted 19.Aug.2026; published 08.Sep.2026.

Please cite as:
Craig M, Walter H, Ali M, Faruque S, Saha S
Linking Women's Empowerment to Mental Health Symptoms: Population-Based Study
JMIR Public Health Surveill 2026;12:e85839
URL: https://publichealth.jmir.org/2026/1/e85839
doi:10.2196/85839

© Maggie Craig, Hannah Walter, Masum Ali, Shahriar Faruque, Sanjib Saha. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 8.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Analgesic Prescribing Trends of Opioids, Nonsteroidal Anti-Inflammatory Drugs, and Paracetamol in Taiwan (2001 2020): Age-Period-Cohort Cross-Sectional Study

Chun-Ji Lin[1], PhD; Chia-Shuan Chang[2], MS; Chih-Yi Wu[3], MS; Jiun-Yi Wang[1], Prof Dr; I-Ching Lin[1,4,5], MD, PhD; Hung-Yi Chiou[3,6], Prof Dr

[1]Department of Healthcare Administration, College of Medical and Health Science, Asia University, 500, Lioufeng Rd., Wufeng, Taichung, Taiwan
[2]Department of Behavioral and Community Health, School of Public Health, University of Maryland, College Park, Maryland, MD, United States
[3]Institute of Population Health Sciences, National Health Research Institutes, 35, Keyan Road, Zhunan, Miaoli, Taiwan
[4]Department of Family Medicine, Asia University Hospital, Taichung, Taiwan
[5]Department of Kinesiology, Health, and Leisure, Chienkuo Technology University, Changhua, Taiwan
[6]School of Public Health, College of Public Health, Taipei Medical University, Taipei, Taiwan

Corresponding Author:
Hung-Yi Chiou, Prof Dr
Institute of Population Health Sciences, National Health Research Institutes, 35, Keyan Road, Zhunan, Miaoli, Taiwan

Abstract

Background: Despite ongoing global changes in analgesic prescribing practices, long-term trends accounting for age, period, and cohort effects remain poorly understood in Asian populations, where opioid and nonopioid prescribing trajectories differ from those observed in Western nations.

Objective: This study aimed to examine the long-term prescribing trends of opioids, nonsteroidal anti-inflammatory drugs (NSAIDs), and paracetamol in Taiwan by age, calendar period, and birth cohort using an age-period-cohort (APC) framework.

Methods: We conducted a population-based repeated cross-sectional study using Taiwan National Health Insurance claims data (2001 - 2020). The cohort included all insured individuals aged 1 to 85 years, with an annual population size of 22.3 to 23.1 million. Exposure variables were structured for APC modeling, comprising 5-year age groups, four 5-year calendar periods (2001 - 2020), and 5-year birth cohorts (1920 - 1924 to 2015 - 2019).

Results: Opioid prescribing prevalence escalated from 123.66 per 100,000 persons in 2001 to 2005 to 706.93 in 2016 to 2020. APC analysis showed a significant positive net drift for opioids (+10.93% per year, 95% CI 10.36%-11.51%; $P<.001$). The estimated magnitude of the opioid prescribing increase was sensitive to the prescription frequency threshold used. In contrast, paracetamol and NSAIDs showed overall declines (net drifts: −2.44%, 95% CI −2.58% to −2.29%, and −0.82%, 95% CI −0.97% to −0.66%, respectively). Notably, local drifts revealed a significant increase in NSAID prescribing only among adolescents aged 11 to 15 years (+0.80% per year, 95% CI 0.34%-1.27%). Generational analysis suggested higher opioid prescribing risk among younger cohorts (rate ratios 3.06 for the 2015 - 2019 cohort), although the increase relative to the 2000 to 2004 reference cohort was not statistically significant.

Conclusions: Reimbursed analgesic prescribing in Taiwan has diverged across drug classes, with a consistent increase in opioid prescribing across all ages and a shift toward NSAID prescribing in younger populations. These claims-based trends should be interpreted as reimbursed prescribing rather than total analgesic use. These findings may inform targeted pharmacovigilance and policies to ensure appropriate pain management while mitigating potential opioid-related risks.

(JMIR Public Health Surveill 2026;12:e92534) doi:10.2196/92534

KEYWORDS

analgesics; population health surveillance; pharmacovigilance; age-period-cohort analysis; real-world data; opioid stewardship; age differences; adolescent analgesics

Introduction

Analgesic prescribing is a core component of pain management and a main objective for pharmacovigilance and population health surveillance. Prescribing patterns are shaped by multiple interacting forces, including demographic aging, evolving clinical guidelines, reimbursement policies, health care access, and medication safety concerns [1,2]. Consequently, long-term monitoring of reimbursed analgesic prescribing at the population

XSL•FO
RenderX

level is essential for identifying changes in prescribing practice and potential safety priorities [2-5].

Over the past two decades, rising opioid-related harms in North America and Europe have prompted broad stewardship initiatives, including guideline revisions, prescribing restrictions, and enhanced monitoring systems, with several settings reporting subsequent stabilization or declines in opioid prescribing [2,6-8]. However, trends have not been uniform across regions. In fact, opioid use in parts of Asia has continued to increase rather than decline [9-12], highlighting the need for context-specific surveillance. Nonopioid analgesics also have distinct trade-offs between clinical efficacy and safety profiles. For example, while some studies found that nonsteroidal anti-inflammatory drugs (NSAIDs) may offer more potent and sustained analgesia than paracetamol (acetaminophen) [13,14], NSAID use is also associated with risks of gastrointestinal bleeding, renal injury, and cardiovascular events [15,16]. For paracetamol, although it is often favored for its relative safety, it still carries a risk of dose-related hepatotoxicity [17]. Analgesic needs and clinical approaches also vary significantly across the life course, ranging from pediatric populations to older adults with multimorbidity [18,19], and are increasingly influenced by generational shifts in attitudes toward pain management. These multifaceted influences highlight the need for a comprehensive assessment of both opioid and nonopioid prescribing to fully capture the evolving landscape of analgesic exposure across different demographic and generational subgroups.

Despite growing attention to analgesic prescribing trends, recent studies have shown that opioid and nonopioid prescribing patterns vary across health systems and population groups [1,20]. However, many studies still focus on a single analgesic class or report only aggregate secular trends, potentially obscuring important heterogeneity across age groups and birth cohorts [7,10,20-22]. Analyses that simultaneously account for age, calendar period, and birth cohort, such as age-period-cohort (APC) approaches, are also needed to disentangle these interrelated temporal dimensions and identify specific age groups or cohorts with disproportionate changes in prescribing risk [23-25]. The limited number of studies jointly considering these factors may inflate observed associations by conflating biological aging with period-specific and generational effects [26]. To address these gaps, we used nationwide Taiwan National Health Insurance (NHI) claims data from 2001 to 2020 to characterize and compare long-term prescribing patterns of opioids, NSAIDs, and paracetamol. Using an APC model, we estimated age, period, and cohort effects to determine whether prescribing trajectories diverged across analgesic classes and identify specific age groups or birth cohorts that may warrant intensified monitoring and targeted stewardship.

Methods

This study followed the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) reporting guidelines for cross-sectional studies. The completed STROBE checklist is provided in Checklist 1.

Study Design and Data Source

A population-based repeated cross-sectional design was used for the current study, and our data were drawn from 2 national databases in Taiwan. For medical use data, we used the Taiwan National Health Insurance Research Database (NHIRD), which is managed by the National Health Insurance Administration. The NHIRD covers more than 99% of the population and contains comprehensive records of outpatient services, hospitalizations, and prescriptions, providing a robust source for prescribing surveillance. For population data, we drew from the Department of Household Registration, Ministry of the Interior.

Ethical Considerations

This study was approved by the Research Ethics Committee of the National Health Research Institutes (EC1111106-E-R2). Because this study used secondary claims data and aggregated population data, the requirement for informed consent was waived by the ethics committee. The NHIRD contains deidentified claims records, and all analyses were conducted using anonymized or deidentified data to protect participant privacy and confidentiality. No participants were contacted, recruited, or compensated in this study.

Study Population and Observation Period

The analytic sample consisted of NHIRD medical claims data from January 1, 2001, to December 31, 2020, providing a continuous 20-year observation period for examining long-term trends in reimbursed analgesic prescribing. Our study included all insured individuals who were aged between 1 and 85 years. The population at risk was estimated using the year-end population counts from the Ministry of the Interior. Depending on the year, the annual study population ranged from 22,266,932 to 23,078,937 from 2001 to 2020. The NHI policy approved equitable access to health care and limits financial incentives related to the choice of prescription drugs. Our study excluded those with missing or unknown gender and birth date data.

Measures

Our study defined prescribed analgesic drugs as medications with recognized pain-relieving effects based on prior studies, clinical practice, and evidence-based pain management guidelines [20]. Specifically, pain-related medications are commonly categorized as opioids, paracetamol, NSAIDs, and gabapentinoids [20]. Although gabapentinoids are officially classified as antiepileptic agents, they are known to produce analgesic effects in clinical settings; however, their use in Taiwan is limited due to restrictions under the NHI system [27]. Therefore, our study excluded gabapentinoids and focused solely on paracetamol, NSAIDs, and opioids.

Analgesic prescriptions were identified using the Anatomical Therapeutic Chemical classification system. In our analysis, opioids corresponded to code N02A, NSAIDs corresponded to code M01A, and paracetamol corresponded to codes N02BE01 and N02BE51. Repeated analgesic prescribing was defined as receiving at least 3 prescriptions from the same analgesic class within a calendar year [28,29].

XSL•FO
RenderX

We grouped age, period, and cohort into 5-year intervals, resulting in 17 age groups (1 - 5, 6 - 10, ..., and 81 - 85 years), 4 period groups (2001 - 2005, 2006 - 2010, 2011 - 2015, and 2016 - 2020), and 20 birth cohort groups (1920 - 1924, ..., and 2015 - 2019). For each age group within each period, we estimated the prevalence per 100,000 population. In addition, we calculated age-standardized prevalence using the population of 2001 to 2005 as the standard population. To obtain age-standardized prevalence per 100,000 persons, the age-specific prevalence estimates for each period were weighted accordingly.

Statistical Analysis

To evaluate temporal trends and age-, period-, and cohort-specific patterns in repeated reimbursed analgesic prescribing, we organized the 2001 to 2020 observation period into four 5-year calendar periods for APC analysis. We then used a 2-step analytical approach. We first conducted a model selection process based on the analysis of deviance as proposed by Clayton and Schifflers [23,24]. We fitted a series of nested APC specifications and used deviance-based tests to assess the incremental contribution of age, period, and cohort effects. The model fit statistics are summarized in Table S1 in Multimedia Appendix 1. Next, given the perfect linear dependency among age, period, and cohort (age = period – cohort), which prevents unique estimation in standard linear regression models, we used the estimable function approach described by Rosenberg et al [25] to address this identification problem. This approach provides interpretable APC summary measures that can be estimated despite the linear dependency among age, period, and cohort rather than relying on uniquely estimated absolute age, period, and cohort effects [25]. Using this approach, we summarized temporal dynamics in prescribing prevalence through 4 APC measures: net drift, local drifts, period rate ratios (RRs), and cohort RRs. Net drift was interpreted as the overall annual percentage change in age-adjusted prescribing prevalence over calendar time, summarizing the combined linear period and cohort trend within the APC framework. Local drifts represented age-specific annual percentage changes, indicating whether prescribing trends increased or decreased differently across age groups. Period RRs represented relative prescribing prevalence in each calendar period compared with the 2001 to 2005 reference period adjusted for age and nonlinear cohort effects. Cohort RRs represented relative prescribing prevalence in each birth cohort compared with the reference birth cohort adjusted for age and nonlinear period effects. Additional age-, period-, and cohort-specific estimates supporting these APC measures are provided in Table S2 in Multimedia Appendix 1. In this study, the 2001 to 2005 period category and the age group of 1 to 5 years were used as the reference categories. Because birth cohort is determined by calendar period and age in the APC framework, this reference age-period combination corresponded to the 2000 to 2004 birth cohort, which was therefore used as the reference cohort. Statistical significance was evaluated using the Wald test with 2-tailed hypotheses and an α value of .05. To assess the influence of the prescription frequency threshold on the APC estimates, we conducted a sensitivity analysis using a broader definition of at least one prescription from the same analgesic class within a calendar year. The results of this sensitivity analysis can be found in Table S3 in Multimedia Appendix 1. All statistical analyses were conducted using SAS (version 9.4; SAS Institute) and R (version 4.3.2; R Foundation for Statistical Computing).

Results

Characteristics and APC Model

Age-specific 5-year period prevalence rates of analgesic prescribing are shown in Table 1. Throughout the study period, NSAIDs and paracetamol were the predominant choice for analgesic prescribing. In the initial period (2001 - 2005), their prevalence rates were substantially higher than those of opioids, ranging from 13,052.95 to 51,952.64 for NSAIDs and from 20,397.55 to 54,805.04 for paracetamol per 100,000 persons across age groups. However, the prescribing patterns shifted distinctively over time across drug classes. Both NSAIDs and paracetamol analgesics showed downward patterns by 2016 to 2020, with NSAIDs dropping to between 12,287.48 and 42,051.40 per 100,000 persons and paracetamol dropping to between 13,722.56 and 47,693.16 per 100,000 persons, respectively. Opioid use escalated substantially across all age groups, with prevalence rates rising from a range of 0.36 to 807.97 per 100,000 persons in 2001 to 2005 to a range of 1.00 to 9942.95 per 100,000 persons in 2016 to 2020.

XSL•FO
RenderX

Table . Age-specific 5-year period prevalence of repeated nonsteroidal anti-inflammatory drug (NSAID), paracetamol, and opioid prescribing among National Health Insurance beneficiaries aged 1 to 85 years in Taiwan (2001 - 2020).

Age group (year)	Prescription prevalence rate per 100,000 persons			
	2001 - 2005	2006 - 2010	2011 - 2015	2016 - 2020
NSAIDs				
1 - 5	51,952.64	48,169.08	52,680.01	41,617.47
6 - 10	24,217.17	25,313.47	28,616.67	21,563.29
11 - 15	14,105.66	14,628.97	17,203.90	15,678.69
16 - 20	13,466.47	12,239.74	13,680.08	13,212.11
21 - 25	13,052.95	12,334.92	12,702.94	12,287.48
26 - 30	15,251.62	13,755.37	15,065.67	14,043.83
31 - 35	18,252.57	15,821.96	17,168.10	17,368.36
36 - 40	19,258.78	16,890.76	17,537.40	17,831.89
41 - 45	19,496.24	17,582.59	17,601.88	16,523.89
46 - 50	20,949.83	18,885.51	19,137.86	17,145.84
51 - 55	22,782.96	20,738.72	20,862.06	19,165.11
56 - 60	26,713.84	22,128.42	22,261.86	20,369.48
61 - 65	30,744.43	26,843.39	24,131.75	22,147.01
66 - 70	34,350.20	31,092.55	29,237.85	23,902.19
71 - 75	35,476.30	33,335.88	31,911.23	27,271.13
76 - 80	33,138.54	32,397.95	31,769.27	28,033.47
81 - 85	43,177.27	40,757.05	43,307.29	42,051.40
Paracetamol				
1 - 5	54,805.04	47,749.79	46,800.10	34,870.73
6 - 10	34,179.53	31,261.83	30,072.42	19,717.46
11 - 15	25,800.41	23,503.86	22,988.54	16,757.48
16 - 20	23,311.37	19,836.47	19,373.07	14,977.87
21 - 25	20,397.55	18,681.90	17,579.56	13,722.56
26 - 30	21,846.40	18,980.07	19,557.73	14,957.33
31 - 35	23,925.21	19,640.98	20,075.26	17,268.34
36 - 40	23,619.40	19,459.89	18,634.64	16,017.20
41 - 45	22,870.09	19,525.99	17,995.14	14,123.04
46 - 50	23,780.99	20,698.83	19,462.23	14,737.30
51 - 55	25,230.82	22,732.19	21,588.89	16,954.16
56 - 60	29,356.79	24,276.50	23,590.93	18,885.78
61 - 65	32,449.95	28,310.23	25,007.26	20,949.41
66 - 70	35,431.97	31,779.13	29,462.70	22,458.44
71 - 75	35,883.83	33,938.91	32,094.42	26,283.92
76 - 80	33,277.46	33,015.64	32,940.62	28,016.85
81 - 85	44,304.00	43,262.79	48,449.17	47,693.16
Opioids				
1 - 5	0.363551	0.79863	0.925144	1.002668
6 - 10	0.592314	1.280099	2.003292	2.496567
11 - 15	2.590302	4.846901	8.33832	8.340548

XSL•FO
RenderX

Age group (year)	Prescription prevalence rate per 100,000 persons			
	2001 - 2005	2006 - 2010	2011 - 2015	2016 - 2020
16 - 20	7.941273	22.24344	35.80282	40.9002
21 - 25	16.31995	45.84773	73.96435	78.85145
26 - 30	33.65626	77.96135	127.8943	132.7993
31 - 35	57.15419	126.1066	205.9964	235.5716
36 - 40	86.86481	200.6227	314.3916	386.2125
41 - 45	121.001	287.2094	463.0504	558.2152
46 - 50	163.247	395.6203	649.4524	803.9249
51 - 55	215.5913	536.5244	870.8949	1085.473
56 - 60	289.2118	718.9867	1127.085	1375.968
61 - 65	375.1389	984.8498	1520.900	1801.301
66 - 70	479.9140	1359.107	2300.359	2444.945
71 - 75	559.7899	1742.142	3196.521	3638.863
76 - 80	586.9455	2023.311	3939.797	4767.857
81 - 85	807.9685	3143.605	7037.366	9942.947

A similar trend was observed for age-standardized prevalence. Specifically, the rates for NSAIDs showed a slight overall decline from 22,222.19 to 19,200.22 per 100,000 persons, and paracetamol also followed a steady downward trajectory from 27,624.88 to 18,271.32 per 100,000 persons. Conversely, opioids surged dramatically, increasing from 123.66 to 706.93 per 100,000 persons over the study period.

On the basis of goodness-of-fit statistics (Table S1 in Multimedia Appendix 1), the full APC model yielded the lowest deviance values among all 3 drug classes (NSAIDs: 50,744.95; paracetamol: 50,184.21; opioids: 1041.15), indicating a significantly superior fit compared with any reduced specifications. The analysis of deviance revealed that removing any single temporal dimension led to a substantial increase in deviance (P<.001 in all cases), suggesting that the reduced models failed to adequately explain the data. This indicates that the observed analgesic prescribing patterns could not be sufficiently characterized by age, period, or cohort alone or by any 2D reduced specification. Therefore, the full APC model was retained to characterize long-term prescribing trends across all 3 temporal dimensions.

Overall Trends (Net Drifts) and Period Effects

Net drifts are shown in Table 2, and period RRs are shown in Table 3 and Figure 1B. Overall, the annual percentage changes varied significantly by drug class. NSAIDs and paracetamol exhibited decreasing trends, with net drifts estimated at −0.82% per year (95% CI −0.97% to −0.66%) for NSAIDs and −2.44% per year (95% CI −2.58% to −2.29%) for paracetamol. Conversely, opioids showed a substantial increasing trend, with the net drift estimated at 10.93% per year (95% CI 10.36%-11.51%). Over the study period, period RRs for NSAIDs relative to the 2001 to 2005 baseline declined, reaching 0.91 (95% CI 0.88-0.93) in 2006 to 2010, 0.94 (95% CI 0.92-0.96) in 2011 to 2015, and 0.86 (95% CI 0.84-0.88) in 2016 to 2020. Similarly, period RRs for paracetamol compared to baseline showed a progressive decline, falling to 0.87 (95% CI 0.85-0.89) in 2006 to 2010, 0.84 (95% CI 0.82-0.86) in 2011 to 2015, and 0.67 (95% CI 0.65-0.68) in 2016 to 2020. Conversely, period RRs for opioids increased substantially in each subsequent period, rising to 2.50 (95% CI 2.39-2.61) in 2006 to 2010, 4.08 (95% CI 3.85-4.33) in 2011 to 2015, and 4.79 (95% CI 4.42-5.18) in 2016 to 2020. Sensitivity analysis using a broader definition of one or more prescriptions showed that the positive opioid net drift remained evident, although attenuated (+6.76% per year, 95% CI 6.49%-7.03%). This finding supports the directional robustness of the increasing opioid prescribing trend but indicates that the estimated magnitude was threshold sensitive (Table S3 in Multimedia Appendix 1).

XSL•FO
RenderX

Table . Estimated net drifts and age-specific local drifts in repeated nonsteroidal anti-inflammatory drug (NSAID), paracetamol, and opioid prescribing among National Health Insurance beneficiaries aged 1 to 85 years in Taiwan (2001 - 2020).

	NSAIDs, % change per year (95% CI)	Paracetamol, % change per year (95% CI)	Opioids, % change per year (95% CI)
Net drift	−0.82 (−0.97 to −0.66)	−2.44 (−2.58 to –2.29)	10.93 (10.36 to 11.51)
Age-specific trends (year)			
1 - 5	−1.34 (−1.75 to −0.92)	−2.94 (−3.37 to −2.52)	8.02 (−4.82 to 22.61)
6 - 10	0.04 (−0.36 to 0.46)	−2.76 (−3.14 to −2.38)	9.45 (1.80 to 17.69)
11 - 15	0.80 (0.34 to 1.27)	−2.41 (−2.80 to −2.02)	9.64 (5.62 to 13.80)
16 - 20	0.39 (−0.07 to 0.86)	−2.55 (−2.94 to −2.17)	10.18 (8.27 to 12.13)
21 - 25	−0.14 (−0.60 to 0.30)	−2.37 (−2.75 to −1.99)	9.93 (8.66 to 11.21)
26 - 30	−0.18 (−0.61 to 0.24)	−2.02 (−2.39 to −1.64)	10.24 (9.31 to 11.18)
31 - 35	−0.06 (−0.46 to 0.32)	−1.91 (−2.27 to −1.55)	10.59 (9.88 to 11.30)
36 - 40	−0.61 (−0.99 to −0.24)	−2.60 (−2.95 to −2.24)	10.81 (10.24 to 11.37)
41 - 45	−1.03 (−1.40 to −0.65)	−3.02 (−3.38 to −2.66)	11.07 (10.59 to 11.55)
46 - 50	−1.12 (−1.49 to −0.74)	−2.91 (−3.27 to −2.55)	11.36 (10.94 to 11.79)
51 - 55	−1.04 (−1.41 to −0.66)	−2.46 (−2.83 to −2.10)	11.28 (10.89 to 11.68)
56 - 60	−1.64 (−2.04 to −1.23)	−2.64 (−3.03 to −2.25)	10.64 (10.25 to 11.02)
61 - 65	−2.17 (−2.58 to −1.75)	−2.83 (−3.23 to −2.42)	10.47 (10.10 to 10.85)
66 - 70	−2.33 (−2.78 to −1.88)	−2.87 (−3.31 to −2.43)	11.06 (10.70 to 11.43)
71 - 75	−1.59 (−2.09 to −1.09)	−1.85 (−2.34 to −1.36)	12.85 (12.46 to 13.23)
76 - 80	−0.81 (−1.41 to −0.22)	−0.72 (−1.30 to −0.13)	15.17 (14.72 to 15.62)
81 - 85	−0.18 (−1.04 to 0.68)	0.26 (−0.57 to 1.11)	18.12 (17.20 to 19.05)

Table . Estimated period rate ratios (RRs) for repeated nonsteroidal anti-inflammatory drug (NSAID), paracetamol, and opioid prescribing among National Health Insurance beneficiaries aged 1 to 85 years in Taiwan (2001 - 2020). Period RRs are estimated relative to the 2001 to 2005 reference period.

	NSAIDs, RR (95% CI)	Paracetamol, RR (95% CI)	Opioids, RR (95% CI)
2001-2005 (reference)	1.00	1.00	1.00
2006-2010	0.91 (0.88-0.93)	0.87 (0.85-0.89)	2.50 (2.39-2.61)
2011-2015	0.94 (0.92-0.96)	0.84 (0.82-0.86)	4.08 (3.85-4.33)
2016-2020	0.86 (0.84-0.88)	0.67 (0.65-0.68)	4.79 (4.42-5.18)

XSL•FO
RenderX

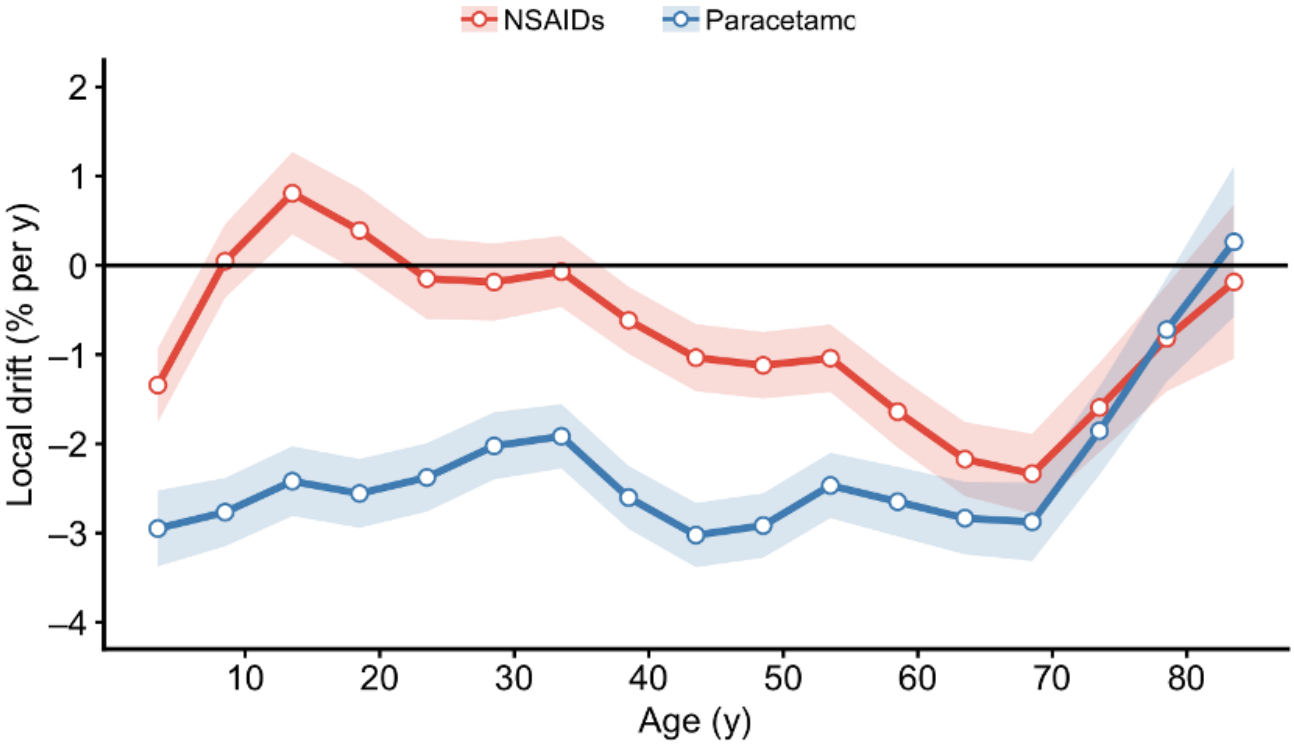
Figure 1. Age-, period-, and cohort-specific patterns in repeated opioid prescribing among National Health Insurance beneficiaries aged 1 to 85 years in Taiwan (2001‑2020). Panel (A) shows age-specific local drifts, expressed as annual percentage changes in opioid prescribing. Panel (B) shows period rate ratios (RRs) relative to the 2001 to 2005 reference period. Panel (C) shows cohort RRs relative to the 2000 to 2004 reference birth cohort. Opioid local drift estimates are displayed separately from those of nonsteroidal anti-inflammatory drugs and paracetamol because their substantially larger magnitude would compress the nonopioid curves and reduce visual readability if plotted on the same scale.

XSL•FO
RenderX

Age-Specific Divergence (Local Drifts)

Different age-specific local drifts were identified for each analgesic class (Figure 1, Figure 2 and Table 2). Among NSAIDs, trends varied across age groups, with local drifts ranging from −2.33% (95% CI −2.78% to −1.88%) to +0.80% (95% CI 0.34%-1.27%) per year. Children and adolescents generally exhibited positive local drifts across age groups. Specifically, estimates were 0.04% (95% CI −0.36% to 0.46%) for the ages of 6 to 10 years, 0.80% (95% CI 0.34%-1.27%) for the ages of 11 to 15 years, and 0.39% (95% CI −0.07% to 0.86%) for the ages of 16 to 20 years. Notably, only the age group of 11 to 15 years showed a statistically significant increase. In contrast, adult and older populations showed declining local drifts (minimum of −2.33%, 95% CI −2.78% to −1.88%), whereas the oldest age group (81 - 85 years) showed a positive but nonsignificant trend.

Figure 2. Age-specific annual percentage changes in repeated nonsteroidal anti-inflammatory drug (NSAID) and paracetamol prescribing among National Health Insurance beneficiaries aged 1 to 85 years in Taiwan (2001 - 2020). The curves represent the estimated local drift (annual percentage change) across age groups for NSAIDs (red) and paracetamol (blue) derived from age-period-cohort models. Shaded ribbons indicate 95% CIs. The solid horizontal black line at 0% marks the threshold of no change; values above this line indicate an increasing trend, whereas values below indicate a decreasing trend.



Paracetamol use showed an overall decline across the life course, with age-specific variation. Specifically, local drifts ranged from −2.94% (95% CI −3.37% to −2.52%) to −2.76% (95% CI −3.14% to −2.38%) among children, from −2.55% (95% CI −2.94% to −2.17%) to −2.41% (95% CI −2.80% to −2.02%) among adolescents, and from −3.02% (95% CI −3.38% to −2.66%) to −1.91% (95% CI −2.27% to −1.55%) among adults aged 31 to 60 years. The magnitude of decline was most pronounced in the age group of 41 to 45 years. Similar to the patterns observed for NSAIDs, the age group of 81 to 85 years exhibited the only positive yet nonsignificant trend.

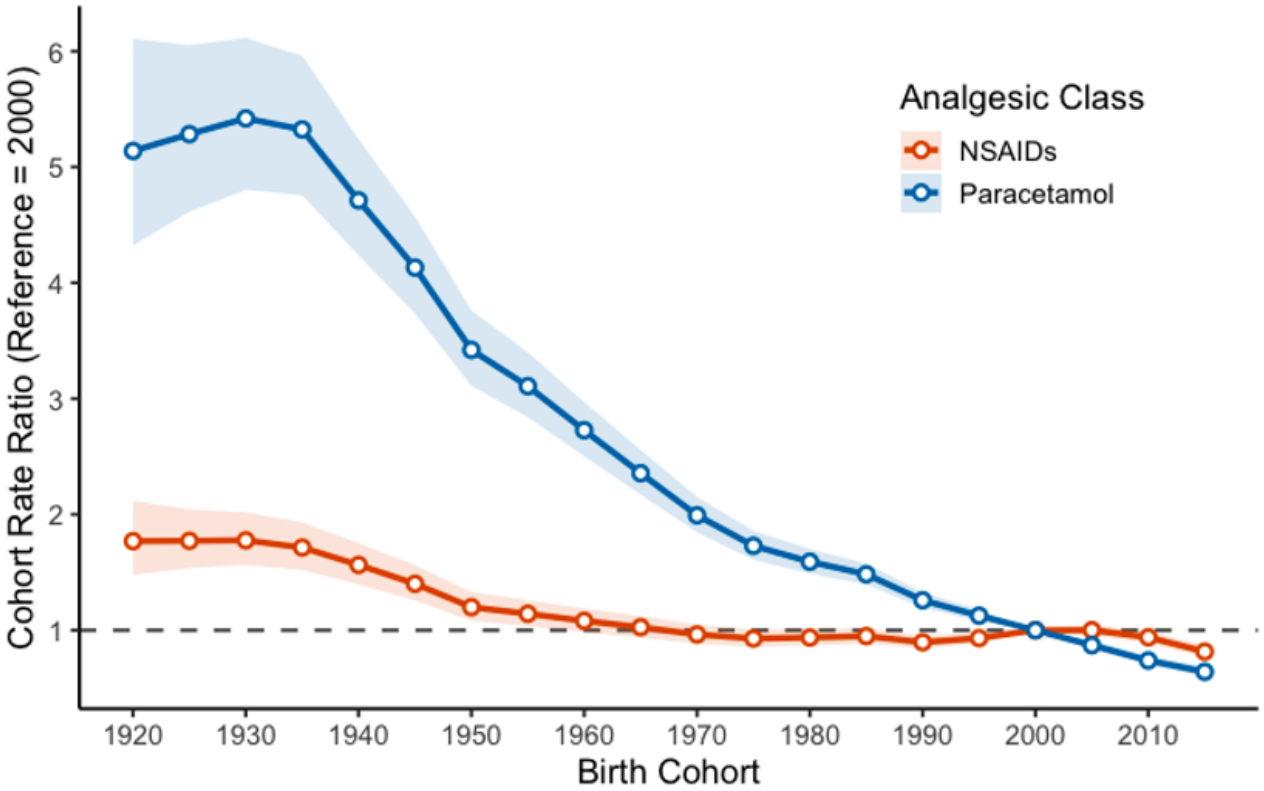
Opioid use showed an overall increasing trend across the age spectrum, with statistically significant effects in nearly all age groups. The only exception was children aged 1 to 5 years, where the increase was not statistically significant. The magnitude of these annual percentage changes generally trended upward with age, reaching 18.12% (95% CI 17.20%-19.05%) in the oldest age group (81 - 85 years).

Generational Shifts

Generational analyses, depicted in Figures 1 and 3 and Table S2 in Multimedia Appendix 1, highlight distinct shifts in prescribing risks over time. For paracetamol, a significant downward trend characterized the generational profile. The RRs dropped substantially from the earliest cohorts (1920-1924), decreasing from 5.13 (95% CI 4.32 - 6.10) to 0.64 (95% CI 0.59 - 0.68) for the most recent cohort (2015-2019); notably, the lack of overlap between these CIs underscores the statistical significance of this decline within the model.

XSL•FO
RenderX

Figure 3. Estimated cohort rate ratios (RRs) for repeated nonsteroidal anti-inflammatory drug (NSAID) and paracetamol prescribing among National Health Insurance beneficiaries aged 1 to 85 years in Taiwan (2001 - 2020). The birth cohort of 2000 to 2004 was selected as the reference group (RR=1.0). The graph presents cohort RRs for repeated reimbursed prescribing in each birth cohort using the 2000 to 2004 birth cohort as the reference group and accounting for age and period components in the age-period-cohort model.



Trends for NSAIDs generally mirrored this downward trajectory but with distinct nuances. While the RRs declined from 1.76 (95% CI 1.47 - 2.11) in the earliest cohorts to 0.81 (95% CI 0.76 - 0.87) for the most recent group, the trend was not entirely monotonic. The 2005 to 2009 NSAID cohort diverged from the decline, presenting a positive risk that, though not statistically significant, marked a notable interruption in the overall downward trend.

Opioids stood apart with a pervasive upward shift. Among cohorts preceding the reference group (born in 2000 - 2004), the relative risk remained significantly lower than the reference yet displayed a distinct upward trajectory, rising from 0.0001 (95% CI 0.0001 - 0.0001) in the 1920 to 1924 cohort to 0.60 (95% CI 0.46 - 0.77) in the 1995 to 1999 cohort. Following the reference cohort, the point estimates were higher among cohorts born after 2005, reaching 3.06 for the 2015 to 2019 cohort; however, these increases were not statistically significant relative to the 2000 to 2004 reference cohort.

Discussion

Principal Findings

To the best of our knowledge, this is the first population-based study using APC analysis to examine the temporal dynamics of analgesic prescribing in Taiwan over a two-decade period. We identified divergent prescribing trends across analgesic classes. While nonopioid analgesics generally exhibited downward trajectories, the decline in NSAID prescribing rate (net drift: −0.82% per year) was notably slower than that of paracetamol (net drift: −2.44% per year), with NSAIDs paradoxically showing increasing trends among children and adolescents. Conversely, opioid prescribing demonstrated a substantial rising trend, with a net drift of 10.93% per year. In addition, our generational analysis revealed a significant overall cohort effect in opioid prescribing, with the cohort RR curve showing an upward pattern across successive birth cohorts. Furthermore, a distinct generational shift was observed in the prescribing of nonopioid analgesics. Among younger generations, prescribing patterns have shifted toward an NSAID-predominant approach; specifically, the likelihood of being prescribed paracetamol has diminished substantially, whereas NSAID prescribing has remained relatively stable or increased.

In contrast to international trends [8,9,21,30], our study identified a continuous and dramatic surge in the likelihood of opioid prescribing in Taiwan. While research using APC to examine opioid prescribing trends across the lifespan remains limited, a study in the United States analyzing individuals aged 16 years and older between 1999 and 2018 identified period effects characterized by an initial increase followed by a subsequent decline [8]. Notably, while the US data indicated a decline in opioid use following its peak around 2015 [8], the likelihood of opioid prescribing in our study continued to escalate without abatement.

Our results align with those of previous Taiwan studies that have identified increasing trends. Specifically, research covering 2002 to 2007 reported an initial rise [31], and a subsequent

XSL•FO
RenderX

study analyzing the 2008 to 2018 period observed a linear increase of less than 20% [10]. However, our results highlighted the importance of applying APC modeling to control for age and cohort effects as the conventional estimates without APC may significantly obscure the true severity of the surge. After controlling for age and cohort effects, the period effect intensified dramatically, with the likelihood of opioid prescribing increasing nearly 5-fold in the most recent period (period RR 4.79) compared to 2001 to 2005.

While prior research indicates that most opioid prescriptions in Taiwan were administered to patients with cancer [10], data from the Taiwan Cancer Registry (2001 - 2020) show that the age-standardized incidence rate of invasive cancer increased by only 21.7% over the study period [32]. This discrepancy suggests that increased prescribing is not solely driven by rising cancer prevalence but likely reflects expanded pain treatment services and evolving standards for noncancer pain management. In addition, although the NHI reimbursement regulations regarding patient eligibility remained largely unchanged between 2001 and 2020, the formulary expanded with the inclusion of 75 new opioid items between 2001 and 2020 [27,33]. That is, the observed increase may not be due to relaxed eligibility criteria but may relate to other factors, such as broader medication options, the expansion of pain management services, high health care accessibility, and a growing emphasis on patient autonomy [10,34,35]. Therefore, continuous surveillance is necessary to ensure reasonable and appropriate opioid prescribing practices [36,37].

Consistent with studies conducted in France and the Netherlands [1,6], our study identified a general decline in nonopioid analgesic prescribing. Notably, this decline differed by drug class, with paracetamol showing a substantially steeper decrease (net drift: −2.44% per year) than NSAIDs (−0.82% per year). Theoretically, these declining trends could be attributed to a reduction in the incidence of diseases requiring analgesia [17,38]. However, this possibility appears unlikely as epidemiological data indicate that the incidence of upper respiratory infections, otitis media, migraine, and tension-type headache remained stable whereas osteoarthritis and low back pain showed an increasing trend in Taiwan throughout the observation period [39]. Another possible explanation lies in the impact of reimbursement policies. According to Article 51 of the National Health Insurance Act, over-the-counter medications are not included in insurance coverage [40]. Following this mandate, the National Health Insurance Administration implemented a policy of gradually phasing out reimbursement for these items since 2005 [41]. Consequently, the decline in nonopioid analgesics likely reflects a shift in acquisition channels.

After controlling for period and cohort effects, our findings indicated that NSAID prescribing showed an increasing trend during adolescence that was statistically significant, specifically in the age group of 11 to 15 years. Specifically, there is mounting evidence suggesting that sleep deficiency is associated with hyperalgesia changes and altered pain processing [42,43]. Given that Taiwanese adolescents have experienced rising rates of sleep problems and inadequate sleep in recent years [44], such chronic sleep deficit among Taiwanese adolescents may facilitate central sensitization, thereby lowering pain thresholds and increasing the demand for analgesia. In addition, obesity and musculoskeletal strain are relatively common among Taiwanese adolescents [45-47], and these conditions may further exacerbate pain sensitivity as obesity-related systemic inflammation can amplify nociceptive responses and increase pain perception [48]. Altogether, the growing use of NSAIDs among younger individuals may reflect their effectiveness in managing inflammation-related pain in the context of increased pain susceptibility.

Furthermore, psychosocial dynamics play a critical role. For adolescents, education demands are high, and falling behind due to absence is a significant concern [49]. Under such pressure, guardians often prioritize antipyretics with superior and prolonged efficacy [50]. This preference aligns with clinical evidence favoring NSAIDs over paracetamol for pediatric pain management [13,14]. Thus, the rising trend in NSAID prescribing may reflect a collective pursuit by adolescents, caregivers, and clinicians for stronger and longer-acting interventions to minimize the disruption of disease on daily routines within a fast-paced, highly competitive society. The proposed mechanisms remain exploratory, and this adolescent-specific NSAID prescribing pattern should be interpreted cautiously given that only the age group of 11 to 15 years reached statistical significance. Nevertheless, this observed rise in NSAID prescribing among adolescents may represent a meaningful deviation warranting continued surveillance.

A central policy implication of applying an APC framework to repeated reimbursed analgesic prescribing is that it helps distinguish broad temporal changes from prescribing patterns concentrated in specific age groups or birth cohorts. In this study, the sustained increase in opioid prescribing, the pronounced increases among older adults, and the adolescent-specific increase in NSAID prescribing indicate where continued surveillance may be particularly relevant within Taiwan's health system. These findings may therefore provide a more targeted basis for future studies using longitudinal NHIRD data to examine clinically meaningful medication-related outcomes, including adverse drug events, opioid-related hospitalizations, and dependence.

Limitations

This study has several limitations. NHIRD claims lack clinical detail (eg, pain severity, duration, and indications) and do not support dose-based metrics. Therefore, our outcomes should be interpreted as repeated reimbursed prescribing events rather than treatment intensity or indication-specific analgesic use. Because this study analyzed population-level analgesic prescribing trends, it could not determine the relative contributions of cancer-related and non–cancer-related pain management to opioid prescribing trends. In addition, because the primary definition required at least 3 prescriptions from the same analgesic class within a calendar year, short-term or episodic analgesic prescribing, including acute postsurgical opioid prescriptions, may have been underestimated. APC analyses are subject to identification constraints, and 5-year groupings may smooth short-term changes and yield less stable boundary cohort estimates; accordingly, we emphasize long-term

XSL•FO
RenderX

patterns. In addition, our APC models were fitted to aggregated age-by-period data; therefore, results should be interpreted at the population level and do not support individual-level inference. Given the ecological nature of this population-level APC analysis, our study could not directly confirm the underlying causes of the observed analgesic prescribing trends. Although we discussed several potential mechanisms and contextual factors, these explanations may not be exhaustive. Moreover, although younger age groups showed generally positive NSAID local drifts, only the age group of 11 to 15 years reached statistical significance, whereas the adjacent age groups of 6 to 10 years and 16 to 20 years did not. Therefore, population-level interpretations of adolescent-specific NSAID prescribing patterns should be made with particular caution. Finally, analyses were restricted to 3 reimbursed analgesic classes and did not capture over-the-counter or out-of-pocket purchases, so declines in reimbursed nonopioid prescribing may partly reflect channel shifts. As this was a population-based repeated cross-sectional study in a single-payer system, causal inference and generalizability to other health systems are limited.

Conclusions

In this population-based APC analysis of reimbursed prescribing in Taiwan (2001 - 2020), analgesic prescribing trends diverged markedly across drug classes: opioid prescribing increased substantially, whereas reimbursed prescribing of paracetamol and NSAIDs declined overall, with a steeper reduction in paracetamol. However, the magnitude of the estimated opioid prescribing increase was sensitive to the prescription frequency threshold used. Age- and cohort-specific estimates revealed pronounced heterogeneity, including rising NSAID prescribing in adolescents and a sustained increase in opioid prescribing in older adults, with evidence of higher opioid prescribing risk in more recent birth cohorts. These findings support continued surveillance and stewardship of opioid prescribing and highlight the need for future studies incorporating clinical indications, dose intensity, and nonreimbursed acquisition channels to clarify underlying drivers.

Acknowledgments

The authors declare the use of generative AI (GenAI) in the research and writing process. According to the Generative AI Delegation Taxonomy (2025), the following tasks were delegated to GenAI tools under full human supervision: proofreading and editing, adapting and adjusting emotional tone, and reformatting. The GenAI tools used were ChatGPT (OpenAI) and Gemini (Google). Responsibility for the final manuscript lies entirely with the authors. GenAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Funding

The authors declared no financial support was received for this work.

Data Availability

The datasets analyzed during the current study are not publicly available due to legal restrictions under Taiwan's National Health Insurance Act and Personal Data Protection Act, which prohibit the public sharing of individual-level medical claims data to protect patient privacy. However, aggregated statistical data generated from these datasets are included in this published article and its supplementary files. Interested researchers can apply for access to the original databases through formal requests directed to the Health and Welfare Data Science Center, Ministry of Health and Welfare, Taiwan.

Authors' Contributions

Conceptualization: CJL (lead), CSC (supporting)
Data curation: CYW
Formal analysis: CJL
Investigation: CJL (lead), CYW (supporting)
Methodology: CJL
Project administration: CJL (lead), CYW (equal)
Resources: HYC (lead), CJL (equal)
Supervision: HYC (lead), CJL (equal)
Validation: CJL
Visualization: CJL (lead), CYW (supporting)
Writing—original draft: CJL (lead), CSC (equal)
Writing—review and editing: CJL (lead), CSC (supporting), CYW (supporting), JYW (supporting), ICL (supporting), HYC (equal)

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Goodness-of-fit statistics of age-period-cohort models, estimated cohort rate ratios for analgesic utilization, and sensitivity analysis results.
[DOCX File, 466 KB - publichealth_v12i1e92534_app1.docx]

Checklist 1
STROBE checklist.
[PDF File, 154 KB - publichealth_v12i1e92534_app2.pdf]

References

1. Veldkamp D, Pooters N, Schers HJ, Akkermans R, Olde Hartman TC, Uijen AA. Changes in analgesic prescriptions in Dutch general practice. Scand J Prim Health Care 2024 Dec;42(4):714-722. [doi: 10.1080/02813432.2024.2387423] [Medline: 39161183]
2. Dowell D, Ragan KR, Jones CM, Baldwin GT, Chou R. CDC Clinical Practice Guideline for Prescribing Opioids for Pain - United States, 2022. MMWR Recomm Rep 2022 Nov 4;71(3):1-95. [doi: 10.15585/mmwr.rr7103a1] [Medline: 36327391]
3. Prescott LF. Paracetamol: past, present, and future. Am J Ther 2000 Mar;7(2):143-147. [Medline: 11319582]
4. Laine L. Gastrointestinal effects of NSAIDs and coxibs. J Pain Symptom Manage 2003 Feb;25(2 Suppl):S32-S40. [doi: 10.1016/s0885-3924(02)00629-2] [Medline: 12604155]
5. Anekar AA, Hendrix JM, Cascella M. WHO analgesic ladder. In: StatPearls: StatPearls Publishing; 2023.
6. Daveluy A, Micallef J, Sanchez-Pena P, et al. Ten-year trend of opioid and nonopioid analgesic use in the French adult population. Br J Clin Pharmacol 2021 Feb;87(2):555-564. [doi: 10.1111/bcp.14415] [Medline: 32496599]
7. Gaertner K, Wildbolz S, Speidel V, Exadaktylos AK, Hautz WE, Müller M. Prevalence and practice of opioid prescription at a Swiss emergency department: 2013-2017. Swiss Med Wkly 2020 Apr 6;150:w20202. [doi: 10.4414/smw.2020.20202] [Medline: 32249923]
8. Shiue KY, Austin AE, Naumann RB, Aiello AE, Marshall SW, Golightly YM. Age, period and cohort-related trends in prescription opioid use in the USA, 1999-2018. J Epidemiol Community Health 2023 Nov;77(11):714-720. [doi: 10.1136/jech-2023-220701] [Medline: 37507219]
9. Cho NR, Chang YJ, Lee D, Kim JR, Ko DS, Choi JJ. Trends in opioid prescribing practices in South Korea, 2009-2019: are we safe from an opioid epidemic? PLoS One 2021;16(5):e0250972. [doi: 10.1371/journal.pone.0250972] [Medline: 33979378]
10. Wang JJ, Chu YR, Teng SF, Chu CC, Ho CH, Chu LL. Prevalence of opioid prescriptions in Taiwan (2008-2018). J Chin Med Assoc 2022 May 1;85(5):603-609. [doi: 10.1097/JCMA.0000000000000720] [Medline: 35353736]
11. Higashi T, Yoshimoto T, Matoba M. Prevalence of analgesic prescriptions among patients with cancer in Japan: an analysis of health insurance claims data. Glob J Health Sci 2012 Oct 12;4(6):197-203. [doi: 10.5539/gjhs.v4n6p197] [Medline: 23121757]
12. Noh Y, Heo KN, Yu YM, Lee JY, Ah YM. Trends in potentially inappropriate opioid prescribing and associated risk factors among Korean noncancer patients prescribed non-injectable opioid analgesics. Ther Adv Drug Saf 2022;13:20420986221091001. [doi: 10.1177/20420986221091001] [Medline: 35509350]
13. de'Angelis GL, Vincenzi F, Fornaroli F, Buonvicino D, Chiarugi A. New perspectives for optimizing fever and pain management in pediatrics: evidence supporting therapeutic awareness in clinical practice. Ital J Pediatr 2025 Aug 15;51(1):255. [doi: 10.1186/s13052-025-02107-3] [Medline: 40817220]
14. van den Anker JN. Optimising the management of fever and pain in children. Int J Clin Pract Suppl 2013 Jan(178):26-32. [doi: 10.1111/ijcp.12056] [Medline: 23163545]
15. Jones R, Rubin G, Berenbaum F, Scheiman J. Gastrointestinal and cardiovascular risks of nonsteroidal anti-inflammatory drugs. Am J Med 2008 Jun;121(6):464-474. [doi: 10.1016/j.amjmed.2008.01.045] [Medline: 18501223]
16. Davies NM, Reynolds JK, Undeberg MR, Gates BJ, Ohgami Y, Vega-Villa KR. Minimizing risks of NSAIDs: cardiovascular, gastrointestinal and renal. Expert Rev Neurother 2006 Nov;6(11):1643-1655. [doi: 10.1586/14737175.6.11.1643] [Medline: 17144779]
17. Freo U, Ruocco C, Valerio A, Scagnol I, Nisoli E. Paracetamol: a review of guideline recommendations. J Clin Med 2021 Jul 31;10(15):3420. [doi: 10.3390/jcm10153420] [Medline: 34362203]
18. Walco GA, Krane EJ, Schmader KE, Weiner DK. Applying a lifespan developmental perspective to chronic pain: pediatrics to geriatrics. J Pain 2016 Sep;17(9 Suppl):T108-T117. [doi: 10.1016/j.jpain.2015.11.003] [Medline: 27586828]
19. Baumbauer KM, Young EE, Starkweather AR, Guite JW, Russell BS, Manworren RC. Managing chronic pain in special populations with emphasis on pediatric, geriatric, and drug abuser populations. Med Clin North Am 2016 Jan;100(1):183-197. [doi: 10.1016/j.mcna.2015.08.013] [Medline: 26614727]
20. Stangeland H, Handal M, Skurtveit SO, et al. Killing pain?: A population-based registry study of the use of prescription analgesics, anxiolytics, and hypnotics among all children, adolescents and young adults in Norway from 2004 to 2019. Eur Child Adolesc Psychiatry 2023 Nov;32(11):2259-2270. [doi: 10.1007/s00787-022-02066-8] [Medline: 36030342]

XSL•FO
RenderX

21. Renny MH, Yin HS, Jent V, Hadland SE, Cerdá M. Temporal trends in opioid prescribing practices in children, adolescents, and younger adults in the US from 2006 to 2018. JAMA Pediatr 2021 Oct 1;175(10):1043-1052. [doi: 10.1001/jamapediatrics.2021.1832] [Medline: 34180978]
22. Dale O, Borchgrevink PC, Fredheim OM, Mahic M, Romundstad P, Skurtveit S. Prevalence of use of non-prescription analgesics in the Norwegian HUNT3 population: impact of gender, age, exercise and prescription of opioids. BMC Public Health 2015 May 2;15:461. [doi: 10.1186/s12889-015-1774-6] [Medline: 25934132]
23. Clayton D, Schifflers E. Models for temporal variation in cancer rates. II: age-period-cohort models. Stat Med 1987 Jun;6(4):469-481. [doi: 10.1002/sim.4780060406] [Medline: 3629048]
24. Clayton D, Schifflers E. Models for temporal variation in cancer rates. I: age-period and age-cohort models. Stat Med 1987 Jun;6(4):449-467. [doi: 10.1002/sim.4780060405] [Medline: 3629047]
25. Rosenberg PS, Check DP, Anderson WF. A web tool for age-period-cohort analysis of cancer incidence and mortality rates. Cancer Epidemiol Biomarkers Prev 2014 Nov;23(11):2296-2302. [doi: 10.1158/1055-9965.EPI-14-0300] [Medline: 25146089]
26. Holford TR. Age–period–cohort analysis. In: Encyclopedia of Biostatistics: John Wiley & Sons; 2005. [doi: 10.1002/0470011815]
27. Latest version of drug reimbursement regulations [Webpage in Chinese]. National Health Insurance Administration, Ministry of Health and Welfare, Taiwan. URL: https://www.nhi.gov.tw/ch/cp-13108-67ddf-2508-1.html [accessed 2026-08-18]
28. Scott IC, Bailey J, White CR, Mallen CD, Muller S. Analgesic prescribing in patients with inflammatory arthritis in England: an observational study using electronic healthcare record data. Rheumatology (Oxford) 2022 Aug 3;61(8):3201-3211. [doi: 10.1093/rheumatology/keab870] [Medline: 34849617]
29. Lindeberg FC, Bell M, Larsson E, Hallqvist L. New prolonged opioid consumption after major surgery in Sweden: a population-based retrospective cohort study. BMJ Open 2023 Apr 26;13(4):e071135. [doi: 10.1136/bmjopen-2022-071135] [Medline: 37185647]
30. Rosner B, Neicun J, Yang JC, Roman-Urrestarazu A. Opioid prescription patterns in Germany and the global opioid epidemic: systematic review of available evidence. PLoS One 2019;14(8):e0221153. [doi: 10.1371/journal.pone.0221153] [Medline: 31461466]
31. Pan HH, Ho ST, Lu CC, Wang JO, Lin TC, Wang KY. Trends in the consumption of opioid analgesics in Taiwan from 2002 to 2007: a population-based study. J Pain Symptom Manage 2013 Feb;45(2):272-278. [doi: 10.1016/j.jpainsymman.2012.02.014] [Medline: 22889859]
32. Health Promotion Administration. URL: https://www.hpa.gov.tw/EngPages/Index.aspx [accessed 2026-08-21]
33. National Health Insurance Administration, Ministry of Health and Welfare, Taiwan. URL: https://www.nhi.gov.tw/en/mp-2.html [accessed 2026-08-21]
34. Yang PC, Shih MJ, Liu YA, et al. Web search trends of implementing the patient autonomy act in Taiwan. Healthcare (Basel) 2020 Sep 21;8(3):353. [doi: 10.3390/healthcare8030353] [Medline: 32967273]
35. Cheng SH, Chiang TL. The effect of universal health insurance on health care utilization in Taiwan. Results from a natural experiment. JAMA 1997 Jul 9;278(2):89-93. [doi: 10.1001/jama.278.2.89] [Medline: 9214512]
36. Shreffler J, Shaw I, Berrones A, Huecker M. Prescription history before opioid overdose death: PDMP data and responsible prescribing. J Public Health Manag Pract 2021;27(4):385-392. [doi: 10.1097/PHH.0000000000001210] [Medline: 32810066]
37. Opioid overdose. World Health Organization. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/opioid-overdose [accessed 2026-08-18]
38. Szamosi S. Recommendation for the substantial choice between various nonsteroidal antiinflammatory drugs based on their safety profile [Article in Hungarian]. Orv Hetil 2018 Nov;159(44):1783-1788. [doi: 10.1556/650.2018.31294] [Medline: 30392407]
39. GBD compare. Institute for Health Metrics and Evaluation. URL: https://vizhub.healthdata.org/gbd-compare [accessed 2026-08-18]
40. National Health Insurance Act. Laws & Regulations Database of The Republic of China (Taiwan). URL: https://law.moj.gov.tw/ENG/LawClass/LawAll.aspx?pcode=L0060001 [accessed 2026-08-21]
41. Hsu JC, Lu CY. The evolution of Taiwan's National Health Insurance drug reimbursement scheme. Daru 2015 Feb 10;23(1):15. [doi: 10.1186/s40199-014-0080-7] [Medline: 25889754]
42. Oh A, Koehler A, Yonker M, Troester M. Sleep disorders and chronic pain syndromes in the pediatric population. Semin Pediatr Neurol 2023 Dec;48:101085. [doi: 10.1016/j.spen.2023.101085] [Medline: 38065632]
43. Haack M, Simpson N, Sethna N, Kaur S, Mullington J. Sleep deficiency and chronic pain: potential underlying mechanisms and clinical implications. Neuropsychopharmacology 2020 Jan;45(1):205-216. [doi: 10.1038/s41386-019-0439-z] [Medline: 31207606]
44. Yang YT, Kaplan KA, Zeitzer JM. A comparison of sleep, depressive symptoms, and parental perceptions between U.S. and Taiwan adolescents with self-reported sleep problems. Sleep Adv 2020;1(1):zpaa004. [doi: 10.1093/sleepadvances/zpaa004] [Medline: 33345187]

XSL•FO
RenderX

45. Chang YC, Hsieh WH, Huang SF, et al. Sex-specific temporal trends in overweight and obese among schoolchildren from 2009 to 2018: an age period cohort analysis. Front Pediatr 2021;9:615483. [doi: 10.3389/fped.2021.615483] [Medline: 34055679]
46. Yang CY, Yeh YC, Cheng MF, Lin MC. The incidence of school-related injuries among adolescents in Kaohsiung, Taiwan. Am J Prev Med 1998 Oct;15(3):172-177. [doi: 10.1016/s0749-3797(98)00046-4] [Medline: 9791634]
47. GBD 2021 Adolescent BMI Collaborators. Global, regional, and national prevalence of child and adolescent overweight and obesity, 1990–2021, with forecasts to 2050: a forecasting study for the Global Burden of Disease Study 2021. Lancet 2025 Mar;405(10481):785-812. [doi: 10.1016/S0140-6736(25)00397-6] [Medline: 40049185]
48. Gold MS, Flake NM. Inflammation-mediated hyperexcitability of sensory neurons. Neurosignals 2005;14(4):147-157. [doi: 10.1159/000087653] [Medline: 16215297]
49. Fegran L, Johannessen B, Ludvigsen MS, et al. Experiences of a non-clinical set of adolescents and young adults living with persistent pain: a qualitative metasynthesis. BMJ Open 2021 Apr 19;11(4):e043776. [doi: 10.1136/bmjopen-2020-043776] [Medline: 33875442]
50. Chang LC, Liu CC, Huang MC. Parental knowledge, concerns, and management of childhood fever in Taiwan. J Nurs Res 2013 Dec;21(4):252-260. [doi: 10.1097/jnr.0000000000000007] [Medline: 24241274]

Abbreviations

APC: age-period-cohort
NHI: National Health Insurance
NHIRD: National Health Insurance Research Database
NSAID: nonsteroidal anti-inflammatory drug
RR: rate ratio
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Mavragani, T Sanchez; submitted 06.Feb.2026; peer-reviewed by ECY Su, YC Hsieh; revised version received 30.Jun.2026; accepted 03.Jul.2026; published 01.Sep.2026.

Please cite as:
Lin CJ, Chang CS, Wu CY, Wang JY, Lin IC, Chiou HY
Analgesic Prescribing Trends of Opioids, Nonsteroidal Anti-Inflammatory Drugs, and Paracetamol in Taiwan (2001‐2020): Age-Period-Cohort Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e92534
URL: https://publichealth.jmir.org/2026/1/e92534
doi:10.2196/92534

© Chun-Ji Lin, Chia-Shuan Chang, Chih-Yi Wu, Jiun-Yi Wang, I-Ching Lin, Hung-Yi Chiou. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 1.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Experts' Opinions on the Sustainable Use of Digital Health Tools for Effective Future Pandemic Preparedness and Response: Questionnaire Study

Kabelo Leonard Mauco[1], PhD; John H Holmes[2], PhD; Anthony Luberti[3,4], MD; Badisa Mosesane[5], BSc

[1]Center for Global Health, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

[2]Department of Biostatistics, Epidemiology, and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

[3]Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

[4]Department of Biomedical and Health Informatics, Children's Hospital of Philadelphia, Philadelphia, PA, United States

[5]Global Health Informatics, Department of Biomedical and Health Informatics, Children's Hospital of Philadelphia, Philadelphia, PA, United States

Corresponding Author:
Kabelo Leonard Mauco, PhD
Center for Global Health
Perelman School of Medicine
University of Pennsylvania
240 John Morgan Building
3620 Hamilton Walk
Philadelphia, PA, 19104-6055
United States
Phone: 1 6465004518
Email: meleomau@gmail.com

Abstract

Background: The COVID-19 pandemic demonstrated the potential role of digital health tools in enhancing pandemic preparedness and response. These tools became essential, supporting not only health care delivery but also decision-making, communication, case identification, contact tracing, surveillance, vaccination rollout, and intervention evaluation. The interest in applying digital health tools to pandemic preparedness and response motivated conversations about digital epidemiology—a field of study that aims to provide insight into health and disease determinants by leveraging diverse digital data sources. In a globalized world, effective preparedness and response to pandemics require coordinated global action.

Objective: This study investigates experts' opinions on strategies for improving global health security through the effective use of digital epidemiology, considering the current landscape of digital determinants of health.

Methods: Epidemiologists, public health specialists, data scientists, and professionals with expertise in various components of digital health were recruited through convenience and snowball sampling methods. Their opinions were elicited using an electronic questionnaire developed by the authors in Research Electronic Data Capture (REDCap; Vanderbilt University). To ensure a global perspective, participants were recruited from Africa, North America, Oceania, and Europe. Thematic analysis and the strengths, weaknesses, opportunities, and threats (SWOT) analysis framework were used to analyze participants' responses.

Results: Most participants were familiar with the concept of digital epidemiology and expressed positive sentiments about its potential in strengthening global health security. Privacy and security, along with ethical and legal considerations, were ranked by most experts as high priority areas that decision-makers and implementers must consider to ensure sustainable integration of digital epidemiology tools in future pandemic preparedness and response. A SWOT analysis of participants' views on the promise of digital epidemiology revealed fewer strengths and more weaknesses compared to other components of the analysis framework.

Conclusions: This study highlights the growing recognition of digital epidemiology as a critical tool for enhancing global health security, particularly using nontraditional data sources and emerging technologies, including artificial intelligence. The study affirms the need for a globally coordinated approach to governance, regulation, and investment in digital health infrastructure to ensure the responsible and effective application of digital innovations in epidemiological practice.

(JMIR Public Health Surveill 2026;12:e84164) doi:10.2196/84164

XSL•FO
RenderX

KEYWORDS

digital health; eHealth; public health informatics; digital epidemiology; pandemic preparedness; pandemic response; digital health readiness; global health; COVID-19; artificial intelligence; AI

Introduction

The COVID-19 pandemic underscored the critical role of digital health tools in enhancing pandemic preparedness and response. It is estimated that the COVID-19 pandemic caused more than 2 million deaths worldwide [1]. A report of the Independent Panel for Pandemic Preparedness and Response concluded that the alert system was too slow, the World Health Organization (WHO) was underresourced, and global political leadership was absent [2]. The study by de Souza et al [3] identified 5 critical areas associated with the global failure in facing the pandemic: deficiency in the global alert system and the fragility of the international health regulations, problems of the international response to the pandemic related to global health governance, the dispersed global adoption of the elimination strategy (zero-COVID), fragile control of the disease with a narrow reading of the associated problems, and finally, global setbacks in achieving the sustainable development goals. The need for coordinated global efforts and effective early detection and timely surveillance for a robust pandemic and epidemic early warning and preparedness has been widely discussed amid the COVID-19 pandemic [4].

Digital health tools became a necessity and their popularity increased during the pandemic not only for the provision of health care services but also to support decision-making, communication and information, case identification, contact tracing, surveillance and monitoring, and rolling out of vaccination programs, as well as evaluation of interventions [5-7]. Digital health technologies illustrated the potential to support diverse needs and approaches for early warning and forecasting of epidemics [8]. Nevertheless, in their report, Al Knawy et al [9] acknowledge that COVID-19 highlighted widespread chronic underinvestment in digital health that hampered public health responses to the pandemic. The often ad hoc adoption and use of digital health during the pandemic further illuminated its potential to transform future pandemic preparedness and response efforts [10].

The discussion on applying digital health tools toward future pandemic preparedness and response has motivated frequent conversations around digital epidemiology. Digital epidemiology has been defined as a field of study that aims to provide insight into health and disease determinants in human populations by building on diverse digital data sources [11]. Furthermore, the study by Tarkoma et al [11] posits that digital epidemiology is envisaged as having the potential to contribute toward global health security through inputs, such as syndromic surveillance, public health surveillance, and early pandemic detection.

In the era of globalization, efficient preparedness and response to future pandemics requires a united global action [12,13]. However, realizing the full potential of digital epidemiology as part of such global efforts becomes elusive due to persistent geopolitical tensions as well as the global digital divide and digital access gap [14]. This study aims to investigate experts' opinions on possible means of improving global health security through the effective use of digital epidemiology within the context of the current global landscape of digital determinants of health.

Methods

Sampling and Recruitment of Study Participants

Epidemiologists, public health specialists, data scientists, and professionals with expertise in the various components of digital health were recruited using convenience sampling. To gather a global perspective, potential study participants were recruited from Africa, North America, Oceania, and Europe.

For recruitment of potential participants from Africa, an initial email was sent to the general membership of the pan-African health informatics association called Health Informatics in Africa. For the United States, initial recruitment of study participants involved sending an email to members of the American Medical Informatics Association. Snowball sampling method was also used. The initial email to potential study participants described the study objectives and the role of a potential study participant. Study data were collected and managed using the Research Electronic Data Capture (REDCap; Vanderbilt University) tool hosted at the Children's Hospital of Philadelphia [15]. The email package also included a link to an electronic questionnaire developed using the REDCap platform. The electronic questionnaire included a consent form that a potential study participant had to agree to before proceeding to answer the questionnaire.

Recruitment of potential study participants was also done at the Medical Informatics Europe 2023 conference held in Göteborg, Sweden. Paper flyers about the study were placed at convenient places throughout the conference venue. In addition to containing details about the study objectives and participant's role, the paper flyers also included a QR code with a link to the electronic questionnaire developed using REDCap.

Recruitment of study participants continued until data saturation was reached, defined for this study as the point at which additional data no longer generated new insights. As an electronic questionnaire was used for this study, the authors aimed to comply with the Checklist for Reporting Results of Internet E-Surveys [16]. The consolidated criteria for reporting qualitative research were followed for reporting of this study. The completed checklist is provided in Multimedia Appendix 1.

Study Questionnaire Design and Development

The initial draft of the study questionnaire was designed based on a review of relevant literature. Following an iterative process, the authors developed a study questionnaire aimed at capturing participants' opinions on effective approaches to improving global health security using digital epidemiology within the

XSL•FO
RenderX

context of the current global landscape of digital determinants of health.

A panel consisting of 5 staff members from the Department of Biomedical and Health Informatics at the Children's Hospital of Philadelphia and the Perelman School of Medicine at the University of Pennsylvania was consulted to ensure content validity of the study questionnaire. The panel received the initial draft questionnaire via email and could make suggestions for improvement where needed. Their suggestions were discussed and agreed upon among the authors and final updates were made on the questionnaire. The structured questionnaire included both closed-ended (Likert-scale responses) and open-ended questions, allowing for analysis of trends as well as insights into experts' recommendations. This approach facilitated a comprehensive evaluation of participants' responses.

Data Analysis

Thematic analysis was used to analyze the qualitative responses from the questionnaire, following an iterative process [17]. All textual data were imported into NVivo software (version 12; Lumivero) to facilitate systematic coding and theme identification. Initial codes were generated based on both predefined study objectives and emerging patterns from the data. Through repeated review and refinement by all authors, broader themes and subthemes were developed to capture key insights. Consensus on the final themes was reached through discussion among all authors.

Descriptive statistics were computed to summarize responses. Participants' responses were analyzed by applying the strengths, weaknesses, opportunities, and threats (SWOT) analysis framework [18] to the digital epidemiology approach. The responses were independently reviewed by 2 coauthors (KLM and BM) and assigned to one of the SWOT categories. Any disagreements or discrepancies in the assignment of responses to SWOT categories were discussed, and consensus on the final categorization of each response was reached by the authors.

Ethical Considerations

This study involved seeking experts' opinion through the use of an electronic questionnaire—what is commonly classified as a minimal risk study, in which the probability and magnitude of harm or discomfort arising from participation were not greater, in and of themselves, than those ordinarily encountered in daily life or during the performance of routine physical or psychological examination or tests. Nevertheless, ethics approval for the study was obtained from the institutional review board at the Children's Hospital of Philadelphia (institutional review board protocol number: 22-020667). Participants received the survey electronically using REDCap, and consent to participate was obtained from all participants through a consent form embedded in the electronic questionnaire (REDCap). Only participants who provided consent were able to proceed and complete the questionnaire. Participants' privacy and confidentiality were ensured throughout the study. REDCap is a secure, web-based platform designed to support electronic data capture for research and operational projects, and it is compliant with the Health Insurance Portability and Accountability Act and the General Data Protection Regulation. The study did not collect any personally identifiable information. All data collected were anonymized and reported in aggregate form only. Participants did not receive any compensation for taking part in the study.

Results

Following the data collection phase (February 2023-February 2024), a total of 22 fully completed questionnaires were retained for analysis, while incomplete responses were excluded. Questionnaires were deemed incomplete if participants failed to respond to any of the mandatory sections or to questions measuring key outcome variables. The study participants represented diverse geographical regions, including Africa (7/22, 32%), the Americas (7/22, 32%), Europe (6/22, 27%), and Oceania (2/22, 9%). A detailed breakdown of this distribution is presented in Table 1.

XSL•FO
RenderX

Table 1. Characteristics of study participants (N=22) and countries represented in a global expert survey on digital epidemiology conducted from February 2023 to February 2024.

Region, subregion, and country	Participants, n (%)
Africa (n=7)	
Eastern Africa	
Rwanda	1 (4.5)
Uganda	2 (9.1)
Malawi	1 (4.5)
Western Africa	
Ghana	1 (4.5)
Central Africa	
Cameroon	1 (4.5)
Southern Africa	
South Africa	1 (4.5)
Americas (n=7)	
North America	
United States	6 (27.3)
Canada	1 (4.5)
Europe (n=6)	
Western Europe	
Switzerland	3 (13.6)
France	1 (4.5)
Netherlands	1 (4.5)
Southern Europe	
Greece	1 (4.5)
Oceania (n=2)	
Australia and New Zealand	
Australia	2 (9.1)

The study participants represented diverse professional backgrounds, with 14 (62%) participants identifying as informaticians, 4 (19%) participants as public health specialists, 3 (14%) participants as medical doctors, and 1 (5%) as a philosopher (Table 2). Participants were required to identify with only 1 mutually exclusive professional category.

Table 2. Occupation distribution of study participants (N=22) in a global expert survey on digital epidemiology conducted from February 2023 to February 2024.

Occupation	Participants, n (%)
Public health	4 (19)
Medical doctor	3 (14)
Informatician	14 (62)
Philosopher	1 (5)

This study aimed to explore experts' opinions on what they would consider as optimal conditions necessary for the use of digital epidemiology to enhance future pandemic preparedness and response. Accordingly, it was essential to assess participants' awareness of the concept of digital epidemiology, as this could influence the study's findings and interpretations. As illustrated in Table 3, most study participants were aware of the concept of digital epidemiology, with more than half of the participants either strongly agreeing or agreeing that they were familiar with digital epidemiology.

XSL•FO
RenderX

Table 3. Awareness of the concept of digital epidemiology among study participants (N=22) in a global expert survey conducted from February 2023 to February 2024.

Response category	Participants, n (%)
Strongly agree	11 (50)
Agree	8 (35)
Neutral	1 (5)
Disagree	1 (5)
Strongly disagree	1 (5)

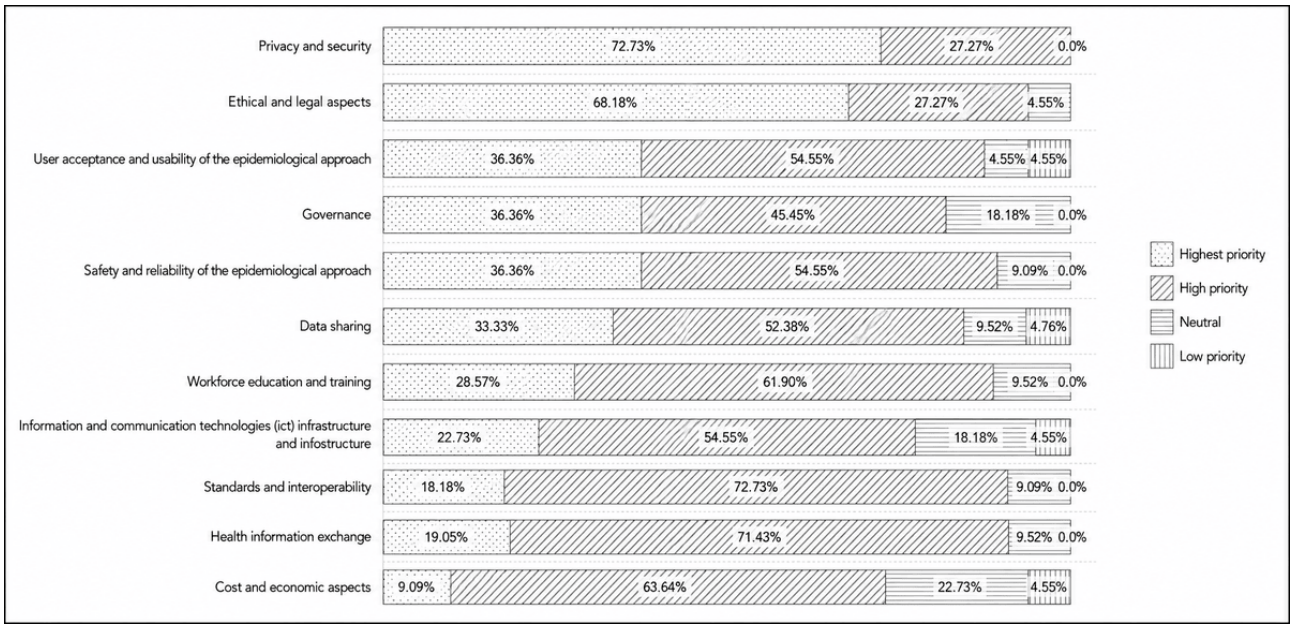
Participants who either strongly agreed or agreed (n=19 participants) that they were aware of the concept of digital epidemiology (Table 3), generally expressed positive sentiments regarding its potential impact and effectiveness in improving global health security (Table 4).

Table 4. Study participants' sentiments regarding the potential impact and effectiveness of digital epidemiology (n=19) in a global expert survey conducted from February 2023 to February 2024.

Statement	Strongly agree, n (%)	Agree, n (%)	Neutral, n (%)	Disagree, n (%)	Strongly disagree, n (%)
Digital data can enhance epidemiology purposes	13 (68)	4 (21)	2 (11)	0 (0)	0 (0)
Digital epidemiology can be effectively applied at a global level toward a coordinated global health surveillance	11 (57)	6 (32)	2 (11)	0 (0)	0 (0)
Digital epidemiology is reliable	4 (21)	12 (63)	2 (11)	1 (5)	0 (0)

Effective integration of digital epidemiology as part of a united global action toward future pandemic preparedness and response would require a certain level of readiness from the global community [7,18]. Key readiness indicators for decision-makers and implementers to consider are outlined in Figure 1. Study participants ranked these indicators based on their perceived importance and priority. In order of importance and priority, from highest to lowest, the indicators were ranked as follows: privacy and security, ethical and legal aspects, user acceptance and usability of the epidemiological approach, governance, safety and reliability of the epidemiological approach, data sharing, workforce education and training, information and communications technology infrastructure and infostructure, standards and interoperability, health information exchange, and costs and economic aspects (Figure 1). Participants also identified associated strengths, weaknesses, opportunities, and threats. It is noteworthy that fewer strengths and more weaknesses were identified relative to the other components of the SWOT analysis (Table 5).

Figure 1. Priority ranking of key readiness indicators for effective integration of digital epidemiology in pandemic preparedness and response, as reported by study participants (n=19) in a global expert survey conducted from February 2023 to February 2024.



XSL•FO
RenderX

Table 5. Strengths, weaknesses, opportunities, and threats (SWOT) analysis of opinions on the effective integration of digital epidemiology in pandemic preparedness and response (n=19) in a global survey of experts conducted from February 2023 to February 2024.

SWOT analysis categories	Themes emerging from participants' responses
Strengths	• Increasing investments in digital technologies by countries • Existence of digital health strategies by countries
Weaknesses	• Limited competent workforce • Weak digital health infrastructure • Digital divide • Poor data quality • Limited data sharing • Digital illiteracy • Limited use of artificial intelligence or machine learning • Weak ethical and legal considerations • Disconnect between public health departments and hospitals
Opportunities	• Global open standards that promote data exchange • Global design • International coordination on ethical considerations • Inclusion of the nonhealth sector
Threats	• Data security • Global economic challenges • Patient acceptance • Ethics • Political interference • Political unawareness

Discussion

Principal Findings

This qualitative study gathered insights from 22 experts (Table 1). Although the sample size may be considered relatively small, it is appropriate for the exploratory nature of the research and emphasizes achieving information power [19]. This approach provides a rich, in-depth, context-specific preliminary evidence rather than statistical generalization. In qualitative research designed to identify patterns across data, recruitment of between 15 and 30 participants has been commonly found to be adequate [20].

Digital epidemiology is a relatively new field that has grown rapidly in recent years, with its humble beginnings traced back to when researchers began mining the increasing amount of internet data for epidemiological purposes [21]. The growth of the field, and thus increased awareness, has been fueled by the growing utility of new data sources and artificial intelligence (AI) in epidemiology [22]. The ad hoc use of digital epidemiology methodologies and techniques during the COVID-19 pandemic likely contributed to the heightened awareness of the field [23]. This may explain why most study participants, especially in light of their professional backgrounds, reported being aware of the concept of digital epidemiology (Table 2). However, it is important to note that the attitude of stakeholders, rather than their awareness, may play a more crucial role in driving sustainable use of digital epidemiology. The technology acceptance model features prominently among the methodologies used to measure attitudes toward technology adoption [24,25]. According to the technology acceptance model, perceived usefulness (PU) is characterized by an individual's belief that using technology improves their job performance, while perceived ease of use (PEOU) refers to their belief that using technology requires minimal effort [24]. Both PU and PEOU are primary drivers that influence attitudes toward using technology [25]. Most study participants agreed that digital data can enhance epidemiological purposes (PU), that digital epidemiology can be effectively applied at a global level toward coordinated global health surveillance (PU), and that digital epidemiology is reliable (PU+PEOU). This is illustrated in Table 4 and reflects participants' positive attitudes toward digital epidemiology, particularly regarding its potential impact and effectiveness in improving global health security.

The Riyadh Declaration on digital health is a global call to action promoting the development of infrastructure and best practices necessary for sharing high-quality, real-time data both locally and globally in the fight against pandemics. Its goal is to enable health systems to access actionable information that can inform timely responses to public health threats [26]. Similarly, the US Centers for Disease Control and Prevention launched the Data Modernization Initiative to enhance the quality, availability, and use of data related to pandemics and public health emergencies [27]. Together, these frameworks serve to guide the integration of digital innovations into national and global health systems as part of coordinated efforts to strengthen preparedness and response to future pandemics. A shared theme emphasized across these frameworks is that leveraging digital epidemiology to address global health threats requires a more robust and integrated global health architecture that is capable of better preparing local, national, regional, and global public health partners to predict, prevent, detect, assess, and respond effectively to emerging threats. This calls for intentional efforts by governments and public health authorities

XSL•FO
RenderX

to mitigate digital exclusion at all levels (global, regional, national, and local) and to address data governance and interoperability challenges inherent in using and sharing diverse data sources [28].

Privacy and security, along with ethical and legal considerations, were ranked by most experts as high priority areas that must be considered by decision-makers and implementers to ensure sustainable integration of digital epidemiology tools in preparedness and response to future pandemics (Figure 1). Digital epidemiology, as previously explained, makes effective use of data not originally intended for epidemiological purposes. Examples of such data include internet searches, social media posts, webpage access logs, mobile phone network data, mobile health apps, wearable devices, sensor generated data, call center records, health and patient registries, and environmental monitoring data [21,23]. As some of these data are generated outside traditional public health systems and those involved in generating such data may not have consented to or even be aware of their use for epidemiological purposes, indeed this raises legitimate concerns related to potential data privacy breaches, security vulnerabilities, and ethico-legal issues. The pertinent literature suggests that data on one's health status have the potential to be reused by third parties to discriminate and restrict individual rights, and that institutional and government access to personal data (eg, geospatial data) could inhibit the exercise of basic freedoms, if individuals feel watched as to what they do or with whom they spend time [29]. In the context of the existing global digital divide and digital access gap, reliance on digital epidemiology to drive pandemic preparedness and response strategies may inadvertently perpetuate existing health disparities, benefiting overrepresented populations in the data, while not providing gains to and even possibly harming populations whose data are underrepresented [30]. In light of the current fragmented regulatory landscape and in the absence of a common set of principles on data privacy, security, and ethics, a global coordinated approach is required to establish frameworks that will ensure safe and responsible use of digital epidemiology as part of a unified global effort toward future pandemic preparedness and response [31]. Findings from a study based on the unified theory of acceptance and use of technology framework indicated that cost did not moderate behavioral intention to adopt the technology [32]. This may help explain why, in this study, experts ranked cost and economic aspects lower among the key readiness factors for the adoption of digital epidemiology (Figure 1).

The WHO's Global Strategy on Digital Health 2020 to 2025 provides a structured road map to guide member states in systematically integrating digital health technologies into their national health systems. According to WHO, the vision of the global strategy is to improve health for everyone, everywhere by accelerating the development and adoption of appropriate, accessible, affordable, scalable, and sustainable person-centric digital health solutions to prevent, detect, and respond to epidemics and pandemics [33]. Existence of such a document is indeed a strength in creating an enabling environment for the use of digital epidemiology in future pandemic preparedness and response. Increasing investments in digital technologies by countries and existence of digital health strategies by countries (Table 5) are part of the strategic objectives advanced by the document. The 4 strategic objectives outlined in the Global Strategy on Digital Health are as follows: (1) promote global collaboration and advance the transfer of knowledge on digital health, (2) advance the implementation of national digital health strategies, (3) strengthen governance for digital health at global, regional, and national levels, and (4) advocate people-centered health systems that are enabled by digital health [32]. These strategic objectives align with concerns highlighted as weaknesses by experts in Table 5. In addition to data privacy, security, and ethics already discussed, experts identified sociopolitical challenges and economic constraints as potential threats to a global approach to digital epidemiology (Table 5). This aligns with the study by Davis [34], which discusses the political determinants of digital health, defining them as the intersecting areas of national and international law and policy, local and global governance, civic engagement, and the influence of transnational commercial interests. The author explains that these factors collectively shape access to, uptake of, and the design of digital health systems [34]. As a case point, global geopolitical tensions and varying digital health maturity levels of nations can significantly impact data sharing in public health, resulting in lapses in global health security.

The full potential of digital epidemiology in achieving coordinated pandemic preparedness and response can indeed be leveraged through advancements in global design and international collaborations on digital health–related issues (Table 5). Opportunities exist in open health data standards [35], which promote interoperability and enable effective data sharing across diverse settings. The WHO's extension of the Global Strategy on Digital Health 2025 to 2027 endeavors to provide critical guidance on AI in health, including the ethics and governance of AI for health, and aims to support member states through global workshops on ethical AI implementation [36] and the code of ethics for health information professionals of the International Medical Informatics Association [37], are but examples of opportunities that offer a pathway toward safe and responsible use of digital epidemiology as part of a unified global effort for future pandemic preparedness and response. As previously explained, digital epidemiology uses data sources beyond the traditional health sector to gain insight into disease patterns and trends. As such, a significant opportunity exists in the continued efforts by global health actors to engage representation from nonhealth sectors in plans for future pandemic preparedness and response (Table 5). These sentiments are emphasized in the WHO Pandemic Agreement which sets out principles, approaches, and tools for better international coordination across a range of areas to strengthen the global health architecture for pandemic prevention, preparedness, and response [38]. The WHO Pandemic Agreement perhaps holds great promise toward significantly contributing to the creation of an enabling environment for the effective role of digital epidemiology in future pandemic preparedness and response efforts.

Limitations

This study is subject to limitations inherent in its design. The use of convenience and purposive sampling methods limits the generalizability of the findings. The sample may have excluded

XSL•FO
RenderX

critical perspectives from underrepresented geographic regions or professional backgrounds. For instance, within the North American region, participants from the United States were heavily overrepresented, with only 1 participant from Canada, thereby excluding perspectives from other countries in the region. The composition of professional backgrounds represented in the sample may have influenced participant responses, potentially skewing perceptions toward more favorable views of digital epidemiology. These notable imbalances may have constrained the diversity and balance of expert opinions captured.

Conclusions

This study explored experts' perspectives on the potential, readiness factors, and challenges associated with integrating digital epidemiology into global pandemic preparedness and response strategies. The findings underscore a growing recognition of digital epidemiology as a critical tool for enhancing global health security, particularly using nontraditional data sources and emerging technologies such as AI. Participants demonstrated high awareness and positive attitudes toward the utility and applicability of digital epidemiology, especially in advancing coordinated global health surveillance. However, the study also identified key barriers that must be addressed to ensure sustainable integration of digital epidemiology into health systems. These barriers include concerns related to data privacy, security, ethics, sociopolitical influences, and digital inequities across and within countries. The study affirms the need for a globally coordinated approach to governance, regulation, and investment in digital health infrastructure. By highlighting the complex interplay between technological, institutional, political, and social determinants of digital epidemiology, this research contributes important insights into how global health actors can leverage digital innovations responsibly and effectively for epidemiological purposes. Future research directions should focus on developing and validating context-specific frameworks that offer clear guidance for the responsible and sustainable integration of digital epidemiology into existing health systems, to support preparedness for and response to future pandemics.

Acknowledgments

This research was supported by the US Fulbright Scholar Program and the Children's Hospital of Philadelphia, Global Health Informatics Program. The authors would like to acknowledge the valuable insights and support from faculty at the Perelman School of Medicine, University of Pennsylvania. OpenAI's GPT-5.2 was used by 1 coauthor for grammar and spelling checks in portions of the manuscript. All text was reviewed and finalized by all authors, who retain full responsibility for its accuracy.

Data Availability

The data used and analyzed during this study are available from the corresponding author upon reasonable request.

Funding

This research was funded by the US Fulbright Scholar Program and the Children's Hospital of Philadelphia's Global Health Informatics Program. Apart from providing financial assistance for research-related activities, the funders had no influence in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Authors' Contributions

KLM contributed to the study concept and design, as well as data collection. KLM, JHH, AL, and BM contributed to data analysis and interpretation of the results. All authors participated in manuscript preparation, critically reviewed the article, and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
COREQ checklist.
[PDF File (Adobe PDF File), 119 KB - publichealth_v12i1e84164_app1.pdf]

References

1. International Monetary Fund. Global prospects and policies. In: World Economic Outlook: A Long and Difficult Ascent. Washington, DC: International Monetary Fund; 2020.
2. Sirleaf EJ, Clark H. Report of the Independent Panel for Pandemic Preparedness and Response: making COVID-19 the last pandemic. Lancet 2021 Jul 10;398(10295):101-103. [doi: 10.1016/s0140-6736(21)01095-3]
3. de Souza LE, Castro MC, Hage Carmo E, Polidoro M. The global failure of facing the pandemic. Glob Health Action 2022 Dec 31;15(1):2124645 [FREE Full text] [doi: 10.1080/16549716.2022.2124645] [Medline: 36285582]

XSL•FO
RenderX

4. Kamalrathne T, Amaratunga D, Haigh R, Kodituwakku L. Need for effective detection and early warnings for epidemic and pandemic preparedness planning in the context of multi-hazards: lessons from the COVID-19 pandemic. Int J Disaster Risk Reduct 2023 Jun 15;92:103724 [FREE Full text] [doi: 10.1016/j.ijdrr.2023.103724] [Medline: 37197332]
5. Williams GA, Fahy N, Aissat D, Lenormand MC, Stüwe L, Zablit-Schmidt I, et al. COVID-19 and the use of digital health tools: opportunity amid crisis that could transform health care delivery. Eurohealth 2022;28(1):29-34 [FREE Full text]
6. Budd J, Miller BS, Manning EM, Lampos V, Zhuang M, Edelstein M, et al. Digital technologies in the public-health response to COVID-19. Nat Med 2020 Aug;26(8):1183-1192. [doi: 10.1038/s41591-020-1011-4] [Medline: 32770165]
7. Ezenwaji CO, Alum EU, Ugwu OP. The role of digital health in pandemic preparedness and response: securing global health? Glob Health Action 2024 Dec 31;17(1):2419694 [FREE Full text] [doi: 10.1080/16549716.2024.2419694] [Medline: 39435565]
8. Digital tools for addressing infectious diseases in the Asia-Pacific region: challenges and opportunities. Asia-Pacific Economic Cooperation. 2021 Dec. URL: https://www.apec.org/docs/default-source/publications/2021/12/digital-tools-for-addressing-infectious-diseases-in-the-asia-pacific-region-challenges-and-opportunities/221_hwg_digital-tools-for-addressing-infectious-diseases-in-the-asia-pacific-region.pdf [accessed 2026-05-31]
9. Al Knawy B, McKillop MM, Abduljawad J, Tarkoma S, Adil M, Schaper L, et al. Successfully implementing digital health to ensure future global health security during pandemics: a consensus statement. JAMA Netw Open 2022 Feb 01;5(2):e220214 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.0214] [Medline: 35195701]
10. Olusanya OA, White B, Melton CA, Shaban-Nejad A. Examining the implementation of digital health to strengthen the COVID-19 pandemic response and recovery and scale up equitable vaccine access in African countries. JMIR Form Res 2022 May 17;6(5):e34363 [FREE Full text] [doi: 10.2196/34363] [Medline: 35512271]
11. Tarkoma S, Alghnam S, Howell MD. Fighting pandemics with digital epidemiology. EClinicalMedicine 2020 Sep;26:100512 [FREE Full text] [doi: 10.1016/j.eclinm.2020.100512] [Medline: 32864592]
12. "COVID-19 shows why united action is needed for more robust international health architecture" - Op-ed article by President Charles Michel, WHO Director General Dr Tedros Adhanom Ghebreyesus and more than 20 world leaders. European Council. 2021. URL: https://www.consilium.europa.eu/en/press/press-releases/2021/03/30/pandemic-treaty-op-ed/ [accessed 2026-05-31]
13. WHO welcomes historic commitment by world leaders for greater collaboration, governance and investment to prevent, prepare for and respond to future pandemics. World Health Organization. 2023 Sep 20. URL: https://tinyurl.com/3c3n3ykd [accessed 2026-05-31]
14. Widening digital gap between developed, developing states threatening to exclude world's poorest from next industrial revolution, speakers tell Second Committee. United Nations. 2023. URL: https://press.un.org/en/2023/gaef3587.doc.htm [accessed 2026-05-31]
15. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
16. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
17. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2008 Jul 21;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
18. van Wijngaarden JD, Scholten GR, van Wijk KP. Strategic analysis for health care organizations: the suitability of the SWOT-analysis. Int J Health Plann Manage 2012;27(1):34-49. [doi: 10.1002/hpm.1032] [Medline: 20603842]
19. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
20. Braun V, Clarke V. Successful Qualitative Research: A Practical Guide for Beginners. Thousand Oaks, CA: SAGE Publication; 2013.
21. Salathé M. Digital epidemiology: what is it, and where is it going? Life Sci Soc Policy 2018 Jan 04;14(1):1 [FREE Full text] [doi: 10.1186/s40504-017-0065-7] [Medline: 29302758]
22. Jiao Z, Ji H, Yan J, Qi X. Application of big data and artificial intelligence in epidemic surveillance and containment. Intell Med 2023 Feb;3(1):36-43 [FREE Full text] [doi: 10.1016/j.imed.2022.10.003] [Medline: 36373090]
23. Fallatah DI, Adekola HA. Digital epidemiology: harnessing big data for early detection and monitoring of viral outbreaks. Infect Prev Pract 2024 Jun 29;6(3):100382 [FREE Full text] [doi: 10.1016/j.infpip.2024.100382] [Medline: 39091623]
24. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319-340. [doi: 10.2307/249008]
25. Rahimi B, Nadri H, Lotfnezhad Afshar H, Timpka T. A systematic review of the technology acceptance model in health informatics. Appl Clin Inform 2018 Jul;9(3):604-634 [FREE Full text] [doi: 10.1055/s-0038-1668091] [Medline: 30112741]
26. Al Knawy B, Adil M, Crooks G, Rhee K, Bates D, Jokhdar H, et al. The Riyadh Declaration: the role of digital health in fighting pandemics. Lancet 2020 Nov 14;396(10262):1537-1539. [doi: 10.1016/s0140-6736(20)31978-4]

XSL•FO
RenderX

27. CDC global digital health strategy: appendices. Centers for Disease Control and Prevention. 2022. URL: https://archive.cdc.gov/#/details?url=https://www.cdc.gov/global-health/media/pdfs/GDHS_Strategy2022_Appendices_REV_508.pdf [accessed 2026-05-31]
28. Li L, Back E, Lee S, Shipley R, Mapitse N, Elbe S, et al. Balancing risks and opportunities: data-empowered-health ecosystems. J Med Internet Res 2025 Mar 25;27:e57237 [FREE Full text] [doi: 10.2196/57237] [Medline: 40132190]
29. Ferretti A, Vayena E. In the shadow of privacy: overlooked ethical concerns in COVID-19 digital epidemiology. Epidemics 2022 Dec;41:100652 [FREE Full text] [doi: 10.1016/j.epidem.2022.100652] [Medline: 36356477]
30. Howe Iii EG, Elenberg F. Ethical challenges posed by big data. Innov Clin Neurosci 2020 Oct 1;17(10-12):24-30 [FREE Full text] [Medline: 33898098]
31. Data privacy, ethics and protection: guidance note on big data for achievement of the 2030 agenda. United Nations Development Group. URL: https://unsdg.un.org/sites/default/files/UNDG_BigData_final_web.pdf [accessed 2026-05-31]
32. Ahmed ER, Zirari B, Tabet Derraz I. The moderating role of cost and convenience on the relationship in the UTAUT model to adopt e-payment services. J Int Technol Inf Manag 2024;32(1):334-363. [doi: 10.58729/1941-6679.1582]
33. Global strategy on digital health 2020-2025. World Health Organization. 2021 Aug 18. URL: https://www.who.int/publications/i/item/9789240020924 [accessed 2026-05-31]
34. Davis SL. Political determinants of digital health: beyond the rainbow. Health Promot Int 2025 Mar 05;40(2):daaf014. [doi: 10.1093/heapro/daaf014] [Medline: 40177937]
35. Kapitan D, Heddema F, Dekker A, Sieswerda M, Verhoeff BJ, Berg M. Data interoperability in context: the importance of open-source implementations when choosing open standards. J Med Internet Res 2025 Apr 15;27:e66616 [FREE Full text] [doi: 10.2196/66616] [Medline: 40232773]
36. Seventy-eighth World Health Assembly – daily update: 23 May 2025. World Health Organization. 2025 May 23. URL: https://www.who.int/news/item/23-05-2025-seventy-eighth-world-health-assembly---daily-update--23-may-2025 [accessed 2026-05-31]
37. The IMIA code of ethics for health information professionals. International Medical Informatics Association. 2016 Aug 28. URL: https://imia-medinfo.org/wp/wp-content/uploads/2015/07/IMIA-Code-of-Ethics-2016.pdf [accessed 2026-05-31]
38. World Health Assembly adopts historic pandemic agreement to make the world more equitable and safer from future pandemics. World Health Organization. 2025 May 20. URL: https://www.who.int/news/item/20-05-2025-world-health-assembly-adopts-historic-pandemic-agreement-to-make-the-world-more-equitable-and-safer-from-future-pandemics [accessed 2026-05-31]

Abbreviations

AI: artificial intelligence
PEOU: perceived ease of use
PU: perceived usefulness
SWOT: strengths, weaknesses, opportunities, and threats
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 15.Sep.2025; peer-reviewed by TS Wu, A Hidki; comments to author 28.Nov.2025; revised version received 09.Apr.2026; accepted 10.Apr.2026; published 16.Jun.2026.

Please cite as:
Mauco KL, Holmes JH, Luberti A, Mosesane B
Experts' Opinions on the Sustainable Use of Digital Health Tools for Effective Future Pandemic Preparedness and Response: Questionnaire Study
JMIR Public Health Surveill 2026;12:e84164
URL: https://publichealth.jmir.org/2026/1/e84164
doi:10.2196/84164
PMID:

©Kabelo Leonard Mauco, John H Holmes, Anthony Luberti, Badisa Mosesane. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 16.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Transforming a Pandemic Tool Into a Public Health Asset: Modified Delphi Study

Mohsen Malekinejad[1], MD, DrPH; Cesar Arturo Mendez-Lizarraga[1], MD, MS, MPH; Nicole Skehan[1], MS; Alyssa Bercasio[1], MS; Rakhee Palekar[2], MD, MPH; William B Lober[3], MD, MS

[1]Institute for Global Health Sciences, University of California, San Francisco, 550, 16th St, 3406, Floor 3, San Francisco, CA, United States

[2]MITRE Corporation, McLean, VA, United States

[3]School of Nursing, University of Washington, Seattle, WA, United States

Corresponding Author:
Mohsen Malekinejad, MD, DrPH
Institute for Global Health Sciences, University of California, San Francisco, 550, 16th St, 3406, Floor 3, San Francisco, CA, United States

Abstract

Background: During the COVID-19 pandemic, over 100 countries successfully deployed smartphone-based exposure notification (EN) apps alongside conventional contact tracing to mitigate the transmission of SARS-CoV-2. These systems enabled millions of COVID-19–positive individuals to anonymously alert community members of a potential exposure. Although evidence on EN program implementation is limited, leveraging expertise from public health professionals, academics, and global implementers from various disciplines can contribute to addressing this research gap.

Objective: This study aimed to explore barriers and facilitators to adoption, identify optimal use cases for these tools, and understand how advancements in technology and policy might maximize EN utility for mitigating other infectious outbreaks.

Methods: We used the Delphi method to develop a comprehensive online survey comprising closed-ended and open-ended questions. We considered consensus was reached when at least 75% of respondents selected “agree” or “strongly agree” for an item or statement. More than 30 global EN experts known to the investigators were invited to participate in the survey, and those who participated were asked to nominate or recruit others deemed eligible via a peer-to-peer method. We used descriptive and content analyses to summarize key findings from the survey. We invited the same experts to attend virtual focus groups to further probe their insights into the survey findings.

Results: Twenty-two experts from 6 countries, including 10 experts from the United States, representing 3 continents, participated. Respondents were involved in the implementation of various EN technologies that encompassed varying levels of integration with existing public health agencies’ infrastructure. Consensus was reached concerning the development of additional functionalities of the system, such as the ability for users to self-report a positive result from a SARS-CoV-2 at-home test, receive additional public health information from the app (ie, information on testing and social services), and receive general venue-based notifications concerning superspreader events. Other items that reached consensus included how the tool could be used for additional epidemiological purposes, such as expanding the technology to other airborne diseases, and the need for integration with surveillance data streams to augment analytics and forecasting.

Conclusions: Global EN systems have the potential to be optimized for public health emergencies beyond COVID-19 and can be used as epidemiological surveillance tools in future disease outbreaks. Decision-makers and experts should revise existing protocols and seek to incorporate new functionality to alleviate the disease burden on both users and public health agencies.

(JMIR Public Health Surveill 2026;12:e77763) doi:10.2196/77763

KEYWORDS

exposure notification; preparedness; response, Delphi; method; pandemic; outbreak

Introduction

Background

The COVID-19 pandemic placed unprecedented demands on the public health workforce. In response, digital contact tracing (DCT) and exposure notification (EN) tools were developed to enhance the efficiency of public health outbreak control, particularly during large case surges [1,2]. While some DCT systems were based on direct reporting, a new generation of more automated, scalable EN systems leveraged smart devices to alert individuals of a potential exposure to the SARS-CoV-2 virus, even without direct contact or relationship with the case. Implementation and evaluation of these EN systems within different jurisdictions were shaped by diverse system

XSL•FO
RenderX

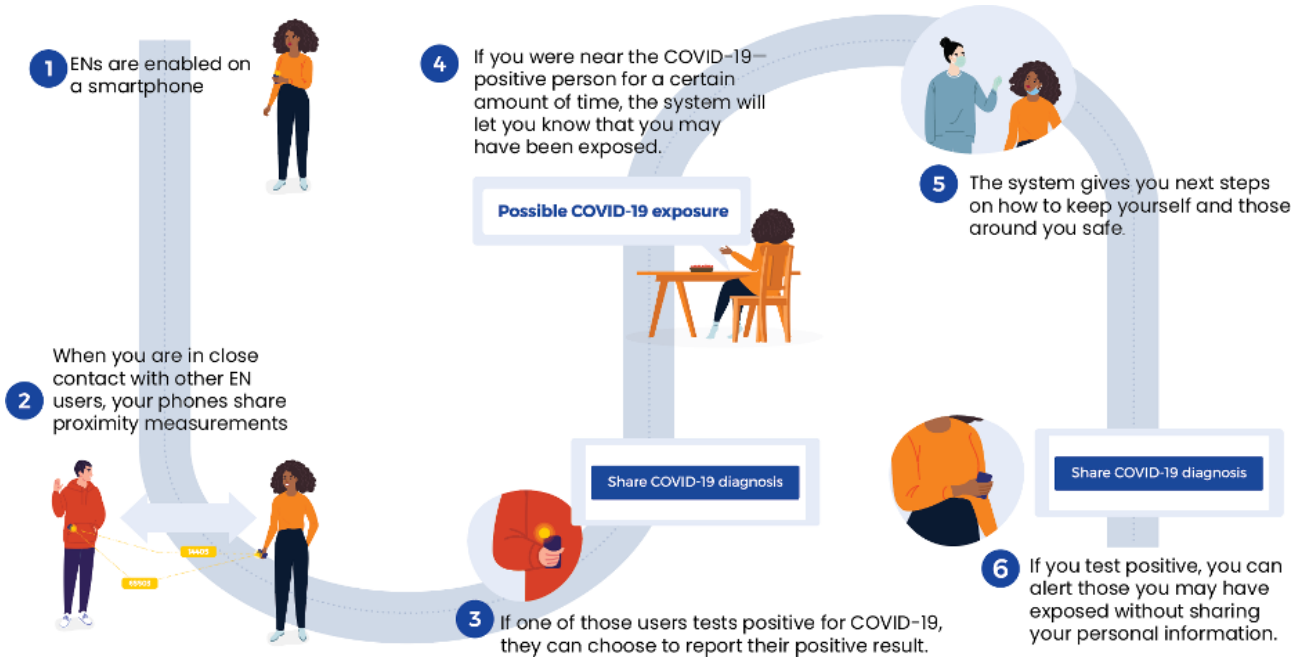
architectures, deployment strategies, and communication protocols, each influenced by political, social, and financial contexts. These tools also varied in their approach to privacy; some collected personal identifiers while others were completely anonymous and privacy-preserving [3].

Proximity Detection Technologies

EN systems work by measuring distances between individuals, using various technologies, such as GPS and Bluetooth Low Energy, to determine if a user was in close contact of a person with a confirmed COVID-19 diagnosis [4]. GPS was used in multiple systems—including those in South Korea and China—because it can capture specific user information [5,6]. However, standard GPS is considered unreliable in determining the exact position of a person and requires a significant amount of battery power.

On the other hand, Bluetooth-based systems could not record an individual's specific location and therefore were unable to identify where a person was exposed; however, they could assess the proximity between 2 individuals and provide greater privacy protection [7,8]. These systems relied on the generation of random digital codes; when 2 devices came close together, the codes were exchanged. By relying on 2 servers, a key server that stored keys from devices whose users reported positive tests and a verification server that checked for genuine positive tests, devices routinely searched for matches among stored codes. If a match occurred, the app alerted the user about a potential exposure, without disclosing who, when, and where the exposure took place [9,10]. Examples of Bluetooth-based EN systems include Singapore's BlueTrace protocol—which was used in Singapore's TraceTogether app and Australia's COVIDSafe app—and Google and Apple's jointly developed EN protocol, Google and Apple Exposure Notification (GAEN) system, which became the most widely adopted EN system during the pandemic. GAEN apps used Bluetooth Low-Energy signal exchanges between EN-enabled smartphones to estimate exposure risk based on duration and proximity to an individual's smartphone who voluntarily disclosed a positive COVID-19 test result (Figure 1) [3,7,11,12].

Figure 1. Overview of mobile exposure notification (EN) apps using the Google and Apple Exposure Notification (GAEN) System. Adapted from the California Department of Public Health and Aronoff-Spencer et al [13].



Other proximity detection technologies used Wi-Fi or ultra-wideband systems, while others deployed a QR-based system or a combination of multiple approaches [8]. Some apps also used a combination of different technologies, such as India's "Aarogya Setu" app, which used both GPS and Bluetooth [14].

Centralized vs Decentralized Data Systems

Global EN systems also differed in whether they used centralized or decentralized protocols. Centralized protocols stored user interactions on a central server, which allowed for more meticulous epidemiologic tracking and analysis. These protocols enabled public health agencies (PHAs) to track second-order contacts and super-spreader events. Singapore's BlueTrace system used a centralized system, as did China's and South Korea's apps [6]. However, centralized protocols were often considered controversial, given that they potentially enabled governments to access personal health data and therefore raised many user privacy concerns. Conversely, decentralized protocols determine and save proximity data and exposure risk entirely on a user's mobile phone to preserve user privacy. While decentralized protocols made epidemiologic analysis difficult, they allowed for greater privacy of user data [7]. Apps created on the GAEN system prioritized user privacy over disease surveillance, using a decentralized protocol, while allowing PHAs to set risk parameters and public health messaging by jurisdiction, at the national or subnational level [13,15].

The Future of EN

Most EN systems relied on government-sponsored servers to enhance cross-jurisdictional interoperability while preserving

XSL•FO
RenderX

user privacy. However, PHAs varied in how they assessed exposure thresholds, further complicating comparisons across implementations [16,17]. While GAEN's privacy-preserving model limited targeted public health messaging and robust impact assessments, modeling studies in settings with meaningful adoption rates—such as the United Kingdom and the states of Washington and California in the United States—suggest that digital EN systems averted significant numbers of new cases [11,12,16-28]. By contrast, non-GAEN systems, which collected identifiable data and location information, offered potential advantages in precision and public health follow-up. These systems could provide more tailored guidance to users and enable PHAs to assess the broader impact of exposures on public health. However, these systems often raised privacy concerns and faced barriers to adoption [2].

As the COVID-19 public health emergency concluded, scientists, policymakers, and public health leaders began to discuss both future and alternative uses of EN systems. Proponents highlighted the unprecedented collaboration between major technology companies and government institutions, resulting in the global deployment of EN systems, with over 300 million users worldwide [20,29-31]. Critics noted that while some countries and states deployed EN systems successfully, many regions either did not adopt the technology or failed to achieve meaningful user uptake. Despite variations in adoption and evaluation, EN systems demonstrated significant potential to augment infectious disease response and a remarkable ability to scale to large numbers of users and link jurisdictions, all while preserving privacy. As health systems prepare for future outbreaks, it is crucial to address gaps in EN design, deployment, and adoption to maximize the impact of these tools. It is also critical to understand the barriers and facilitators to their use to guide the development of policies and technologies that strengthen national and global health capacities.

We developed a modified Delphi study and engaged experts to explore barriers and facilitators to adoption, identify optimal use cases for these tools, and understand how advancements in technology and policy might maximize EN utility for mitigating other infectious disease outbreaks. By learning from the successes and limitations of EN systems during the COVID-19 pandemic, we seek to inform the design and implementation of more effective outbreak response tools for future use.

Methods

Study Design Overview

We used a modified Delphi Type 3 approach to gather expert knowledge on EN systems, using a comprehensive survey designed to capture both quantitative and qualitative data [32]. Appendix A in Multimedia Appendix 1 summarizes the study characteristics [33,34].

Curation of Ideas to Inform Survey Development

We curated initial ideas based on the authors' direct experiences implementing an EN system in California, as well as insights gathered from meetings of the Digital Notification Alliance (DNA), a US-based community of EN practitioners active during the COVID-19 pandemic, with participation from representatives of 16 states. To further develop our framework, we conducted a targeted literature search in PubMed and Google Scholar for relevant peer-reviewed articles and white papers during March-April 2023. The search focused on studies describing the implementation, evaluation, or adoption of EN systems and, when contextually relevant, other DCT tools during the COVID-19 pandemic. We used combinations of keywords such as "exposure notification," "digital contact tracing," "implementation," "barriers," "facilitators," "COVID-19," and "evaluation." We searched for papers published between 2020 and 2023 in English. Publications were included if they discussed real-world implementation experiences, adoption barriers, or lessons learned from EN systems. The references identified through this process—together with relevant white papers and reports from our professional networks—were used to structure survey items and domains [3,5,9-11,17,20-22,24,26,28,29,35]. Additionally, we contacted our network of EN implementers across the United States and internationally to inquire about relevant publications and gray literature.

Survey Development

Based on curated inputs, we developed a Delphi survey to collect expert insights on best practices, implementation challenges, and future directions for EN systems, with a focus on those using the GAEN protocol during the COVID-19 pandemic. The survey comprised six major sections: (1) participant consent and survey introduction; (2) expert demographics and professional background; (3) perceived successes, barriers, and facilitators of EN deployment; (4) perspectives on key system functionalities, disease applicability, macro-level governance, and considerations around privacy, equity, and communication; (5) opinions on potential uses of EN technologies beyond infectious disease outbreaks; and (6) participant demographics and peer referrals.

The final survey included approximately 100 questions, incorporating multiple-choice (including Likert-scale and rank-order items) and open-ended formats. Likert-scale items assessed the perceived importance or level of agreement on various implementation dimensions, while open-ended items allowed for elaboration and identification of additional considerations. The survey was administered via Qualtrics [36]. The instrument is available in Appendix B in Multimedia Appendix 1.

Participant Recruitment

Using a combination of convenience and peer-referral methods, we recruited participants from our extensive network of US-based and international EN experts with whom we had collaborated and exchanged scientific ideas throughout the pandemic. We initially identified participants through the DNA network group to share lessons learned and best practices in EN implementation. Through the DNA network, we distributed invitations via its membership listserv, solicited referrals (snowball sampling), and leveraged its endorsement to broaden participation beyond our immediate collaborators. We also invited key authors of published papers on EN prior to the inception of our study. In total, 56 eligible experts representing PHAs, academia, and the technology industry were invited to

XSL•FO
RenderX

participate. Eligibility criteria included being at least 18 years of age and having experience in the development, implementation, or evaluation of EN systems during the COVID-19 pandemic. Of the 56 identified experts, a total of 30 initially expressed interest in participating.

Pilot Testing and Data Collection

Using QualtricsXM, we conducted a pilot test with 4 individuals to evaluate the survey's logistics, length, content, and flow [36]. Based on feedback, we made minor revisions and then distributed the survey via email to 30 experts who had previously expressed interest in participating. The survey was launched in fall 2024, with up to 3 reminder emails sent at 7- to 10-day intervals. All responses were anonymous and untraceable after the pilot phase. Data from the pilot were included in the final data analysis.

Data Analysis

We conducted descriptive analysis of the quantitative data using Microsoft Excel. For Likert-scale questions, we adopted the 75% agreement threshold that was suggested in the systematic review of reporting Delphi studies to establish a working definition of "large majority" for our study [34]. This threshold was used to identify items that were deemed more important or to represent a higher level of agreement among experts, but it should not be seen as a consensus threshold. For items or statements that required ranking, we used a weighted score to determine the order based on the number of responses for each question:

Weighted score=[(number of first place votes)×1.0]+[(number of second place votes)×0.5] +[(number of third place votes)×0.33]+[(number of fourth place votes)×0.25] +[(number of fifth place votes)×0.2]+[(number of sixth place votes)×0.16] +[(number of seventh place votes)×0.14]+[(number of eighth place votes)×0.12] +[(number of ninth place votes)×0.11]

For the qualitative data, 2 researchers conducted a thematic analysis to identify recurring themes, which were used to supplement our analysis of the quantitative data. This allowed us to further elaborate on expert experiences and opinions, from which we could infer major key points and generate actionable recommendations. Examples are shown in Table 1.

Table . Thematic analysis of open-ended question responses[a].

Themes and questions	Responses	Subtheme
Lessons from the COVID-19 pandemic		
Do you think your EN[b] program achieved goals and objectives?	"It did so by complementing conventional contact tracing; compared to the latter, it led to an estimated 5% - 10% increase in the number of cases detected."	Public health
The next EN application		
Would you consider continuing EN programs if COVID-19 resources were available?	"EN systems should ideally run continuously, even when there is no immediate need. That way, it can be ready to scale up quickly when needed."	Public health
Do you have suggestions as to what governance principles and procedures should be in place prior to the next implementation of EN systems?	"International coordination between public health agencies and tech companies is necessary to iron out interests and articulate norms that can guide decisions."	Public policy
Are there ways to improve the opt-in sharing of information?	"There should be levels of information sharing for the comfortability of opting into the system..., but there should be no requirement for it."	Technical features

[a]The themes included in the survey are shown in the first column. Based on question types and responses, 3 subthemes emerged: technical features of EN applications, public health, and public policy issues.

[b]EN: exposure notification.

Expert Review and Validation of Findings and Final Recommendations

Following the initial round of data analysis, all expert participants were invited to a live, 2-hour virtual presentation to review findings and provide additional insights. Two sessions were conducted, and discussions were recorded with participants' permission. One researcher took detailed notes from the sessions and analyzed participant feedback. Comments and suggestions from these sessions were integrated into the final interpretation of results. The final recommendations were developed through a multistep process. Topics and ideas that were deemed important or reached a high level of agreement in the quantitative survey were prioritized. These were further enriched by qualitative feedback provided by participants and subsequently validated through the live expert reviews and discussions.

Ethical Considerations

Our study protocol was reviewed and approved by the Committee on Human Research at the University of California, San Francisco (UCSF), through an expedited review process (IRB number: 23-39886; reference number: 398640). The study did not include personal or sensitive questions and focused instead on gathering expert opinions about a digital health product. All data were stored securely within UCSF's protected

XSL•FO
RenderX

data systems and reported only in aggregate form to safeguard participant privacy. Participants could choose to share their general geographic location for national-level analysis. Those who provided informed consent were invited to complete the survey and received a US $50 gift card for participating in at least 1 round of data collection. Individuals who also took part in the live discussion received an additional US $100 gift card.

Results

Demographics

A total of 22 individuals from 3 World Health Organization (WHO) Regions (Americas, Europe, and Western Pacific), representing 6 countries, participated in our Delphi. We asked participants to select their involvement with EN systems during the pandemic, as all had carried out more than one specific activity (Table 2). The most common roles included data analysis (n=19, 86.3%), coordination with public and private partners (n=17, 77.2%), and elaborating policies and guidelines for decision-makers (n=14, 63.6%; Table 3).

XSL•FO
RenderX

Table . Participants' primary country for exposure notification (EN) implementation, years of experience with EN, and other characteristics of EN implementation (N=22).

Characteristics	Values, n (%)
Primary country in which participants were involved in EN[a] implementation	
United States of America	10 (45.5)
Belgium	2 (9.1)
United Kingdom	2 (9.1)
Germany	1 (4.5)
Singapore	1 (4.5)
Italy	1 (4.5)
Unknown	5 (22.7)
Years of experience in EN systems	
1 - 3 y	11 (50.0)
3 or more years	10 (45.4)
Not reported	1 (4.5)
Type of EN system first implemented	
Customized GAEN[b] app	8 (36.3)
ENX (Exposure Notification Express)	5 (22.7)
Fully custom app	5 (22.7)
Mixed or hybrid model	1 (4.5)
Other	2 (9.1)
Unknown	1 (4.5)
Integration of EN systems to PHA[c] public health response	
Not entirely integrated	9 (40.9)
Fully integrated	7 (31.8)
Stand-alone systems	2 (9.1)
Other	1 (4.5)
Unknown	3 (13.6)
Main implementer of EN systems in jurisdiction	
Mixed model	11 (55.0)
Health departments	7 (35.0)
Contract with private entity	2 (10.0)
Not reported	2 (9.1)

[a]EN: exposure notification.

[b]GAEN: Google and Apple Exposure Notification system.

[c]PHA: public health agencies.

XSL•FO
RenderX

Table . Participants' areas of expertise and roles adopted in the exposure notification (EN) implementation[a].

Areas of expertise and roles adopted	Values (n)
Areas of expertise	
Public health	15
Epidemiology	14
EN[b] systems and technology	12
Policy	5
Health communications	4
Engineering	4
Research	10
Other	8
Patient care	2
Responder's role in EN implementation	
Data analysis	19
Coordination with public and private partners	17
Health communication media	13
Program implementation	12
User experience	11
Interface design	8
Systems development	12
Policies and guidelines elaboration for decision-makers	14

[a]Respondents (N=22) were asked to select all options that applied to their profile.

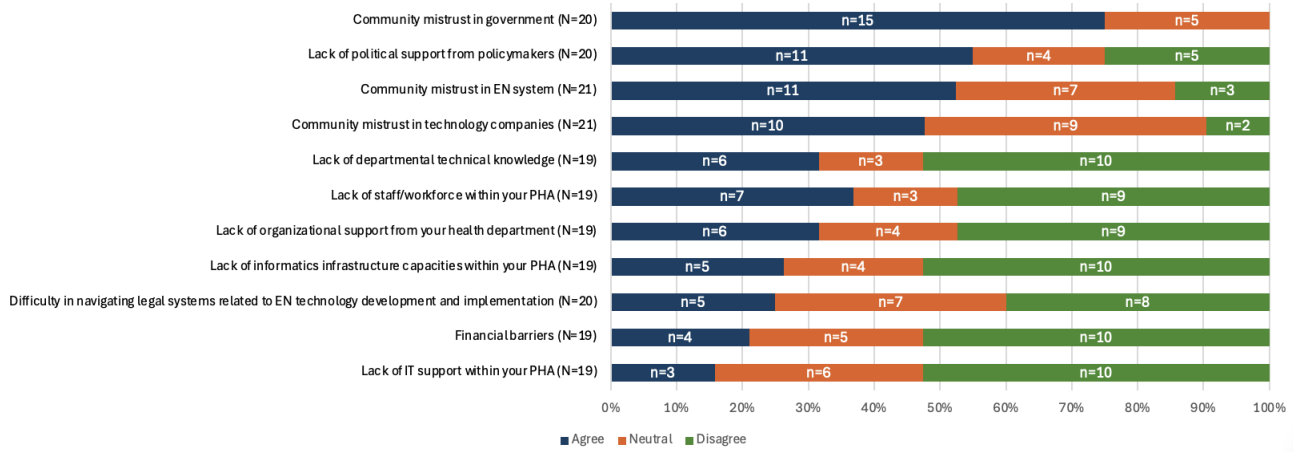
[b]EN: exposure notification.

Lessons From EN System Implementation During the COVID-19 Pandemic

Overview

When inquired about specific barriers to EN program implementation, the large majority of participants (15/20, 75% of respondents) identified "community mistrust in the government" as a major barrier (Figure 2). Items voted by at least half of the participants as barriers to implementation included the following: the lack of political support from policymakers (11/20, 55%) and community mistrust in EN systems (11/21, 52.4%). Other items voted by less than 50% of respondents as barriers included, but were not limited to, the following: community mistrust in technology companies (10/21, 47.6%), lack of staff workforce within PHAs (7/19, 36.8%), and lack of departmental technical knowledge (6/19, 31.6%).

Figure 2. Barriers to exposure notification (EN) program implementation during COVID-19. Response statements have been shortened for practical reasons. Agreement and disagreement percentages are aggregates of "agree," "strongly agree," "disagree," and "strongly disagree," respectively. Statements are ranked from the highest to the lowest percentage of agreement.



XSL•FO
RenderX

Implementation Barriers

Respondents pointed out a combination of factors such as "lack of awareness, privacy concerns, government mistrust, and lack of understanding of how the app worked" as barriers to PHA implementation. For others, a bigger budget could have translated into "continuous development, better privacy features, and relevant data analysis." In the case of the United States, the lack of a centralized system caused a "significant burden on jurisdictions." In addition, we inquired about ideal resources (framed as a "Wishlist") for better programmatic and health outcomes. Respondents indicated a desire for higher budgets for increasing user engagement, stronger analytical resources for tracking relevant metrics, access to testing without symptoms, full support from decision-makers, an opt-out modality, central collection of outcome data (eg, test results for notified contacts), better integration between testing and treatment resources (as well as with other agencies), more app promotion, and a "multilateral dialogue on EN system design principles with tech giants with a deference to public health."

Program Goals

When inquired whether participants thought their EN programs achieved proposed goals and objectives, some stressed that they did so by "alerting hundreds of thousands of people of exposures," while others expressed that the system notified "far more contacts per exposure than traditional contact tracing." Conversely, some believed they did not achieve their goals and objectives because they faced little adoption, late integration into their health authorities' processes, and "lack of public communication about the processes of the app."

Epidemiological and Clinical Considerations for Future EN Systems

With respect to the applicability of proximity-based EN apps for infectious diseases, the large majority of respondents labeled airborne diseases as "very applicable" (17/21, 80.9%), while fomite-transmitted diseases were considered "very applicable" by less than half (9/20, 45%). On the other hand, blood-borne diseases and sexually transmitted infections were predominantly considered as "not very applicable" by 80% (16/20) and 65% (13/20) of respondents, respectively. Concerning the incorporation of diseases into future EN apps, the large majority of participants considered COVID-19 (20/22, 90.9%) and measles (17/22, 77.3%). Other diseases that were considered by at least half of the respondents were influenza (15/22, 68.2%) and tuberculosis (13/22, 59.1%). Additional pathogens that were suggested in the written comments by participants included respiratory syncytial virus, meningococcal meningitis, Disease X, HIV, hepatitis A, Mpox, and other sexually transmitted infections.

Finally, we asked participants to rank key epidemiological and clinical characteristics of pathogens that should be considered for future app developments, which resulted in the following order of priorities: (1) transmissibility of the pathogen (incubation period, infectiousness, etc; 82 points); (2) available protective actions (75 points); (3) disease burden (71 points); (4) level of community immunity (67 points); and (5) economic availability (or feasibility) to take protective actions (65 points).

Public Health and Implementation Considerations for Future EN Systems

EN Value for the Future

When asked if there was any value in continuing EN programs after the end of the COVID-19 pandemic as a public emergency of international concern, respondents agreed that there is a need to be ready: "waiting for an emergency will always be too late...even modest benefit with ongoing use is worthwhile." Pandemic preparedness and response require a set of continuous actions, and for EN systems, "ideally they should run continuously...ready to scale up quickly when needed." A challenge highlighted was the need to demonstrate its value to secure the necessary resources and political support, with one participant noting, "we need to demonstrate value to enable resource allocation and political support."

Strategies to Improve EN Adoption

Clear communication—such as the benefits for the individual and community, goals, and privacy issues—is key, as is open communication about any uncertainties (ie, "letting the wider public know what we don't know"). On communicating with users about their exposure, messages need to be accurate and readable. As shared by the group, "information has to be credible and accessible." Another suggested strategy was to include venue-based exposure alerts. On how to best leverage experience from implementors, a shared answer was to create standardized guides for PHA EN deployment, covering design, development, implementation, evaluation, and other aspects encompassing a public health intervention.

Communication Campaign

When participants were asked to select topics that should be considered for public-facing communication campaigns, topics deemed "very important" by the large majority were as follows: reminders to use the system when testing positive (18/22, 81.8%) and that users should directly benefit from using the system (17/22, 77.2%). Items that were selected by more than 50% but were below the large majority threshold included how to use the system (15/22, 68.1%), community benefits from using the system (13/22, 59.1%), and user privacy protection (13/22, 59.1%). Marketing strategies for EN technologies through public agencies included using social media and communicating to the broader public while advertising the app as "a tool for citizen science" to strengthen societal preparedness. About reducing technological disparities, "using free standalone devices" that can provide text-based notifications and inform of actions that communities can take was identified as the most common approach. Finally, when inquired about the utility and importance of a comprehensive and multifunctional EN app for future public health threats, participants stated that "it may lose focus on what the main priority...and users may become overwhelmed," while a participant suggested that EN apps should be limited to infectious disease monitoring and control.

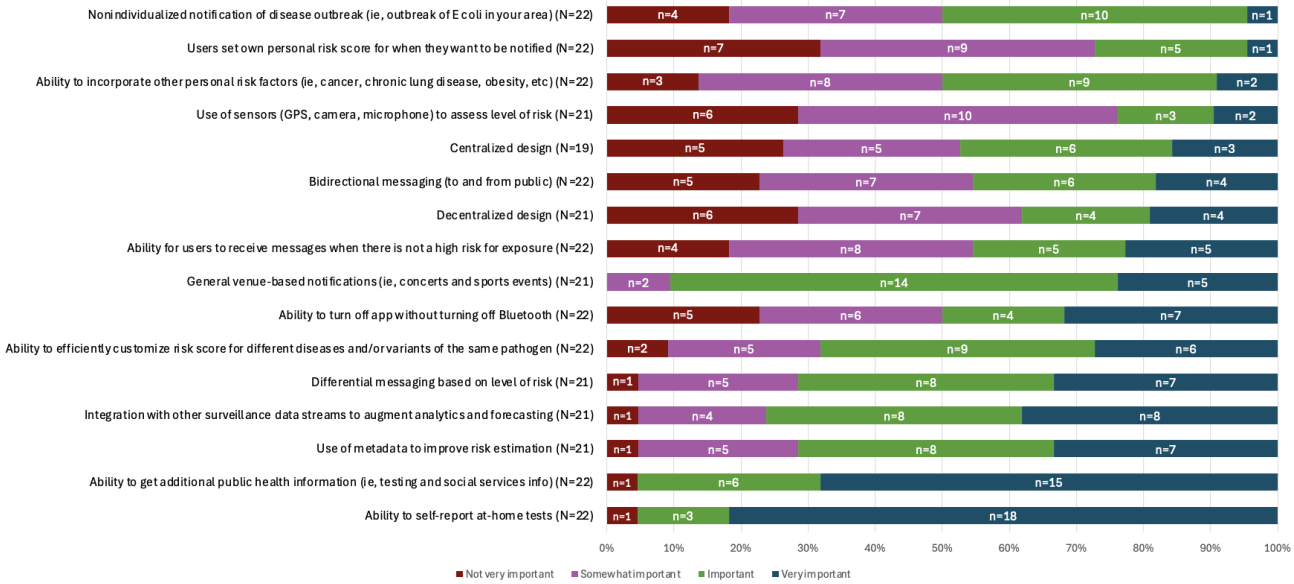
XSL•FO
RenderX

Technical Features Considerations for Future EN Systems

General Technical Features for Future App Development

Participants were asked to label the importance of functionalities and characteristics for future EN systems. The ability to self-report at-home tests was the only item that was considered "very important" by the large majority (18/22, 81.8%) of respondents, while the ability to get additional public health information was considered "very important" by 68.2% (15/22). Also, the venue-based notifications feature was labeled as "important" by more than half (14/21, 66.7%) of participants. Features like the use of sensors to assess the level of risk and the ability to set a personal risk score for notifications were categorized as "somewhat important" by almost half of the participants. No feature was deemed "not very important" by one-third or more of respondents. The level of importance of all functionalities is shown below (Figure 3).

Figure 3. System functionalities and characteristics for the next exposure notification (EN) app are shown together by their level of importance on a 4-point Likert scale. Items are ranked based on the percentages obtained for "very important" from respondents (n=22).



Specific Technology for Proximity Detection

Half of the participants considered mixed EN signaling technologies to be the best solution for future developments, followed by only Bluetooth Low Energy (5/22, 22.7%) and ultra-wideband (2/22, 9.1%). Mixed EN technology examples included combining Bluetooth Low Energy with other methods (ie, GPS or venue-based alerts), and participants commonly wrote down "Bluetooth plus Ultra-wideband system" as another option. Nearly 20% (4/22) of respondents did not offer a specific opinion.

When participants were asked to rank strategies to ease user burden, the order is as follows: (1) easier methods for activation (97 points), (2) increased access to other public health resources (86 points), (3) more accessible language (85 points), (4) different user workflows (73 points), (5) new methods for direct user communication (50 points), and (6) new EN hardware (29 points). Furthermore, the group was divided on whether the use of custom hardware or devices should (10/21, 47.7%) or should not (11/21, 52.3%) be incorporated. Further insight on this revealed that it has the potential "to be quickly deployed in outbreaks in populations without smartphones" but made no sense in settings where mobile phone coverage is high, as "it seems unnecessary to ask them to carry another device."

Privacy

Out of 4 options (accuracy of EN, perceived risk by users and their willingness to adopt, scalability tradeoff, and system evaluation capabilities), we asked participants to identify the option that is more important to consider for future EN development in relation to the level of privacy. Privacy was the only characteristic deemed as "more important" by the large majority against the perceived risk by users and their willingness to adopt (18/22, 81.8%). Privacy was voted as "less important" by more than half of participants and less than the large majority for the following: accuracy of EN (13/22, 59.1%) and system evaluation capabilities (12/21, 57.1%). Finally, privacy was neither more nor less important for the scalability tradeoff (50% vs 50%; N=20). The use of technology with a Bluetooth-based protocol that does not require information about the location of the user was considered advantageous, as "it preserved privacy while not creating a huge burden on user batteries—unlike GPS—...with the downside of not fully understanding the location of cases." Concerning how users should share information, participants felt that "there should not be a requirement for data sharing" and that users should be provided with layers or levels of data sharing that users can choose to opt into, allowing users to exercise autonomy.

Institutional and Governance Considerations for the Future EN Systems

Concerning where to place EN programs in organizational structures, more than half opted for a hybrid model (eg, sharing responsibilities between emergency response and data informatics; 13/22, 59.1%), leaving out emergency response (5/22, 22.7%), data systems-information (2/22, 9.1%), and contact tracing programs (2/22, 9.1%) as potential sites. Other

XSL•FO
RenderX

proposed locations included disease surveillance and biosecurity within the military structure. Another response suggested a more restrained approach: "discussing the benefits and challenges with agencies, as it requires an interdisciplinary approach."

On governance for EN programs, some acknowledge that it should be on a country level and that "a regulatory framework should be in place before the next public health emergency." A considerable proportion of participants (13/21, 61.9%) expressed that implementation levels should be nationwide, with an option for states to customize.

International and national institutions that should be reached out to for future collaborations include the WHO (14/34, 41.2%), National Institutes of Health (9/34, 26.4%), and Federal Emergency Management Agency (8/34, 23.5%), as well as other agencies such as Centers for Disease Control and Prevention, Homeland Security, Food and Drug Administration, European Center for Disease Prevention and Control, and central hubs from around the world. When asked to rank institutions that should be tasked with building trust, the following order emerged: (1) public health departments (79 points), (2) EN system-technologies (66 points), (3) state government (64 points), (4) technology companies (54 points), and (5) federal governments (52 points). With respect to organizing working groups (such as DNA in the United States), respondents expressed that these groups should continue to meet and plan for the next steps, as "it fosters relationships with private and public partners." The most effective way to use these groups is to leverage their expertise in designing next-generation tools.

The importance of working with smartphone manufacturers was emphasized by the large majority of the group (17/22, 77.2%), as some pointed out that in addition to tapping into their technical expertise, "it helps build relationships and improve operations that benefit actors." When examining the possibility of working with other companies aside from Google and Apple, participants expressed that mobile phone and app companies should be included, though "many settings are regulated by these two companies...both need to open up their ecosystem and infrastructure that allows others to work," highlighting the support from big tech companies.

Discussion

Summary of Key Findings

To our knowledge, this is the first study to systematically capture insights from experts on the future of EN systems. Although others convened group discussions for similar purposes, those attempts were not as comprehensive as ours, were not peer-reviewed, and were convened or led by industry personnel [35]. Our study suggests that during the COVID-19 pandemic, the main barriers to EN program implementation included community mistrust in technology companies and mistrust in EN systems, followed by a lack of political support from policymakers. Airborne diseases, such as measles, influenza, respiratory syncytial virus, and tuberculosis, were considered the most appropriate group of diseases to incorporate into EN systems. Additional features to consider for the next app include pathogen transmissibility, protective measures, and the burden of disease in the community. Some of the most important features and system characteristics for future apps were the capacity to self-report at-home tests, access to additional public health information, general venue notifications, and differential messaging based on risk level. Finally, for public policy purposes, the study highlights the need to further discuss the benefits and challenges with involved agencies, as it requires collaboration within and across institutions, and efforts should be made to foster working groups to enhance EN systems.

Public Health and Policy Implications and Recommendations

This study serves multiple purposes. The first is to describe barriers experienced during the COVID-19 pandemic that require action for improving performance and facilitating the implementation of future EN systems. The second is to provide a landscape of improvements to be considered for the next batch of EN systems. These improvements build upon the experience participants shared and make a call for intersectoral collaboration across public health professionals, the technology industry, and policymakers with key roles in advancing preparedness and response against infectious diseases across all government levels, civil society, and academia. Third, it advocates the preservation of current technologies to enhance public health functions by expanding its scope of work to other infectious diseases and by promoting working groups to leverage expertise (regardless of geographical dimensions), develop guidelines, and foster networks across sectors. Finally, it lays out recommendations for decision-makers, public health professionals, and the technology industry for enhancing the next generation of EN systems as follows:

First, efforts to strengthen EN systems should address key implementation and adoption barriers. Legal barriers should be identified and resolved first to facilitate timely technological development and deployment. At the same time, community trust should be built via close partnership with public health authorities at local, state, and national levels. Furthermore, close collaboration with technology companies is critical to develop strategic work plans aimed at strengthening user trust and increasing app adoption.

Second, to enhance the functionality and effectiveness of EN apps, participants recommended incorporating several core epidemiological and clinical features. These include information about the transmissibility and natural history of the pathogen, clear guidance on protective actions that individuals can take—both pharmacological and nonpharmacological—as well as information about available economic resources to support those actions. Including contextual information about disease burden was also considered important. In addition, experts supported expanding EN systems beyond COVID-19 to other high-burden pathogen groups, including airborne diseases, vaccine-preventable diseases, and sexually transmitted infections.

Third, several system-level and user-level characteristics were identified as priorities for future development. EN systems should enable general venue notifications and provide differential messaging based on individual risk level. The use of metadata to improve risk estimation and the integration of

XSL•FO
RenderX

other surveillance data sources to enhance analytics and forecasting were also recommended. At the user level, functionality should allow individuals to self-report at-home test results, request additional public health information, incorporate individual-level risk factors, and receive notifications that clearly indicate exposure risk level (eg, high or low). Participants also emphasized that users should be able to turn off the app regardless of Bluetooth status. To improve access, EN systems should incorporate easier methods for app activation, expand access to other public health resources, and improve accessibility by offering multiple language options.

Finally, to support sustainable implementation, experts recommended fostering structured working groups that include local and national public health departments and other relevant organizations. These groups should create subcommittees focused on specific topics, use standardized approaches to estimate key performance indicators, organize workshops to equip implementation teams with practical tools, and systematically document and share lessons learned. Finally, planning, implementation, and evaluation of EN systems should involve coordinated engagement with public health departments, EN technology developers, state and local governments, technology companies, and, when appropriate, federal agencies.

Comparison to Prior Work

Although no prior peer-reviewed study has examined the same breadth of EN implementation barriers, disease applicability, future system functionality, and governance priorities in one place, some studies addressed certain aspects of our findings. For instance, other studies also noted that adoption of EN and DCT apps was influenced by privacy concerns, trust, perceived usefulness, and usability [6,37]. Furthermore, another paper emphasized that EN deployment depended on close collaboration among technology companies, developers, and public health authorities, while fragmented rollout and limited data collection complicated effectiveness assessment [7]. Our findings also reinforce the conclusions from the best-practice guidance and implementation experiences from several European countries: EN apps should rapidly connect users to testing, quarantine guidance, and other public health resources, and risk-stratified messaging and actionable recommendations may improve both engagement and public health utility [8]. Implementation experiences in Finland also suggested that EN systems function best when integrated into broader public health response activities rather than operating as stand-alone notification tools [38]. Our findings on future system characteristics suggest integration with surveillance data—extending prior work that proposed standardized evaluation metrics for EN systems [13] and suggested that EN activity may support short-term outbreak forecasting and surveillance functions [39]. Finally, our recommendation for structured working groups and stronger cross-sector coordination is reinforced by lessons from collaborative work in the United States during the COVID-19 pandemic, such as the MIT Lincoln Laboratory initiatives [35] and DNA [40], which emphasized iterative assessment, governance mechanisms, and sustained collaboration among PHAs, policymakers, and technology partners.

Caveats and Limitations

Although this study is the first to systematically gather and synthesize data to outline actionable steps for strengthening EN systems, the findings should be interpreted with caution. First, our sampling strategy and modest sample size limit generalizability, particularly because most participants were based in the United States and only about half reported direct implementation experience.

Second, this study employed a type 3 Delphi approach to gather expert perspectives on enhancing EN systems as part of infectious disease response. By design, our modified Delphi approach included only a single round, which constrained our ability to capture a more iterative or comprehensive assessment. Unlike a type 4 Delphi that seeks structured consensus over multiple rounds, our goal was to surface areas of opportunity and identify priorities as health systems prepare for emerging and reemerging infectious disease outbreaks. Additional work is needed to further refine these priorities and inform next steps for advancing EN technologies.

Finally, although our analysis was not explicitly structured around an implementation science framework, the barriers and facilitators identified map closely onto commonly used domains—such as contextual readiness, stakeholder engagement, and system integration—and can guide future research that applies models like the CFIR (Consolidated Framework for Implementation Research) or RE-AIM (Reach, Effectiveness, Adoption, Implementation, and Maintenance) to digital public health interventions.

Conclusion

EN systems have the potential to play a broader role in supporting ongoing public health responses to outbreaks beyond COVID-19. Our findings suggest that there are opportunities to adapt EN technologies into more flexible tools for future outbreaks. Achieving this mandate will require revisiting existing protocols, improving core functionalities, and addressing practical barriers that limit usability and integration with routine public health activities and cross-sectoral coordination.

Despite its limitations, our study helped identify priorities as a starting point for a roadmap to strengthen EN systems as part of preparedness for emerging and re-emerging infectious diseases. With system improvements, stronger coordination between technology and public health leaders, and continued evaluation, EN technologies can evolve into a more reliable and sustainable component of modern outbreak response.

Acknowledgments

The authors thank the California Department of Public Health, University of California-San Francisco's Institute for Global Health Sciences Master of Science Program. The authors used generative AI tools, including Grammarly and ChatGPT (OpenAI,

XSL•FO
RenderX

version 4.0), for limited language editing and proofreading under full human supervision. AI was not used for data analysis, interpretation, or generation of scientific content. All responsibility for the final manuscript rests with the authors. Use of AI was limited to selected sections of the Introduction and Discussion.

Funding

This project was supported by the Resource Allocation Program (RAP) award, a University of California, San Francisco, internal funding mechanism.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: MM
Data curation: CAM-L, NS
Formal analysis: CAM-L, NS
Methodology: MM
Project administration: NS, AB
Resources: AB
Supervision: MM
Visualization: CAM-L
Writing – original draft: MM, CAM-L
Writing – review & editing: NS, RP, WBL

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary study quality characteristics and Delphi survey questions.
[DOCX File, 76 KB - publichealth_v12i1e77763_app1.docx]

References

1. Coronavirus disease (COVID-19): contact tracing. World Health Organization. URL: https://www.who.int/news-room/questions-and-answers/item/coronavirus-disease-covid-19-contact-tracing [accessed 2025-01-10]
2. Amicosante AMV, Rosso A, Bernardini F, et al. COVID-19 contact tracing strategies during the first wave of the pandemic: systematic review of published studies. JMIR Public Health Surveill 2023 Jun 23;9:e42678. [doi: 10.2196/42678] [Medline: 37351939]
3. Nebeker C, Kareem D, Yong A, Kunowski R, Malekinejad M, Aronoff-Spencer E. Digital exposure notification tools: a global landscape analysis. PLOS Digit Health 2023 Sep;2(9):e0000287. [doi: 10.1371/journal.pdig.0000287] [Medline: 37656671]
4. Alghamdi NS, Alghamdi SM. The role of digital technology in curbing COVID-19. Int J Environ Res Public Health 2022 Jul 7;19(14):8287. [doi: 10.3390/ijerph19148287] [Medline: 35886139]
5. Kim H. COVID-19 apps as a digital intervention policy: a longitudinal panel data analysis in South Korea. Health Policy 2021;125(11):1430-1440. [doi: 10.1016/j.healthpol.2021.07.003] [Medline: 34301408]
6. Bannister-Tyrrell M, Chen M, Choi V, Miglietta A, Galea G. Systematic scoping review of the implementation, adoption, use, and effectiveness of digital contact tracing interventions for COVID-19 in the Western Pacific Region. Lancet Reg Health West Pac 2023 May;34:100647. [doi: 10.1016/j.lanwpc.2022.100647] [Medline: 37256207]
7. Bair H, Wanger JD, Shah NR. A brief history of exposure notification during the COVID-19 pandemic in the United States, 2020-2021. Public Health Rep 2022;137(2_suppl):23S-28S. [doi: 10.1177/00333549221099533] [Medline: 35674242]
8. O'Connell J, Abbas M, Beecham S, et al. Best practice guidance for digital contact tracing apps: a cross-disciplinary review of the literature. JMIR mHealth uHealth 2021 Jun 7;9(6):e27753. [doi: 10.2196/27753] [Medline: 34003764]
9. Exposure notification. Apple Developer. URL: https://developer.apple.com/exposure-notification/ [accessed 2025-04-28]
10. Privacy-preserving contact tracing. Apple. URL: https://www.apple.com/covid19/contacttracing [accessed 2025-04-28]
11. Pegollo L, Maggioni E, Gaeta M, Odone A. Characteristics and determinants of population acceptance of COVID-19 digital contact tracing: a systematic review. Acta Biomed 2021 Dec 10;92(S6):e2021444. [doi: 10.23750/abm.v92iS6.12234] [Medline: 34889313]
12. Elmokashfi A, Sundnes J, Kvalbein A, et al. Nationwide rollout reveals efficacy of epidemic control through digital contact tracing. Nat Commun 2021 Oct 11;12(1):5918. [doi: 10.1038/s41467-021-26144-8] [Medline: 34635661]

XSL•FO
RenderX

13. Aronoff-Spencer E, Nebeker C, Wenzel AT, et al. Defining key performance indicators for the California COVID-19 Exposure Notification System (CA Notify). Public Health Rep 2022;137(2_suppl):67S-75S. [doi: 10.1177/00333549221129354] [Medline: 36314660]
14. Prakash AV, Das S. Explaining citizens' resistance to use digital contact tracing apps: a mixed-methods study. Int J Inf Manage 2022 Apr;63:102468. [doi: 10.1016/j.ijinfomgt.2021.102468] [Medline: 36540570]
15. Apple and Google partner on COVID-19 contact tracing technology. News from Google. URL: https://blog.google/inside-google/company-announcements/apple-and-google-partner-covid-19-contact-tracing-technology/ [accessed 2025-01-11]
16. Burdinski A, Brockmann D, Maier BF. Understanding the impact of digital contact tracing during the COVID-19 pandemic. PLOS Digit Health 2022 Dec;1(12):e0000149. [doi: 10.1371/journal.pdig.0000149] [Medline: 36812611]
17. Pozo-Martin F, Beltran Sanchez MA, Müller SA, Diaconu V, Weil K, El Bcheraoui C. Comparative effectiveness of contact tracing interventions in the context of the COVID-19 pandemic: a systematic review. Eur J Epidemiol 2023 Mar;38(3):243-266. [doi: 10.1007/s10654-023-00963-z] [Medline: 36795349]
18. Hossain AD, Jarolimova J, Elnaiem A, Huang CX, Richterman A, Ivers LC. Effectiveness of contact tracing in the control of infectious diseases: a systematic review. Lancet Public Health 2022 Mar;7(3):e259-e273. [doi: 10.1016/S2468-2667(22)00001-9] [Medline: 35180434]
19. Unim B, Zile-Velika I, Pavlovska Z, et al. The role of digital tools and emerging devices in COVID-19 contact tracing during the first 18 months of the pandemic: a systematic review. Eur J Public Health 2024 Jul 1;34(Supplement_1):i11-i28. [doi: 10.1093/eurpub/ckae039] [Medline: 38946444]
20. Sujarwoto S, Maharani A. Facilitators and barriers to the adoption of mHealth apps for COVID-19 contact tracing: a systematic review of the literature. Front Public Health 2023;11:1222600. [doi: 10.3389/fpubh.2023.1222600] [Medline: 38145061]
21. Thomas Craig KJ, Rizvi R, Willis VC, Kassler WJ, Jackson GP. Effectiveness of contact tracing for viral disease mitigation and suppression: evidence-based review. JMIR Public Health Surveill 2021 Oct 6;7(10):e32468. [doi: 10.2196/32468] [Medline: 34612841]
22. Barrat A, Cattuto C, Kivelä M, Lehmann S, Saramäki J. Effect of manual and digital contact tracing on COVID-19 outbreaks: a study on empirical contact data. J R Soc Interface 2021 May;18(178):20201000. [doi: 10.1098/rsif.2020.1000] [Medline: 33947224]
23. Salathé M, Althaus C, Anderegg N, et al. Early evidence of effectiveness of digital contact tracing for SARS-CoV-2 in Switzerland. Swiss Med Wkly 2020 Dec;150:w20457. [doi: 10.4414/smw.2020.20457] [Medline: 33327003]
24. Daniore P, Nittas V, Ballouz T, et al. Performance of the Swiss digital contact-tracing app over various SARS-CoV-2 pandemic waves: repeated cross-sectional analyses. JMIR Public Health Surveill 2022 Nov 11;8(11):e41004. [doi: 10.2196/41004] [Medline: 36219833]
25. Vogt F, Haire B, Selvey L, Katelaris AL, Kaldor J. Effectiveness evaluation of digital contact tracing for COVID-19 in New South Wales, Australia. Lancet Public Health 2022 Mar;7(3):e250-e258. [doi: 10.1016/S2468-2667(22)00010-X] [Medline: 35131045]
26. Wymant C, Ferretti L, Tsallis D, et al. The epidemiological impact of the NHS COVID-19 app. Nature 2021 Jun;594(7863):408-412. [doi: 10.1038/s41586-021-03606-z] [Medline: 33979832]
27. Chow BWK, Lim YD, Poh RCH, et al. Use of a digital contact tracing system in Singapore to mitigate COVID-19 spread. BMC Public Health 2023 Nov 16;23(1):2253. [doi: 10.1186/s12889-023-17150-0] [Medline: 37974135]
28. Jeon S, Rainisch G, Harris AM, et al. Estimated cases averted by COVID-19 digital exposure notification, Pennsylvania, USA, November 8, 2020–January 2, 2021. Emerg Infect Dis 2023 Feb;29(2):426-430. [doi: 10.3201/eid2902.220959] [Medline: 36639132]
29. Alanoca S, Guetta-Jeanrenaud N, Ferrari I, Weinberg N, Çetin RB, Miailhe N. Digital contact tracing against COVID-19: a governance framework to build trust. Int Data Priv Law 2021 May 5;11(1):3-17. [doi: 10.1093/idpl/ipab001]
30. Health at a glance: Latin America and the Caribbean 2020. : OECD Publishing; 2020 URL: https://www.oecd.org/content/dam/oecd/en/publications/reports/2020/06/health-at-a-glance-latin-america-and-the-caribbean-2020_4f138987/6089164f-en.pdf [accessed 2022-12-08]
31. How contact tracing and warning apps helped during the COVID-19 pandemic. European Commission. URL: https://commission.europa.eu/strategy-and-policy/coronavirus-response/travel-during-coronavirus-pandemic/contact-tracing-and-warning-apps-during-covid-19_en [accessed 2025-01-11]
32. Niederberger M, Spranger J. Delphi technique in health sciences: a map. Front Public Health 2020;8:457. [doi: 10.3389/fpubh.2020.00457] [Medline: 33072683]
33. Humphrey-Murto S, Varpio L, Gonsalves C, Wood TJ. Using consensus group methods such as Delphi and nominal group in medical education research. Med Teach 2017 Jan;39(1):14-19. [doi: 10.1080/0142159X.2017.1245856] [Medline: 27841062]
34. Diamond IR, Grant RC, Feldman BM, et al. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. J Clin Epidemiol 2014 Apr;67(4):401-409. [doi: 10.1016/j.jclinepi.2013.12.002] [Medline: 24581294]

XSL•FO
RenderX

35. Automated exposure notification for COVID-19. MIT Lincoln Laboratory. URL: https://www.ll.mit.edu/r-d/publications/automated-exposure-notification-covid-19 [accessed 2025-04-28]
36. Qualtrics XM. URL: https://www.qualtrics.com/ [accessed 2025-04-28]
37. Oyibo K, Sahu KS, Oetomo A, Morita PP. Factors influencing the adoption of contact tracing applications: systematic review and recommendations. Front Digit Health 2022;4:862466. [doi: 10.3389/fdgth.2022.862466] [Medline: 35592459]
38. Pihlajamäki M, Wickström S, Puranen K, Helve O, Yrttiaho A, Siira L. Implementing and maintaining a SARS-CoV-2 exposure notification application for mobile phones: the Finnish experience. JMIR Public Health Surveill 2023 Jul 13;9:e46563. [doi: 10.2196/46563] [Medline: 37440286]
39. Aronoff-Spencer E, Mazrouee S, Graham R, et al. Exposure notification system activity as a leading indicator for SARS-CoV-2 caseload forecasting. PLOS One 2023;18(8):e0287368. [doi: 10.1371/journal.pone.0287368] [Medline: 37594936]
40. Digital Notification Alliance (DNA). URL: https://dna-enx.org/ [accessed 2026-05-08]

ABBREVIATIONS

DCT: digital contact tracing
DNA: Digital Notification Alliance
EN: exposure notification
GAEN: Google and Apple Exposure Notification system
PHAs: public health agencies
UCSF: University of California, San Francisco
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 19.May.2025; peer-reviewed by G Carot-Sans, S Mazrouee; revised version received 31.Jul.2026; accepted 02.Aug.2026; published 16.Sep.2026.

Please cite as:
Malekinejad M, Mendez-Lizarraga CA, Skehan N, Bercasio A, Palekar R, Lober WB
Transforming a Pandemic Tool Into a Public Health Asset: Modified Delphi Study
JMIR Public Health Surveill 2026;12:e77763
URL: https://publichealth.jmir.org/2026/1/e77763
doi:10.2196/77763

© Mohsen Malekinejad, Cesar Arturo Mendez-Lizarraga, Nicole Skehan, Alyssa Bercasio, Rakhee Palekar, William B Lober. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 16.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluation of the Finnish Exposure Notification App "Koronavilkku" According to Key Performance Indicators: Cross-Sectional Study

Wioleta Kitowska[1,2], MSc; Mika Pihlajamäki[2], MSc; Sara Wickström[2], MA; Kaija Tolmunen[3], MSc; Otto Helve[2], MD, PhD; Juha Mykkänen[2], PhD; Timothée Dub[2], MD; Lotta Siira[2], PhD

[1]Field Epidemiology Path (EPIET), ECDC Fellowship Programme, European Centre for Disease Prevention and Control, Stockholm, Sweden
[2]Finnish Institute for Health and Welfare (THL), Mannerheimintie 166, Helsinki, Uusimaa, Finland
[3]Gofore Oyj, Helsinki, Finland

Corresponding Author:
Lotta Siira, PhD
Finnish Institute for Health and Welfare (THL), Mannerheimintie 166, Helsinki, Uusimaa, Finland

Abstract

Background: Mobile exposure notification apps (ENAs) can play a significant role in future pandemics. Identifying lessons learned and areas for improvement from such apps, particularly within specific social and cultural contexts, is therefore crucial.

Objective: This study aimed to evaluate the Finnish ENA "Koronavilkku," according to chosen key performance indicators to assess the ENA's use and performance during the COVID-19 pandemic in Finland, and to identify strengths and improvement areas.

Methods: Using available data on Finland's ENA, we defined metrics for four key performance indicators: ENA use extent, enablers and barriers to use, the ENA's influence on user behavior, and overall acceptability. We performed an evaluation study combining system data and data from a survey of 4061 respondents conducted in April 2022. A sentiment analysis was performed on open-ended responses.

Results: Koronavilkku ENA peak coverage reached approximately 53% of active smartphone users shortly after launch. Use declined over time, particularly among individuals diagnosed with COVID-19. User engagement remained high—among ENA users who received a token, the proportion that entered their token into the ENA was more than 50% during the whole study period. Males, individuals older than 34 years, and those living alone were less likely to use the ENA, while those with higher education and households with children were more likely to be ENA users. Almost 70% (64,642/94,280) of SARS-CoV-2–positive users reported not receiving a token from health authorities. Users reported higher adherence to pandemic guidance than nonusers, with prevalence ratios ranging from 1.24 (95% CI 1.19 - 1.29) to 1.65 (95% CI 1.45 - 1.91). Descriptive analyses suggested higher reported adherence among partially compliant users after exposure notifications, whereas little change was observed among noncompliant individuals. Overall acceptability was relatively high, with 68% of respondents using it at least once. However, 49% of those users eventually discontinued use. Our sentiment analysis of open-ended feedback suggests that users expressed more positive attitudes toward the ENA compared to nonusers. However, a substantial portion of the feedback reflected neutral or negative sentiments.

Conclusions: Based on observed ENA use trends, we recommend maintaining long-term public interest and engagement in ENAs during prolonged emergencies. Tailored communication should be developed encouraging uptake among demographic groups with lower adoption rates. Initial studies during app rollout could inform these strategies. A major barrier to the ENA's performance was token issuance; earlier and more systematic automation of token distribution in Finland could have enabled more timely exposure notifications and reduced bottlenecks. Future ENAs should define evaluation criteria during development, with periodic assessments measuring effectiveness and informing improvements. A deeper analysis of open-ended feedback provided for the Koronavilkku ENA using advanced language models could provide additional insights into user perceptions and concerns. By addressing these areas, future digital contact tracing tools may improve uptake, acceptability, and integration into pandemic response efforts.

(JMIR Public Health Surveill 2026;12:e89472) doi:10.2196/89472

KEYWORDS

COVID-19; pandemic; cross-sectional; Finland; digital health; mobile app; preparedness

XSL•FO
RenderX

Introduction

Finland reported its first imported case of COVID-19 in late January 2020, and by mid-March the number of confirmed cases had risen to over 400. In response, on March 16 the government declared a state of emergency, introducing wide-ranging measures including the closure of schools and public facilities, restrictions on public gatherings, and limitations on cross-border travel. Some of these measures were gradually eased by May 2020, although restrictions were reintroduced in autumn as case numbers increased again. Finland began its vaccination campaign in late December 2020, and by November 2021, approximately 80% of the population aged 12 years and older had been vaccinated. As vaccination coverage expanded and the epidemiological situation improved, restrictions were progressively lifted, with all major measures ending by the end of June 2022 [1-3].

Mobile exposure notification apps (ENAs) using digital proximity tracing were introduced during the COVID-19 pandemic in many countries worldwide. In Finland, the national ENA "Koronavilkku" was launched on August 31, 2020, and discontinued on June 1, 2022. The aim of the Finnish ENA was intended to supplement conventional contact tracing and serve as an early warning system for the individual user in case of potential exposure to a SARS-CoV-2–positive person in order to provide them with behavioral guidance mitigating the chance of further transmission [4]. ENA use was voluntary and free of charge.

The ENA was developed using a decentralized approach. Finnish legislation and the European architecture and protocols used for ENAs placed strong emphasis on privacy, data security, and data safety. Due to these constraints, very limited performance information was gathered through the ENA to describe its potential effectiveness under the stated aims [4]. In addition to the ENA, a dedicated system was provided to all health care provider organizations in Finland to enable token issuance for infected app users, which allowed users to activate exposure notifications to potential contacts. The organizations, in turn, granted system access to health professionals who sent tokens to ENA users.

Several studies have evaluated the uptake, acceptability, and potential impact of ENAs during the COVID-19 pandemic, with mixed findings. Modeling and empirical studies suggest that even moderate uptake may contribute to reducing transmission, although effectiveness depends on multiple factors, including user engagement and timeliness of notifications [5-10]. More recent research has emphasized variability in how effectiveness is defined and measured, as well as the importance of contextual factors in determining impact [11]. Research across European countries has also highlighted variability in adoption and acceptability, with sociodemographic factors, trust, and perceived usefulness influencing uptake [12-15]. Implementation challenges, such as delays in notification processes and integration with public health systems, can limit the overall performance of these tools [16,17]. However, comprehensive evaluations combining system-level performance indicators with user-reported experiences remain limited. Our aim was to evaluate the Finnish ENA according to chosen key performance indicators (KPIs) to gain insight into the ENA's use and performance during the COVID-19 pandemic in Finland, as well as to identify strengths and areas for improvement.

Methods

Overview

We evaluated the ENA during its period of use (August 31, 2020-June 1, 2022) based on the following set of KPIs:

1. extent of ENA use,
2. enablers and barriers to ENA use,
3. influence of the ENA on user behavior, and
4. general acceptability of the ENA.

Operational Definitions

Active ENA use: the device has the ENA installed, is operational, and communicates at least once every 24 hours with the backend server to retrieve keys (see "key-matching process").

Keys: randomly generated, encrypted identifiers exchanged through Bluetooth Low Energy between participating devices that are near one another, according to defined distance and duration criteria. These keys provided the basis for anonymous "matching" of devices and the subsequent triggering of exposure notifications.

Key-matching process: phones regularly downloaded keys from users who had received a positive SARS-CoV-2 result and entered their token into the ENA from the backend server and checked them against their own history of contacts for matches. If a match was found, it triggered the exposure notification mechanism based on the calculated exposure risk value (ERV).

Token: one-time code issued by the system used by health care providers upon receiving a positive COVID-19 diagnosis. The token was issued manually to ENA users by health care professionals, although automatic issuance was introduced in some parts of Finland from June 2021. When entered into the ENA by the user, the user's keys were sent to the backend server. Entering a token locked the ENA and it no longer tracked exposures.

Exposure notification: a message issued by the ENA alerting the user that they had recently been in proximity to an individual who has received a positive SARS-CoV-2 finding and has shared their result via the app.

Data Collection

We accessed system activity data from the ENA provider and the ENA's backend. In Finland, the ENA providers were the Social Insurance Institution (Kela) and the Finnish Institute for Health and Welfare (THL). Variables retrieved included: number of downloads (data available for August 31, 2020-February 8, 2021), number of active users (data available for December 7, 2020-April 24, 2022), number of tokens issued (data available for September 7, 2020-May 29, 2022), and number of tokens used (data available for September 7, 2020-May 29, 2022).

XSL•FO
RenderX

Furthermore, we used data from the cross-sectional online survey conducted from April 13 to April 22, 2022, by a Finnish polling company among respondents aged 15‐79 years, based on a convenience sample of the company's and its subcontractors' panelists. The survey included 4061 respondents, corresponding to approximately 0.07% of the Finnish population (~5.5 million). Weights were assigned to all respondents based on the distribution of the Finnish population with regard to sex, age, and geographic area (data from Statistics Finland – population by sex and age and major regions in mainland Finland as of December 31, 2020).

Variables from the survey included respondents' sociodemographic characteristics, ENA use, ENA removal, reasons for nonactivation, reasons for removal, receipt of exposure notifications, token receipt, speed of token delivery, use of tokens, adherence to recommendations after notification, reasons for discontinuation of adherence to recommendations, consulting a health care provider after exposure notification, general behavior during the pandemic, COVID-19 diagnosis, and overall feedback on the ENA.

We used data on population size by sex, age, region, and year during 2020‐2022 from Statistics Finland. We retrieved the daily number of COVID-19 cases reported to the National Infectious Diseases Register (NIDR) from August 31, 2020, to June 1, 2022 [18].

Koronavilkku Survey Variables

We created the following terms based on variables from the Koronavilkku survey:

- ENA user or nonuser: we divided respondents into two main groups: ENA users and nonusers based on the question "Have you downloaded the Koronavilkku application on your phone?" ENA users were further classified as reporting current or previous use of the ENA. Nonusers were classified as reporting never using the ENA or downloading the ENA but never activating it.
- Exposure notification guidance*:* along with the exposure notification, the ENA gave guidance that was in line with general national pandemic guidelines (Figure 1). Based on the guidance, we used the following terms: self-observation of well-being, avoiding social interactions, hand hygiene, wearing a mask, avoiding errands, cough hygiene, working from home.

XSL•FO
RenderX

Figure 1. Guidance shown to users of the Finnish exposure notification app upon receipt of an exposure notification from March 4, 2021, to June 16, 2021. The instructions were updated in line with the national guidelines [4].

XSL•FO
RenderX

The survey included two questions about adherence to the above-mentioned exposure notification guidance: general behavior during the pandemic and after receiving an exposure notification. Respondents were able to mark their adherence to each of the seven points on a scale of "never, part-time, entire time, don't recall/unsure."

Data Analysis

We assessed each of the four KPIs (extent of ENA use, enablers and barriers to ENA use, ENA influencing user behavior, and general acceptability) according to a set of predefined metrics specified in Multimedia Appendix 1. The analysis was performed using R software (version 4.2.1; R Foundation for Statistical Computing) in RStudio (Posit Software, PBC). *P* values less than .05 were considered significant.

This study used two main data sources for analysis: (1) ENA backend data, used to describe the extent of ENA use; and (2) data from the Koronavilkku survey, used to assess enablers and barriers to ENA use, ENA influence on user behavior, and general acceptability. The data sources for each metric are specified in Multimedia Appendix 1.

Survey weights based on age, sex, and region were applied to descriptive analyses of survey data (excluding sentiment analysis) to improve representativeness of the Finnish population. Prevalence ratios (PRs) were estimated using log-binomial regression models fitted to unweighted data. A large language model (LLM) was used for sentiment analysis of open-ended feedback in the survey. The sentiment analysis was performed using Python 3 (version 3.10.12, Python Software Foundation) and the Llama model version 3.1 70B (model identifier: llama3.1:70b-instruct-q4_K_M; Meta). Finnish input texts were analyzed natively without translation. The response was classified using a prompt that asked for feedback to be in one of four categories: positive, neutral, negative, and not understandable feedback. Prompt rules asked it to return only one of these categories and never to output the original clause or anything other than the classes.

Sentiment Analysis Validation

To compare Llama classifications to human classifications, we randomly sampled 100 feedback responses, and one reviewer added manual annotations about human-evaluated sentiment. Overall agreement was 63% (Cohen κ=0.486) and considered "moderate." The main pattern found was that Llama classified positives and negatives well but misclassified many human-labeled neutrals as "not understandable feedback." These were mostly policy or feature suggestions without clear opinion on the current version or just statements that the users did not have anything to say.

The validation results can be found in Table S1 in Multimedia Appendix 2.

Generative AI

Generative AI was used solely as an analytic tool for sentiment classification (LLM) and was not used to draft or edit the paper.

Ethical Considerations

The app user data were collected and analyzed as part of THL's mandate; therefore, no ethical approval was required for this study. Upon being recruited to the survey study, participants gave their informed consent. The participants did not receive remuneration for their participation; the polling company had a monthly lottery, where participants were eligible to win prizes, such as gift cards. All analysis was performed on anonymized data; no personal identifying data was used. The legal basis for the ENA and THL's role related to it is detailed in the Communicable Diseases Act 1227/2016 HE 101/2020.

Results

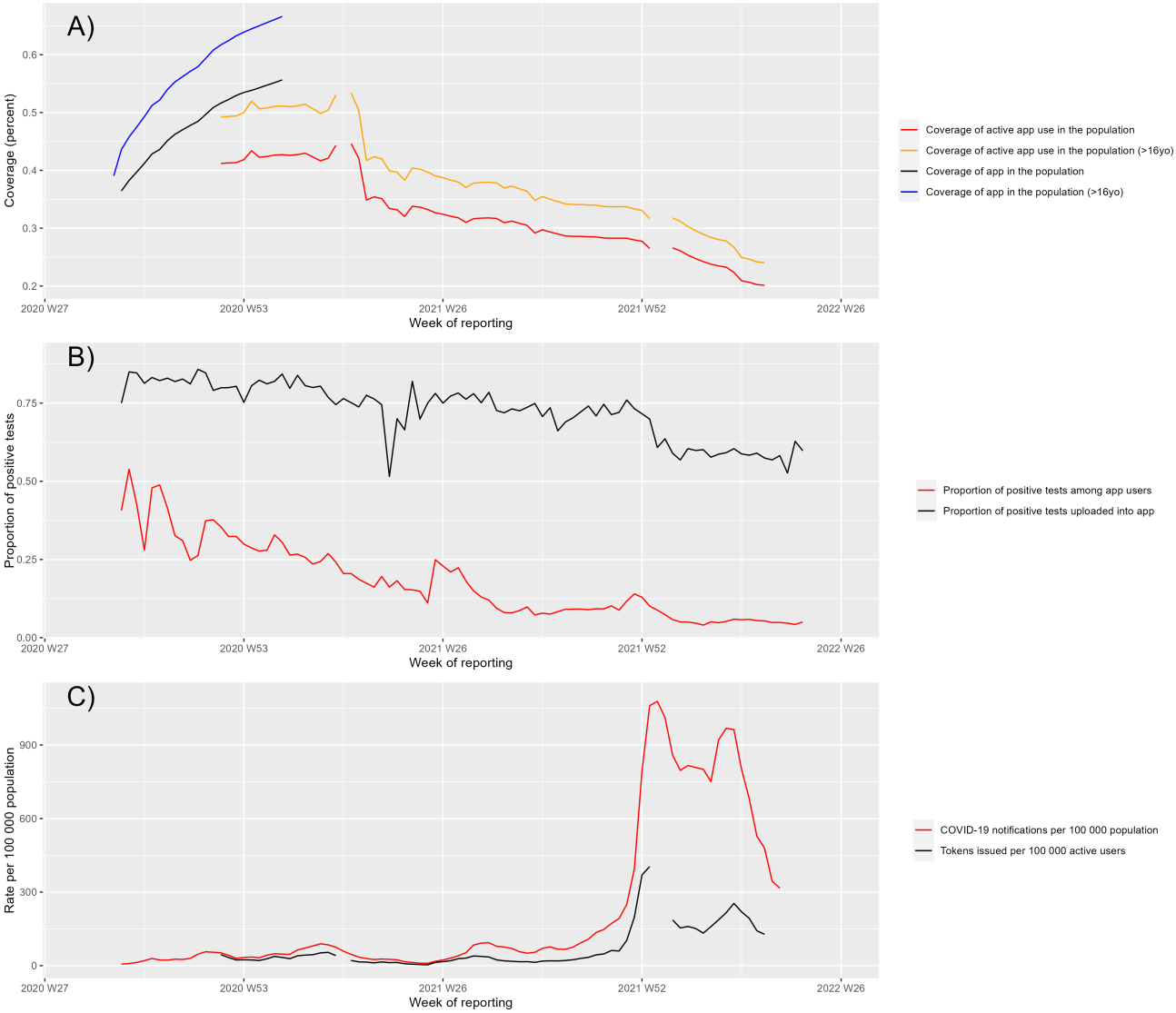
Extent of ENA Use

The extent of ENA use was assessed based on backend ENA data.

We used the number of downloads as a proxy for active ENA use at the introduction of the app, after which we switched to monitoring active ENA use. In late 2020, the ENA was actively used by around 40% - 50% of the Finnish population. In 2021, the active coverage peaked around weeks 12 - 14 at 53% among the Finnish population older than 16 years old (45% in the general population), after which it declined to less than 30% in early 2022 (Figure 2A).

XSL•FO
RenderX

Figure 2. (A) Proportion of the population that have downloaded and actively use the exposure notification app in Finland August 31, 2020, to April 24, 2022. Results shown for the whole population of Finland and for the population of Finland >16 years old. (B) Proportion of positive SARS-CoV-2 tests among ENA users among all SARS-CoV-2 positive tests in Finland August 31, 2020, to April 24, 2022; proportion of positive SARS-CoV-2 tests uploaded into the ENA (used tokens) among all ENA users that received a token (issued tokens) in Finland August 31, 2020, to April 24, 2022. (C) Rate of SARS-CoV-2 notifications per 100,000 population and rate of tokens issued per 100,000 active users.



At all times during the pandemic, fewer than 50% of all people who tested positive for COVID-19 both used the ENA and entered their tokens. The proportion was highest at the launch of the ENA but declined to less than 30% before the start of 2021. In total, 94,280 tokens were issued during the ENA's lifetime, of which 69% (64,642/94,280) were entered into the ENA (Figure 2B).

The weekly proportion of ENA users who received a token and entered it into the ENA to alert other users of potential exposure reached 86% (586/683) in late 2020 and did not dip below 50% at any point. During the period of ENA use, the mean was 73% (SD 9%) and the median 75% (IQR 67%-80%), indicating high user engagement (Figure 2B).

The rate of positive tests among ENA users relative to the rate of positive tests reported in the general population was similar at the beginning, when the COVID-19 notification rate was relatively low. The rate of tokens issued did not scale in the same way as the notification rate during its highest peaks (Figure 2C).

Enablers and Barriers to ENA Use

Enablers and barriers to ENA use were assessed based on data from the Koronavilkku survey.

Sociodemographic Characteristics of ENA Users vs. Non-ENA Users

The prevalence of ENA nonusers was higher among males than females (PR 0.93, 95% CI 0.89 - 0.97) and in the age groups above 34 years. There were no significant regional differences between ENA users and nonusers (Table 1).

XSL•FO
RenderX

Table . Sociodemographic characteristics of exposure notification app users vs non-exposure notification app users based on a survey conducted in Finland from April 13 to 22, 2022 (n=4061, weighted n=4334). Results weighted according to age, sex, and regional distribution of the Finnish population.

Characteristics	Non-ENA[a] users, (N=1321), n (%)	Non-ENA users, Weighted (n=1407)[b]	ENA users, (N=2740), n (%)	ENA users, Weighted (n=2927)[b]	PR[c]	95% CI	*P* value
Sex							
Female	617 (47)	46% (43.3 - 48.8)	1441 (53)	51.9% (50 - 53.8)	Ref.	__[d]	—
Male	704 (53)	54% (51.2 - 56.7)	1299 (47)	48.1% (46.2 - 50)	0.93	0.89 - 0.97	<.001
Age group (years)							
15 - 24	140 (11)	11.3% (9.7 - 13.2)	397 (14)	15.2% (13.8 - 16.6)	Ref.	—	—
25 - 34	189 (14)	13.8% (12.1 - 15.8)	495 (18)	17.4% (16.1 - 18.9)	0.98	0.91 - 1.05	.54
35 - 49	394 (30)	25.3% (23.1 - 27.6)	741 (27)	22.8% (21.4 - 24.4)	0.88	0.83 - 0.94	<.001
50 - 64	354 (27)	27.2% (24.9 - 29.7)	639 (23)	23.3% (21.8 - 25)	0.87	0.81 - 0.93	<.001
65 - 79	244 (18)	22.3% (19.9 - 24.9)	468 (17)	21.2% (19.6 - 23)	0.89	0.83 - 0.96	.002
Region							
Southern Finland	292 (22)	20.9% (18.8 - 23.2)	594 (22)	21% (19.5 - 22.6)	Ref.	—	—
Helsinki-Uusimaa	385 (29)	28.4% (26 - 30.9)	910 (33)	32.6% (30.9 - 34.4)	1.05	0.99 - 1.11	.11
Western Finland	332 (25)	25.8% (23.5 - 28.3)	659 (24)	24.4% (22.8 - 26)	0.99	0.93 - 1.06	.8
Northern and Eastern Finland	312 (24)	24.9% (22.5 - 27.4)	577 (21)	22% (20.4 - 23.6)	0.97	0.91 - 1.04	.34
Education							
Primary or national school	212 (16)	16.5% (14.6 - 18.7)	283 (10)	10.8% (9.6 - 12)	Ref.	—	—
High school	180 (14)	13.5% (11.7 - 15.5)	417 (15)	15.1% (13.8 - 16.5)	1.22	1.12 - 1.34	<.001
Vocational, technical or business school	435 (33)	32.5% (30 - 35.1)	679 (25)	24.6% (23 - 26.3)	1.07	0.98 - 1.17	.16
University of applied sciences	173 (13)	12.7% (11 - 14.6)	431 (16)	15.1% (13.8 - 16.5)	1.25	1.14 - 1.37	<.001
College	137 (10)	10.8% (9.2 - 12.7)	348 (13)	13.2% (12 - 14.6)	1.26	1.14 - 1.38	<.001
University or college	184 (14)	14% (12.2 - 16)	582 (21)	21.2% (19.7 - 22.8)	1.33	1.22 - 1.45	<.001
Occupation							
Employee	386 (30)	28.5% (26.1 - 31.1)	727 (27)	25.3% (23.7 - 26.9)	Ref.	—	—
Functionary	70 (5.4)	5.2% (4.1 - 6.5)	288 (11)	10% (9 - 11.2)	1.23	1.15 - 1.31	<.001
Higher functionary, expert	90 (7)	6.7% (5.5 - 8.2)	339 (13)	11.9% (10.7 - 13.1)	1.21	1.13 - 1.29	<.001
Leading position, entrepreneur	80 (6.2)	6.1% (4.9 - 7.6)	250 (9.2)	9% (7.9 - 10.1)	1.16	1.07 - 1.25	<.001

XSL•FO
RenderX

Characteristics	Non-ENA[a] users, (N=1321), n (%)	Non-ENA users, Weighted (n=1407)[b]	ENA users, (N=2740), n (%)	ENA users, Weighted (n=2927)[b]	PR[c]	95% CI	*P* value
Pensioner	326 (25)	28.2% (25.7 - 30.9)	584 (22)	25% (23.3 - 26.8)	0.98	0.92 - 1.05	.59
Stay-at-home-parent	31 (2.4)	2.4% (1.7 - 3.4)	33 (1.2)	1.2% (0.8 - 1.7)	0.79	0.6 - 0.98	.06
Student	108 (8.4)	8.3% (6.9 - 10)	286 (11)	10.5% (9.4 - 11.8)	1.11	1.03 - 1.19	.005
Unemployed	193 (15)	14.2% (12.4 - 16.1)	187 (6.9)	6.6% (5.8 - 7.6)	0.75	0.67 - 0.84	<.001
Other	5 (0.4)	0.4% (0.1 - 0.9)	15 (0.6)	0.5% (0.3 - 0.9)	1.15	0.82 - 1.39	.29
Unknown	32	—	31	—	—	—	—
Household situation							
Single	532 (40)	40.2% (37.6 - 43)	917 (33)	33.9% (32.1 - 35.7)	Ref.	—	—
Couple	298 (23)	23.2% (21 - 25.6)	684 (25)	25.6% (24 - 27.4)	1.10	1.04 - 1.16	.001
Household with children	257 (19)	18% (16 - 20.1)	685 (25)	23% (21.4 - 24.5)	1.15	1.09 - 1.21	<.001
Other adult household	234 (18)	18.6% (16.5 - 20.8)	454 (17)	17.5% (16.1 - 19)	1.04	0.97 - 1.11	.22

[a]ENA: exposure notification app.

[b]weighted % (95% CI).

[c]PR: prevalence ratio.

[d]N/A: not applicable.

ENA users had a higher level of education more often than nonusers. Compared to those that had primary or national school-level education, ENA use was more prevalent among those that finished high school (PR 1.22, 95% CI 1.12 - 1.34), a university of applied sciences (PR 1.25, 95% CI 1.14 - 1.37), college (PR 1.26, 95% CI 1.14 - 1.38), or university or college (PR 1.33, 95% CI 1.22 - 1.45; Table 1).

Compared to employees, ENA use was more prevalent among those holding functionary, higher functionary, and leadership positions as well as among students. Nonusers were more often unemployed (PR 0.75, 95% CI 0.67 - 0.84; Table 1).

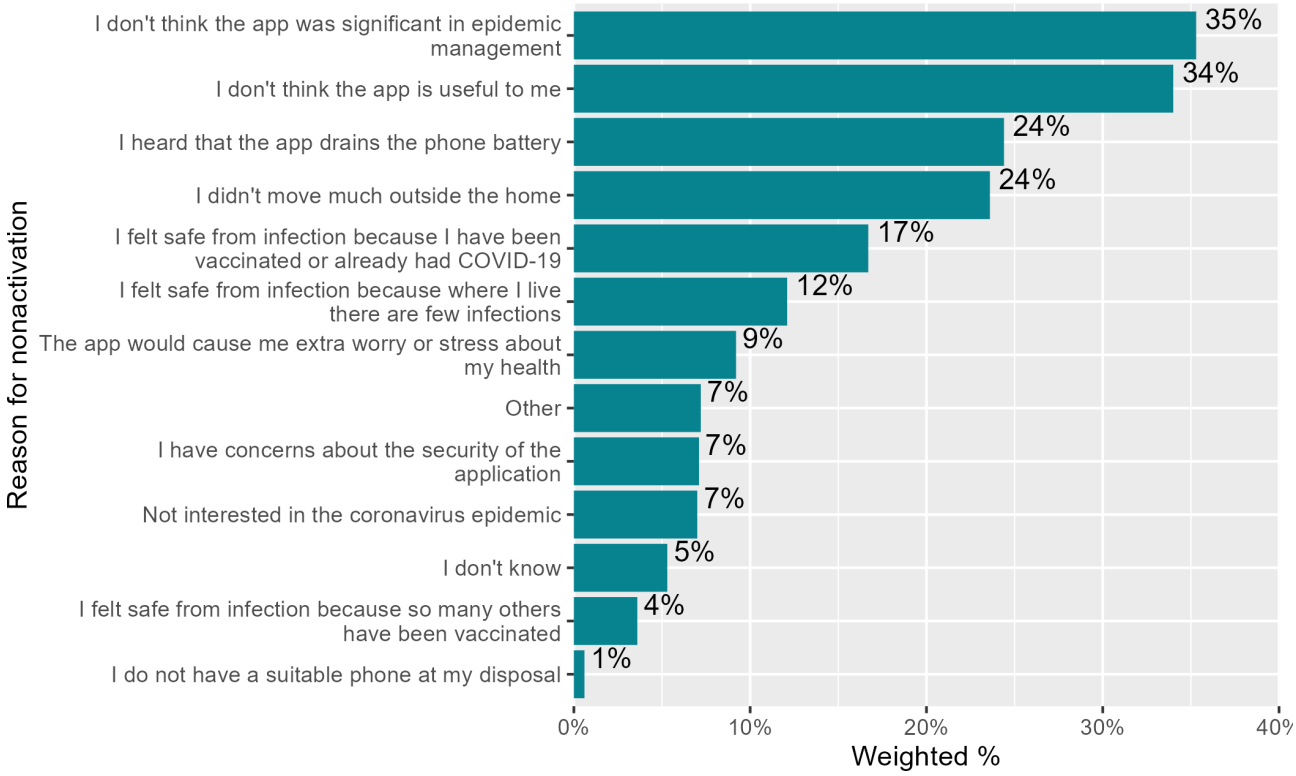
ENA users were more often couples (PR 1.10, 95% CI 1.04 - 1.16) or living in households with children (PR 1.15, 95% CI 1.09 - 1.21; Table 1).

Reasons for Nonactivation of the ENA

The most frequent reasons for not activating the ENA after download were perceived lack of usefulness of the ENA for pandemic control (35%) or for oneself (34%). Around 24% of respondents indicated phone battery life and staying at home as reasons for nonactivation (Figure 3).

XSL•FO
RenderX

Figure 3. Weighted frequency distribution of reasons for not activating the exposure notification app (n=149, weighted n=163). Data from survey conducted in Finland from April 13 - 22, 2022. Results weighted according to age, sex, and regional distribution of the Finnish population. Exact estimates with 95% CIs are provided in Table S2 in Multimedia Appendix 2.



General Behavior During Pandemic of ENA Users vs Non-ENA Users

ENA users reported at least part-time adherence to general pandemic guidance more often than nonusers, ranging from avoiding errands (PR 1.24, 95% CI 1.19 - 1.29) to wearing a mask (PR 1.65, 95% CI 1.45 - 1.91; Table 2). However, these associations may reflect self-selection and cannot be interpreted as causal.

XSL•FO
RenderX

Table . Adherence to exposure notification guidance of exposure notification app users compared to non-exposure notification app users (n=4061, weighted n=4334). Results weighted according to age, sex, and regional distribution of the Finnish population.

Exposure notification guidance	Non-ENA[a] users[b]	Non-ENA users adhering (yes)[c], n (%)	Non-ENA users adhering weighted (yes)[d]	ENA users[b]	ENA users adhering (yes)[c], n (%)	ENA users adhering weighted (yes)[d]	PR[e]	95% CI	*P* value (chi-square test)
Working from home	1254/1333	452 (36)	35.5% (32.9 - 38.2)	2646/2821	1488 (56)	55.1% (53.1 - 57)	1.30	1.24 - 1.36	<.001
Self-observation of well-being	1281/1365	937 (73)	72.8% (70.3 - 75.3)	2687/2869	2307 (86)	85.1% (83.7 - 86.5)	1.35	1.26 - 1.46	<.001
Avoiding social interactions	1297/1381	986 (76)	76.3% (73.9 - 78.6)	2705/2889	2340 (87)	86.3% (84.9 - 87.6)	1.30	1.21 - 1.41	<.001
Wearing a mask	1310/1395	1142 (87)	87.6% (85.7 - 89.3)	2716/2900	2594 (96)	95.6% (94.7 - 96.3)	1.65	1.45 - 1.91	<.001
Hand hygiene	1298/1383	1182 (91)	91.5% (89.8 - 92.9)	2709/2893	2601 (96)	96.1% (95.3 - 96.8)	1.43	1.25 - 1.65	<.001
Avoiding errands	1297/1383	232 (18)	18% (16 - 20.3)	2693/2878	849 (32)	31.2% (29.5 - 33)	1.24	1.19 - 1.29	<.001
Cough hygiene	1288/1372	1183 (92)	92.1% (90.5 - 93.4)	2700/2884	2584 (96)	95.7% (94.9 - 96.4)	1.31	1.16 - 1.5	<.001

[a]ENA: exposure notification app.

[b]n/weighted N.

[c]Reported at least part time adherence.

[d]weighted % (95% CI).

[e]PR: prevalence ratio.

Token Use

Based on survey data, of those ENA users reporting a positive SARS-CoV-2 result (n=881, weighted n=904), close to 70% did not receive a token and 1% reported being diagnosed abroad. Of the 30% receiving a token (n=259, weighted n=269), most (79%) reported entering it into the ENA. For approximately 72% of users, tokens were delivered in under 48 hours from diagnosis, and of those, 51% were received within 24 hours. However, delay in token delivery by 3 - 4 days was reported by 17%, 5 - 7 days by 3%, and over 7 days by 2%. Six percent did not recall the speed of token delivery.

These survey-based token-use estimates are not directly comparable with token use proportions calculated based on ENA backend data, as the backend metrics reflect all eligible users over the full study period, whereas survey metrics reflect self-reported experiences in a cross-sectional respondent sample.

Influence of ENA on User Behavior

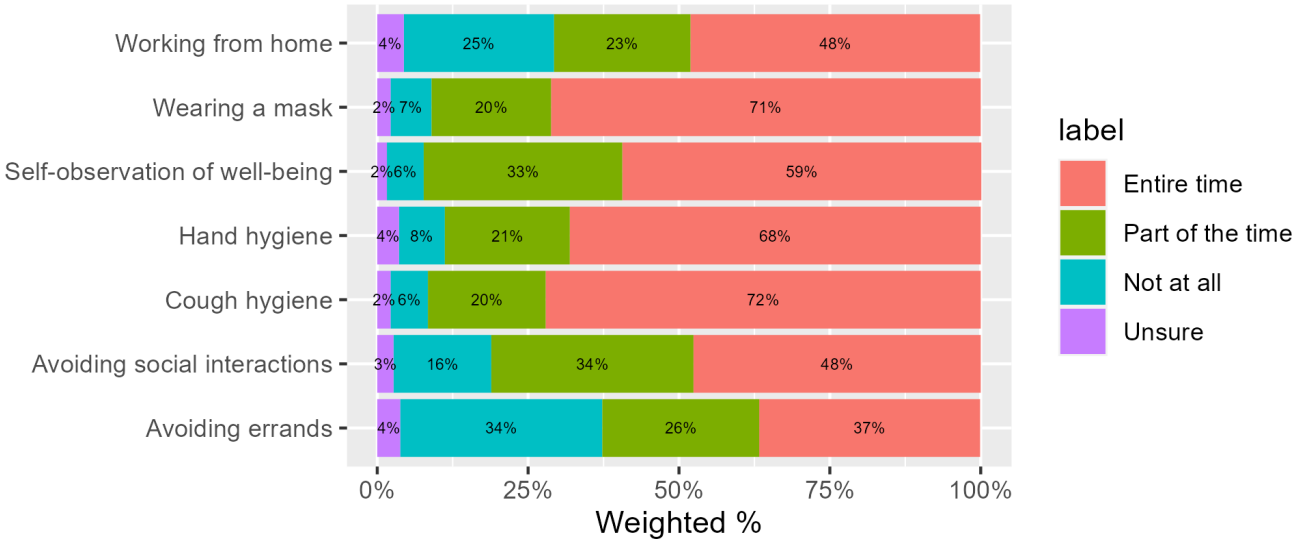
Influence of ENA on user behavior was assessed based on data from the Koronavilkku survey.

Adherence to Recommendations After Notification

The majority of ENA users who received a notification of potential exposure adhered at least part-time to all guidance provided in the ENA. The least adhered to recommendations were “to avoid running errands” (34% reported no adherence), “work from home” (25% no adherence), and “avoid social interactions” (16% no adherence; Figure 4).

XSL•FO
RenderX

Figure 4. Weighted frequency distribution of adherence to pandemic behavioral recommendations after receiving an exposure notification in the exposure notification app (n=644, weighted n=670). Data from survey conducted in Finland from April 13 to 22, 2022. Results weighted according to age, sex, and regional distribution of the Finnish population. Exact estimates with 95% CIs are provided in Table S3 in Multimedia Appendix 2.



Change in Adherence to Recommendations in General During Pandemic vs After Notification

Looking at ENA users (n=2740, weighted n=2927), Bowker's test of symmetry showed significant differences between general adherence to pandemic recommendations versus after receiving an exposure notification for working from home (χ^2_3=75.5; *P*<.001), self-observation of well-being (χ^2_3=18.3; *P*<.001), avoiding social interactions (χ^2_3=114.6; *P*<.001), wearing a mask (χ^2_3=26.9; *P*<.001), and avoiding errands (χ^2_3=74.4; *P*<.001).

Adherence to hand hygiene (χ^2_3=0.6; *P*=.90) and cough hygiene (χ^2_3=0.6; *P*=.90) guidelines showed no differences.

Further analysis of transition patterns in adherence to pandemic recommendations suggested the following:

1. More transitions from none to full-time adherence for "avoiding errands" (odds ratio [OR] 6.08).
2. Transitions from part-time to full-time adherence were more frequent than the reverse for "working from home" (OR 3.53), "self-observation" (OR 2.07), "avoiding social interactions" (OR 6.05), "wearing a mask" (OR 2.5), "avoiding errands" (OR 4.04).
3. Transitions from part-time to no adherence were observed for "working from home" (OR 5.61), "avoiding social interactions" (OR 2.19).

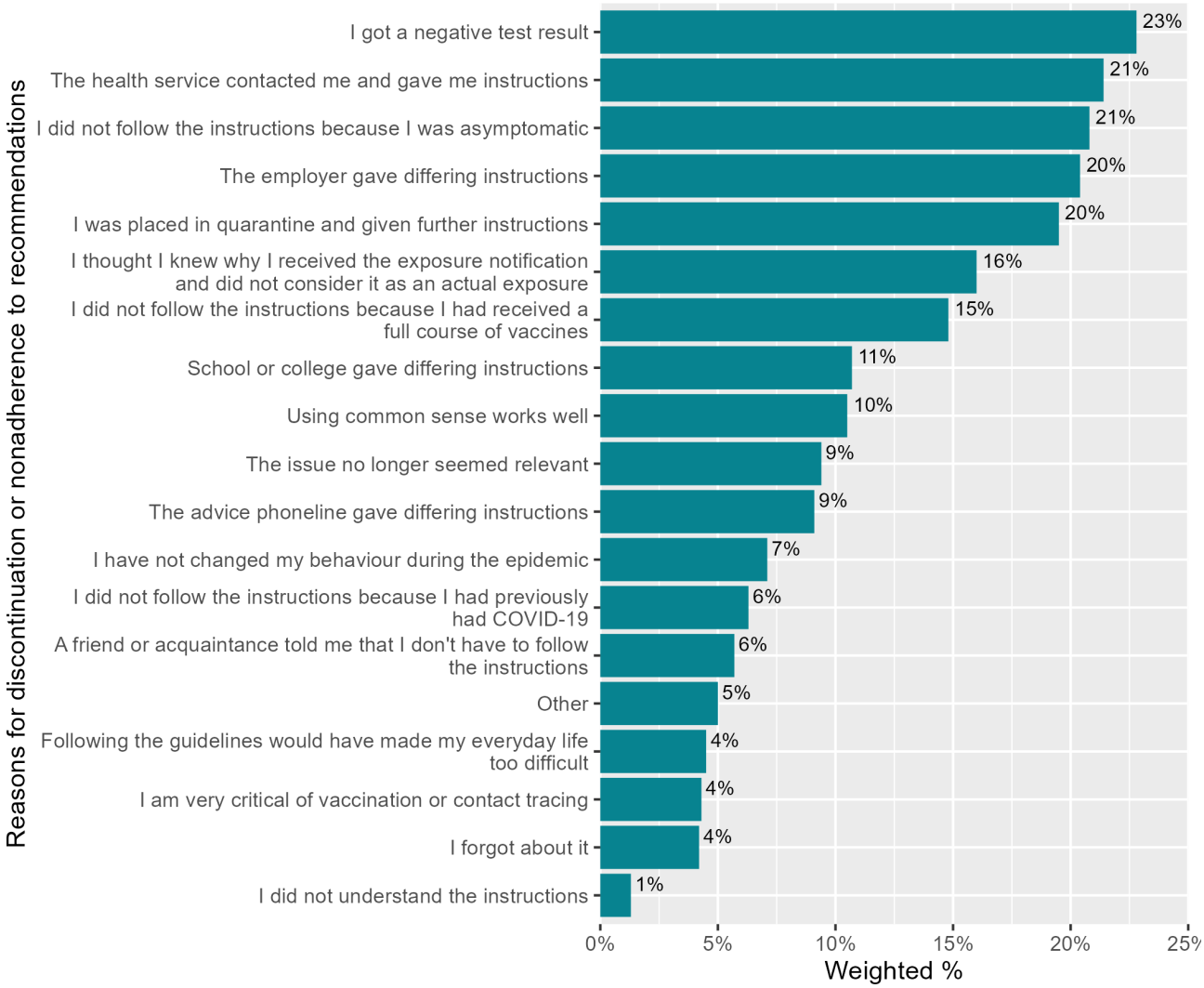
These transition ORs are descriptive and should not be interpreted as inferential tests of specific pairwise changes.

Reasons for Discontinuation of Adherence to Recommendations

A negative test was the most stated reason for not adhering, partly adhering, or discontinuing adherence to exposure notification guidance (23%). Other commonly reported reasons included receiving instructions from health care (21%), being asymptomatic (21%), receiving instructions from employer (20%), or being put in mandatory quarantine as defined by the Communicable Diseases Act (20%; Figure 5).

XSL•FO
RenderX

Figure 5. Weighted frequency distribution of reasons for discontinuation or nonadherence to pandemic behavioral recommendations after receiving an exposure notification in the exposure notification app (n=534, weighted n=556). Data from survey conducted in Finland from April 13 to 22, 2022. Results weighted according to age, sex, and regional distribution of the Finnish population. Exact estimates with 95% CIs are provided in Table S4 in Multimedia Appendix 2.



Contacting Health Care Provider

Among ENA users receiving an exposure notification (n=644, weighted n=670), more than 50% contacted health care afterwards either by phone (25%) or through the Finnish electronic service platform providing symptom checks and guidance called “Omaolo” (DigiFinland Oy; 27%). Around 45% did not report contacting health care and the remaining 3% were unsure.

Acceptability of the ENA

The ENA’s acceptability was assessed based on data from the Koronavilkku survey.

ENA Use

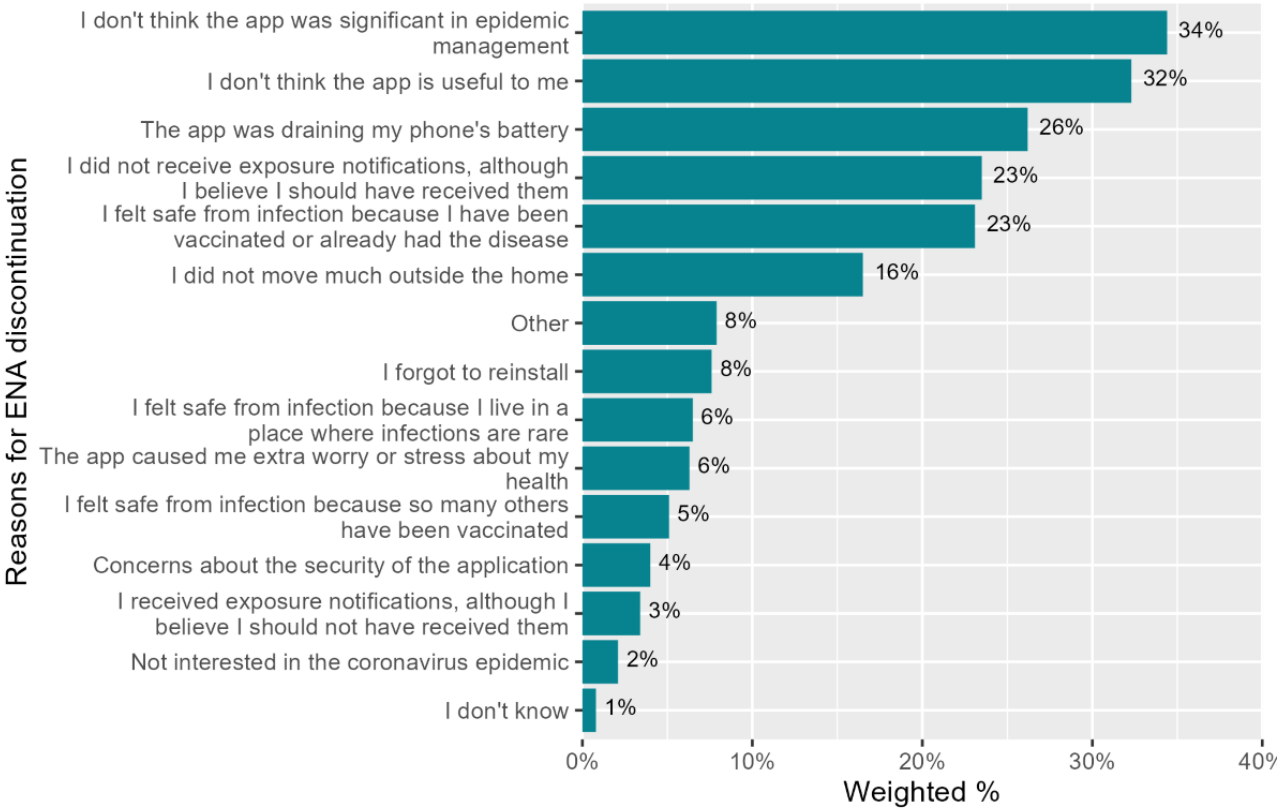
Among 4061 respondents (weighted n=4334), 33% never downloaded or activated the ENA. Another 34% downloaded and actively used it at the time of the survey, while 33% downloaded the ENA but later removed it or did not activate it again on a new mobile phone.

Reasons for Removal

Among former ENA users, the most common reasons for discontinuation of ENA use included not seeing any use for pandemic control (34%) or to oneself (32%). Negative effect on battery life was the third most common reason (26%; Figure 6).

XSL•FO
RenderX

Figure 6. Weighted frequency distribution of reasons for discontinuation of ENA use (n=1370, weighted n=1444). Data from Koronavilkku survey conducted in Finland from April 13 - 22, 2022. Results weighted according to age, sex, and regional distribution of the Finnish population. Exact estimates with 95% CIs are provided in Table S5 in Multimedia Appendix 2. ENA: exposure notification app.

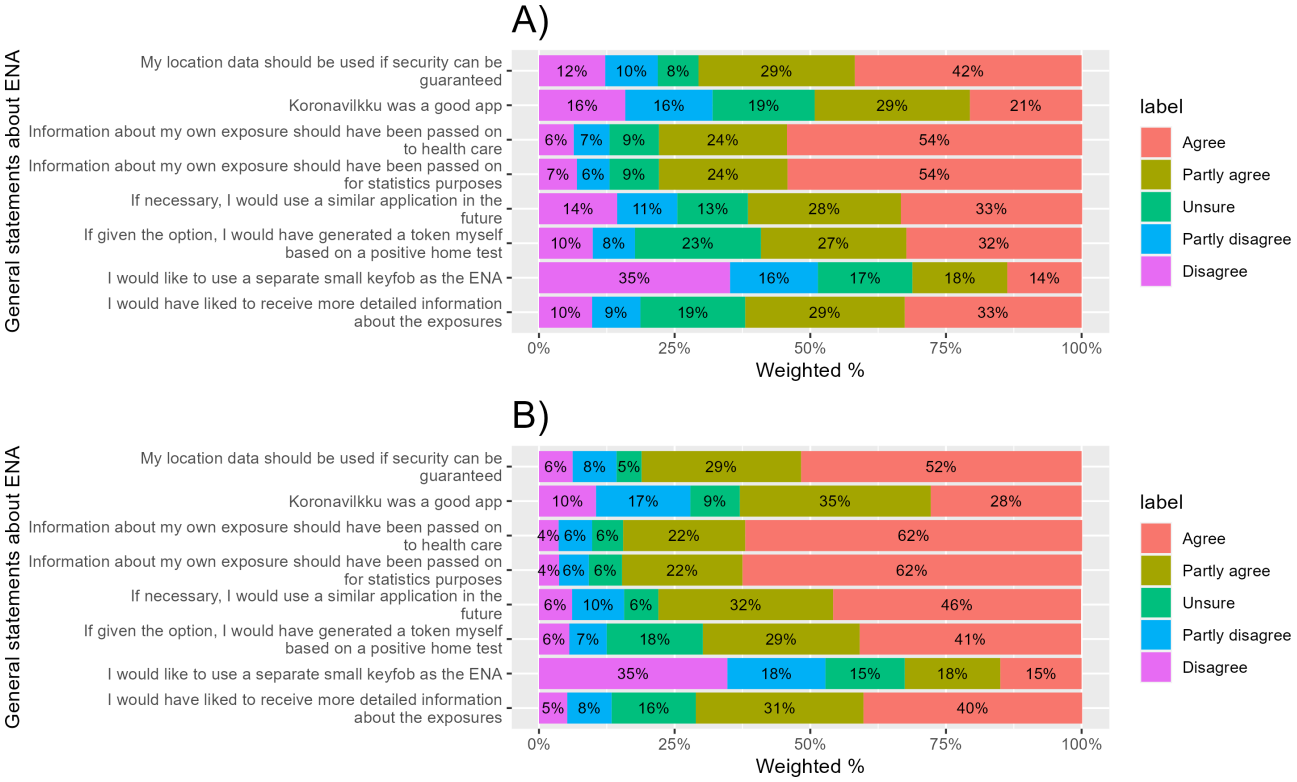


Overall Feedback

Opinions on the quality of the ENA were divided. Almost half (49%) of the respondents at least partly agreed with the statement "Koronavilkku was a good application." Almost two-thirds of respondents (62%) at least partly agreed that they would use the ENA if it was relaunched (Figure 7A).

XSL•FO
RenderX

Figure 7. Weighted frequency distribution of attitude toward general statements about the ENA among (A) all respondents (n=4061, weighted n=4334); (B) among ENA users (n=2740, weighted n=2927). Data from Koronavilkku survey conducted in Finland from April 13 - 22, 2022. Results weighted according to age, sex, and regional distribution of the Finnish population. Exact estimates with 95% CIs are provided in Table S6 in Multimedia Appendix 2. ENA: exposure notification app.



Among ENA users, this percentage was higher: 63% at least partly agreed that it was a good ENA, and 78% stated that they would use it again (Figure 7B).

Between 62% - 78% of respondents stated they would have accepted more data (such as additional information about exposure or localization data) being collected by the ENA or being sent to notify health authorities and for situation monitoring or statistical analyses (Figure 7A). The percentage of respondents who at least partly agreed with all statements was higher among ENA users (Figure 7B).

Sentiment Analysis

Sentiment analysis using an LLM was performed for open-ended feedback provided by 813 respondents, 557 (69%) of whom were ENA users.

In comparison to the survey respondents, open-ended feedback was provided more often by males, aged 15 - 24 years, and persons living in Helsinki-Uusimaa and Western Finland regions (Table 3).

XSL•FO
RenderX

Table . Characteristics of respondents giving open feedback vs those responding to survey overall.

Characteristics	Open feedback respondents (n=813), n (%)	Survey respondents (N=4061), n (%)	*P* value[a]
ENA[b] use	557 (69)	2740 (67)	.51
Sex			<.001
Female	341 (42)	2058 (51)	
Male	472 (58)	2003 (49)	
Age group (years)			<.001
15 - 24	148 (18)	537 (13)	
25 - 34	139 (17)	684 (17)	
35 - 49	224 (28)	1135 (28)	
50 - 64	183 (23)	993 (24)	
65 - 79	119 (15)	712 (18)	
Region			<.001
Southern Finland	162 (20)	886 (22)	
Helsinki-Uusimaa	300 (37)	1295 (32)	
Western Finland	207 (25)	991 (24)	
Northern and Eastern Finland	144 (18)	889 (22)	

[a]Pearson chi-squared test with simulated *P* value (based on 2000 replicates).

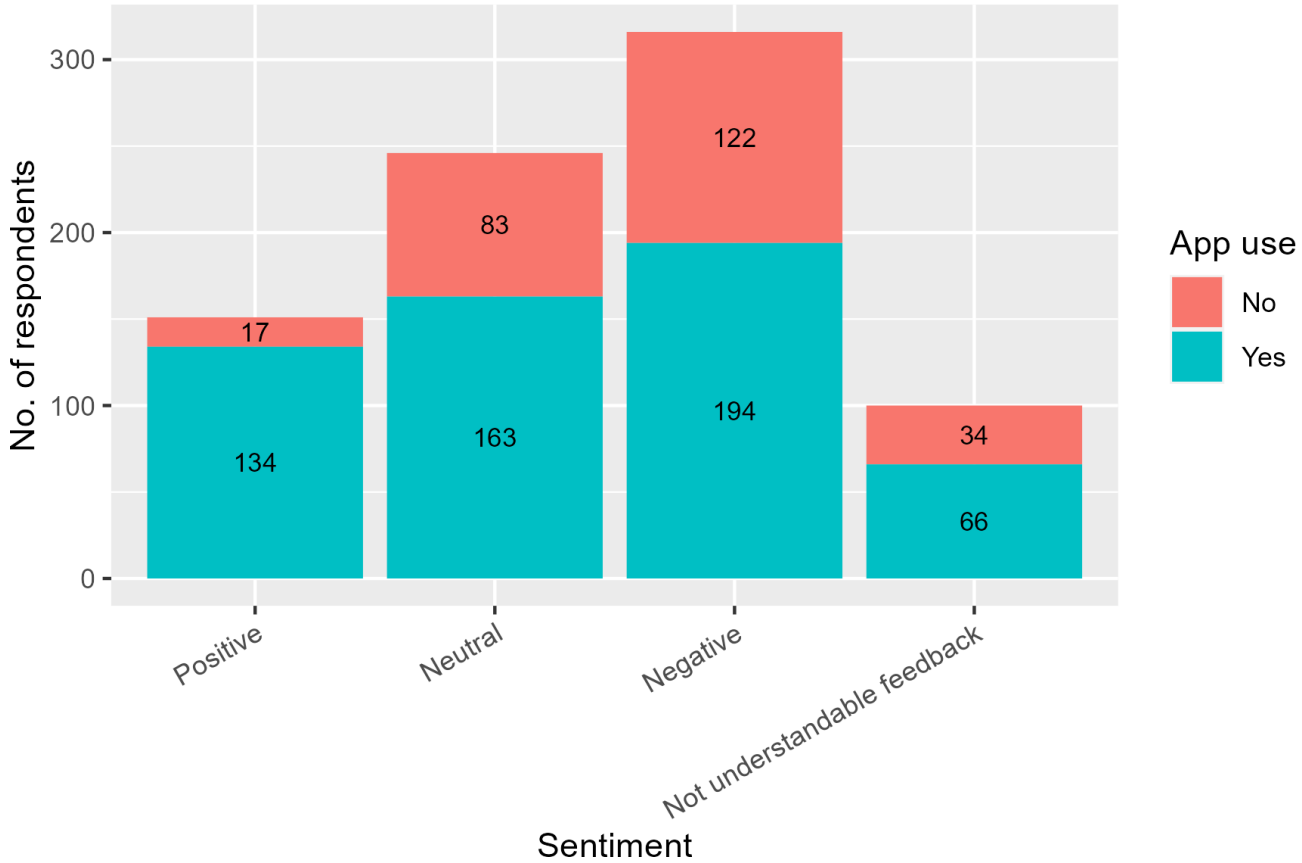
[b]ENA: exposure notification app.

Sentiment analysis determining the emotional tone of the open-ended feedback shows that most provided feedback had a negative sentiment (316/813, 39%), while positive-oriented feedback was provided by 19% (151/813) of respondents. Overall, 12% (100/813) of the provided feedback was not understandable by the model.

A higher percentage of positive responses were provided by ENA users than non-ENA users (24%, 134/557 vs 6.6%, 17/256). Nonuser comments more often expressed negative or neutral sentiments (48%, 122/256 and 32%, 83/256, respectively) than ENA user comments (35%, 194/557, and 29%, 163/557, respectively; Figure 8).

XSL•FO
RenderX

Figure 8. General sentiment analysis of open-ended feedback according to four categories (positive, neutral, negative, and not understandable feedback; n=813).



Discussion

Principal Findings

This study evaluated the Finnish ENA Koronavilkku using predefined KPIs related to uptake, user characteristics, behavioral influence, and acceptability. We found that the ENA achieved high initial uptake but declined over time, with relatively high engagement among users who received tokens. Survey results indicated differences in ENA use across demographic groups and higher self-reported adherence to guidance among ENA users compared to nonusers. While exposure notifications were associated with changes in adherence among some users, barriers such as token issuance and declining engagement over time limited overall system performance.

Comparison With Prior Work

Extent of ENA Use

The Koronavilkku ENA saw widespread adoption during its initial months, with peak active coverage reaching approximately 53% of smartphone users and 45% of the general population. However, as the pandemic progressed, token use declined, especially among people diagnosed with COVID-19, a key target group, with fewer than half of notified COVID-19 cases registered in the ENA, and this proportion fell over time. This trend may reflect reduced app use, reduced token issuance, or a combination of both. Despite this, overall user engagement remained relatively high throughout the ENA's operational period.

ENA effectiveness is closely tied to its level of adoption. Studies indicate that even modest uptake (15% - 30%) can reduce infections [5,6], while others provide more conservative estimates, indicating that 70% - 80% coverage among smartphone users (56% in the general population) is required to control the pandemic effectively [7,8]. Achieving 80% ENA coverage has been linked to a reduction in COVID-19 mortality by up to 85% [9]. Despite these varying results, one could extrapolate that ENA effectiveness is complex and influenced by multiple factors, with partial benefits seen already at lower coverage rates of around 20% [10]. Notably, Koronavilkku achieved the highest active ENA user coverage among European Union and European Economic Area countries [19].

Barriers and Enablers to ENA Use

Our findings identified demographic disparities in ENA use. Males, individuals 35 or older, and those living alone were less likely to use the ENA, while females, those with higher education, and households with children were more likely to be ENA users. However, studies from other European countries revealed differing and sometimes contradictory results. A French study found no significant link between sociodemographic factors and ENA use, except for financial deprivation [12], while a German study reported that males and those 65 years and older were more likely to use the national ENA [13]. Results from the latter study associating higher education levels with increased ENA use were in line with our findings [13]. On the other hand, a different German study indicated that those who were more unwilling to install the national ENA and use it correctly were men and those aged between 30 and 59 years

XSL•FO
RenderX

old [14]. These differences suggest that country-specific cultural and contextual differences also play a significant role in ENA use.

A major barrier to the ENA's performance was token issuance. Over two-thirds of COVID-19-positive users did not receive a token in connection with contact tracing by health care professionals. However, among those who did, most entered their token into the ENA, suggesting strong user motivation when the system functioned properly. Initially, tokens were issued manually by health care contact tracing teams, but high workloads likely caused delays. Recognizing this, Finland implemented partial automation of token issuance in some regions between June and October 2021 [4]. Despite this, issuing exposure tokens remained a significant challenge, particularly during high-incidence periods such as the Omicron wave in late 2021-early 2022. Wider automation of token issuance based on positive results from local systems or a national patient data repository instead of providing a separate system to health care providers would have likely increased the number of issued tokens and shortened token delivery times. However, these possibilities were constrained by a lack of readily available system integrations, risks of not getting complete and reliable result data from the national repository, and short validity periods of tokens. The process for automated token issuance has been described in detail elsewhere [4].

Timely exposure notifications are crucial for effective mitigation. Ferretti et al [20] found that reducing quarantine delays from three to two days significantly curtailed SARS-CoV-2 spread, provided a large portion of the population used the national ENA. Similarly, a Swiss study showed that ENA users quarantined earlier than those not warned by the ENA, though delays in token issuance, linked to manual contact tracing bottlenecks, remained a challenge [16]. A study from Singapore demonstrated that automating public health responses expedited the issuance of quarantine orders compared to manual contact tracing [17].

Our study also explored reasons for nonactivation of the ENA. Consistent with findings from other studies, the most commonly cited reason for not using or uninstalling the ENA was a perceived lack of usefulness [15,21]. Additionally, concerns about battery drain were mentioned as a reason for removing the ENA [21]. These barriers emphasize the need for improved communication around the ENA's intended benefits and for technical solutions to mitigate the effect on battery life.

ENA Influence on User Behavior

Most users who received exposure notifications followed at least some pandemic guidelines recommended by the ENA. ENA users reported adherence at least part-time to general pandemic guidance more often than nonusers, which may reflect self-selection, with ENA users being more health-conscious in general compared to those who did not download the ENA.

Notifications were associated with higher reported adherence among partially compliant users but showed little association among noncompliant individuals. A UK study similarly found high self-reported compliance with self-isolation recommendations following exposure notifications [22].

The only recommendation with declining adherence was "working from home." This likely reflects job constraints (remote work was not feasible for all occupations) or legal arrangements under which official isolation or quarantine was treated as sick leave, with no obligation to work remotely, even though it was highly recommended by the government [23,24]. Over half of notified users sought health care guidance, which may indicate the ENA played a role in prompting further preventive actions. However, about 20% cited a negative COVID-19 test result as their reason for not fully adhering to guidelines.

General Acceptability of the ENA

The ENA's general acceptability was relatively high, with 68% of respondents using it at least once. However, nearly half eventually stopped using it, particularly as the pandemic prolonged. Similarly, users from other countries also reported their ENAs as trustworthy, reliable, and user-friendly [25]. Nevertheless, a study in Belgium found that features focused on informing citizens or facilitating test appointments were viewed more favorably than those related to access and control [15]. In addition, ENA acceptability has been linked to trust in government, scientists, and health care professionals [12].

A substantial proportion of users expressed interest in enhanced features such as automated exposure notifications to health care providers and improved data tracking. However, it is crucial to balance these functionalities with user privacy concerns that can also evolve over time. In the development of the Finnish ENA, addressing such concerns was a priority. At no point did the ENA process or have access to the identities of users or of persons they encountered, irrespective of infection status. Furthermore, location data from encounters were neither collected nor used, despite the technical feasibility of doing so. Notably, a German study found that public acceptance of privacy-invading technologies, such as tracking, was higher at the beginning of the pandemic, but this acceptance waned as the pandemic continued [26].

Our sentiment analysis of open-ended feedback suggests that users generally expressed more positive attitudes toward the ENA compared to nonusers. While a considerable portion of the feedback reflected neutral or negative sentiments, this should be interpreted with caution. Research on response bias suggests that people are more likely to provide open-ended feedback when they have strong opinions or concerns (either very positive or very negative), whereas users with more moderate views may be less inclined to respond without specific prompting [27-29]. As a result, the overall sentiment of the user base may be more positive than the open-ended responses alone suggest. Additionally, there appear to be differences between those who chose to provide open feedback and the broader survey respondents, suggesting that certain demographics were more inclined to share their views. In contrast, a study conducted in Ireland found that the overall perception of their COVID-19 ENA was mostly positive (90%), although they conducted their sentiment analysis manually [30].

XSL•FO
RenderX

Limitations

Several limitations should be acknowledged and considered when interpreting the observed results. First, the survey was conducted late in the pandemic, when at-home testing with over-the-counter SARS-CoV-2 diagnostic tests had become more common [31,32]. At that time, laboratory testing was limited [31,32], and few exposure tokens were being issued due to the requirement for a laboratory-confirmed result. Additionally, the ENA was discontinued on June 1, 2022, but many users had already uninstalled it in the preceding months, which may have introduced recall bias. Furthermore, we were only able to analyze prevalence in this cross-sectional study, not infer causal relationships. Although the survey sample represents a small proportion of the Finnish population, weighting was applied based on age, sex, and region to improve representativeness. However, as participation was voluntary, the sample may overrepresent individuals with higher interest in health or digital tools, potentially limiting generalizability. Social desirability bias could have influenced responses to seem more positive, particularly regarding adherence to pandemic guidelines. Despite this, similar patterns of ENA adoption and user behavior have been reported in other European countries, including Germany and the United Kingdom, where uptake and use were associated with sociodemographic factors, trust in public health systems, and user perceptions [13,22].

The LLM used for sentiment analysis has its own drawbacks. First, the model was unable to categorize all responses, with some classified as “not understandable feedback.” Second, each response was assigned a single sentiment category, even though some answers may have contained a mix of sentiments. Third, it faced challenges in accurately interpreting tone and complex language structures. Furthermore, while recent advancements have significantly improved the model’s performance with the Finnish language, translation and interpretation issues can still occur.

The survey did not collect information on the reasons for not using the ENA, only reasons for not activating it after downloading. This additional information could have been useful to identify potential problem areas and opportunities for improvement.

Finally, questions about adherence to health guidelines after receiving an exposure notification referred only to the most recent notification, limiting our ability to assess changes over time. Our analysis relied on a proxy, comparing general behavior during the pandemic to behavior following an exposure notification, but this approach does not provide a comprehensive picture of behavioral shifts.

Conclusions and Recommendations

This study provides insights into the implementation and performance of a national ENA during a prolonged public health emergency.

Based on observed ENA use trends, regular information campaigns are recommended to maintain long-term public interest and engagement in ENAs during prolonged emergencies.

Furthermore, tailored communication should be developed to encourage uptake among demographic groups with lower adoption rates. Initial studies during ENA rollout could inform these strategies.

Earlier and more systematic automation of token distribution across all regions is critical to enable timely exposure notifications and reduce manual bottlenecks. This should be facilitated by ensuring integrated availability of high-quality laboratory results data from systems using open standard interfaces instead of having to create such mechanisms during the crisis.

Future ENAs should define evaluation criteria during development, with periodic assessments to measure effectiveness and inform improvements. Collecting data on reasons for not downloading the ENA, as well as the timing of installations and removals, could offer valuable insights into retention patterns. Surveys should include questions on factors that would encourage installation, reinstallation, and positive feedback to identify best practices.

From a hardware perspective, nonphone-based contact tracing solutions, such as wearable devices or hardware tokens, could help address battery-life limitations and potentially improve uptake. A deeper analysis of open-ended feedback provided for the Koronavilkku ENA using more advanced language models could provide additional insights into user perceptions and concerns.

Overall, these findings highlight that the success of digital contact tracing tools depends not only on technological functionality, but also on system integration, public trust, and sustained user engagement. Addressing these factors will be essential to support their effective use in future pandemic preparedness and response efforts.

Acknowledgments

We thank all the respondents who generously dedicated their time and effort to respond to our questionnaires.

We extend gratitude to the ECDC EPIET/EUPHEM Fellowship frontline coordinator Barbara Schimmer (based at the National Institute for Public Health and the Environment – RIVM in the Netherlands) for her guidance, input, and continuous support and encouragement.

Funding

The ENA project was funded by the Finnish Government (LTAE III/2020) and a European Union (EU) grant (FINLAND-LC-01590941). Funders had no role in writing or approving this manuscript. WK has been a fellow of the ECDC

XSL•FO
RenderX

Fellowship Programme, supported financially by the European Centre for Disease Prevention and Control (ECDC). The views and opinions expressed herein do not state or reflect those of ECDC. ECDC is not responsible for the data and information collation and analysis and cannot be held liable for conclusions or opinions drawn. The article processing charge will be covered by the Finnish Institute for Health and Welfare.

Data Availability

Data are available on reasonable request to the authors, subject to approval and applicable restrictions.

Authors' Contributions

WK conceived and designed the study, and analyzed and interpreted the data. LS and TD supervised the study and contributed throughout the process. MP performed the sentiment analysis. WK coordinated, drafted, and finalized the manuscript. All authors contributed to the manuscript and approved the final version.

Conflicts of Interest

WK is a fellow of the ECDC Fellowship Programme, supported financially by the European Centre for Disease Prevention and Control. The views and opinions expressed herein do not state or reflect those of ECDC. ECDC is not responsible for the data and information collation and analysis and cannot be held liable for conclusions or opinions drawn.

Multimedia Appendix 1
Key performance indicator metrics.
[PDF File, 157 KB - publichealth_v12i1e89472_app1.pdf]

Multimedia Appendix 2
Sentiment analysis validation, exposure notification app use, adherence to behavioral recommendations, and attitudes toward the exposure notification app.
[DOCX File, 44 KB - publichealth_v12i1e89472_app2.docx]

References

1. COVID-19 cases in the infectious diseases registry. The National Institute for Health and Welfare (THL). URL: https://sampo.thl.fi/pivot/prod/en/epirapo/covid19case/fact_epirapo_covid19case [accessed 2025-09-04]
2. Press releases on the government's decisions. Finnish Government. URL: https://valtioneuvosto.fi/en/decisions/press-releases [accessed 2025-09-04]
3. Tiirinki H, Tynkkynen LK, Sovala M, et al. COVID-19 pandemic in Finland - preliminary analysis on health system response and economic consequences. Health Policy Technol 2020 Dec;9(4):649-662. [doi: 10.1016/j.hlpt.2020.08.005] [Medline: 32874860]
4. Pihlajamäki M, Wickström S, Puranen K, Helve O, Yrttiaho A, Siira L. Implementing and maintaining a SARS-CoV-2 exposure notification application for mobile phones: the Finnish experience. JMIR Public Health Surveill 2023 Jul 13;9:e46563. [doi: 10.2196/46563] [Medline: 37440286]
5. Wymant C, Ferretti L, Tsallis D, et al. The epidemiological impact of the NHS COVID-19 app. Nature 2021 Jun;594(7863):408-412. [doi: 10.1038/s41586-021-03606-z] [Medline: 33979832]
6. Abueg M, Hinch R, Wu N, et al. Modeling the effect of exposure notification and non-pharmaceutical interventions on COVID-19 transmission in Washington state. NPJ Digit Med 2021 Mar 12;4(1):49. [doi: 10.1038/s41746-021-00422-7] [Medline: 33712693]
7. Hernández-Orallo E, Calafate CT, Cano JC, Manzoni P. Evaluating the effectiveness of COVID-19 bluetooth-based smartphone contact tracing applications. Appl Sci 2020;10(20):7113. [doi: 10.3390/app10207113]
8. Hinch R, Probert W, Nurtay A, et al. Effective configurations of a digital contact tracing app: a report to NHSX. London: NHSX; 2020 URL: https://api.semanticscholar.org/CorpusID:216028666 [accessed 2026-08-17]
9. Grimm V, Mengel F, Schmidt M. Extensions of the SEIR model for the analysis of tailored social distancing and tracing approaches to cope with COVID-19. Sci Rep 2021 Feb 18;11(1):4214. [doi: 10.1038/s41598-021-83540-2] [Medline: 33603113]
10. Jenniskens K, Bootsma MCJ, Damen J, et al. Effectiveness of contact tracing apps for SARS-CoV-2: a rapid systematic review. BMJ Open 2021 Jul 12;11(7):e050519. [doi: 10.1136/bmjopen-2021-050519] [Medline: 34253676]
11. Quach HL, Walsh EI, Hoang TNA, Terrett RNL, Vogt F. Effectiveness of digital contact tracing interventions for COVID-19: a systematic scoping review. Public Health 2025 May;242:146-156. [doi: 10.1016/j.puhe.2025.02.019] [Medline: 40068321]
12. Touzani R, Schultz E, Holmes SM, et al. Early acceptability of a mobile app for contact tracing during the COVID-19 pandemic in France: national web-based survey. JMIR Mhealth Uhealth 2021 Jul 19;9(7):e27768. [doi: 10.2196/27768] [Medline: 34086589]

XSL•FO
RenderX

13. Grill E, Eitze S, De Bock F, et al. Sociodemographic characteristics determine download and use of a Corona contact tracing app in Germany-Results of the COSMO surveys. PLoS ONE 2021;16(9):e0256660. [doi: 10.1371/journal.pone.0256660] [Medline: 34473733]
14. Blom AG, Wenz A, Cornesse C, et al. Barriers to the large-scale adoption of a COVID-19 contact tracing app in Germany: survey study. J Med Internet Res 2021 Mar 2;23(3):e23362. [doi: 10.2196/23362] [Medline: 33577466]
15. Walrave M, Waeterloos C, Ponnet K. Reasons for nonuse, discontinuation of use, and acceptance of additional functionalities of a COVID-19 contact tracing App: cross-sectional survey study. JMIR Public Health Surveill 2022 Jan 14;8(1):e22113. [doi: 10.2196/22113] [Medline: 34794117]
16. Ballouz T, Menges D, Aschmann HE, et al. Adherence and association of digital proximity tracing app notifications with earlier time to quarantine: results from the Zurich SARS-CoV-2 cohort study. Int J Public Health 2021;66:1603992. [doi: 10.3389/ijph.2021.1603992] [Medline: 34471402]
17. Chow BWK, Lim YD, Poh RCH, et al. Use of a digital contact tracing system in Singapore to mitigate COVID-19 spread. BMC Public Health 2023 Nov 16;23(1):2253. [doi: 10.1186/s12889-023-17150-0] [Medline: 37974135]
18. Koronavilkku-kokonaisuuden viikkoraportti [Web page in Finnish]. Finnish Institute for Health and Welfare (THL). URL: https://www.thl.fi/episeuranta/vilkku/koronavilkku_viikkoraportti.html [accessed 2026-05-07]
19. European Commission, Directorate-General for Communications Networks, Content and Technology Digital Contact Tracing Study – Study on Lessons Learned, Best Practices and Epidemiological Impact of the Common European Approach on Digital Contact Tracing to Combat and Exit the COVID-19 Pandemic: Publications Office of the European Union; 2022. [doi: 10.2759/146050]
20. Ferretti L, Wymant C, Kendall M, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. Science 2020 May 8;368(6491):eabb6936. [doi: 10.1126/science.abb6936] [Medline: 32234805]
21. Montagni I, Roussel N, Thiébaut R, Tzourio C. Health care students' knowledge of and attitudes, beliefs, and practices toward the French COVID-19 App: cross-sectional questionnaire study. J Med Internet Res 2021 Mar 3;23(3):e26399. [doi: 10.2196/26399] [Medline: 33566793]
22. Dowthwaite L, Fischer J, Perez Vallejos E, et al. Public adoption of and trust in the NHS COVID-19 contact tracing app in the United Kingdom: quantitative online survey study. J Med Internet Res 2021 Sep 17;23(9):e29085. [doi: 10.2196/29085] [Medline: 34406960]
23. Communicable diseases act [Web page in Finnish]. Finlex. 2016. URL: https://www.finlex.fi/en/legislation/2016/1227?language=fin&highlightId=1156046&highlightParams=%7B%22type%22%3A%22BASIC%22%2C%22search%22%3A%221227%2F2016%22%7D [accessed 2026-08-23]
24. Government resolution VM/2021/205 [Webpage in Finnish]. Ministry of Finance Finland. 2021. URL: https://vm.fi/paatos?decisionId=0900908f80778508 [accessed 2025-08-10]
25. Bente BE, van 't Klooster J, Schreijer MA, et al. The Dutch COVID-19 contact tracing app (the CoronaMelder): usability study. JMIR Form Res 2021 Mar 26;5(3):e27882. [doi: 10.2196/27882] [Medline: 33724198]
26. Kozyreva A, Lorenz-Spreen P, Lewandowsky S, et al. Psychological factors shaping public responses to COVID-19 digital contact tracing technologies in Germany. Sci Rep 2021 Sep 21;11(1):18716. [doi: 10.1038/s41598-021-98249-5] [Medline: 34548550]
27. Lewicka M, Czapinski J, Peeters G. Positive - negative asymmetry or 'When the heart needs a reason'. Euro J Social Psych 1992 Sep;22(5):425-434. [doi: 10.1002/ejsp.2420220502]
28. Poncheri R, Lindberg J, Thompson L, Surface E. A comment on employee surveys: negativity bias in open-ended responses. Organ Res Methods 2007;11(3):614-630. [doi: 10.1177/1094428106295504]
29. Miller AL, Dumford AD. Open-ended survey questions: item nonresponse nightmare or qualitative data dream? Surv Pract 2014 Oct 1;7(5):1-11. [doi: 10.29115/SP-2014-0024]
30. Rekanar K, O'Keeffe IR, Buckley S, et al. Sentiment analysis of user feedback on the HSE's Covid-19 contact tracing app. Ir J Med Sci 2022 Feb;191(1):103-112. [doi: 10.1007/s11845-021-02529-y] [Medline: 33604836]
31. THL has updated its instructions regarding coronavirus testing [Web page in Finnish]. Finnish National Institute for Health and Welfare THL. URL: https://www.sttinfo.fi/tiedote/69928530/thl-on-paivittanyt-koronavirustestaukseen-liittyvia-ohjeitaan?publisherId=69817778 [accessed 2026-08-23]
32. Mild coronavirus symptoms usually do not require testing – avoid contact for five days and follow regional guidelines [Web page in Finnish]. Finnish National Institute for Health and Welfare THL. URL: https://www.sttinfo.fi/tiedote/69929563/lievien-koronavirusoireiden-takia-ei-yleensa-tarvitse-kayda-testissa-valta-kontakteja-viiden-paivan-ajan-ja-noudata-alueellisia-ohjeita?publisherId=69817778 [accessed 2026-08-23]

Abbreviations

ENA: exposure notification app
ERV: exposure risk value
KPI: key performance indicator
LLM: large language model

XSL•FO
RenderX

NIDR: National Infectious Diseases Register
OR: odds ratio
PR: prevalence ratio

Edited by A Mavragani, T Sanchez; submitted 16.Dec.2025; peer-reviewed by TN Susilawati, Y Hu; revised version received 12.May.2026; accepted 12.Jul.2026; published 21.Sep.2026.

Please cite as:
Kitowska W, Pihlajamäki M, Wickström S, Tolmunen K, Helve O, Mykkänen J, Dub T, Siira L
Evaluation of the Finnish Exposure Notification App "Koronavilkku" According to Key Performance Indicators: Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e89472
URL: https://publichealth.jmir.org/2026/1/e89472
doi:10.2196/89472

© Wioleta Kitowska, Mika Pihlajamäki, Sara Wickström, Kaija Tolmunen, Otto Helve, Juha Mykkänen, Timothée Dub, Lotta Siira. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 21.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Automated Daily Phone Surveys in Older Adults: Feasibility Study in the Multi-Ethnic Study of Atherosclerosis

Sally J Kim, BA; Cyanna McGowan, MPH; Shaina J Alexandria, PhD; Grace D Achepohl, MS; Lauren Hoffer, BS; Norrina B Allen, MPH, PhD; Mandy Wong, MS; Kiarri N Kershaw, MPH, PhD

Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, 680 North Lake Shore Drive, Suite 1400, Chicago, IL, United States

Corresponding Author:
Kiarri N Kershaw, MPH, PhD
Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, 680 North Lake Shore Drive, Suite 1400, Chicago, IL, United States

Abstract

Background: An ecological momentary assessment (EMA) is a methodological framework designed to capture real-time data in a participant's natural environment, yet its feasibility in older adults, particularly in a large, multisite study, remains underexplored.

Objective: This study aims to assess the feasibility of implementing automated daily phone surveys in a large, geographically diverse cohort of older adults.

Methods: A total of 1283 participants from the Multi-Ethnic Study of Atherosclerosis (MESA) were enrolled across 6 US field sites. Participants were scheduled to receive 1 automated phone survey per day for 7 consecutive days following a clinic visit. Feasibility was assessed through a dual lens of participant engagement (survey adherence, phone survey duration, and cancellations) and administrative implementation (scheduling errors).

Results: Of the 1272 participants who received at least 1 of the 7 phone surveys, high adherence was observed, with 81.84% (1041/1272) completing 5 to 7 surveys. Age was the only demographic characteristic significantly associated with completion (P=.002), as younger participants demonstrated higher adherence. While the median survey duration was 3 (IQR 2‑4) minutes, 26.84% (2371/8835) of surveys lasted 5 minutes or longer. However, adherence remained remarkably robust even among participants experiencing these extended durations. Participant-initiated cancellations during the 7-day monitoring period were rare, occurring among 21 of 1272 (1.65%) participants who began monitoring. Administrative scheduling errors were also uncommon, accounting for 37 of the 8835 (0.42%) total scheduled surveys not being delivered.

Conclusions: Findings support the feasibility of automated daily phone surveys for collecting repeated real-world data in large, multisite studies of older adults. Future studies should incorporate standardized participant feedback to identify barriers to engagement and support more inclusive, participant-centered adaptations of automated phone-based EMA protocols.

(JMIR Public Health Surveill 2026;12:e86333) doi:10.2196/86333

KEYWORDS

ecological momentary assessment; older adults; daily diary; feasibility; phone-based survey; aging research; Multi-Ethnic Study of Atherosclerosis

Introduction

Ecological momentary assessment (EMA) encompasses a diverse methodological framework designed to capture real-time data on thoughts, feelings, and behaviors within the context of a participant's natural environment. By using repeated sampling, EMA offers a more dynamic picture of daily life that effectively minimizes the recall biases often associated with traditional laboratory-based or retrospective methods [1]. EMA exists on a spectrum regarding its procedural flexibility: sampling protocols can be structured as signal-contingent (random prompts), event-contingent (participant-initiated), or interval-contingent (fixed schedules). The frequency and duration of EMA vary significantly across research designs; protocols may range from multiple assessments within a single day to a single assessment over many days, and monitoring periods are similarly variable in length [2-5]. These assessments can also be deployed through a wide variety of interfaces, ranging from mobile apps and wearable sensors to paper diaries and telephone-based systems [1,6-8].

While older adults have been increasingly included in EMA research and recent research suggests that EMA compliance in older adults is generally promising, the broader body of evidence remains mixed and limited [9-18]. A 2023 systematic review

XSL•FO
RenderX

and meta-analysis by Yao et al [17] indicates a combined compliance rate of 86.41%, yet completion rates are significantly influenced by participant-level factors and much of the existing evidence is limited by methodological constraints. Studies are predominantly characterized by small sample sizes, often less than a sample size of 100, and limited multisite recruitment [9-18]. Consequently, while existing research provides a foundational understanding of EMA feasibility, these findings may not fully generalize to the unique complexities of studies involving a larger, geographically diverse sample of older adults.

To address these limitations and accommodate the significant scale and geographic distribution of the Multi-Ethnic Study of Atherosclerosis (MESA), we implemented an EMA framework designed for both participant acceptability and operational sustainability. Specifically, among the diverse sampling strategies categorized under EMA, we used the daily diary approach, which is a fixed-interval method involving a single assessment at a consistent time each day [1]. This method was selected to support large-scale implementation because its fully automated, phone-based interface enabled data collection across a multisite cohort of more than 1000 older adults while minimizing burden on participants and research staff.

To evaluate the feasibility of this approach, we used a dual lens that considered both participant engagement and the administrative requirements of the study. Participant engagement feasibility reflected whether participants completed and continued with the daily phone surveys. Participant engagement was evaluated through descriptive participant characteristics, survey adherence metrics, and participant-initiated cancellation of the phone surveys. These measures allowed us to characterize the stress reactivity ancillary study sample relative to the broader MESA Exam 7 cohort, identify patterns across survey adherence groups, describe the extent to which participants completed the phone surveys and their experience, and assess active requests to stop future phone surveys. Administrative implementation feasibility was assessed by quantifying scheduling errors that disrupted the full delivery of phone surveys during the 7-day monitoring period. Specifically, we tracked the number of participants impacted by at least one scheduling error and the total volume of missed surveys relative to the overall study target. This study aims to assess the feasibility of automated daily phone surveys in a large, multisite study of older adults. Our findings contribute to the growing body of work on EMA use in older adults and offer practical guidance for implementing automated phone-based protocols in large-scale, multisite studies.

Methods

Study Population

The MESA is an ongoing, community-based cohort study designed to investigate the prevalence and progression of subclinical cardiovascular disease among adults. The study enrolled 6814 racially and ethnically diverse men and women aged 45 to 84 years who were free of known cardiovascular disease at enrollment. Participants were recruited from 6 field centers across the United States: New York, NY; Baltimore, MD; Forsyth County, NC; St. Paul, MN; Chicago, IL; and Los Angeles County, CA. Baseline (Exam 1) data collection took place from 2000 to 2002, with multiple examinations conducted thereafter among participants from Exam 1: Exam 2 (2002 - 2004), Exam 3 (2004 - 2005), Exam 4 (2005 - 2007), Exam 5 (2010 - 2011), Exam 6 (2016 - 2018), and Exam 7 (2022 - 2024). These participants were followed longitudinally to identify updates on health, interim hospitalizations, cardiovascular events, and mortality. Detailed information on the MESA study and recruitment procedures has been published previously [19] and is available on the study website [20].

Exam 7 consisted of core examination components and 7 ancillary studies. Participants from the original MESA cohort who remained active at follow-up were invited to complete the core Exam 7 visit. All participants completing the core exam were considered for ancillary studies unless they met ancillary-specific exclusion criteria. A total of 2276 MESA participants successfully completed Exam 7 [20]. Of this initial cohort, 1834 individuals were identified as eligible for recruitment into ancillary studies. Among those eligible, 1524 participants provided informed consent to participate in the ancillary studies. Within this consenting group, 1283 individuals agreed to participate and were enrolled specifically in the MESA Stress Reactivity Ancillary Study (ancillary study). Reasons for attrition included ineligibility at the screening phase, declining to participate in ancillary research generally, or declining the specific protocols required for the stress reactivity ancillary study.

This paper primarily focuses on participants and data from the MESA Stress Reactivity Ancillary Study. Participants were eligible if they agreed to wear a continuous ambulatory heart monitor (Cardiac Insight, Inc) and complete daily phone-based surveys over a 7-day monitoring period. Exclusion criteria included allergy to skin adhesives, refusal to shave the upper left chest, and the presence of implanted electronic devices such as pacemakers. The stress reactivity ancillary study component of the Exam 7 visit was estimated to take 30 minutes and included consent, heart monitor application, and participant instructions. Although enrollment in the stress reactivity ancillary study included both heart monitor wear and daily phone surveys, the present analysis focuses specifically on the feasibility of implementing the automated daily phone survey protocol.

Automated Daily Phone Surveys Protocol

Given the large, multisite structure of MESA and the average age of its participant population (mean 77.3, SD 7.50 years), selecting an EMA method that balanced acceptability, scalability, and staff sustainability was a priority. To gauge possible data collection options, an off-cycle survey covering a variety of topics was administered to MESA participants. Among the 1381 respondents, 39.03% (539/1381) indicated willingness to use a mobile device for data collection. Given the limited interest in using mobile devices for data collection in the larger MESA cohort, an alternative EMA delivery method was needed as opposed to smartphone app-based EMA.

Our study drew on Almeida's daily telephone interview approach [21], in which participants completed brief telephone interviews that took approximately 10 to 15 minutes each

XSL•FO
RenderX

evening for 8 consecutive days. However, for the stress reactivity ancillary study, live interviewer-administered calls were not considered practical considering the number of participants in MESA. Additionally, since the willingness to use mobile devices for data collection was low, an alternative approach that did not depend on mobile device use was needed. This led to the selection of a call-based method that could also accommodate participants who relied on landlines. Collectively, these considerations led to a modified version of the Almeida approach that used a centralized, automated system to deliver the daily phone surveys.

The automated phone-based system was implemented using REDCap, a data collection and management tool developed by Vanderbilt University, together with the third-party voice delivery service Twilio Voice (Twilio Inc). REDCap was used to store participants' phone survey scheduling preferences, daily stress survey questions, and responses, while Twilio served as the platform for generating automated outbound calls. At the clinic visit (day 0), field center staff entered the information needed to schedule the automated daily phone surveys in REDCap, including the date of the first call, the participant's preferred call time, preferred language (English, Spanish, Mandarin, or Cantonese), time zone, and phone number. This 1-time entry triggered the automated survey call sequence for the full 7-day monitoring period, so that the field center staff did not need to manually reschedule calls for each day of the monitoring period. This initial entry was critical because it governed both the first scheduled call and the subsequent daily calls over the 7-day monitoring period. Incorrect entry could therefore disrupt the delivery of the full survey sequence and affect participants' opportunity to receive or complete surveys.

The first call was intended to be scheduled for the day after the clinic visit. Participants selected a preferred call time after 4:00 PM based on their availability, as this time window was intended to increase the likelihood that participants would be at home and available to answer the phone surveys if they indicated to receive the daily phone surveys at their landline. The preferred call time had to be entered into REDCap in military format for the automated calls to go out after 4:00 PM. For instance, if a clinic visit occurred on January 1, 2023, and the participant selected 4:30 PM as their preferred call time, the first call should be entered into REDCap as January 2, 2023, 16:30.

The daily stress survey was programmed in REDCap as a survey instrument, and participants heard the survey questions from this instrument during the automated phone calls delivered through the Twilio Voice integration. Therefore, once the participants' scheduling information was entered, REDCap transmitted both the survey content and call settings to Twilio, which placed the automated call. REDCap rendered each survey question, Twilio delivered the questions by phone, participants entered responses using their phone keypad, and the responses were returned to and stored in REDCap. This process was repeated until all survey questions were completed, after which the call ended [22]. For participants who selected a language other than English, the daily stress survey questions and call script were translated from the English version and delivered in the participant's preferred language: Spanish, Mandarin, or Cantonese. Non-English calls used text-to-speech audio through the same automated workflow using REDCap and Twilio. No additional study-specific modifications were made to the automated system for tonal languages. This approach accommodated participant language and scheduling preferences while reducing staff burden and supporting standardized data collection across multiple sites.

From the participant's perspective, once this 1-time setup was initiated at the clinic visit by the field center staff, participants received 1 phone survey per day for 7 consecutive days, beginning the day after their clinic visit. The daily stress surveys were delivered each day at the participants' preferred time. Participants responded to these phone-based surveys using their phone keypad. Most questions had "Yes" or "No" response options, with responses recorded as "1" for Yes and "0" for No. Some questions included branching logic, where a Yes response triggered follow-up questions answered on a numeric scale, typically ranging from 1 to 4 but extending up to 7. Each survey was expected to be completed within 5 minutes. If a participant missed the initial call, the system automatically initiated a follow-up call 15 minutes later. Up to 2 additional follow-up calls were sent at 15-minute intervals for a maximum of 3 phone calls per day across the 7-day monitoring period. If all 3 attempts were missed, the phone survey was considered incomplete for that day. Surveys were also classified as incomplete if a participant terminated the call early before responding to all survey questions. A survey was considered complete if the participant answered the call and progressed through all the programmed questions. The surveys were not designed to allow partial completion.

Procedures to Support Survey Delivery and Protocol Adherence

Several procedures were put in place to support accurate protocol implementation and maximize the likelihood that participants received all 7 scheduled phone surveys. Participants using cell phones were encouraged to program the Twilio number into their contacts under "MESA Study" to minimize the risk of calls being missed or marked as spam. For those using landlines, the caller ID displayed "MESA Study." During the initial clinic visit, participants also completed a demo phone survey to familiarize themselves with the automated call system. They were asked to dial a practice number and complete a brief test phone survey using the same response format as the scheduled calls, with field center staff available to answer any questions and provide assistance during this practice session.

To support accurate protocol implementation and ensure quality control, the centralized coordinating team (hereinafter "reading center") for the stress reactivity ancillary study received real-time notifications after the REDCap scheduling forms were completed by the field center staff. Upon receiving the notification, the reading center staff reviewed the entered information for accuracy and completeness. If any discrepancies or missing data were identified, the reading center staff contacted the appropriate field center to request corrections. Once the information was corrected, the call schedule was reprogrammed accordingly via the Twilio platform. Because the first phone survey was scheduled for any time after 4:00 PM the day after the initial clinic visit, field center staff were

XSL•FO
RenderX

expected to complete the scheduling form as close to the clinic visit as possible. This allowed the reading center staff enough time to review the information and request or make corrections before the first scheduled phone survey. All communication between the reading center and field centers was documented in Smartsheet (Smartsheet Inc), a centralized work management platform used to track and coordinate study operations.

Field center staff were also instructed to conduct a check-in call with participants 2 days after the clinic visit and following the first scheduled phone survey to ensure the participant had not encountered any issues with receiving or completing the survey. If the first survey was missed, field center staff were expected to work with the participants to identify the reason and provide support as needed. In parallel, the reading center staff monitored survey completion status in REDCap for each participant daily. If a participant missed 2 consecutive surveys, the reading center staff contacted the corresponding field center staff to coordinate outreach and troubleshoot ongoing issues. Participant feedback, when provided, was documented in Smartsheet. At the conclusion of the 7-day monitoring period, the total number of completed surveys per participant was tracked in REDCap.

Feasibility

To assess the feasibility of our EMA approach, we considered both the participant experience and the administrative requirements of the study. Participant engagement feasibility was evaluated using three components: (1) descriptive participant characteristics, (2) survey adherence, and (3) participant-initiated cancellation of the phone surveys. Descriptive participant characteristics were examined for 2 purposes. First, we compared participants enrolled in the MESA Stress Reactivity Ancillary Study with MESA Exam 7 participants who were not enrolled in the ancillary study. Second, among ancillary study participants included in the adherence analysis, we examined demographic and socioeconomic characteristics across survey completion groups to explore patterns in participant engagement, with particular attention to participants with high adherence and zero adherence.

Survey adherence was defined as the number of daily phone surveys completed during the 7-day monitoring period. Participants were categorized into 4 completion groups: 0 surveys, 1‑2 surveys, 3‑4 surveys, and 5‑7 surveys. Completion of 5‑7 surveys was considered high adherence because it represented completion of most of the daily phone surveys. This categorization was informed by EMA studies that commonly define high adherence at the participant level as completion of approximately 80% of prompts, although this threshold is not formally established [13]. At the opposite end, participants with zero adherence were defined as participants completing 0 surveys. Means and SDs were reported for continuous variables, and frequencies were reported for categorical variables across the 4 completion groups. These analyses were descriptive in nature, and no formal hypothesis testing was performed. We also characterized participant engagement by analyzing survey duration, defined as the time elapsed from phone survey initiation to completion, to assess the burden and tolerability of the daily phone surveys.

In addition to survey adherence, we also assessed participant-initiated cancellation of the daily phone surveys as another indicator of participant engagement feasibility. Participant-initiated cancellations were defined as participant requests to stop the daily phone survey and were categorized as either premonitoring or during monitoring cancellations. Premonitoring cancellations were defined as requests to opt out of the phone surveys before the start of the 7-day monitoring period. During monitoring, cancellations referred to requests made after the first phone survey went out and at any point during the 7-day monitoring window. Participants could request cancellation by notifying field center staff, including at the clinic visit or after the monitoring period had begun.

Participants who canceled before monitoring began did so before the first scheduled phone survey was delivered and therefore did not receive any daily phone surveys. These premonitoring cancellations were retained as a separate subgroup because they represented participant-initiated requests to stop the phone survey component before survey delivery began, which differs from missed or incomplete surveys due to nonresponse or administrative scheduling errors. Therefore, premonitoring cancellations were not included in the survey adherence analysis. Participants who canceled after the monitoring period began were included in adherence analyses based on the number of surveys completed before cancellation. There was no standardized protocol requiring staff to systematically ask for or record reasons for cancellations. As a result, cancellation reasons were available only in some cases and were not analyzed formally. When available, these reasons were used descriptively to provide context for possible participant burden, lack of interest, or challenges with the automated phone survey system.

Administrative implementation feasibility was defined as the extent to which study staff successfully followed the automated daily phone survey protocol according to study procedures. A scheduling error was defined as any staff entry error that resulted in one or more expected daily phone surveys not being sent as intended. Therefore, to evaluate administrative implementation feasibility, we calculated (1) the number and proportion of participants affected by at least one scheduling error, and (2) the number and proportion of expected daily phone surveys that were not sent because of such errors.

Ethical Considerations

The Multi-Ethnic Study of Atherosclerosis was approved by the University of Washington institutional review board under STUDY00014523, with reliance agreements from participating institutions. All participants provided written informed consent for both the core and ancillary study components. Participants received US $50 compensation for participation in the stress reactivity ancillary study. Data used in this study were deidentified prior to distribution, with HIPAA (Health Insurance Portability and Accountability Act)–defined identifiers removed [20]. Access was limited to approved investigators and governed by MESA data use requirements, including safeguards against participant reidentification, restrictions on disclosure of potentially identifying information, and prohibitions on further data distribution.

XSL•FO
RenderX

Statistical Analysis

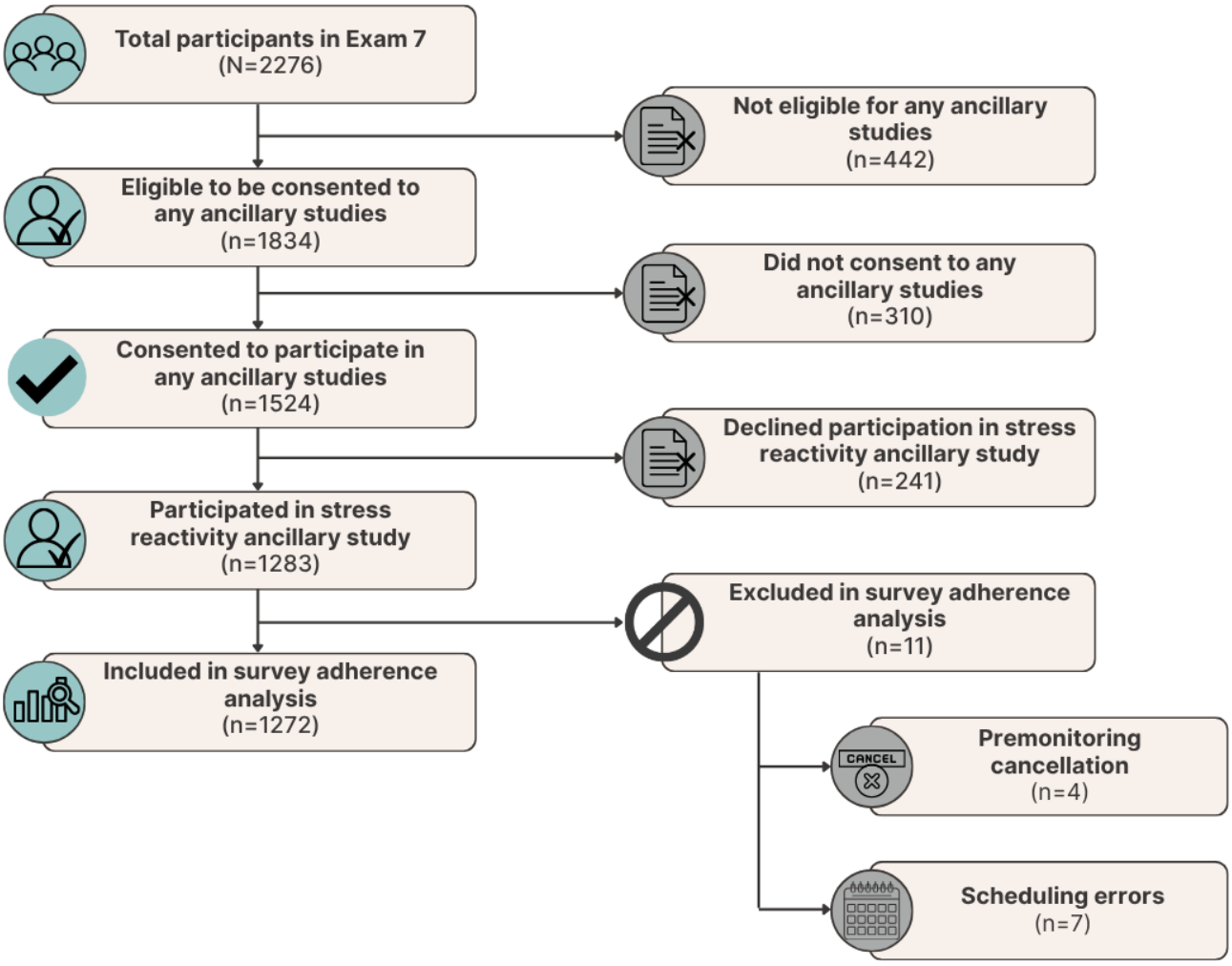
All analyses were descriptive and aimed at evaluating the feasibility of implementing automated EMA phone surveys in a large cohort of older adults. Descriptive statistics were generated to compare the demographics and socioeconomic characteristics of the MESA Exam 7 participants with those of our stress reactivity ancillary study participants. Frequencies and percentages were used to summarize scheduling errors, survey completion rates, and participant-initiated cancellations. All analyses were conducted using SAS (SAS Institute Inc). Data were compiled from REDCap and Smartsheet, and participant-level survey completion logs were used to calculate all outcome measures.

Results

Study Enrollment and Data Volume Summary

A total of 1283 participants were enrolled in the MESA Stress Reactivity Ancillary Study from the primary Exam 7 cohort of 2276 individuals. Participant flow from initial Exam 7 completion to final stress reactivity ancillary study enrollment is summarized in Figure 1. While all 1283 participants were included in initial demographic comparisons, the analysis of survey adherence focused on the 1272 individuals who received at least 1 phone survey. A total of 11 participants were excluded from the adherence analysis: 4 individuals who initiated cancellations prior to the monitoring period and 7 individuals who failed to receive any surveys due to administrative scheduling errors. A total of 8835 phone surveys were scheduled across the stress reactivity ancillary study's enrollment period. This aggregate volume of participants and scheduled surveys serves as the foundation for our evaluation of participant engagement and administrative implementation feasibility.

Figure 1. This flowchart illustrates participant recruitment, eligibility, and enrollment from the larger Exam 7 cohort into the stress reactivity ancillary study, and inclusion in the survey adherence analysis.



Participant Engagement Feasibility: Participant Characteristics and Survey Adherence

A total of 1283 participants were enrolled in the MESA Stress Reactivity Ancillary Study from the larger MESA Exam 7 cohort (N=2276). To evaluate the representativeness of the ancillary study sample, characteristics were compared between enrolled and nonenrolled participants (Table 1). Enrolled participants were significantly younger than those who did not enroll (mean age 75.96, SD 7.07 years vs 79.07, SD 7.75 years; P<.001).

XSL•FO
RenderX

Significant differences were also observed regarding socioeconomic status. Participants in the ancillary study had higher levels of educational attainment (*P*<.001) and higher annual household incomes (*P*=.01). Specifically, 27.95% (355/1270) of enrolled participants held a graduate degree compared to 24.67% (240/973) of nonenrolled participants, and a higher proportion of the enrolled group reported an income exceeding US $100K (25.06% vs 23.33%). Despite these differences, the enrolled sample remained demographically representative of the larger Exam 7 cohort in terms of gender (*P*=.23) and racial and ethnic distribution (*P*=.88), with similar proportions of African American (338/1283, 26.34%), Chinese (141/1283, 10.99%), Hispanic (282/1283, 21.98%), and non-Hispanic White (522/1283, 40.69%) participants.

Table . Demographic and socioeconomic characteristics of participants in the MESA[a] Exam 7 cohort and the subset of participants enrolled in the MESA Stress Reactivity Ancillary Study across the 6 US field sites.

Participant characteristics	MESA Exam 7 participants (N=2276)	MESA Stress Reactivity Ancillary Study participants (enrolled; n=1283)	MESA Stress Reactivity Ancillary Study participants (nonenrolled; n=993)	*P* value[b]
Age (years), mean (SD)	77.3 (7.50)	75.96 (7.07)	79.07 (7.75)	<.001
Sex, n (%)				.23
Female	1249 (54.88)	690 (53.78)	559 (56.29)	
Male	1027 (45.12)	593 (46.22)	434 (43.71)	
Race and ethnicity, n (%)				.88
African American	571 (25.09)	338 (26.34)	233 (23.46)	
Chinese	268 (11.78)	141 (10.99)	127 (12.79)	
Hispanic	516 (22.67)	282 (21.98)	234 (23.56)	
Non-Hispanic White	921 (40.47)	522 (40.69)	399 (40.18)	
Education, n (%)				<.001
High school diploma or less	626 (27.50)	305 (24.02)	321 (32.99)	
Some college, technical, and associate's degree	589 (25.88)	354 (27.87)	235 (24.15)	
College degree	433 (19.02)	256 (20.16)	177 (18.19)	
Some graduate school or graduate degree	595 (26.14)	355 (27.95)	240 (24.67)	
Income level (US $), n (%)				.01
<$25,000	442 (19.42)	228 (18.67)	214 (23.78)	
$25,000 to <$50,000	535 (23.51)	310 (25.39)	225 (25.00)	
$50,000 to <$100,000	628 (27.59)	377 (30.88)	251 (27.89)	
>$100,000	516 (22.67)	306 (25.06)	210 (23.33)	

[a]MESA: Multi-Ethnic Study of Atherosclerosis.

[b]*P* values compare characteristics of the MESA Exam 7 cohort with those enrolled in the stress reactivity ancillary study.

Descriptive analysis of survey completion groups revealed that adherence was largely consistent across demographic and socioeconomic characteristics (Table 2). Age was the only variable significantly associated with survey completion (*P*=.002), with participants in Group 1, the noncompletion group, being significantly older than those in Group 4, the high-adherence group (mean age 79.11 vs 75.74 years). High adherence (5 - 7 surveys) remained consistently above 77% across all racial and ethnic subgroups, with the highest rates observed among African American participants (442/520, 85%). However, variation was noted by language. Participants completing surveys in Mandarin (43/56, 76.79%) or Cantonese (22/31, 70.97%) demonstrated lower high-adherence rates than those using English (874/1058, 82.61%) or Spanish (102/126, 80.95%). Completion rates were similar across socioeconomic tiers, with high adherence rates between 78% and 84% for all education and income levels. Noncompletion was uncommon across all subgroups, generally ranging between 4% and 10%.

XSL•FO
RenderX

Table . Distribution of daily phone survey completion groups by demographic and socioeconomic characteristics among participants in the MESA[a] Stress Reactivity Ancillary Study.

Participant characteristics	Overall[b] (N=1272)	Number of phone surveys completed			
		Group 1[c]: 0 surveys (n=66)	Group 2[c]: 1 - 2 surveys (n=50)	Group 3[c]: 3 - 4 surveys (n=115)	Group 4[c]: 5 - 7 surveys (n=1041)
Age (years)[d], mean (SD)	75.92 (7.05)	79.11 (7.92) [d]	76.42 (7.09)	75.51 (7.58)	75.74 (6.89)
Sex, n (%)					
Female	684 (53.73)	33 (4.82)	28 (4.09)	56 (8.19)	567 (82.89)
Male	588 (46.23)	33 (5.61)	22 (3.74)	59 (10.03)	474 (80.61)
Race and ethnicity, n (%)					
African American	520 (40.85)	22 (4.23)	20 (3.85)	36 (6.92)	442 (85.00)
Chinese	135 (10.60)	10 (7.41)	6 (4.44)	15 (11.11)	104 (77.04)
Hispanic	337 (26.47)	17 (5.04)	13 (3.86)	35 (10.39)	272 (80.71)
Non-Hispanic White	280 (22.01)	17 (6.07)	11 (3.93)	29 (10.36)	223 (79.64)
Language of call, n (%)					
English	1058 (83.18)	50 (4.73)	39 (3.69)	95 (8.98)	874 (82.61)
Spanish	126 (9.91)	9 (7.14)	7 (5.56)	8 (6.35)	102 (80.95)
Chinese (Mandarin)	56 (4.40)	4 (7.14)	2 (3.57)	7 (12.50)	43 (76.79)
Chinese (Cantonese)	31 (2.44)	3 (9.68)	2 (6.45)	4 (12.90)	22 (70.97)
Education, n (%)					
High school diploma or less	300 (23.81)	13 (4.33)	8 (2.67)	27 (9.00)	252 (84.00)
Some college, technical, and associate's degree	351 (27.88)	20 (5.70)	18 (5.13)	38 (10.83)	275 (78.35)
College degree	255 (20.24)	15 (5.88)	8 (3.14)	22 (8.63)	210 (82.35)
Some graduate school or graduate degree	353 (28.02)	18 (5.10)	15 (4.25)	28 (7.93)	292 (82.72)
Income level (US $), n (%)					
<$25,000	224 (18.48)	15 (6.70)	7 (3.13)	18 (8.04)	184 (82.14)
$25,000 to <$50,000	308 (25.41)	13 (4.22)	13 (4.22)	33 (10.71)	249 (80.84)
$50,000 to <$100,000	374 (30.88)	18 (4.81)	13 (3.48)	33 (8.82)	310 (82.89)
>$100,000	305 (25.17)	15 (4.92)	13 (4.26)	24 (7.87)	253 (82.95)

[a]MESA: Multi-Ethnic Study of Atherosclerosis.

[b]Values are presented as n (%). Percentages in the Overall column are column percentages.

[c]Values are presented as n (%). Percentages in Groups 1 - 4 are row percentages.

[d]Significant at the .05 level.

Beyond demographic and socioeconomic trends, we also assessed the practical experience of survey completion. Among the 1272 participants who received at least 1 phone survey during the 7-day monitoring period, the median survey completion time was 3 (IQR 2 - 4) minutes, which was consistent with the anticipated duration. However, we evaluated extended durations as a proxy for participant burden and tolerability. While the phone surveys were designed to be less than 5 minutes, 2371 of the total 8835 (26.84%) surveys lasted 5 minutes or longer. This was a common experience among the participants in the survey adherence analysis, with 73.19% (931/1272) of participants completing at least 1 survey that met this threshold. For these phone surveys, the mean completion time was approximately 7 (SD 5.03) minutes. Despite the increased time commitment, these participants demonstrated high adherence, completing a median of 6 (IQR 3-7) surveys.

XSL•FO
RenderX

More extreme durations were rare but followed a distinct demographic pattern. Surveys lasting 10 minutes or more accounted for 2.60% (230/8835) of total phone surveys and occurred among 15.25% (194/1272) of participants. Phone surveys exceeding 20 minutes were exceptionally infrequent, representing only 0.62% (53/8835) of all surveys. We observed that participants who required more time to complete the phone surveys tended to be older than the average participant, with the mean age increasing incrementally alongside survey duration from 76.09 (SD 7.01) years in the 5-minute or longer category to 77.73 (SD 7.31) years in the 20-minute or longer category. Notably, while participants experiencing these longer durations were more likely to miss at least 1 individual phone survey, they consistently completed most of their scheduled phone surveys across the 7-day monitoring period, maintaining a median of 6 completed surveys across both duration categories (IQR for over 10 minutes: 5-7; IQR for over 20 minutes: 4-7).

Participant Engagement Feasibility: Cancellations

Participant-initiated cancellations occurred both before and during the monitoring period. Of the 1283 participants enrolled in the stress reactivity ancillary study, premonitoring cancellations only accounted for 0.31% (4/1283) of participants. These early withdrawals were primarily associated with an inability to comply with other broader aspects of the ancillary study rather than issues specific to the automated phone system. During monitoring, cancellations accounted for 1.65% (21/1272) of participants who received at least 1 phone survey. Of these individuals, 23.81% (5/21) had completed at least 1 survey before deciding to discontinue.

Administrative Implementation Feasibility

To evaluate administrative implementation feasibility, we assessed operational accuracy by identifying scheduling errors that disrupted the intended 7-day survey sequence. Among the 1283 participants enrolled in the stress reactivity ancillary study, only 7 (0.55%) individuals were affected by scheduling errors that prevented their phone surveys from being initiated. As a result, these participants did not receive any of the daily phone surveys during the 7-day monitoring period. An additional 17 (1.33%) participants received fewer than the 7 intended phone surveys because of scheduling errors during the initial scheduling process at the clinic visit. Additionally, only 0.42% (37/8835) were not delivered because of scheduling errors. These findings indicate that scheduling errors affected only a small proportion of participants and expected surveys, supporting the administrative implementation feasibility of the automated phone-based daily surveys across a large, multisite research infrastructure.

Discussion

Principal Findings

By evaluating both participant engagement and administrative implementation, this study supports the feasibility of automated daily phone surveys for collecting real-time data in a large, geographically diverse cohort of older adults. Comparing the ancillary study sample with the broader MESA Exam 7 cohort provided important context for interpreting the generalizability of these feasibility findings. The lack of significant differences in gender distribution (P=.23) and racial and ethnic composition (P=.88) suggests that the ancillary study sample was broadly similar to the Exam 7 cohort across these demographic characteristics. However, participants enrolled in the ancillary study were younger (P<.001) and had higher levels of education (P<.001) and household income (P=.01) than those who were not enrolled. Because age was the only participant characteristic significantly associated with survey completion (P=.002), the high adherence rate of 81.84% observed in this study may be higher than what would be expected in older cohorts or in populations with fewer socioeconomic resources. Therefore, while our findings support the feasibility of automated daily phone surveys, additional tailoring may be needed to support consistent engagement among older adults who may face greater barriers to participation.

Participant engagement remained high, with 81.84% (1041/1272) of the participants who received at least 1 survey demonstrating high adherence by completing 5 - 7 surveys. Statistical analysis revealed that age was the only demographic variable significantly associated with survey completion (P=.002). Although adherence was consistently high across racial and ethnic groups, descriptive variation was observed by survey language, as Mandarin- and Cantonese-speaking participants showed slightly lower rates of high adherence compared to English- and Spanish-speaking participants. The study used translated scripts but did not implement additional system-level modifications specific to tonal languages. Because formal feedback was not systematically collected from Mandarin- and Cantonese-speaking participants, we cannot determine whether language delivery, automated voice quality, cultural adaptation, or other factors contributed to these differences. These findings suggest that future automated daily phone surveys may benefit from additional usability testing and language-specific tailoring for individuals who speak other languages. At the same time, offering the phone surveys in English, Spanish, Mandarin, and Cantonese according to participant preference may have contributed to the enhanced engagement across this cohort.

This feasibility study also offers insights regarding participant burden and tolerability. While surveys were designed to take no more than 5 minutes, 26.84% (2371/8835) of all phone surveys exceeded this duration. Despite these extended survey durations, adherence remained high: participants who experienced at least 1 survey lasting longer than 5 minutes still completed a median of 6 phone (IQR 3-7) surveys. These findings suggest that the automated phone survey protocol was generally tolerable for older adults, even when individual surveys exceeded the intended duration. The simplified response format, consisting primarily of Yes or No keypad entries and brief numeric scales, may have contributed to the high adherence observed.

A small subset of participants in Group 1 provided feedback on why they did not respond to any or missed the phone surveys: 8 participants reported never receiving the phone surveys, 5 cited hearing difficulties, and 4 mentioned phone-related issues such as a disconnected number or accidentally blocking the study line. These observations suggest that noncompletion

XSL•FO
RenderX

among this subgroup should be explored further to uncover specific barriers to participation and to determine whether additional tailoring in the automated phone-based survey approach could enhance engagement and accessibility.

The use of premonitoring and during monitoring cancellation metrics provided a nuanced view of study attrition. Both forms of participant-initiated cancellation were rare, with 4 premonitoring cancellations and 21 during monitoring cancellations, totaling 25 participants, or 1.95% (25/1283) of the 1283 ancillary study participants. A small subset of participants provided reasons for canceling the phone surveys. Most premonitoring withdrawals were associated with an inability to comply with broader study components rather than the phone surveys themselves, even though participants were encouraged to continue the daily phone surveys regardless of their participation in other study components. Other individuals reported finding the automated calls a nuisance, preferring live interviewer calls for privacy, or feeling stressed while waiting for the scheduled automated phone calls. Despite these instances, the low overall rate of attrition supports the high acceptability of the automated daily phone surveys within an older adult population.

Beyond participant-facing features, the protocol optimized administrative implementation feasibility and staff sustainability by automating the initial scheduling process, which significantly reduced the daily administrative burden on field center staff and provided a notable advantage for large-scale studies where coordination across multiple sites is complex. This reduction in daily tasks allowed staff to prioritize proactive participant follow-up, troubleshoot technical issues, and support overall protocol adherence. Together, these findings help fill an important gap in the EMA literature and support the feasibility of using automated phone surveys in a large, multisite population of older adults.

Limitations

Despite the promising results, several limitations should be considered. First, although the majority of participants completed most or all of the surveys, a small subset either completed none or requested early cancellation. These individuals may have experienced barriers such as limited comfort with phone technology, hearing impairments, or other unmeasured factors. However, because our protocol did not include a formal process for collecting feedback from participants who were disengaged, we were unable to determine the reasons for noncompletion. Incorporating brief follow-up interviews or satisfaction surveys could provide valuable insights for future research. Additionally, we did not assess participants' perceptions of the survey experience, such as satisfaction or perceived burden. These perspectives are essential for interpreting adherence and for refining EMA protocols to maximize feasibility and acceptability among older adults. Finally, although the delivery of phone surveys was fully automated, the initial scheduling process relied on manual entry by field staff. While errors were rare, this step remains a potential source to further streamline workflows. Future efforts might explore opportunities to automate or standardize scheduling processes, informed by field staff feedback, to further minimize the risk of human error, especially in multisite studies where coordination is complex.

Conclusion

Our findings suggest that daily EMA surveys by phone can be both practical and scalable for use in large, multisite studies involving older adults. High adherence rates, as well as low cancellation and operational error, suggest this approach is not only feasible but well-tolerated in a real-world research setting. These results provide a promising model for incorporating EMA into aging research. Future studies should explore strategies to improve accessibility among subgroups with lower adherence (eg, due to hearing problems) and evaluate the validity and predictive value of phone-based EMA data. As EMA continues to evolve as a tool for real-time data capture, our findings offer actionable insights for designing and deploying EMA in older adult populations, with implications for broader adoption in multisite health research.

Acknowledgments

The authors thank the other investigators, the staff, and the participants of the Multi-Ethnic Study of Atherosclerosis (MESA) study for their valuable contributions. A full list of participating MESA investigators and institutions is available on the MESA study website [19]. This paper has been reviewed and approved by the MESA Publications and Presentations Committee. A generative language model (ChatGPT, version 4) was used in a limited capacity to assist with proofreading and clarity editing. Responsibility for the final manuscript rests entirely with the authors.

Funding

This research was supported by contracts 75N92020D00001, HHSN268201500003I, N01-HC-95159, 75N92020D00005, N01-HC-95160, 75N92020D00002, N01-HC-95161, 75N92020D00003, N01-HC-95162, 75N92020D00006, N01-HC-95163, 75N92020D00004, N01-HC-95164, 75N92020D00007, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC95168, and N01-HC-95169 from the National Heart, Lung, and Blood Institute; by grants UL1- TR-000040, UL1-TR-001079, and UL1-TR-001420 from the National Center for Advancing Translational Sciences (NCATS); and by grants R01AG067557, R01AG058969, and R01AG054069 from the National Institute on Aging. The funders had no role in study design, data collection, reporting or the decision to submit for publication.

XSL•FO
RenderX

Data Availability

The datasets generated or analyzed during this study are available through the National Heart, Lung, and Blood Institute's Biologic Specimen and Data Repositories Information Coordinating Center (BioLINCC) and the National Institutes of Health Database of Genotypes and Phenotypes (dbGaP) [23,24].

Authors' Contributions

KNK, SJK, and GDA conceptualized the manuscript. SJA, CM, and MW curated the data. SJK conducted the formal analysis with input from CM. SK drafted the original manuscript with the support of KNK. All authors critically reviewed and approved the final version of the manuscript.

Conflicts of Interest

NBA received NIH funding from NHLBI for MESA. All other authors declare no conflicts of interest.

References

1. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
2. Burke LE, Shiffman S, Music E, et al. Ecological momentary assessment in behavioral research: addressing technological and human participant challenges. J Med Internet Res 2017 Mar 15;19(3):e77. [doi: 10.2196/jmir.7138] [Medline: 28298264]
3. Wrzus C, Neubauer AB. Ecological momentary assessment: a meta-analysis on designs, samples, and compliance across research fields. Assessment 2023 Apr;30(3):825-846. [doi: 10.1177/10731911211067538] [Medline: 35016567]
4. Businelle MS, Hébert ET, Shi D, et al. Investigating best practices for ecological momentary assessment: nationwide factorial experiment. J Med Internet Res 2024 Aug 12;26(3):e50275. [doi: 10.2196/50275] [Medline: 39133915]
5. Smyth JM, Jones DR, Wen CKF, Materia FT, Schneider S, Stone A. Influence of ecological momentary assessment study design features on reported willingness to participate and perceptions of potential research studies: an experimental study. BMJ Open 2021 Jul 30;11(7):e049154. [doi: 10.1136/bmjopen-2021-049154] [Medline: 34330860]
6. Compernolle S, Vetrovsky T, Maes I, et al. Older adults' compliance with mobile ecological momentary assessments in behavioral nutrition and physical activity research: pooled results of four intensive longitudinal studies and recommendations for future research. Int J Behav Nutr Phys Act 2024 Aug 26;21(1):92. [doi: 10.1186/s12966-024-01629-z] [Medline: 39187862]
7. Ramsey AT, Wetherell JL, Depp C, Dixon D, Lenze E. Feasibility and acceptability of smartphone assessment in older adults with cognitive and emotional difficulties. J Technol Hum Serv 2016;34(2):209-223. [doi: 10.1080/15228835.2016.1170649] [Medline: 27683018]
8. Nam S, Dunton GF, Ordway MR, et al. Feasibility and acceptability of intensive, real-time biobehavioral data collection using ecological momentary assessment, salivary biomarkers, and accelerometers among middle-aged African Americans. Res Nurs Health 2020 Sep;43(5):453-464. [doi: 10.1002/nur.22068] [Medline: 32856310]
9. Crawford JL, English T, Braver TS. Incorporating ecological momentary assessment into multimethod investigations of cognitive aging: promise and practical considerations. Psychol Aging 2022 Feb;37(1):84-96. [doi: 10.1037/pag0000646] [Medline: 35113616]
10. Cain AE, Depp CA, Jeste DV. Ecological momentary assessment in aging research: a critical review. J Psychiatr Res 2009 Jul;43(11):987-996. [doi: 10.1016/j.jpsychires.2009.01.014] [Medline: 19272611]
11. Maher JP, Rebar AL, Dunton GF. Ecological momentary assessment is a feasible and valid methodological tool to measure older adults' physical activity and sedentary behavior. Front Psychol 2018;9:1485. [doi: 10.3389/fpsyg.2018.01485] [Medline: 30158891]
12. Moore RC, Kaufmann CN, Rooney AS, et al. Feasibility and acceptability of ecological momentary assessment of daily functioning among older adults with HIV. Am J Geriatr Psychiatry 2017 Aug;25(8):829-840. [doi: 10.1016/j.jagp.2016.11.019] [Medline: 28065496]
13. Rullier L, Atzeni T, Husky M, et al. Daily life functioning of community-dwelling elderly couples: an investigation of the feasibility and validity of Ecological Momentary Assessment. Int J Methods Psychiatr Res 2014 Jun;23(2):208-216. [doi: 10.1002/mpr.1425] [Medline: 24375556]
14. Stone AA, Schneider S, Smyth JM. Evaluation of pressing issues in Ecological Momentary Assessment. Annu Rev Clin Psychol 2023 May 9;19:107-131. [doi: 10.1146/annurev-clinpsy-080921-083128] [Medline: 36475718]
15. Charness N, Boot WR. Aging and information technology use: potential and barriers. Curr Dir Psychol Sci 2009;18(5):253-258. [doi: 10.1111/j.1467-8721.2009.01647.x]
16. Smith A. Older adults and technology use: usage and adoption. : Pew Research Center; 2014 Apr 3 URL: http://www.pewinternet.org/2014/04/03/usage-and-adoption/ [accessed 2026-06-28]
17. Yao L, Yang Y, Wang Z, Pan X, Xu L. Compliance with ecological momentary assessment programmes in the elderly: a systematic review and meta-analysis. BMJ Open 2023 Jul 12;13(7):e069523. [doi: 10.1136/bmjopen-2022-069523] [Medline: 37438069]

XSL•FO
RenderX

18. Bild DE, Bluemke DA, Burke GL, et al. Multi-Ethnic Study of Atherosclerosis: objectives and design. Am J Epidemiol 2002 Nov 1;156(9):871-881. [doi: 10.1093/aje/kwf113] [Medline: 12397006]
19. The Multi-Ethnic Study of Atherosclerosis. MESA. 2025. URL: https://www.mesa-nhlbi.org [accessed 2025-10-15]
20. Overview and protocol. MESA. 2026. URL: https://mesa-nhlbi.org/about/overview [accessed 2026-06-28]
21. Almeida DM, Wethington E, Kessler RC. The daily inventory of stressful events. Assessment 2002 Mar;9(1):41-55. [doi: 10.1177/1073191102091006]
22. Gupta G. Improve participant engagement in redcap healthcare research with Twilio SMS & voice. Twilio. 2023 Jul 10. URL: https://www.twilio.com/en-us/blog/improve-engagement-healthcare-research-redcap-sms-voice [accessed 2026-06-28]
23. BioLINCC - biologic specimen and data repository information coordinating center. National Institutes of Health. 2019. URL: https://biolincc.nhlbi.nih.gov/home/ [accessed 2026-07-09]
24. Database of genotypes and phenotypes. National Institutes of Health. 2026. URL: https://dbgap.ncbi.nlm.nih.gov/home/ [accessed 2026-07-09]

Abbreviations

EMA: ecological momentary assessment
HIPAA: Health Insurance Portability and Accountability Act
MESA: Multi-Ethnic Study of Atherosclerosis

Edited by A Mavragani, T Sanchez; submitted 22.Oct.2025; peer-reviewed by MR Frumkin, W Song; revised version received 06.May.2026; accepted 30.May.2026; published 16.Jul.2026.

Please cite as:
Kim SJ, McGowan C, Alexandria SJ, Achepohl GD, Hoffer L, Allen NB, Wong M, Kershaw KN
Automated Daily Phone Surveys in Older Adults: Feasibility Study in the Multi-Ethnic Study of Atherosclerosis
JMIR Public Health Surveill 2026;12:e86333
URL: https://publichealth.jmir.org/2026/1/e86333
doi:10.2196/86333

© Sally J Kim, Cyanna McGowan, Shaina J Alexandria, Grace D Achepohl, Lauren Hoffer, Norrina B Allen, Mandy Wong, Kiarri N Kershaw. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 16.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Multidimensional Discrimination Toward Single-Parent Families and Its Association With Depressive Symptoms of Parents: Cross-Sectional Study in South Korea

Seong-Uk Baek[1], MD; Jin-Ha Yoon[2], Prof Dr

[1]Graduate School, Yonsei University College of Medicine, Seodaemungu, Republic of Korea

[2]Department of Preventive Medicine, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemungu, Republic of Korea

Corresponding Author:
Jin-Ha Yoon, Prof Dr
Department of Preventive Medicine, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemungu, Republic of Korea

Abstract

Background: Discrimination toward single-parent families (SPFs) is prevalent at structural and individual levels.

Objective: This study examined the association between perceived discrimination toward SPFs and parental depressive symptoms in South Korea.

Methods: This study included a nationally representative sample of 3300 single mothers (n=2205, 66.8%) and fathers (n=1095, 33.2%). Single parents' perceptions of discrimination toward SPFs were measured using eight items evaluating the discrimination toward both participants and their children, which were categorized into four groups (lowest, low, high, and highest). Depressive symptoms were evaluated using the 9-item Patient Health Questionnaire. To examine the association between discrimination toward SPFs and depressive symptoms in single mothers and fathers, logistic regression models were used, and adjusted odds ratios and 95% CIs were calculated.

Results: Of all participants, 11.7% (n=386) reported depressive symptoms. The prevalence of depressive symptoms was 7.7% (57/744), 6.4% (44/684), 8.2% (54/659), and 21.7% (155/714) among individuals with the lowest, low, high, and highest levels of discrimination, respectively. Compared to those experiencing the lowest level of discrimination, the highest level of discrimination was associated with 5.10-fold (95% CI 3.33 - 7.79) and 6.12-fold (95% CI 2.80 - 13.39) higher odds of depressive symptoms among single mothers and fathers, respectively. Further analyses demonstrated that discrimination directed toward both oneself and one's children was associated with depressive symptoms.

Conclusions: Discrimination against SPFs was prevalent in Korea and associated with depressive symptoms in both single mothers and fathers.

(JMIR Public Health Surveill 2026;12:e83771) doi:10.2196/83771

KEYWORDS

family; mental health; single father; single mother; single parenthood; stigma

Introduction

In recent decades, sociocultural shifts have increased societal interest in diverse family structures. These family forms, which often include civil partnerships, same-sex parent families, and single-parent families (SPFs), challenge the traditional and often narrowly defined concept of a family as consisting of a heterosexual couple and their biological children [1]. This shift reflects diverse ways in which individuals form meaningful familial relationships. In Korea, approximately 1.5 million households consisted of SPFs in 2023, accounting for 6.9% of total households [2]. In light of the growing interest in SPFs, the Korean government has implemented policies aimed at assisting child-rearing and providing financial resources for SPFs [3,4].

The mental health of SPF members has garnered substantial public health interest, as both single parents and their children are more vulnerable to various risk factors for mental health deterioration, including material deprivation, dual burden of work and childcare, and lack of social support [1,5,6]. For instance, previous studies have demonstrated that various factors, such as socioeconomic status, work-life balance, and social acceptance and support, can be associated with the mental health of single parents [5,7,8]. Indeed, single parents are more likely to have mood disorders and depressive symptoms than couple parents [9,10].

SPFs face various forms of discrimination at the institutional, structural, and individual levels [11]. In South Korea, influenced by Confucianism, which upholds the traditional family structure as the norm [12], cultural acceptance and institutional support

XSL•FO
RenderX

for SPFs have progressed more slowly than in Western societies [13]. Parents of SPFs often experience stigmatization and discrimination in their daily lives, which negatively affects their mental well-being and overall health. For example, single mothers, particularly those who are divorced, face negative perceptions and stigmatization from neighbors and even close acquaintances [5,14]. Moreover, single parents are frequently stigmatized as having a lower socioeconomic status, rendering them more vulnerable to institutional discrimination in public and professional spheres than partnered parents [5,14,15]. Such stigmatization and discrimination can limit access to resources, employment opportunities, and social support systems [13].

Discrimination, irrespective of its nature, is a well-documented risk factor for poor mental health; however, the specific link between discrimination experienced by SPFs and depressive symptoms has been scarcely investigated. Experiencing discrimination can act as a major trigger for depressive symptoms by increasing chronic stress and lowering self-esteem [16]. Furthermore, social isolation and loneliness caused by discrimination can mediate the link between discrimination and depressive symptoms [17].

The existing literature on the mental health of single parents has some limitations. While studies have found that discrimination toward SPFs is linked to poor mental health among parents [7,18,19], most studies have primarily focused on single mothers, leaving important gaps in understanding sex-based differences in discrimination experiences and their association with depressive symptoms among single parents [20]. Therefore, based on a nationwide survey conducted in Korea, the objective of this study was to examine the association between perceived discrimination toward SPFs and the depressive symptoms of both single mothers and fathers by comprehensively assessing everyday discrimination experiences using a multidimensional discrimination measurement instrument. We hypothesized that perceived discrimination toward SPFs would be positively associated with depressive symptoms among single parents.

Methods

Study Population

The study sample was collected from the Single-Parent Family Survey, administered by the Korean Women's Development Institute under the Ministry of Gender Equality and Family in Korea. The survey was conducted between August and November 2021. To ascertain a representative sample, the selection process involved probability proportional sampling of 331 administrative districts across South Korea, each containing at least eight single-parent households using the data from the 2019 Population and Housing Census. Systematic sampling was conducted within each district to select 8 to 15 households, and the household heads were surveyed. The survey response rate was 86.5%. If a selected household head refused to participate or remained unavailable despite repeated contact attempts, the household was classified as nonresponsive and replaced with an adjacent reserve household (13.5%). Consequently, 3300 parents of SPFs were included in the survey. In the survey, a single-parent household was defined as one in which the head of the household was divorced, widowed, never married, or separated, and the household included the head and at least one of their children aged 18 years or younger. A standardized survey weight was calculated for each individual to enhance the representativeness of the sample. Data collection involved structured one-on-one household interviews by trained interviewers conducted in compliance with COVID-19 prevention guidelines.

Discrimination Toward SPFs

The perceived discrimination toward SPFs was evaluated in eight contexts. Specifically, participants were asked, "How often do you and your child experience unfair treatment or discrimination as a single-parent family? Please provide your responses from the perspectives of yourself and your child for each of the following contexts." The first five items measured perceived discrimination toward participants in the following contexts: (1) neighborhood or community, (2) school or childcare facilities, (3) family and relatives, (4) workplace, and (5) public institutions. The remaining three items assessed participants' perceived discrimination against their child in the following contexts: (1) neighborhood or community, (2) school or childcare facilities, and (3) family and relatives. To facilitate the participants' understanding and ensure standardized responses, 2 to 3 examples were provided for each item, as detailed in Table S1 in Multimedia Appendix 1. Each item was rated on a 4-point Likert scale (1 "Never experienced discrimination," 2 "Tended not to experience discrimination," 3 "Tended to experience discrimination," and 4 "Experienced severe discrimination"). Furthermore, participants could specify whether they had disclosed their single-parent status in each context rather than responding on a 4-point Likert scale. In such cases, the corresponding item was considered missing and replaced using multiple imputations. The total discrimination score ranged from 8 to 32, with higher scores indicating greater perceived discrimination toward SPFs. The Cronbach α was 0.93 for the sample. Based on the quartile values of the discrimination scores in the sample, participants were categorized into four groups: lowest (scores 8 - 10), low (scores 11 - 14), high (scores 15 - 17), and highest (scores 18 - 32) levels of discrimination.

Depressive Symptoms

The Korean version of the 9-item Patient Health Questionnaire, a validated tool comprising nine items that evaluate mood and vitality over the past two weeks, was used to assess depressive symptoms [21,22]. Each item is scored on a scale ranging from 0 to 3, resulting in a total score between 0 and 27. Based on previous literature, a total score of 10 or more was classified as indicative of depressive symptoms [21,23].

Covariates

The following sociodemographic features of the participants were considered confounders in the analyses: sex (male, female), age (<40 y, 40 - 49 y, ≥50 y), education level (middle school or below, high school, college or above), marital status (divorced, others—separated, widowed, never married), monthly household income (<₩1,000,000, ₩1,000,000 - ₩1,999,999, ₩2,000,000 - ₩2,999,999, and ≥₩3,000,000; an exchange

XSL•FO
RenderX

rate of ₩1000=US $0.67 is applicable), employment status (employed, unemployed), the number of children (one, two, three, or more), and physical activity. Physical activity was assessed using the following question: "How often do you currently engage in regular physical exercise (eg, walking, hiking, gym workouts, soccer, cycling, swimming)?" Participants who reported engaging in regular physical exercise 1 day or more per week were classified as physically active (yes), whereas all others were classified as no.

Data Analysis

For the descriptive analysis, the sociodemographic characteristics of the study participants were examined according to their perceived discrimination levels. The prevalence of depressive symptoms was explored according to study variables. Moreover, the response frequency for each item was examined to analyze the distribution of discrimination responses for each context.

Multivariate logistic regression models were used to evaluate the association between discrimination of SPFs and depressive symptoms. Subsequently, discrimination was separated into two types: discrimination toward parents and discrimination toward children. Specifically, we assessed how an increase in the summed scores for the five items assessing discrimination toward parents and an increase in the three items assessing discrimination toward their children were associated with depressive symptoms. Adjusted odds ratios (ORs) and 95% CIs were determined. All analyses were not only conducted for the overall sample but also stratified by the sex of the parent to account for potential sex differences. Missing values were handled using multiple imputations, with 20 complete datasets generated through the multiple imputation by chained equations technique. Pooled estimates were calculated and presented. Statistical analyses and visualizations were conducted using R software (version 4.6.2; R Foundation for Statistical Computing). The complex survey design that involves systematic sampling was addressed using the "svydesign" and "svyglm" functions in the R package *survey* [24]. Subsequently, the results were combined using the *mice* package. We also assessed the robustness of our findings by performing a sensitivity analysis of all cases (n=2801). Finally, considering the wide range of scores in the highest category of perceived discrimination, it was further divided into 18 to 24 and 25 to 32 to examine whether a dose-response pattern could be observed.

Ethical Considerations

This study was a secondary data analysis based on anonymized raw data from the Single-Parent Family Survey, which is publicly available at MicroData Integrated Service [25] without specific permission. To protect participants' privacy and confidentiality, all data were anonymized and deidentified before the publication of the dataset by the Korean Women's Development Institute. Therefore, the data provided to the authors were fully anonymized. Informed consent was provided by all participants during the original data collection procedure, and participants received no compensation for completing the survey. The Institutional Review Board (IRB) of the Yonsei Health System approved the study protocol (IRB application 2024-3489-002; IRB granting 4-2024-1607) and waived the requirement for additional informed consent because the study involved secondary analysis of existing data.

Results

Table 1 shows the characteristics of the 3300 study participants. The sample consisted of 1095 (33.2%) male parents and 2205 (66.8%) female parents. Compared to those who experienced the lowest level of discrimination, individuals who reported higher levels of discrimination were more likely to be younger, have a marital status other than divorced, have lower income levels, be employed, and have only one child. However, no clear difference was noted in the proportions of male and female participants across the levels of perceived discrimination.

Table 2 presents the prevalence of depressive symptoms. Of the 3300 participants, 11.7% (n=386) had depressive symptoms. The prevalence of depressive symptoms was 7.7% (57/744), 6.4% (44/684), 8.2% (54/659), and 21.7% (155/714) among those with the lowest, low, high, and highest levels of discrimination, respectively. Furthermore, the prevalence of depressive symptoms was higher among women, those with low educational levels, those with low income levels, unemployed individuals, and individuals with three or more children.

XSL•FO
RenderX

Table . Sociodemographic features of the study participants. Single-Parent Family Survey, 2021, South Korea (cross-sectional study).

	Total (N=3300), n (%)	Perceived discrimination (n=2801)[a], n (%)				*P* value
		Lowest (n=744)	Low (n=684)	High (n=659)	Highest (n=714)	
Sex						.60
Male	1095 (33.2)	256 (34.4)	224 (32.7)	236 (35.8)	236 (33.1)	
Female	2205 (66.8)	488 (65.6)	460 (67.3)	423 (64.2)	478 (66.9)	
Age (years)						.003
<40	722 (21.9)	153 (20.6)	139 (20.3)	155 (23.5)	167 (23.4)	
40 - 49	2036 (61.7)	433 (58.2)	434 (63.5)	390 (59.2)	453 (63.4)	
≥50	542 (16.4)	158 (21.2)	111 (16.2)	114 (17.3)	94 (13.2)	
Education level						<.001
Middle school or below	120 (3.6)	27 (3.6)	19 (2.8)	36 (5.5)	20 (2.8)	
High school	1778 (53.9)	441 (59.3)	315 (46.1)	344 (52.2)	422 (59.1)	
College or above	1402 (42.5)	276 (37.1)	350 (51.2)	279 (42.3)	272 (38.1)	
Marital status						.047
Divorced	2684 (81.3)	591 (79.4)	551 (80.6)	535 (81.2)	606 (84.9)	
Others	616 (18.7)	153 (20.6)	133 (19.4)	124 (18.8)	108 (15.1)	
Income level (₩[b])						<.001
<1,000,000	118 (3.6)	35 (4.7)	16 (2.3)	27 (4.1)	22 (3.1)	
1,000,000 - 1,999,999	1030 (31.2)	274 (36.8)	168 (24.6)	179 (27.2)	226 (31.7)	
2,000,000 - 2,999,999	1252 (37.9)	245 (32.9)	293 (42.8)	267 (40.5)	268 (37.5)	
≥3,000,000	900 (27.3)	190 (25.5)	207 (30.3)	186 (28.2)	198 (27.7)	
Employment status						.02
Employed	2680 (81.2)	585 (78.6)	577 (84.4)	549 (83.3)	596 (83.5)	
Unemployed	620 (18.8)	159 (21.4)	107 (15.6)	110 (16.7)	118 (16.5)	
Number of children						<.001
One	1842 (55.8)	367 (49.3)	375 (54.8)	380 (57.7)	417 (58.4)	
Two	1248 (37.8)	306 (41.1)	284 (41.5)	232 (35.2)	261 (36.6)	
Three or more	210 (6.4)	71 (9.5)	25 (3.7)	47 (7.1)	36 (5.0)	
Physical activity						.002
Yes	1472 (44.6)	322 (43.3)	293 (42.8)	323 (49.0)	277 (38.8)	
No	1828 (55.4)	422 (56.7)	391 (57.2)	336 (51.0)	437 (61.2)	

[a]Participants with missing values were excluded (n=499).

[b]An exchange rate of ₩1000=US $0.67 is applicable.

XSL•FO
RenderX

Table . Prevalence of depressive symptoms according to study variables. Single-Parent Family Survey, 2021, South Korea (cross-sectional study; N=3300).

	Depressive symptoms, n (%)		*P* value
	Yes	No	
Perceived discrimination			<.001
Lowest (n=744)	57 (7.7)	687 (92.3)	
Low (n=684)	44 (6.4)	640 (93.6)	
High (n=659)	54 (8.2)	605 (91.8)	
Highest (n=714)	155 (21.7)	559 (78.3)	
Missing values (n=499)	76 (15.2)	423 (84.8)	
Sex			.003
Male (n=1095)	102 (9.3)	993 (90.7)	
Female (n=2205)	284 (12.9)	1921 (87.1)	
Age (years)			.04
<40 (n=722)	66 (9.1)	656 (90.9)	
40 - 49 (n=2036)	257 (12.6)	1779 (87.4)	
≥50 (n=542)	63 (11.6)	479 (88.4)	
Education level			<.001
Middle school or below (n=120)	19 (15.8)	101 (84.2)	
High school (n=1778)	245 (13.8)	1533 (86.2)	
College or above (n=1402)	122 (8.7)	1280 (91.3)	
Marital status			.80
Divorced (n=2684)	312 (11.6)	2372 (88.4)	
Others (n=616)	74 (12.0)	542 (88.0)	
Income level (₩)[a]			<.001
<1,000,000 (n=118)	33 (28.0)	85 (72.0)	
1,000,000 - 1,999,999 (n=1030)	190 (18.4)	840 (81.6)	
2,000,000 - 2,999,999 (n=1252)	122 (9.7)	1130 (90.3)	
≥3,000,000 (n=900)	41 (4.6)	859 (95.4)	
Employment status			<.001
Employed (n=2680)	218 (8.1)	2462 (91.9)	
Unemployed (n=620)	168 (27.1)	452 (72.9)	
Number of children			.03
One (n=1842)	220 (11.9)	1622 (88.1)	
Two (n=1248)	131 (10.5)	1117 (89.5)	
Three or more (n=210)	35 (16.7)	175 (83.3)	
Physical activity			<.001
Yes (n=1472)	229 (15.6)	1243 (84.4)	
No (n=1828)	157 (8.6)	1671 (91.4)	

[a]An exchange rate of ₩1000=US $0.67 is applicable.

Table 3 presents the response frequencies for the discrimination items. Approximately 13.4% to 18.3% of the participants reported tending to experience or experiencing severe discrimination. Specifically, 1.4% to 2.6% of the participants reported experiencing severe discrimination.

XSL•FO
RenderX

Table . Response frequencies on the discrimination questionnaire among the study participants. Single-Parent Family Survey, 2021, South Korea (cross-sectional study; N=3300).

	Never experienced discrimination, n (%)	Tended not to experience discrimination, n (%)	Tended to experience discrimination, n (%)	Experienced severe discrimination, n (%)	Did not disclose being a single parent, n (%)
Neighborhood or community	898 (27.2)	1450 (43.9)	476 (14.4)	75 (2.3)	401 (12.2)
Schools or childcare facilities	1135 (34.4)	1390 (42.1)	519 (15.7)	85 (2.6)	171 (5.2)
Family and relatives	1335 (40.5)	1304 (39.5)	457 (13.8)	78 (2.4)	126 (3.8)
Workplace	1244 (37.7)	1346 (40.8)	394 (11.9)	66 (2.0)	250 (7.6)
Public institutions	1240 (37.6)	1481 (44.9)	396 (12.0)	46 (1.4)	137 (4.1)
Neighborhood or community (child)	961 (29.1)	1355 (41.1)	545 (16.5)	51 (1.5)	388 (11.8)
Schools or childcare facilities (child)	1192 (36.1)	1385 (42.0)	505 (15.3)	47 (1.4)	171 (5.2)
Family and relatives (child)	1274 (38.6)	1444 (43.8)	417 (12.6)	41 (1.2)	124 (3.8)

Table 4 shows the association between perceived discrimination toward SPFs and parental depressive symptoms. Compared with those with the lowest discrimination level, those with the low (adjusted OR 1.64, 95% CI 1.05 - 2.57), high (adjusted OR 1.74, 95% CI 1.12 - 2.69), and highest (adjusted OR 5.33, 95% CI 3.65 - 7.76) levels were more likely to have depressive symptoms. Furthermore, the highest discrimination level was associated with 6.12-fold (95% CI 2.80 - 13.39) and 5.10-fold (95% CI 3.33 - 7.79) increases in the odds of depressive symptoms in men and women, respectively.

Table . Association between discrimination toward single-parent families and depressive symptoms of parents. Single-Parent Family Survey, 2021, South Korea (cross-sectional study; n=3300). The models were adjusted for sex (overall model), age, education, marital status, income, employment status, number of children, and physical activity.

	Overall, adjusted OR[a] (95% CI)	Male, adjusted OR (95% CI)	Female, adjusted OR (95% CI)
Perceived discrimination			
Lowest	Reference	Reference	Reference
Low	1.64 (1.05 - 2.57)	1.94 (0.75 - 5.00)	1.56 (0.94 - 2.59)
High	1.74 (1.12 - 2.69)	1.75 (0.75 - 4.11)	1.79 (1.07 - 2.97)
Highest	5.33 (3.65 - 7.76)	6.12 (2.80 - 13.39)	5.10 (3.33 - 7.79)
Continuous scale			
8-point increase	3.26 (2.62 - 4.05)	3.65 (2.27 - 5.85)	3.14 (2.46 - 4.01)

[a]OR: odds ratio.

Figure 1 shows the associations between domains of perceived discrimination and depressive symptoms among single parents. A 5-point increase in discrimination toward oneself, as measured by the summed score, was associated with a 2.11-fold (95% CI 1.46 - 3.05), 1.95-fold (95% CI 0.82 - 4.61), and 2.12-fold (95% CI 1.41 - 3.19) increase in the odds of depressive symptoms in the overall, male, and female samples, respectively. Similarly, a 3-point increase in discrimination toward one's children, as measured by the summed score, was associated with a 1.55-fold (95% CI 1.06 - 2.25), 1.89-fold (95% CI 0.73 - 4.87), and 1.48-fold (95% CI 0.98 - 2.22) increase in the odds of depressive symptoms in the overall, male, and female samples, respectively.

XSL•FO
RenderX

Figure 1. Associations between domains of perceived discrimination and depressive symptoms among single parents. The odds of depressive symptoms were estimated for a 5-point increase in discrimination toward oneself and a 3-point increase in discrimination toward one's children. Single-Parent Family Survey, 2021, South Korea (cross-sectional study).

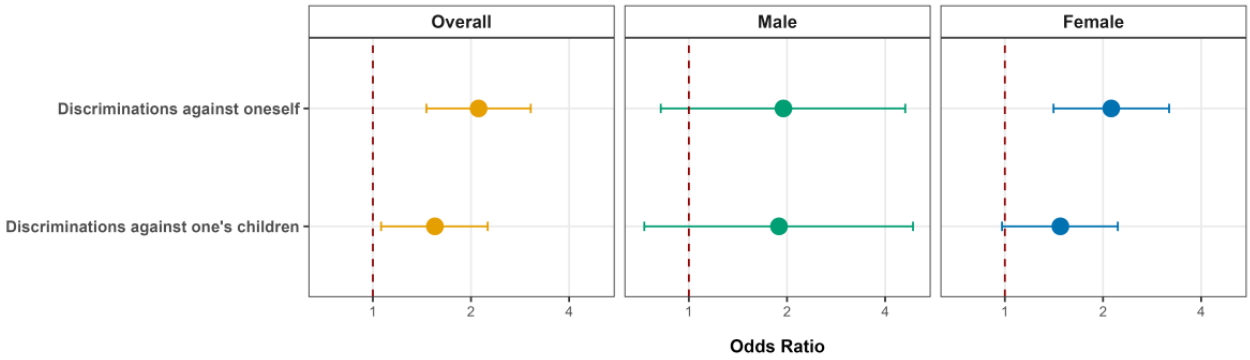


Table S2 in Multimedia Appendix 1 presents the association between discrimination toward SPFs and depressive symptoms in the complete cases. The findings of the sensitivity analysis were consistent with those of the main analysis, showing a similar positive association between discrimination experience and depressive symptoms. The results from the analyses further stratifying the highest category showed that, although the CIs were wide, notably elevated ORs were observed at the highest levels of discrimination, suggesting a dose-response pattern for this exposure category (Table S3 in Multimedia Appendix 1).

Discussion

This study found an association between perceived discrimination toward SPFs and depressive symptoms in single parents. Single mothers and fathers who experienced the highest levels of discrimination were more likely to report depressive symptoms than those who experienced the lowest levels of discrimination. Furthermore, discrimination directed towards both parents and their children was positively associated with parents' depressive symptoms.

Descriptive analysis revealed that 11.7% (386/3300) of the participants presented with depressive symptoms. This figure is notably higher than the prevalence of depressive symptoms measured by the 9-item Patient Health Questionnaire in nationally representative Korean samples, which ranged from 2.8% to 5.2% [26,27]. Furthermore, while a substantial number of single parents reported experiencing discrimination across various domains, clear sex differences in their levels of discrimination were not evident.

Our findings align with those of the existing literature, showing an association between experiences of discrimination and mental health problems such as depressive symptoms, suicidal behaviors, and anxiety [28,29]. Furthermore, the results of our study are consistent with those of previous research, showing that discrimination toward minority family structures is associated with poor mental health outcomes in affected individuals. For instance, discrimination targeting families with same-sex marriage has been linked to adverse mental health outcomes for themselves and their children [30,31]. Moreover, perceived discrimination negatively correlated with happiness among unmarried mothers [19]. Although various forms of discrimination, including those based on sex and race/ethnicity, have consistently been shown to be associated with mental health problems [28,29], this study adds to the literature by exploring the relationship between discrimination directed toward SPFs and depressive symptoms.

Complex mechanisms may explain the association between perceived discrimination toward SPFs and depressive symptoms of single parents. Multiple pathways may link the experience of discrimination to mental health deterioration, and the existing literature suggests that discrimination can induce psychological, physiological, and behavioral responses that contribute to poor health outcomes [16]. Prolonged exposure to discrimination can lead to chronic psychological distress by inducing anxiety regarding exclusion or by reducing self-esteem [32,33]. The persistence of stress can eventually result in the development of depressive symptoms. Moreover, physiologically, chronic stress arising from the experience of discrimination was linked to the dysregulation of the hypothalamic-pituitary-adrenal axis, which manifests as increased cortisol levels [34,35]. Hypothalamic-pituitary-adrenal axis overactivity plays a key role in the pathophysiology of depressive symptoms [36]. Finally, the discrimination experienced by SPFs may lead to behavioral changes. For instance, experiencing chronic stress may induce individuals to rely on coping mechanisms, such as smoking or alcohol consumption, which can contribute to the deterioration of mental health [37].

This study had certain limitations. First, the causal effects of discrimination toward SPFs on depressive symptoms could not be identified, as the analyses were based on cross-sectional survey data. For instance, the possibility of reverse causality, in which depressive symptoms influence the perception of discrimination [38], could not be ruled out in this cross-sectional study. Therefore, longitudinal studies are needed to examine whether experiences of discrimination can predict the onset of depressive symptoms over time. Second, key variables, including depressive symptoms, were assessed using self-reported measures, which may be subject to measurement errors, such as recall bias. Future studies should consider using medical records or physician diagnoses to verify the association between discrimination and objective clinical diagnoses. Third, certain factors that could influence both discrimination and depressive symptoms, such as psychiatric history, personality, social support, neighborhood income levels, alcohol use, smoking status, and other health behaviors, were not accounted

XSL•FO
RenderX

for due to the lack of available information. We acknowledge that the failure to account for these potential unmeasured confounders may have resulted in an overestimation of the association between discrimination experiences and depressive symptoms. Fourth, our discrimination survey questionnaire aimed to capture experiences of discrimination toward SPFs across various daily contexts. However, as it relies on subjective recall, it may be subject to recall bias. Additionally, the survey responses demonstrated high reliability, the instrument has not undergone formal validation. Because the measure was not externally validated, it may be conceptually limited and may not comprehensively capture discrimination toward SPFs. Additionally, the subjective nature of the measure may introduce systematic bias if individuals with depressive symptoms differentially report higher levels of perceived discrimination in comparison to those without [38]. Therefore, future research should establish a theoretical framework for understanding discrimination experiences among SPFs and develop a validated measurement tool to overcome these limitations. Fifth, although this study addressed missing data using multiple imputation, the assumption of missing at random could not be verified. We acknowledge that missingness may have been influenced by unobserved socioeconomic or health-related factors; therefore, the results derived from multiple imputation may still be subject to bias. Sixth, although this study aimed to explore sex-specific associations between discrimination toward SPFs and depressive symptoms, the relatively small sample size of single fathers resulted in wider CIs and less stable estimates compared with those observed among single mothers. Future epidemiological studies may address this limitation by including a larger sample of single fathers.

Despite these limitations, this study contributes to existing literature in several ways. First, the study sample was selected through systematic sampling, representing a nationally representative population, which enhances the generalizability of the findings. Second, our study provides novel insights by analyzing the scarcely studied association between discrimination toward SPFs and mental health.

This study showed that single parents experience discrimination toward both themselves and their children owing to their single-parent status, which is closely associated with depressive symptoms. An association between perceived discrimination toward SPFs and depressive symptoms was noted in both single mothers and fathers.

Acknowledgments

This study used data from the 2021 Single-Parent Family Survey, conducted by the Korea Women's Development Institute (KWDI) under the Prime Minister's Office of the Republic of Korea. The authors gratefully acknowledge the researchers at KWDI for providing access to the data and for their efforts in conducting the survey. We also extend our sincere thanks to the anonymous participants of the 2021 Single-Parent Family Survey. During the preparation of this work, the authors used ChatGPT from OpenAI for enhancing grammar and proofreading under full human supervision. Responsibility for the final manuscript lies entirely with the authors.

Funding

The authors declared that no financial support was received for this work.

Data Availability

The raw data of the Single-Parent Family Survey, 2021 were anonymized by the researchers at the Korean Women's Development Institute. The raw data are publicly available at the home page of MicroData Integrated Service [25].

Authors' Contributions

Conceptualization: SUK
Methodology: SUK
Software: SUK
Formal analysis: SUK
Investigation: SUK
Resources: SUK
Data curation: SUK
Writing – original draft: SUK
Project administration: SUK
Validation: JHY
Supervision: JHY
Writing – review & editing: JHY

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Survey questionnaire on the perceived discrimination toward single-parent families, sensitivity analyses based on the complete sample, and the results of the analysis further stratifying the highest category.
[DOCX File, 21 KB - publichealth_v12i1e83771_app1.docx]

Checklist 1
STROBE checklist.
[PDF File, 169 KB - publichealth_v12i1e83771_app2.pdf]

References

1. Jensen TM, Sanner C. A scoping review of research on well‐being across diverse family structures: rethinking approaches for understanding contemporary families. J Family Theo Revie 2021 Dec;13(4):463-495. [doi: 10.1111/jftr.12437]
2. The lives of men and women in 2023: a statistical overview [Web Page in Korean]. Ministry of Gender Equality and Family. 2023 Sep 6. URL: https://www.mogef.go.kr/nw/rpd/nw_rpd_s001d.do?mid=news405&bbtSn=709517 [accessed 2025-01-03]
3. Jang MS, Lee YH. Policy needs & improvements for single-parent families childcare [Article in Korean]. Korean Fam Resour Manag Assoc 2016;20(2):141-163 [FREE Full text]
4. Chin M, Lee J, Lee S, Son S, Sung M. Family policy in South Korea: development, current status, and challenges. J Child Fam Stud 2012 Feb;21(1):53-64. [doi: 10.1007/s10826-011-9480-1]
5. Vo T, Canty L. Global mental health experiences of single mothers: a mixed methods research synthesis. J Adv Nurs 2023 Jan;79(1):68-82. [doi: 10.1111/jan.15461] [Medline: 36226874]
6. Strickland JR. The relationship of food insecurity and mental health in single-parent households: a literature review. URNCST J 2024;8(1-12):1-13. [doi: 10.26685/urncst.554]
7. Kareem OM, Oduoye MO, Bhattacharjee P, et al. Single parenthood and depression: a thorough review of current understanding. Health Sci Rep 2024 Jul;7(7):e2235. [doi: 10.1002/hsr2.2235] [Medline: 38983687]
8. Dharani MK, Balamurugan J. The psychosocial impact on single mothers' well-being - a literature review. J Educ Health Promot 2024 Apr 29;13:148. [doi: 10.4103/jehp.jehp_1045_23] [Medline: 38784275]
9. Kim KJ, Yoo JH. Single mothers and mental health in South Korea: the Seventh Korea National Health and Nutrition Examination Survey, 2016. Korean J Fam Pract 2019 Aug 20;9(4):331-335. [doi: 10.21215/kjfp.2019.9.4.331]
10. Subramaniam M, Prasad RO, Abdin E, Vaingankar JA, Chong SA. Single mothers have a higher risk of mood disorders. Ann Acad Med Singap 2014 Mar;43(3):145-151. [Medline: 24714708]
11. Choi S, Byoun SJ, Kim EH. Unwed single mothers in South Korea: Increased vulnerabilities during the COVID-19 pandemic. Int Soc Work 2020 Sep;63(5):676-680. [doi: 10.1177/0020872820941040]
12. Park IH, Cho LJ. Confucianism and the Korean Family. J Comp Fam Stud 1995 Mar 1;26(1):117-134. [doi: 10.3138/jcfs.26.1.117]
13. Park H, Choi J, Jo H. Living arrangements of single parents and their children in South Korea. Marriage Fam Rev 2016 Feb 17;52(1-2):89-105. [doi: 10.1080/01494929.2015.1073653]
14. Rusyda HM, Lukman ZM, Subhi N, et al. Coping with difficulties: social inequality and stigmatization on single mothers with low income household. Pertanika J Soc Sci Humanit 2011;19:157-162 [FREE Full text]
15. Lauster N, Easterbrook A. No room for new families? A field experiment measuring rental discrimination against same-sex couples and single parents. Soc Probl 2011 Aug;58(3):389-409. [doi: 10.1525/sp.2011.58.3.389]
16. Williams DR, Lawrence JA, Davis BA, Vu C. Understanding how discrimination can affect health. Health Serv Res 2019 Dec;54(Suppl 2):1374-1388. [doi: 10.1111/1475-6773.13222] [Medline: 31663121]
17. Han S, Hu Y, Wang L, et al. Perceived discrimination and mental health symptoms among persons living with HIV in China: the mediating role of social isolation and loneliness. AIDS Care 2021 May;33(5):575-584. [doi: 10.1080/09540121.2020.1847246] [Medline: 33233938]
18. Kim A, Jeon S, Song J. Self-stigma and mental health in divorced single-parent women: mediating effect of self-esteem. Behav Sci (Basel) 2023 Sep 6;13(9):744. [doi: 10.3390/bs13090744] [Medline: 37754022]
19. Thasleema A, Rajan SK. Perceived discrimination and happiness among tribal unmarried mothers. Psychol Stud 2022 Sep;67(3):332-339. [doi: 10.1007/s12646-022-00668-5]
20. Kim JH. The effect of discrimination on depression in single-parent household heads [Article in Korean]. J Fam Resour Manag Policy Rev 2024 Aug 31;28(3):17-27. [doi: 10.22626/jkfrma.2024.28.3.002]
21. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613. [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
22. Park SJ, Choi HR, Choi JH, Kim KW, Hong JP. Reliability and validity of the Korean version of the Patient Health Questionnaire-9 (PHQ-9) [Article in Korean]. Anxiety Mood 2010;6(2):119-124 [FREE Full text]
23. Levis B, Benedetti A, Thombs BD, DEPRESsion Screening Data (DEPRESSD) Collaboration. Accuracy of Patient Health Questionnaire-9 (PHQ-9) for screening to detect major depression: individual participant data meta-analysis. BMJ 2019 Apr 9;365:l1476. [doi: 10.1136/bmj.l1476] [Medline: 30967483]

XSL•FO
RenderX

24. Lumley T. Analysis of complex survey samples. J Stat Soft 2004 Apr 15;9(8):1-19. [doi: 10.18637/jss.v009.i08]
25. MicroData Integrated Service. URL: https://mdis.kostat.go.kr/index.do [accessed 2026-04-13]
26. Kim SY, Yoo DM, Min C, Choi HG. Assessment of the difference in depressive symptoms of the Korean adult population before and during the COVID-19 pandemic using a community health survey. J Affect Disord 2022 Mar 1;300:130-136. [doi: 10.1016/j.jad.2021.12.107] [Medline: 34965394]
27. Lee EJ, Kim SJ. Prevalence and related factors of depression before and during the COVID-19 pandemic: findings from the Korea National Health and Nutrition Examination Survey. J Korean Med Sci 2023 Mar 13;38(10):e74. [doi: 10.3346/jkms.2023.38.e74] [Medline: 36918028]
28. Vines AI, Ward JB, Cordoba E, Black KZ. Perceived racial/ethnic discrimination and mental health: a review and future directions for social epidemiology. Curr Epidemiol Rep 2017 Jun;4(2):156-165. [doi: 10.1007/s40471-017-0106-z] [Medline: 28920011]
29. Vargas SM, Huey SJ, Miranda J. A critical review of current evidence on multiple types of discrimination and mental health. Am J Orthopsychiatry 2020;90(3):374-390. [doi: 10.1037/ort0000441] [Medline: 31999138]
30. Donnelly R, Robinson BA, Umberson D. Can spouses buffer the impact of discrimination on depressive symptoms? An examination of same-sex and different-sex marriages. Soc Ment Health 2019;9(2):192-210. [doi: 10.1177/2156869318800157] [Medline: 31223514]
31. Mazrekaj D, Jin Y. Mental health of children with gender and sexual minority parents: a review and future directions. Humanit Soc Sci Commun 2023;10(1):509. [doi: 10.1057/s41599-023-02019-9]
32. Urzúa A, Ferrer R, Godoy N, et al. The mediating effect of self-esteem on the relationship between perceived discrimination and psychological well-being in immigrants. PLoS ONE 2018;13(6):e0198413. [doi: 10.1371/journal.pone.0198413] [Medline: 29927968]
33. Brandt L, Liu S, Heim C, Heinz A. The effects of social isolation stress and discrimination on mental health. Transl Psychiatry 2022 Sep 21;12(1):398. [doi: 10.1038/s41398-022-02178-4] [Medline: 36130935]
34. Lee DB, Peckins MK, Miller AL, et al. Pathways from racial discrimination to cortisol/DHEA imbalance: protective role of religious involvement. Ethn Health 2021 Apr;26(3):413-430. [doi: 10.1080/13557858.2018.1520815] [Medline: 30198761]
35. Mikulska J, Juszczyk G, Gawrońska-Grzywacz M, Herbet M. HPA axis in the pathomechanism of depression and schizophrenia: new therapeutic strategies based on its participation. Brain Sci 2021 Sep 30;11(10):1298. [doi: 10.3390/brainsci11101298] [Medline: 34679364]
36. Vreeburg SA, Hoogendijk WJG, van Pelt J, et al. Major depressive disorder and hypothalamic-pituitary-adrenal axis activity: results from a large cohort study. Arch Gen Psychiatry 2009 Jun;66(6):617-626. [doi: 10.1001/archgenpsychiatry.2009.50] [Medline: 19487626]
37. Rosario M, Schrimshaw EW, Hunter J. Cigarette smoking as a coping strategy: negative implications for subsequent psychological distress among lesbian, gay, and bisexual youths. J Pediatr Psychol 2011 Aug;36(7):731-742. [doi: 10.1093/jpepsy/jsp141] [Medline: 20123704]
38. Liu H, Yang TC. Examining the reciprocity between perceived discrimination and health: a longitudinal perspective. Popul Res Policy Rev 2022 Aug;41(4):1757-1777. [doi: 10.1007/s11113-022-09712-8] [Medline: 35979183]

Abbreviations

IRB: Institutional Review Board
OR: odds ratio
SPF: single-parent family

Edited by A Mavragani, T Sanchez; submitted 08.Sep.2025; peer-reviewed by AR Hoque, ME Arayici; revised version received 24.Feb.2026; accepted 23.Mar.2026; published 23.Apr.2026.

Please cite as:
Baek SU, Yoon JH
Multidimensional Discrimination Toward Single-Parent Families and Its Association With Depressive Symptoms of Parents: Cross-Sectional Study in South Korea
JMIR Public Health Surveill 2026;12:e83771
URL: https://publichealth.jmir.org/2026/1/e83771
doi:10.2196/83771

© Seong-Uk Baek, Jin-Ha Yoon. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 23.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium,

XSL•FO
RenderX

provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Barriers and Opportunities to Include Underrepresented Population Groups in Vaccine Trials: Cross-Sectional, Observational, Online Survey Study From the VACCELERATE Research Network

Mine Durusu Tanriover[1,2], MD; Dimitrios Poulimeneas[3], BSc, MPH, PhD; Grammatiki-Christina Tsopela[3], MSc, MD; Ioannis Kopsidas[3], MD, PhD; Christos Argyropoulos[4], MEng, MSc, PhD; Romina Di Marzo[5]; Olena Valdenmaiier[6], MSc; Stine Finne Jakobsen[6], MSc, PhD; Paula Valle Simon[7,8], PhD; Antonio Javier Carcas-Sansuan[7,8], MD; Ole F Olesen[5], PhD; Oliver A Cornely[9,10,11,12], MD, PhD; Zoi Dorothea Pana[13], MSc, MD, PhD; Theoklis Zaoutis[3], MSCE, MD, PhD; Murat Akova[2,14], MD

[1]Department of Internal Medicine, Faculty of Medicine, Hacettepe Üniversitesi Aşı Enstitüsü Sıhhiye, Ankara, Turkey

[10]Department of Internal Medicine, Center for Integrated Oncology Aachen Bonn Cologne Duesseldorf (CIO ABCD) and Excellence Center for Medical Mycology (ECMM), Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany

[11]Clinical Trials Centre Cologne (ZKS Köln), Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany

[12]Center for Molecular Medicine Cologne (CMMC), Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany

[13]School of Medicine, Department of Basic & Clinical Sciences, University of Nicosia, Nicosia, Cyprus

[14]Department of Infectious Diseases and Clinical Microbiology, Faculty of Medicine, Hacettepe University, Ankara, Turkey

[2]Vaccine Institute, Hacettepe University, Ankara, Turkey

[3]Collaborative Center for Clinical Epidemiology and Outcomes Research (CLEO), Athens, Greece

[4]Chemical Engineering Department, King Fahd University of Petroleum & Minerals (KFUPM), Dhahran, Saudi Arabia

[5]European Vaccine Initiative (EVI), Heidelberg, Germany

[6]Centre of Excellence for Health, Immunity and Infections, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark

[7]Hospital La Paz Institute for Health Research (IdiPAZ), Madrid, Spain

[8]Servicio Madrileño de Salud, Madrid, Spain

[9]Cologne Excellence Cluster on Cellular Stress Responses in Aging-Associated Diseases (CECAD), Institute of Translational Research, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany

Corresponding Author:
Mine Durusu Tanriover, MD
Department of Internal Medicine, Faculty of Medicine, Hacettepe Üniversitesi Aşı Enstitüsü Sıhhiye, Ankara, Turkey

Abstract

Background: Despite the vast growth of vaccine studies during the SARS-CoV-2 pandemic, clinical trials failed to adequately represent diverse societal groups, resulting in the underrepresentation of specific populations. Understanding the factors hampering participation in vaccine clinical trials is essential to better identify structural, ethical, and communication barriers and to improve inclusive strategies for broader and more equitable participation in future vaccine research.

Objective: This study aimed to identify the perceived barriers to participation in vaccine trials among pregnant and lactating women, children aged younger than 18 years, and adults aged older than 65 years, as reported by professionals with expertise in vaccines or vaccine trials.

Methods: An online questionnaire was developed to gather personal information, group-specific barriers to vaccine trial participation, and suggestions to overcome these barriers. Data are presented as absolute (n/N) and relative frequencies (%).

Results: A total of 115 respondents, the majority (n=73, 63.5%) of whom were working in the scientific community, completed the online survey. Challenges in recruiting children were identified due to “safety or efficacy concerns,” “difficulties about ethics and regulatory issues,” and “lack of targeted information and communication.” Challenges in recruiting pregnant and lactating women were primarily “ethics and regulatory requirements,” “safety issues,” and “lack of prioritization or interest.” “Lack of information and communication channels adapted to the specific target group,” along with “lack of prioritization,” were the main challenges in recruiting older participants. Provision of health-related incentives, including but not limited to access to new treatments and receiving expert medical care, seems to be the top-rated motivation to participate in vaccine clinical trials.

Conclusions: The main challenges in recruiting pregnant and lactating women and children in vaccine trials involve safety and efficacy concerns, as well as lengthy ethical and regulatory processes. For older adults, key issues include poor communication

XSL•FO
RenderX

channels tailored to their needs, limited information, lack of prioritization, funding, infrastructure, and industry interest. Across all underrepresented groups, low awareness of and poor communication about research opportunities were major barriers. Additionally, mobility issues affected older adults, while lack of motivation and incentives affected children, and low health literacy and provider uncertainty impacted pregnant and lactating women. Improving communication infrastructure and enhancing communication strategies with clear, tailored messages to build trust and motivate participation are essential to improve inclusion in vaccine research.

(JMIR Public Health Surveill 2026;12:e89025) doi:10.2196/89025

KEYWORDS

barriers; COVID-19; vaccinations; pandemic preparedness; vaccine trials; volunteer registry

Introduction

Ensuring a diverse representation of population groups in clinical trials is essential for advancing health equity, as these groups may have different risk factors and variable responses to medical products [1]. Despite the rapid expansion of vaccine studies during the SARS-CoV-2 pandemic, clinical trials failed to adequately represent diverse societal groups, resulting in the underrepresentation of specific populations [2]. Underrepresented groups could include people who have inadequate access to health care for any reason, including unreachable or inadequate vaccination systems, health care provider discrimination, lack of health care provider recommendations, or legal restrictions [3,4]. A cross-sectional study examining trials registered at ClinicalTrials.gov identified that among 230 vaccine trials based in the United States, Black or African American individuals (constituting 10.6% of all participants) and adults aged 65 years or older (constituting 12.1% of the participants) were underrepresented compared with US census data [5]. On the other hand, certain populations with access to health care may still be underrepresented in clinical trials largely because of the exclusion of these population groups right at the initial stage of vaccine trial design. For instance, only 8 (9.5%) clinical trial protocols among the 84 phase II and phase III vaccine clinical trials reporting from 2009 to 2019 explicitly included pregnant people [6]. A semistructured qualitative interview conducted in the United Kingdom showed that although pregnant women acknowledged the importance of vaccine trials to demonstrate the safety and effectiveness of vaccines, they were reluctant about their own involvement in such trials [7].

We cannot underestimate the impact of increasing vaccine hesitancy on the underrepresentation of certain populations in vaccine clinical trials. For instance, Willis et al [8], who surveyed 1205 Arkansas adults in July and August of 2020 to evaluate COVID-19 vaccine hesitancy, showed that among Black or African American populations, 50% of respondents showed hesitancy toward receiving the COVID-19 vaccine, compared to 18.37% of White respondents.

Recently, Daniel et al [9] investigated the enablers and barriers to participation in vaccine trials. Safety concerns, institutional distrust and misinformation, sociocultural and religious norms, and logistical burdens were the most prominent barriers overall. Safety concerns were the predominant barrier among pregnant and pediatric participants. Recruiting children can be particularly challenging and may even lead to early trial termination due to failure to meet recruitment targets [10].

Strategies to better ensure diverse representation in vaccine clinical trials, particularly in preventive vaccine studies, are needed. Understanding the factors hampering participation in vaccine clinical trials for certain population groups may enhance future access to and engagement with a diverse selection of volunteers in clinical trials and facilitate better promotion of tools such as the VACCELERATE Volunteer Registry [11-17]. In addition to pinpointing the most underrepresented population groups, it is also crucial to identify the barriers that prevent professionals from involving these population groups in vaccine trials.

Within the VACCELERATE consortium, a task force was assigned to launch, disseminate, and analyze an extended online survey targeted at professionals with expertise in vaccines or vaccinations, with the aim of identifying barriers to participation in vaccine trials among pregnant and lactating women, children (aged <18 y), and older adults (aged >65 y) in the European region. For this reason, an internal consultation addressed to the National Coordinators was conducted to identify the challenges of vaccine trial participation of the underrepresented population groups [11]. Although the underrepresented populations could be diverse, based on the results of the initial survey, as well as a consultation with experts from the European Centre for Disease Prevention and Control and the World Health Organization, the task force selected (1) children (aged <18 y), (2) pregnant and lactating women, and (3) older adults (aged >65 y) to be the primary focus of the extended survey. Hence, this study focuses specifically on children, pregnant and lactating women, and older adults for the purposes of the analysis.

This study aimed to identify the perceived barriers to participation in vaccine trials among pregnant and lactating women, children aged younger than 18 years, and adults aged older than 65 years, as reported by professionals with expertise in vaccines or vaccine trials.

Methods

Study Design and Population

VACCELERATE is a pan-European clinical research network established to accelerate vaccine trials and pandemic preparedness [18]. The project aimed to streamline and harmonize clinical trial efforts across Europe. One of the work packages was dedicated to designing and implementing a European Union–wide, dynamic, harmonized, and sustainable

XSL•FO
RenderX

volunteer registry for phases II and III vaccine clinical trials with an initial focus on the COVID-19 pandemic and for the future expansion to any forthcoming European epidemic or pandemic. Within this work package, the study was designed to identify the causes of underrepresentation of pregnant and lactating women, children aged younger than 18 years, and adults aged older than 65 years as perceived by professionals in the vaccine field. It was not intended to explore the reasons why these population groups are not targeted at the trial design stage or to compare the magnitude of differences and the impact of perceived barriers across the underrepresented groups. Indeed, we did not inquire on the reasons why most of the vaccine trials were excluding these population groups and how we could get sponsors to design trials that are more inclusive and diverse. Instead, we have assumed that trials are open to children, pregnant and lactating women, and older adults.

An online 11-item questionnaire, divided into 3 sections, was developed by the authors. The full questionnaire can be found in Multimedia Appendix 1. The authors were engaged in several rounds of discussion to list the most appropriate response options for each group, as analyzing the results of the open-ended questions in this heterogeneous group of underrepresented populations would have been particularly challenging. Section A consisted of 6 questions, aiming to capture personal information of the survey respondents, such as (1) the capacity in which respondents filled out the survey (expert opinion or organization representative), (2) the underrepresented population groups that the respondents had experience with (children, pregnant and lactating women, and/or older adults), (3) affiliation of the respondent (civil society, funder, industry, and public health agencies), (4) geographical area of expertise, (5) highest level of expertise (spanning from junior to executive), and (6) total years of experience in vaccines and vaccination. Section B aimed to capture group-specific barriers to vaccine trial participation of certain population groups. The questions of section B followed a smart-branching design based on the answers of the respondents in section A. More specifically, in questions 7, 8, and 9 (referring to children, pregnant and lactating women, and older adults, respectively), survey respondents were asked to agree or disagree with statements provided regarding (1) the reasons they are unable to recruit volunteers from the given population group and (2) the main reasons the given population group does not participate in clinical trials as the professionals perceive from their side. Section C was titled “Overcoming Barriers” and involved 2 questions. Question 10 was a free text question, asking professionals for their opinion on what could improve equal access and representation of underrepresented population groups in COVID-19 vaccine research. Finally, question 11 asked on motivators of individuals to participate in a volunteer registry.

The VACCELERATE consortium included National Coordinators from Austria, Belgium, Cyprus, the Czech Republic, Denmark, France, Germany, Greece, Hungary, Ireland, Italy, Lithuania, the Netherlands, Poland, Portugal, Serbia, Slovakia, Spain, Sweden, Israel, Norway, Switzerland, and Turkey. A pilot study was done with the National Coordinators to ensure that the online platform for the survey was working appropriately. As these coordinators were all experts in vaccine research from 23 different countries, this pilot study also ensured that the questions were valid and sound before the survey questionnaire was circulated more widely. No patient or trial participant was targeted either in the pilot phase or in the actual survey.

To reach as many professionals as possible, National Coordinators were asked to list relevant agencies and organizations (ie, public health authorities, academia, nongovernmental organizations, and research development departments) in their network. After mapping down the potential participants, we created an extensive list of possibly eligible professionals. We used this list to disseminate our survey and attract more professionals with opportunity sampling, given that effective communication and collaboration with professionals are key to survey success. Invitation letters (Multimedia Appendix 2) included a statement to point out that the survey targeted professionals with expertise in vaccines or vaccinations to enroll qualified respondents.

Ethical Considerations

Ethics Approval

The study was conducted in accordance with the Declaration of Helsinki and approved by the Hacettepe University Ethics Committee (approval 2023/06 - 30; approved on April 4, 2023).

Informed Consent

Informed consent was obtained from all survey respondents prior to answering the questions of the survey. Respondents could withdraw at any point before submission by simply closing their web browser.

Privacy and Confidentiality

The survey was hosted by LimeSurvey, a platform fully compliant with the General Data Protection Regulation, and was accessible via the VACCELERATE stakeholder survey link [19]. All collected data were anonymized; no traceable information was collected.

After clicking on submission, there was no way of identifying the respondent or associating their answers with them. As such, the answers could not be revoked once submitted. No identifying information of the participants was collected or published.

Participant Compensation

Participants did not receive any compensation for participation.

Statistical Analysis

Data are presented as absolute (n/N) and relative frequencies (%). Statistical analysis was performed using Microsoft Excel. Free-text data are presented as raw answers and categorized.

Results

Overview

As a first step, a list was prepared, including contact information of professionals and institutions to serve as the basis of the dissemination process. Afterward, the list was refined to contain 219 email addresses from 26 countries. Invitation letters for the professionals to participate in the online survey were initially

XSL•FO
RenderX

sent on June 8, 2023. The invitation letter and the link to the survey were also distributed through personal communication networks. The initial deadline to reach the targeted number of at least 100 responses was set on September 30, 2023; however, as the target number of responses could not be reached, the deadline was extended 3 times until November 7, 2023.

Demographics and Expertise Area of Respondents

A total of 115 survey respondents with expertise in the vaccine field in the European region completed the survey. Most respondents (n=91, 79.1%) stated that they answered the survey on their own behalf, while 24 (20.9%) respondents indicated that they responded on behalf of or were appointed by an organization. The distribution of countries where the institution, company, or organization of the survey respondent is based is presented in Table 1.

Table . The distribution of countries where the institution, company, or organization of the survey respondent is based.

Country	European region	Values, n (%)
Turkey	Southern Europe	29 (25.2)
Sweden	Northern Europe	16 (13.9)
Greece	Southern Europe	10 (8.7)
Spain	Southern Europe	8 (7)
The Netherlands	Western Europe	6 (5.2)
Austria	Western Europe	5 (4.3)
Norway	Northern Europe	5 (4.3)
Cyprus	Southern Europe	4 (3.5)
Germany	Western Europe	4 (3.5)
Switzerland	Western Europe	4 (3.5)
Australia	N/A[a]	3 (2.6)
Belgium	Western Europe	3 (2.6)
Italy	Southern Europe	3 (2.6)
Slovenia	Southern Europe	3 (2.6)
Denmark	Northern Europe	2 (1.7)
France	Western Europe	2 (1.7)
Lithuania	Northern Europe	2 (1.7)
Albania	Southern Europe	1 (0.9)
Canada	N/A	1 (0.9)
Ireland {Republic}	Northern Europe	1 (0.9)
Poland	Eastern or Central Europe	1 (0.9)
Singapore	N/A	1 (0.9)
United States	N/A	1 (0.9)

[a]NA: nonapplicable.

Table 2 shows the population groups that the respondents work toward improving the health of, and Table 3 exhibits the study areas or groups that are most relevant to the respondents' areas of expertise.

XSL•FO
RenderX

Table . The distribution of the underrepresented groups that the survey respondents had experience with (N=115).

Groups[a]	Values, n (%)
All	38 (33)
Adults aged >65 y	51 (44.3)
Pregnant women	31 (27)
Lactating women	25 (21.7)
Children aged <5 y	35 (30.4)
Children aged 5 - 11 y	25 (21.7)
Children aged 12 - 17 y	28 (24.3)

[a]Experience among groups may be overlapping.

Table . Distribution of the area of expertise of the survey respondents.

Area of expertise	Values, n (%)
Civil Society (eg, NGOs[a], patient advocacy groups, etc)	9 (7.8)
Funder (providers of nonrepayable funds such as public and other traditional R&D[b] funders)	1 (0.9)
Industry	11 (9.6)
Policy and decision makers	7 (6.1)
Public health agency, including European (HERA[c], ECDC[d], EC[e]) and international (WHO[f])	21 (18.3)
Regulatory Authority (including EMA[g])	1 (0.9)
Scientific community (academia, RIs[h] supporting vaccine R&D or other relevant global health fields)	73 (63.5)
Vaccine development alliances (biomedical R&D RIs [National, European, International, PDPs[i], and other major initiatives])	5 (4.3)
I do not wish to specify	5 (4.3)

[a]NGO: nongovernment organizations.
[b]R&D: research and development.
[c]HERA: Health Emergency Preparedness and Response Authority.
[d]EDCD: European Centre for Disease Prevention and Control.
[e]EC: European Commission.
[f]WHO: World Health Organization.
[g]EMA: European Medicines Agency.
[h]RI: research infrastructure.
[i]PDP: product development partnership.

The expertise of the respondents was relatively equally divided among the 3 groups, whereas 1 in 3 respondents reported having experience with all groups. While 21.7% (25/115) of the respondents declared expertise at the entry and associate levels (<5 y) in the field of vaccines and vaccination, 39.1% (45/115) had expertise at the midsenior level (≥5 y), 27% (31/115) were at the director level, and 12.2% (14/115) were at executive levels. Overall, 44.3% (51/115) of the respondents had an experience of 15 years or more in the fields of vaccine and vaccination. Most (73/115, 63.5%) of the respondents were working in the scientific community, whereas 18.3% (21/115) and 9.6% (11/115) of them were working in public health agencies and the pharmaceutical industry, respectively. The geographical areas to which the respondents' expertise is most related are presented in Table 4. There was a heterogeneous representation of the Southern Europe or Mediterranean region, followed by Western Europe, which constituted the most represented geographical regions.

XSL•FO
RenderX

Table . Distribution of the geographical area where the survey respondents' expertise in vaccines is or are mostly related[a].

Geographical area	Values, n (%)
Western Europe	39 (33.9)
Eastern Europe	17 (14.8)
Central Europe	11 (9.6)
Northern Europe	19 (16.5)
Southern Europe or Mediterranean	50 (43.5)

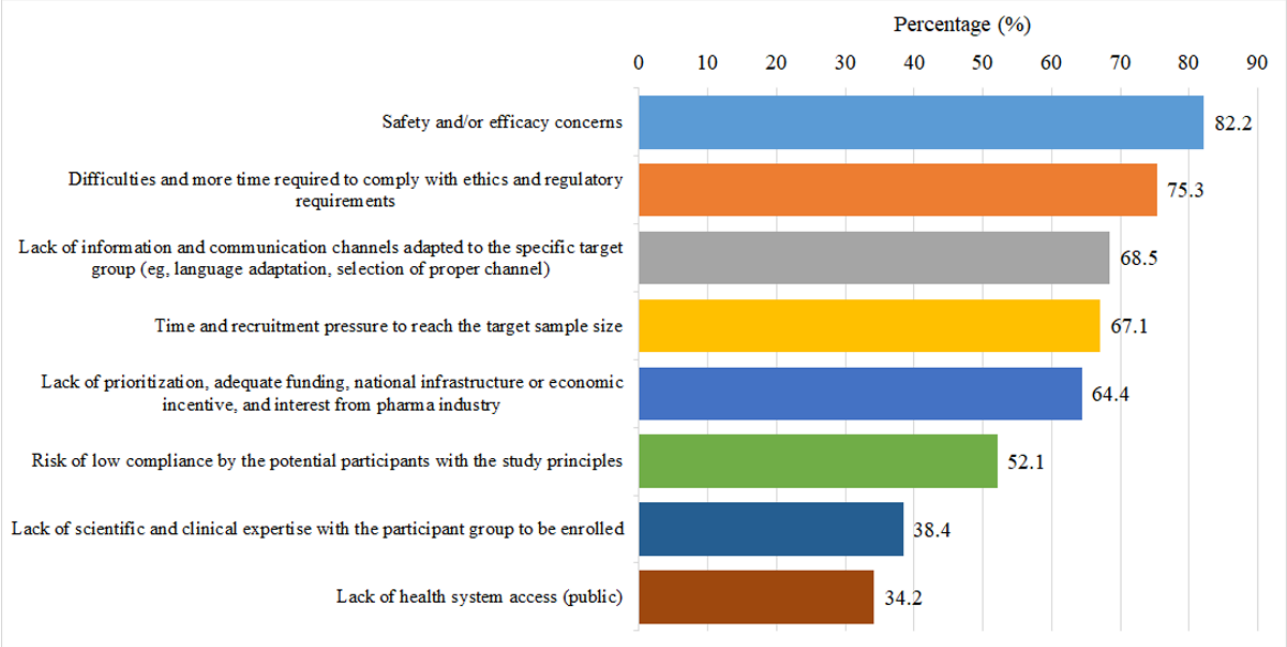
[a]Some respondents reported more than one geographical area.

Reported Challenges in Recruiting Certain Population Groups in Vaccine Trials

Pediatric Population

Figure 1 shows the challenges in recruiting children as participants; the 3 most pertinent being "safety and/or efficacy concerns" (60/73, 82.2%), "difficulties and more time required to comply with ethics and regulatory requirements" (55/73, 75.3%), and "lack of information and communication channels adapted to the specific target group" (50/73, 68.5%).

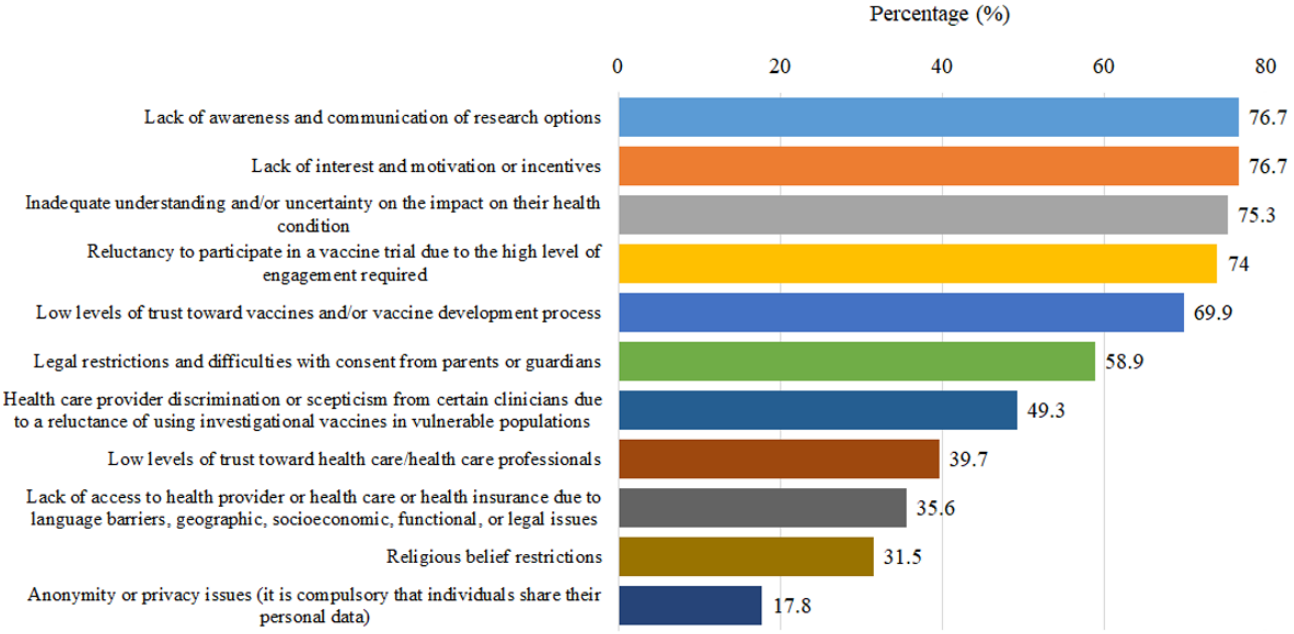
Figure 1. Distribution of reasons providers are unable to find or not willing to recruit a sufficient number of children participants in vaccine trials from the survey respondents' perspective.



The top 3 reasons why children are not participating in vaccine trials, as perceived by the responding professionals, were emphasized as "lack of awareness and communication of research options" and "lack of interest and motivation or incentives" (56/73, 76.7% for both), "inadequate understanding and/or uncertainty on the impact on their health condition" (55/73, 75.3%) and "reluctancy to participate in a vaccine trial due to the high level of engagement required" (54/73, 74%; Figure 2).

XSL•FO
RenderX

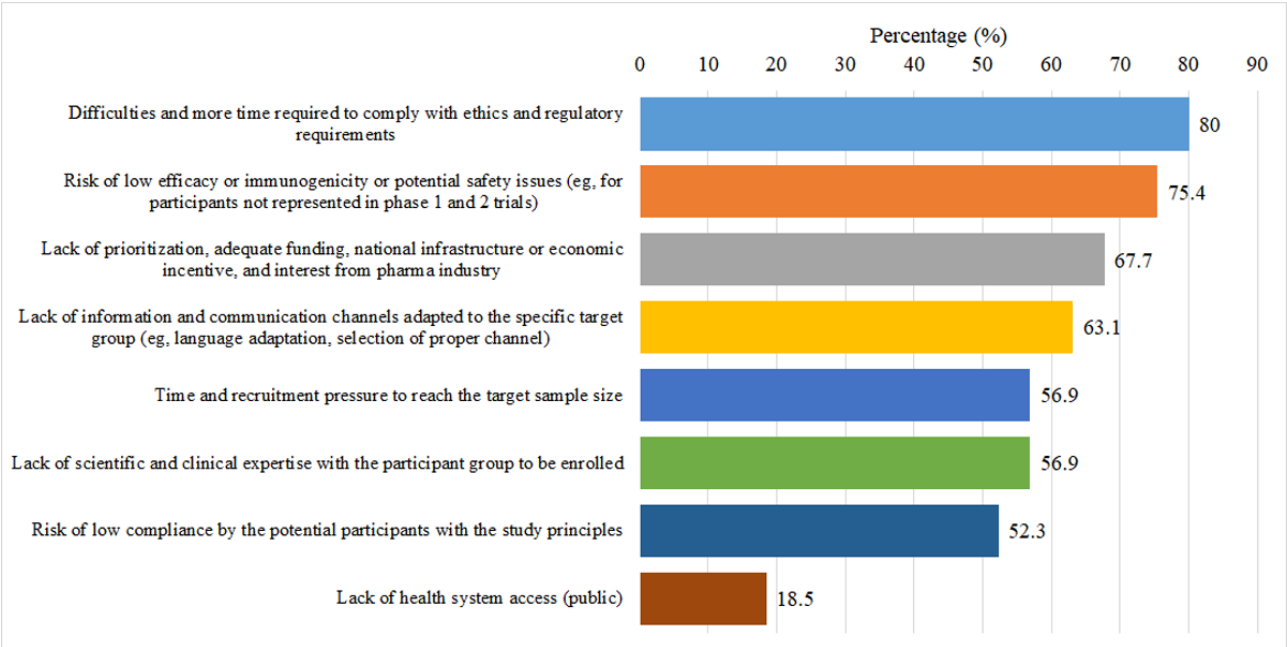
Figure 2. Distribution of the main reasons that children are not participating in vaccine trials from the survey respondents' perspective.



Pregnant and Lactating Women

The main challenges in recruiting pregnant and lactating women in vaccine trials were "difficulties and additional time required to comply with ethics and regulatory requirements" (52/65, 80%), "risk of low efficacy or immunogenicity or potential safety issues (eg, for participants not represented in phase I and II trials)" (49/65, 75.4%), and "lack of prioritization, adequate funding, national infrastructure or economic incentive, and interest from pharma industry" (44/65, 67.7%; Figure 3).

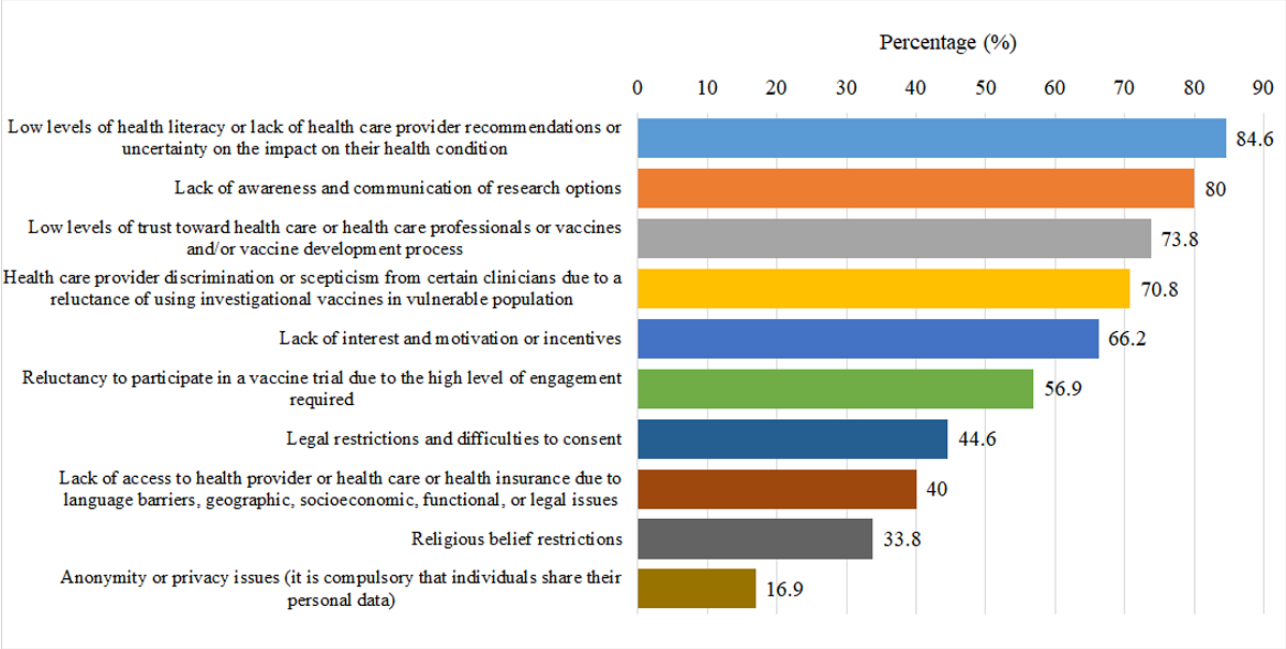
Figure 3. Distribution of reasons providers are unable to find or not willing to recruit a sufficient number of pregnant and lactating women participants in vaccine trials from the survey respondents' perspective.



"Low levels of health literacy or lack of health care provider recommendations or uncertainty on the impact on their health condition" (55/65, 84.6%), "lack of awareness and communication of research options" (52/65, 80%), and "low levels of trust toward health care, health care professionals, vaccines and/or vaccine development process" (48/65, 73.8%) were the main reasons why pregnant and lactating women are not participating in vaccine trials (Figure 4).

XSL•FO
RenderX

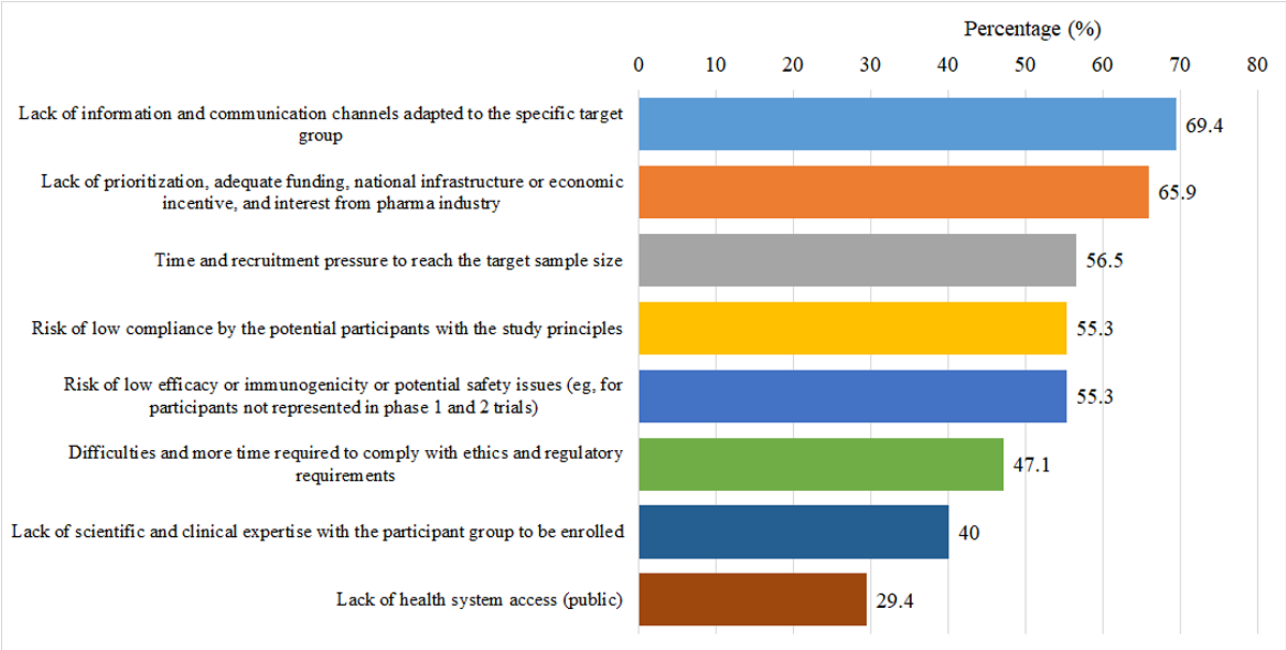
Figure 4. Distribution of the main reasons that pregnant and lactating women are not participating in vaccine trials from the survey respondents' perspective.



Older Population

Figure 5 depicts the challenges in recruiting older adult participants. The 3 most frequent challenges were "lack of information and communication channels adapted to the specific target group" (59/85, 69.4%), "lack of prioritization, adequate funding, national infrastructure, or economic incentive and interest from pharma industry" (56/85, 65.9%), and "time and recruitment pressure to reach the target sample size" (48/85, 56.5%).

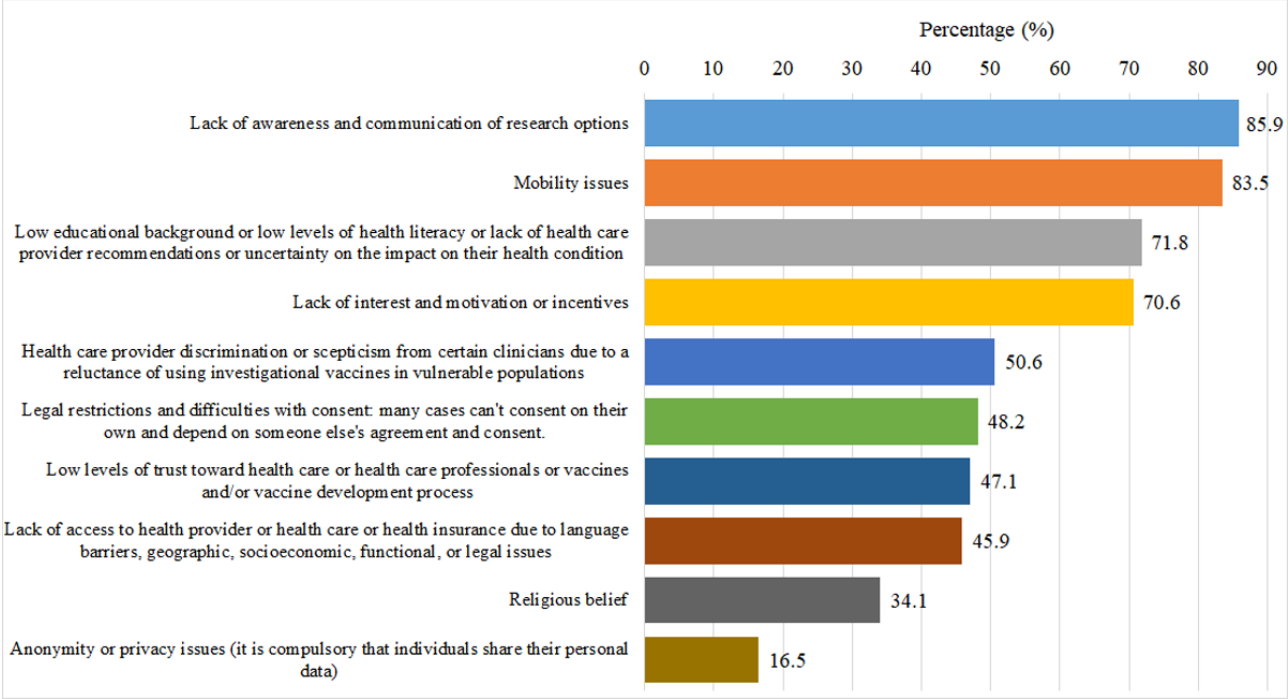
Figure 5. Distribution of reasons providers are unable to find or not willing to recruit a sufficient number of older adult participants in vaccine trials from the survey respondents' perspective.



The main barriers for which older adults are not participating in vaccine trials were "lack of awareness and communication of research options" (73/85, 85.9%), "mobility issues" (71/85, 83.5%), and "low educational background, low levels of health literacy, lack of health care provider recommendations or uncertainty on the impact on their health condition" (61/85, 71.8%; Figure 6).

XSL•FO
RenderX

Figure 6. Distribution of the main reasons that older adults are not participating in vaccine trials from the survey respondents' perspective.



Reported Facilitators in Recruiting Certain Population Groups in Vaccine Trials

The responses to the free-text question asking about how to improve equal access or improved representation in COVID-19 vaccine trials were mainly segregated into 5 categories: access and participation, collaboration and trust, information and communication, safety and compensation, and others. Factors that would motivate the participation in vaccine trials or studies are presented in Table 5.

Table . Distribution of the domains from the perspective of the survey respondents that motivate the participation of volunteers in a volunteer registry for vaccine trials or studies.

Motivation for participation	Values, n (%)
Provision of health-related incentives (including access to new treatments, getting expert medical care)	82 (71.3)
Provision of monetary incentives	59 (51.3)
Rapid access to disease-specific related information (eg, COVID-19 or Monkeypox)	64 (55.7)
Personal empowerment for active participation in the fight against an urgent public health threat	70 (60.9)
Sharing an opinion that influences the design of trials or studies	43 (37.4)

Discussion

The recent COVID-19 pandemic demonstrated that new vaccine development against an emergent pathogen can prevent the loss of millions of people, restore the livelihoods of communities, and ensure safe travel. However, real-life evaluation of vaccine efficacy can be challenging due to the difficulties of recruiting different segments of the population and ensuring representation of the wider community with different sociodemographic characteristics. We have shown from the professionals' perspective that recruitment of children, pregnant and lactating women, and older adults had specific challenges that required tailored approaches.

Unfortunately, there has been limited exploration of these challenges in vaccine trials from the perspective of other professionals to compare with the perspectives of our cohort. This is indeed one of the strengths of this study to provide novel findings where there is a knowledge gap. Studies have either explored the challenges based on race or ethnicity and primary language [20,21] or challenges in therapeutic trials such as cancer trials [22]. Therefore, we can only comment on the findings of other studies that report on the reflections of participants (either themselves or the caregivers of participants).

Recently, Meskell et al [23] published a Cochrane review on the factors that influence an individual's choice to participate in a vaccine trial during a pandemic or epidemic. The trial team emerged as a significant factor, impacting how trial information was conveyed, the design of the trial for participant convenience, the provision of financial incentives, and access to additional support services. Personal considerations also played a role, including concerns regarding vaccine side effects, efficacy, and

XSL•FO
RenderX

potential effects on daily life. Personal incentives encompassed access to vaccines, enhanced health outcomes, improved disease understanding, and the easing of pandemic or epidemic restrictions. Societal benefits included community assistance and contribution to scientific progress, often driven by memories of families and friends lost to the disease. Family and societal influences were also influential in potential participants' decision-making processes. In another study, public information about COVID-19 vaccines and vaccine trials in Europe was compiled from official sources (eg, governments, public agencies, and nongovernment organizations). While information for the public was abundant across Europe, there was a large variation in number, type, and target audiences for the information among countries. Notably, little or no information was found for population groups that are typically underrepresented in vaccine clinical trials [12].

In this study, we surveyed professionals with expertise in vaccines or vaccinations to understand, for each underrepresented group, (1) the reasons why the providers are unable to find or unwilling to recruit participants from the underrepresented groups in vaccine trials and (2) the reasons why the individuals from the underrepresented groups are not participating in vaccine trials. The challenges in recruiting children and pregnant and lactating women were mainly shaped around "safety or efficacy concerns," "difficulties about ethics and regulatory issues," and "lack of targeted information and communication." A rising threat of vaccine hesitancy among parents can be shaped around safety concerns. Caudal et al [24] illustrated that primary concerns revolved around vaccine adjuvants, the potential for short- and long-term adverse effects, and the risk of developing a disease or disability due to vaccination. Detoc et al [25] disclosed from the participant's point of view that having a favorable opinion about vaccines was the only independent factor associated with agreement to enroll in a preventive vaccine trial. Acceptance of participation in vaccine trials was closely linked to a generally positive attitude toward vaccines. Notably, information provided by physicians was pivotal in the decision-making process for 70.2% of those who accepted participation. Similarly, parents whose children participated in the phase III clinical trial of the Oral Cholera Vaccine—Simplified disclosed logistical issues, costs, misinformation, and concerns about trial procedures and potential side effects as the major barriers [26]. Vaccine hesitancy poses a potential obstacle to recruitment in vaccine trials; therefore, enhancing altruism and ensuring the quality of information provided are crucial factors in fostering acceptance of participation in such trials.

Meanwhile, the reasons why children and pregnant and lactating women are not participating were mainly stated around "lack of awareness and communication," "inadequate understanding and uncertainty," and "trust issues." Social capital, defined by Putnam, refers to "features of social organization such as networks, norms, and trust, that facilitate coordination for mutual benefit" [27]. Terraneo et al [28] found an association between social capital and the inclination to participate as volunteers in COVID-19 vaccine trials. These results may also be pertinent in encouraging parents and pregnant women to engage in vaccine trials. At the same time, fine-tuning the ethical and regulatory processes may not only accelerate recruitment but also help the sponsors to design trials that are not excluding these population groups right at the beginning. Interestingly, survey respondents stated "health care provider discrimination or skepticism due to a reluctance of using investigational vaccines" as a challenge in recruiting pregnant and lactating women, which corresponds to a lack of scientific and clinical knowledge not only among underrepresented populations but also among the health care providers. Upskilling the health care providers about vaccines and vaccinology or seeking researchers with expertise in pregnancy would be helpful to motivate the participation of pregnant and lactating women in vaccine trials.

Enrollment of older people has been a challenge during COVID-19 vaccine studies. However, it is crucial to ensure adequate representation of older individuals in clinical trials assessing preventive approaches to gain a comprehensive understanding of both the benefits and risks of interventions across the entire population. "Lack of information and communication channels adapted to the specific target group'' was especially apparent as the main challenge in recruiting older adult participants. A French digital national platform, namely Covireivac, was set up in June 2020 [29]. An analysis of this registry underscored that computer-based recruitment methods might not be suitable for older individuals, despite the presence of a toll-free hotline and third-party registration options on the platform [30]. Conversely, in Germany, following public announcements regarding a planned phase I trial of the vaccine candidate MVA-SARS-2-S against SARS-CoV-2, volunteers successfully reached out to the study team via email. Remarkably, one-third of these volunteers were aged older than 60 years [31]. These data show that effective and culturally tailored combinations of communication strategies can increase the participation of volunteers among older age groups.

Among the main reasons why older adults are not participating in trials from the health care professionals' perspective were also "lack of awareness and communication" and "low levels of health literacy" similar to the previous groups. However, a new challenge, "mobility issues," was added as a challenge specific to this target group. Frailty among the older adults not only impacts mobility but also the ability to engage, and this may pose a significant barrier. Limited mobility and lack of social support for older adults may be one of the most important challenges in their participation in clinical trials but also the easiest challenge to overcome by careful planning and resource allocation. However, among a survey of 400 older adults living in Atlanta, Georgia, two-thirds of the participants admitted that they were willing to participate in vaccine clinical trials; however, one-third had had limitations due to physical, mental, or emotional problems [32]. The need to travel to the clinical research site was also salient as a barrier.

Provision of health-related incentives, including but not limited to access to new treatments and receiving expert medical care, seems to be the top-rated motivation for volunteers to participate in COVID-19 vaccine trials. However, it should also be considered that monetary incentives, in particular, might lead to an overrepresentation of financially disadvantaged groups [33].

XSL•FO
RenderX

Heterogeneous results with regard to the barriers for different underrepresented population groups demonstrate the need for tailored policies for each group. Andrasik et al [34] suggest that all clinical research sites are required to have active community advisory groups with clear lines of communication to clinical research site staff and leadership. To align all the stakeholders in vaccine development and implementation studies on the same agenda, the gap between the scientific community, public health community, and the policymakers should be closed. Hence, to convert the outputs of this task into societal impact, the results will be shared with diverse groups of stakeholders who play a role in vaccine studies. This can help to overcome some of the obstacles to enroll underrepresented population groups in clinical vaccine studies just by fine-tuning the processes such as ethics and regulatory processes.

The first limitation of the study was the narrower assessment of the underrepresented populations, focusing only on children, pregnant and lactating women, and older adults, while omitting other underrepresented groups (eg, socioeconomically disadvantaged populations, ethnic minorities, migrants, or individuals with disabilities). Another limitation was the fact that the extended survey was addressed to a very specific target population, that is, professionals with hands-on experience in vaccines and/or vaccinations. Given the lack of a systematic measurement of these professionals across countries, we could not perform representative sampling. Furthermore, we did not assess the perceptions of the patients or trial participants themselves. Future studies should elaborate on our findings and exhaustively explore means to reach results from more diverse and representative samples. A possible direction toward this goal is communication with national and international professional and expert organizations. Another limitation is that the survey covered responses from 23 countries, although Turkey, Sweden, and Greece (29, 16, and 10 respondents, respectively) constituted 47.8% of the represented countries. Most of the responses reflected the personal, not institutional, views of the respondents. Moreover, 81.7% (94/115) of the respondents were either from the scientific or public health community (73/115, 63.5% and 21/115, 18.3%, respectively), indicating the gap in the countries and the professional groups that could be reached. However, there was relatively good representation in terms of geographical regions and the population groups with which the survey respondents work, although the results cannot be generalized beyond Europe. Finally, we have used different response options instead of string responses for each underrepresented group in the survey for the sake of a quantitative descriptive analysis, but this might have been a potential source of bias as a few of the response options were not identical across groups. The responses should be interpreted as insights from experts in the vaccine field rather than reflections of trial participants. The results can therefore not directly point out the perspectives of participants from underrepresented populations but rather guide the design of future studies to get personal insights from these populations.

In conclusion, the major challenges stated by the providers for recruiting pregnant and lactating women and children were around safety or efficacy concerns and difficulties, and more time required to comply with ethics and regulatory requirements. On the other hand, the major challenges in recruiting participants from the older population were the lack of information and communication channels adapted to the specific target group and lack of prioritization, adequate funding, national infrastructure, or economic incentive and interest from pharma industry. While the lack of awareness and communication of research options was a common major perceived barrier for the participation of all the underrepresented groups, mobility issues emerged as another prominent factor for the older people. Lack of interest and motivation or incentives was another major perceived barrier for the children, while low levels of health literacy or lack of health care provider recommendations or uncertainty on the impact on their health condition were highlighted for the pregnant and lactating women. Improvement in communication infrastructure and channels to convey effective, easy-to-understand, and tailored messages to build trust in vaccines and convey the research options to motivate people to participate in vaccine trials seems to be a crucial step to increase the participation of the underrepresented groups.

Acknowledgments

This work is presented on behalf of the VACCELERATE consortium [18]. The authors declare that no generative artificial intelligence tools were used in any portion of the manuscript preparation.

Funding

This work was supported by the European Union Horizon 2020 research and innovation program under grant 101037867 (VACCELERATE). The research leading to these results was conducted as part of the VACCELERATE consortium. For further information, please refer to the official website of the VACCELERATE network (https://vaccelerate.eu/).

Data Availability

The dataset is available from the authors upon reasonable request.

Authors' Contributions

Conceptualization: ZDP, DP, GCT, OAC, and IK
Data curation, preparation of stakeholders' list: all authors
Formal analysis: MDT

XSL•FO
RenderX

Funding acquisition: ZDP, OAC
Investigation: MDT, DP, MA, CA, SFJ, PVS, AJC-S, IK
Methodology, questionnaire development, review: SFJ, PVS, AJC-S, DP, MA
Project administration: RDM, DP
Supervision: ZDP, MA, TZ
Validation: MDT, DP, MA
Visualization: MDT
Writing – original draft preparation: MDT
Writing – review & editing: all authors

Conflicts of Interest

MDT received a research grant within the scope of the HORIZON 2020 grant agreement. OAC reports grants or contracts from BMBF, Cidara, DZIF, EU-DG RTD, F2G, Gilead, MedPace, MSD, Mundipharma, Octapharma, Pfizer, Scynexis; consulting fees from Abbvie, AiCuris, Basilea, Biocon, Boston Strategic Partners, Cidara, Seqirus, Gilead, GSK, IQVIA, Janssen, Matinas, MedPace, Menarini, Molecular Partners, MSG-ERC, Mundipharma, Noxxon, Octapharma, Pardes, Partner Therapeutics, Pfizer, PSI, Scynexis, Seres, Shionogi, The Prime Meridian Group; speaker and lecture honoraria from Abbott, AbbvieAbbVie, Akademie für Infektionsmedizin, Al-Jazeera Pharmaceuticals/Hikma, amedes, AstraZeneca, Deutscher Ärzteverlag, Gilead, GSK, Grupo Biotoscana/United Medical/Knight, Medscape/WebMD, MedUpdate, MSD, Moderna, Mundipharma, Noscendo, Paul-Martini-Stiftung, Pfizer, Sandoz, Seqirus, Shionogi, streamedup!, Touch Independent, Vitis; payment for expert testimony cidaraCidara; participation on a DRC or DSMB for Cidara, IQVIA, Janssen, MedPace, PSI, Pulmocide. AJC-S received a grant with code ICI21-00065 from Instituto de Salud Carlos III. Ministerio de Educación, Innovación y Universidades. CA is the associate editor for *JMIR Public Health and Surveillance* and the *Journal of Medical Internet Research*. The remaining authors declare no conflicts of interest.

Multimedia Appendix 1
VACCELERATE survey question. The questionnaire was developed specifically for this study.
[PDF File, 430 KB - publichealth_v12i1e89025_app1.pdf]

Multimedia Appendix 2
VACCELERATE invitation letter.
[DOCX File, 161 KB - publichealth_v12i1e89025_app2.docx]

References

1. Kelsey MD, Patrick-Lake B, Abdulai R, et al. Inclusion and diversity in clinical trials: actionable steps to drive lasting change. Contemp Clin Trials 2022;116:106740. [doi: 10.1016/j.cct.2022.106740] [Medline: 35364292]
2. Xiao H, Vaidya R, Liu F, Chang X, Xia X, Unger JM. Sex, racial, and ethnic representation in COVID-19 clinical trials: a systematic review and meta-analysis. JAMA Intern Med 2023 Jan 1;183(1):50-60. [doi: 10.1001/jamainternmed.2022.5600] [Medline: 36469312]
3. Flanagan SM, Hancock B. Reaching the hard to reach--lessons learned from the VCS (voluntary and community sector). A qualitative study. BMC Health Serv Res 2010 Apr 8;10(1):92. [doi: 10.1186/1472-6963-10-92] [Medline: 20377850]
4. Ozawa S, Yemeke TT, Evans DR, Pallas SE, Wallace AS, Lee BY. Defining hard-to-reach populations for vaccination. Vaccine (Auckl) 2019 Sep;37(37):5525-5534. [doi: 10.1016/j.vaccine.2019.06.081]
5. Flores LE, Frontera WR, Andrasik MP, et al. Assessment of the inclusion of racial/ethnic minority, female, and older individuals in vaccine clinical trials. JAMA Netw Open 2021 Feb 1;4(2):e2037640. [doi: 10.1001/jamanetworkopen.2020.37640] [Medline: 33606033]
6. Minchin J, Harris GH, Baumann S, Smith ER. Exclusion of pregnant people from emergency vaccine clinical trials: a systematic review of clinical trial protocols and reporting from 2009 to 2019. Vaccine (Auckl) 2023 Aug 7;41(35):5159-5181. [doi: 10.1016/j.vaccine.2023.06.073] [Medline: 37442686]
7. Youssef E, Calvert A, Greening V, et al. Understanding the barriers and enablers to participation in vaccine trials in a pregnant population from diverse ethnic background in an inner-city UK hospital. PLoS One 2024;19(10):e0312799. [doi: 10.1371/journal.pone.0312799] [Medline: 39475877]
8. Willis DE, Andersen JA, Bryant-Moore K, et al. COVID-19 vaccine hesitancy: race/ethnicity, trust, and fear. Clin Transl Sci 2021 Nov;14(6):2200-2207. [doi: 10.1111/cts.13077] [Medline: 34213073]
9. Daniel N, Smith C, Miah N, et al. Enablers and barriers to participation in vaccine trials: a narrative synthesis. Vaccine (Auckl) 2026 Feb 15;73:128183. [doi: 10.1016/j.vaccine.2025.128183] [Medline: 41475153]
10. CanSino Biologics Inc. A study to evaluate the safety and immunogenicity of Ad5-nCoV and Ad5-nCoV-IH in coronavac immunized children and adolescents. ClinicalTrials.gov. 2023 Jun 15. URL: https://clinicaltrials.gov/study/NCT05169008 [accessed 2025-12-04]

XSL•FO
RenderX

11. Poulimeneas D, Koniordou M, Kousi D, et al. The challenges of vaccine trial participation among underserved and hard-to-reach communities: an internal expert consultation of the VACCELERATE consortium. Vaccines (Basel) 2023 Nov 29;11(12):1784. [doi: 10.3390/vaccines11121784] [Medline: 38140188]
12. Luís C, Di Marzo VR, Kaur M, et al. Vaccination rates in Europe are not associated with online media intensity. JCOM 2022;21(5):A05. [doi: 10.22323/2.21050205]
13. Salmanton-García J, Stewart FA, Heringer S, et al. VACCELERATE Volunteer Registry: a European study participant database to facilitate clinical trial enrolment. Vaccine (Auckl) 2022 Jul 29;40(31):4090-4097. [doi: 10.1016/j.vaccine.2022.05.022] [Medline: 35659449]
14. Salmanton-García J, Wipfler P, Valle-Simón P, et al. VACCELERATE Site Network: real-time definition of clinical study capacity in Europe. Vaccine (Auckl) 2023 Jun 13;41(26):3915-3922. [doi: 10.1016/j.vaccine.2023.05.006] [Medline: 37210309]
15. Argyropoulos CD, Leckler J, Salmanton-García J, et al. Enhancing public health communication regarding vaccine trials: design and development of the Pan-European VACCELERATE toolkit. JMIR Public Health Surveill 2023 Apr 3;9:e44491. [doi: 10.2196/44491] [Medline: 36878478]
16. Themistocleous S, Argyropoulos CD, Vogazianos P, et al. Perspectives of European Patient Advocacy Groups on Volunteer Registries and Vaccine Trials: VACCELERATE survey study. JMIR Public Health Surveill 2024 Apr 4;10:e47241. [doi: 10.2196/47241] [Medline: 38573762]
17. Malik S, Dorothea ZP, Argyropoulos CD, et al. Data interoperability in COVID-19 vaccine trials: methodological approach in the VACCELERATE project. JMIR Med Inform 2025 Mar 7;13:e65590. [doi: 10.2196/65590] [Medline: 40056469]
18. New publication: self-testing for respiratory viruses. VACCELERATE. URL: https://vaccelerate.eu [accessed 2025-12-04]
19. VACCELERATE stakeholder survey. LimeSurvey. URL: https://lime.omega-cro.com.tr/index.php/422312?newtest=Y&lang=en [accessed 2025-12-04]
20. Heffernan ME, Barrera L, Guzman ZR, et al. Barriers and facilitators to recruitment of underrepresented research participants: perspectives of clinical research coordinators. J Clin Transl Sci 2023;7(1):e193. [doi: 10.1017/cts.2023.611] [Medline: 37745931]
21. Price KA, Warsame R, Toure EA, et al. Qualitative study of health care professional perspectives on recruiting minority patients in cancer trials. Health Equity 2025;9(1):53-59. [doi: 10.1089/heq.2024.0184] [Medline: 40123842]
22. Yousafi S, Rangachari P, Holland ML. Barriers to recruitment and retention among underrepresented populations in cancer clinical trials: a qualitative study of the perspectives of clinical trial research coordinating staff at a cancer center. J Healthc Leadersh 2024;16:427-441. [doi: 10.2147/JHL.S488426] [Medline: 39502080]
23. Meskell P, Biesty LM, Dowling M, et al. Factors that impact on recruitment to vaccine trials in the context of a pandemic or epidemic: a qualitative evidence synthesis. Cochrane Database Syst Rev 2023 Sep 1;9(9):MR000065. [doi: 10.1002/14651858.MR000065.pub2] [Medline: 37655964]
24. Caudal H, Briend-Godet V, Caroff N, Moret L, Navas D, Huon JF. Vaccine distrust: investigation of the views and attitudes of parents in regard to vaccination of their children. Ann Pharm Fr 2020 Jul;78(4):294-302. [doi: 10.1016/j.pharma.2020.03.003] [Medline: 32434681]
25. Detoc M, Launay O, Dualé C, et al. Barriers and motivations for participation in preventive vaccine clinical trials: experience of 5 clinical research sites. Vaccine (Auckl) 2019 Oct 16;37(44):6633-6639. [doi: 10.1016/j.vaccine.2019.09.048] [Medline: 31543417]
26. Chapagain RH, Adhikari S, Kunwar K, et al. Parental motivations and perceived barriers to participating in pediatric vaccine clinical trials: findings from the OCV-S trial in Nepal. Vaccine X 2024 Dec;21:100585. [doi: 10.1016/j.jvacx.2024.100585] [Medline: 39686923]
27. Putnam RD. Social capital and public affairs. Bulletin of the American Academy of Arts and Sciences 1994 May;47(8):5 [FREE Full text] [doi: 10.2307/3824796]
28. Terraneo M, Quaglia V, Nuvolati G, et al. Social capital and willingness to participate in COVID-19 vaccine trials: an Italian case-control study. BMC Public Health 2022 Nov 21;22(1):2137. [doi: 10.1186/s12889-022-14562-2] [Medline: 36411406]
29. Deverez volontaire pour tester les vaccins!. COVIREIVAC. URL: www.covireivac.fr [accessed 2025-12-04]
30. Bonneton M, Sambourg J, Luong Nguyen LB, et al. A national platform for COVID-19 vaccine studies recruitment in France: Covireivac volunteer's characteristics. Hum Vaccin Immunother 2022 Nov 30;18(6):2109364. [doi: 10.1080/21645515.2022.2109364] [Medline: 36018773]
31. Shamsrizi P, Kramer FJ, Addo MM, Fathi A. Characterization of individuals interested in participating in a phase I SARS-CoV-2 vaccine trial. Vaccines (Basel) 2021 Oct 19;9(10):1208. [doi: 10.3390/vaccines9101208] [Medline: 34696316]
32. Raheja D, Davila EP, Johnson ET, Deović R, Paine M, Rouphael N. Willingness to participate in vaccine-related clinical trials among older adults. Int J Environ Res Public Health 2018 Aug 14;15(8):1743. [doi: 10.3390/ijerph15081743] [Medline: 30110896]
33. Kalbaugh CA, Kalbaugh JM, McManus L, Fisher JA. Correction: Healthy volunteers in US phase I clinical trials: sociodemographic characteristics and participation over time. PLoS ONE 2022;17(1):e0263362. [doi: 10.1371/journal.pone.0263362] [Medline: 35081171]

XSL•FO
RenderX

34. Andrasik MP, Broder GB, Wallace SE, et al. Increasing Black, Indigenous and People of Color participation in clinical trials through community engagement and recruitment goal establishment. PLoS ONE 2021;16(10):e0258858. [doi: 10.1371/journal.pone.0258858] [Medline: 34665829]

Edited by A Mavragani, T Sanchez; submitted 09.Dec.2025; peer-reviewed by H Anandan, L Souckova; revised version received 12.Feb.2026; accepted 25.Feb.2026; published 07.Apr.2026.

Please cite as:
Tanriover MD, Poulimeneas D, Tsopela GC, Kopsidas I, Argyropoulos C, Di Marzo R, Valdenmaiier O, Jakobsen SF, Valle Simon P, Carcas-Sansuan AJ, Olesen OF, Cornely OA, Pana ZD, Zaoutis T, Akova M
Barriers and Opportunities to Include Underrepresented Population Groups in Vaccine Trials: Cross-Sectional, Observational, Online Survey Study From the VACCELERATE Research Network
JMIR Public Health Surveill 2026;12:e89025
URL: https://publichealth.jmir.org/2026/1/e89025
doi:10.2196/89025

© Mine Durusu Tanriover, Dimitrios Poulimeneas, Grammatiki-Christina Tsopela, Ioannis Kopsidas, Christos Argyropoulos, Romina Di Marzo, Olena Valdenmaiier, Stine Finne Jakobsen, Paula Valle Simon, Antonio Javier Carcas-Sansuan, Ole F. Olesen, Oliver A Cornely, Zoi Dorothea Pana, Theoklis Zaoutis, Murat Akova. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 7.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Physical Therapy Provision After COVID-19 in the Netherlands: Retrospective Cohort Study

Renee Veldkamp[1], PhD; Peter Spreeuwenberg[1], Drs; Arie Cornelis Verburg[2,3], PhD; Thomas Johannes Hoogeboom[2,4,5], PhD; Rob A de Bie[6], PhD; Willemijn M Meijer[1], PhD

[1]Nivel, Netherlands Institute for Health Services Research, Otterstraat 118, Utrecht, The Netherlands

[2]Radboud Institute for Health Sciences, IQ Health, Radboud University Medical Center, Nijmegen, The Netherlands

[3]Association for Quality in Physical Therapy (SKF), Zwolle, The Netherlands

[4]Department of Rehabilitation, Physiotherapy Science and Sports, UMC Utrecht Brain Center, University Medical Center Utrecht, Utrecht, The Netherlands

[5]Research Group Innovation of Human Movement Care, University of Applied Sciences Utrecht, Utrecht, The Netherlands

[6]CAPHRI School for Public Health and Primary Care, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, The Netherlands

Corresponding Author:
Renee Veldkamp, PhD
Nivel, Netherlands Institute for Health Services Research, Otterstraat 118, Utrecht, The Netherlands

Abstract

Background: Many patients have been treated by physical therapists after an infection with COVID-19. The policy changes and unique situation of the COVID-19 pandemic necessitate a retrospective analysis.

Objective: This study examined physical therapy provision for adult patients recovering from COVID-19 in primary care in the Netherlands over the first 2.5 years of the pandemic. The primary aim was to examine changes in the amount of health care provision and associated factors. Furthermore, we aimed to gain insight into the patient population and changes over time, as well as the use of advised outcome measures in physical therapy after COVID-19.

Methods: We conducted a retrospective cohort study using routine health care data from the electronic health records of physical therapy practices in Dutch primary care. In total, 71,152 COVID-19 treatment trajectories initiated between March 1, 2020, and December 18, 2022, were included. Multilevel generalized linear mixed model analyses examined changes over time in the patient population and in the amount of physical therapy within treatment trajectories (ie, duration in weeks and number of consultations). Outcome measures were analyzed using descriptive statistics.

Results: Completed treatment trajectories had an overall duration of 24 weeks with 28 consultations, with a decreasing trend in the amount of physical therapy provided over time ($P<.001$). Also, a parabolic effect of age ($P<.001$) was found, showing a maximum number of consultations and duration of the treatment trajectory among older patients aged 60 to 79 years, with smaller quantities for younger and older patients. Treatment trajectories for women were 1.5 weeks longer than those for men (β 1.5, SE 0.12; $P<.001$) but did not include more consultations ($P=.76$). In 83% (59,184/71,152) of treatment trajectories, a baseline score on at least one of the included advised outcome measures was registered. Median (IQR) baseline scores were 420 (340-490) m on the 6-minute walking test (n=27,035); 13 (10-17) seconds on the 5-times sit-to-stand test (n=2337); 5.8 (5 - 6.3) points on the Fatigue Severity Scale (n=995); 8 (7-10) points on the Patient-Specific Functional Scale (PSFS; n=55,074); 10 (9-12) points on the Short Physical Performance Battery (n=4665); and 29 (22-38) and 30 (24-40) kg of grip strength, respectively, on the left (n=12,306) and right (n=12,293) hand. Difference scores could be examined for the 6-minute walking test, grip strength, and the PSFS.

Conclusions: This study demonstrates the potential of routinely collected health care data to describe real-world physical therapy provision for COVID-19 in Dutch primary care. Patients showed severe perceived fatigue, reduced functional capacity, and lower-limb strength at the start of physical therapy. The amount of physical therapy declined over the first 2.5 years of the pandemic, differed by age and gender, and was generally lower than the amount reimbursed by the Dutch government through basic health insurance.

(JMIR Public Health Surveill 2026;12:e88633) doi:10.2196/88633

KEYWORDS

COVID-19; physical therapy; electronic health records; routine health care data; Dutch primary care

XSL•FO
RenderX

Introduction

Since the start of the SARS-CoV-2 pandemic in early 2020, millions of people have contracted COVID-19 [1]. Patients who were admitted to the hospital could develop acute respiratory distress syndrome, muscle weakness, or post–intensive care syndrome [2-4]. Furthermore, it is estimated that between 3% and 6% of people continue to experience symptoms beyond the acute phase of the infection, known as postacute COVID-19 syndrome [5-8]. Common persistent symptoms include fatigue, dyspnea, memory problems, posttraumatic stress disorder, sleep disorders, anxiety, depression, and joint pain [9-11], substantially impacting general health, activities of daily living, quality of life, and working capacity of patients [9,12-14]. Considering the impact of COVID-19 and postacute COVID-19 syndrome on motor and cognitive functions, many patients have been treated by allied health care providers, mostly by physical and exercise therapists (hereafter, physical therapists) [15]. Physical therapists play an important role in supporting respiratory function, guiding active mobilization, ensuring a good load vs load capacity [16], and providing therapy for impairments in movement-related functioning affecting daily activities or participation [17].

Therefore, a temporary policy change was introduced in the Netherlands during the early months of the pandemic to enhance the accessibility to physical therapy after a COVID-19 infection under certain conditions [18]. Usually, in the Netherlands, physical therapy is included in the obligatory basic health insurance package only for a selection of diagnoses. Physical therapy for other diagnoses is paid out-of-pocket or reimbursed through voluntary supplementary health insurance. As of July 18, 2020, basic health insurance temporarily covered up to 50 physical therapy sessions for patients with COVID-19 over a 6-month period, with the option for a medically justified extension [18]. Furthermore, as COVID-19 was a new condition, treatment guidelines had to be developed based on the experiences of the first patients with COVID-19 and the opinions of experts. Over time, many studies have provided more insight into the biological and clinical aspects of the condition, and several adjustments to treatment guidelines were made [17]. However, it is not known how the policy regarding physical therapy for COVID-19 in primary care and the changing circumstances over time influenced the physical therapy use of patients with COVID-19. With such policy changes and unique circumstances, it is necessary to look back at how this was handled in order to learn from them and increase future pandemic preparedness.

Therefore, the overall aim of this study was to gain insight into the real-world physical therapy provision for adult patients recovering from COVID-19 in Dutch primary care over the first 2.5 years of the pandemic. In the Netherlands, several registries are available that receive routine health care data from the electronic health records (EHRs) of primary care physical therapy practices. Data from these registries can be used to build a retrospective cohort of the patients who sought physical therapy for symptoms of COVID-19, using real-world data. These data therefore offer the opportunity to look back on the period of the COVID-19 pandemic and learn from it. More specifically, the primary aim of this study was to examine the amount of physical therapy provided for patients after a COVID-19 infection over time and to identify which factors (ie, patient characteristics and course of the pandemic) were associated with it. The second aim was to provide insight into changes over time in the patient population starting with physical therapy after COVID-19 and in the clinical profile of these patients. The final aim was to examine the use of outcome measures as recommended in Dutch guidelines concerning physical therapy for COVID-19 [17].

Methods

Study Design and Setting

This retrospective cohort study was set up as part of a greater nationwide project in the Netherlands to evaluate the recovery of patients receiving allied health care treatment in primary care practices after a SARS-CoV-2 infection, the ParaCOV study (The Dutch Consortium Allied Healthcare COVID-19) [19]. Routine health care data from EHRs collected in the databases of (1) Nivel, the Netherlands Institute for Health Services Research, Primary Care Database (Nivel-PCD) [20], (2) the Dutch National Data Registry (in Dutch: LDK) of the Dutch Association for Quality in Physical Therapy (in Dutch: SKF), (3) the National Registry Exercise Therapy (in Dutch: LDO), and (4) the National Physical Therapy Database (in Dutch: LDF) of the Royal Dutch Society for Physical Therapy (in Dutch: KNGF) were combined. Together, these databases contain information on approximately 1 million unique patients receiving physical therapy in primary care for a broad range of diagnoses. The data in these registries are structured in the same way, enabling a smooth process of combining the datasets. To avoid duplicates, data from physical therapy practices affiliated with data collection via Nivel (Nivel-PCD, LDK, and LDO), which are all collected via the same route, were used as a starting point and supplemented with data from nonoverlapping practices affiliated with LDF. To define comorbidities (see *Outcomes*), data from approximately 500 general practices collected in Nivel-PCD, covering around 10% of the Dutch population, and the Nivel-PCD physical therapy practices were linked for the subset of patients for whom information was available in both of these databases.

Ethical Considerations

Data collected for the registries were pseudonymized at the source (within the practices), leaving out direct identifying information such as names and address. This study has been approved in accordance with the governance code of Nivel-PCD under number NZR-00322.046. The use of EHRs for research purposes is allowed under certain conditions. When these conditions are fulfilled, neither obtaining informed consent from patients nor approval by a medical ethics committee is obligatory for this type of observational study containing no directly identifiable data (Art 24 GDPR [General Data Protection Regulation] Implementation Act jo art 9.2 sub j GDPR [21]).

Cohort

In the EHRs of physical therapists, data on the provided care and accompanying information are registered per patient and

XSL•FO
RenderX

per diagnosis in a so-called treatment trajectory, wherein the diagnosis is registered using the diagnostic code for allied health care professionals (DCSPH-code [22]). Inclusion criteria for the cohort of physical therapy treatment trajectories in this study were (1) DCSPH-code 9363 (ie, COVID-19); (2) starting March 1, 2020, or later; (3) consisting of at least 2 consultations, that is, registered contacts on separate dates (see also *Outcomes*); and (4) adult patients (18 y and older). Physical therapists were only allowed to use this diagnostic code when a patient had a referral from a general practitioner or medical specialist indicating that the patient was in need of physical therapy after a COVID-19 infection. Data were available up to February 20, 2023. However, to reduce the chance of missing treatment trajectories based on not yet available data in the respective databases and to avoid the drop in starting treatment trajectories over the Christmas holidays, only treatment trajectories starting the latest on December 18, 2022, were included in the final cohort.

Additionally, the study period was divided into start weeks, running from Monday to Sunday, with March 1, 2020, falling into start week 1 and the week of December 12 to 18, 2022, being the last start week (start week 147). Each treatment trajectory was assigned to the week in which that treatment trajectory started, based on the date of the first consultation.

The amount of physical therapy provided can only be analyzed on treatment trajectories that have ended. Therefore, for those analyses, a subset of treatment trajectories was defined. A treatment trajectory was defined as being closed when the therapist registered it as such or when there had not been a consultation for more than 60 days. Based on the median duration of closed treatment trajectories in the full dataset, treatment trajectories starting in the last 22 start weeks of the final cohort were excluded to minimize the effect of including only shorter closed treatment trajectories near the end of the data collection. Therefore, this subset consisted of closed treatment trajectories starting the latest in start week 134 (September 18, 2022).

Data are provided per treatment trajectory, and the assumption was made that all treatment trajectories could be assigned to individual patients (as 95.4%, 64,668/67,825 of patients had only 1 registered treatment trajectory over time).

Outcomes

Patient Characteristics

Patient characteristics, such as “age,” “gender,” “duration of symptoms,” and “number of comorbidities,” were collected at the start of treatment. “Duration of symptoms” was registered by the physical therapist at the start of the treatment trajectory and refers to the duration in weeks from the start of the symptoms to the first visit of the patient at the physical therapist in three categories: (1) first visit within 0 to 2 weeks, (2) first visit within 3 to 12 weeks, and (3) first visit more than 12 weeks after the start of the symptoms for which the patient visits the physical therapist. The number of comorbidities (0, 1, or more than 1) was defined for the subset of patients for whom data were available from (1) Nivel-PCD general practices and/or (2) registered outcome measures in the treatment trajectory. Per treatment trajectory, it was determined whether the patient was registered as having diabetes, obesity, heart failure, cardiovascular diseases and hypertension, chronic pulmonary disease, chronic kidney damage, and liver disease, in at least one of the data sources, or none of those comorbidities in either source. See Multimedia Appendix 1 for the definition of the comorbidities.

Population Characteristics

The subtype of the SARS-CoV-2 infection was unknown at the individual patient level but was preassigned to each treatment trajectory based on the most dominant virus subtype at the population level 9 weeks before the treatment start. The most dominant virus subtype was defined as the virus type that had the highest percentage of patients in that week, according to the Dutch National Institute for Public Health and the Environment (in Dutch: RIVM [23]). A delay of 9 weeks before the start of the treatment trajectory was chosen based on the median time between the start of the infection and the start of the physical therapy treatment trajectory, as defined in the data from the national prospective cohort study on allied health care in patients recovering from COVID-19 [19].

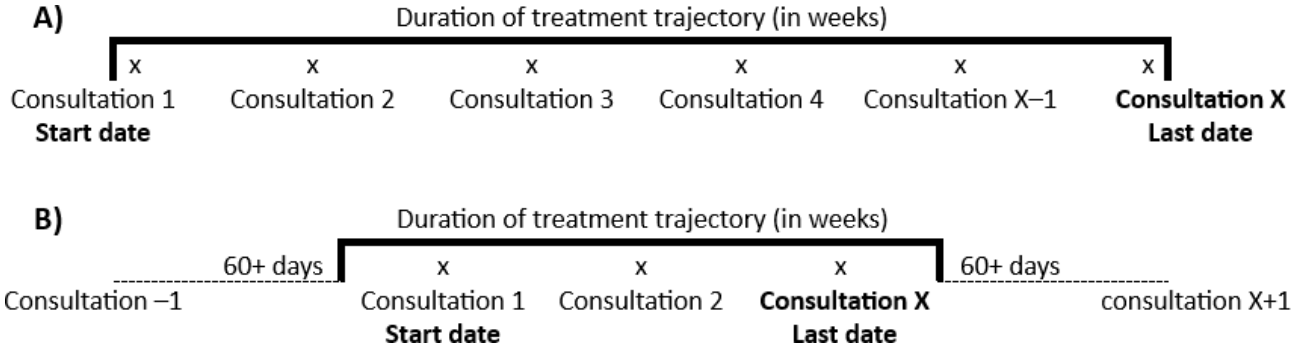
Amount of Health Care Provision Per Treatment Trajectory

The provided health care per treatment trajectory is registered as so-called care products. A consultation was defined as a date with a registered care product within the treatment trajectory, irrespective of the number of care products registered within the treatment trajectory on that specific date. For example, when 2 care products (eg, an “intake” and “treatment”) for a patient were registered on the same date in the same treatment trajectory, this counted as one consultation. Based on this, the variables were defined as follows (Figure 1):

- Start date: The first date in the treatment trajectory with a consultation (Figure 1A). An exception was made for treatment trajectories in which there was a gap of more than 60 days between the first and second consultations, while after the second consultation, there were more consultations within less than 60 days. In that case, the second consultation was taken as the start date of the treatment trajectory, assuming that the actual treatment started then (Figure 1B).
- Last date: The last date in the treatment trajectory with a consultation (Figure 1A), unless there was a gap of more than 60 days since the penultimate consultation, while the penultimate consultation was part of a streak of consultations (Figure 1B). Here, the assumption was made that a sole registered care product after 60 days was not part of the actual treatment but had more administrative or aftercare reasons.
- Duration of the treatment trajectory: The number of weeks between the assigned start date and the last date of the treatment trajectory (the bold line in Figure 1).
- Number of consultations: The number of consultations within the determined treatment trajectory. In the final cohort, only treatment trajectories consisting of at least 2 consultations were included.

XSL•FO
RenderX

Figure 1. Visualization of determined variables regarding physical treatment trajectories for COVID-19: start date, last date, number of consultations within, and duration in weeks of a treatment trajectory. The treatment trajectory is determined based on the registered consultations. The first consultation defines the start date and the last consultation, the last date of the treatment trajectory (A), in which first and last consultations that are separated by at least 60 days from the respective next and previous consultations are not considered (B). The duration of the treatment trajectory and the number of consultations within it are the number of weeks between the start date and last date (bold line) and the total number of consultations (consultation X).



Outcome Measures

Many different outcome measures can be registered in the EHRs and sent to registries [24]. Therefore, a selection of outcome measures was based on a combination of the position statement provided by the KNGF [17], guidelines for physical therapists provided by the ParaCOV-consortium [19], and data availability in the registries [25]. This concerned 4 functional therapist-assessed outcome measures and 2 patient-reported outcomes. The 5-times sit-to-stand test (5TSTS) was recommended for functional strength. In the 5TSTS, the patient is asked to get up and sit down in a chair 5 times in a row as fast as possible without using their hands [26]. The 6-minute walking test (6MWT) and Short Physical Performance Battery (SPPB) were (optionally) recommended for functional capacity. The 6MWT measures the maximum distance a patient can walk within 6 minutes [27]. The SPPB consists of 3 parts: balance, gait speed, and 5TSTS, on which 0 to 4 points can be scored [28]. The 5TSTS is performed as described above, and the gait speed task consists of walking 4 meters at a normal pace; in both subtasks, the time is recorded and scored. The balance tests include, in this particular order, side-by-side stand, semitandem stand, and tandem stand. Grip strength was recommended to be measured with a hand-held dynamometer as an indicator of overall upper-body strength [29].

Patient-reported outcomes were the Fatigue Severity Scale (FSS) for patient-reported fatigue [30] and the Patient-Specific Functional Scale (PSFS) for functional status [31]. In the PSFS, the patient is asked to rate the extent to which they feel limited in 3 self-selected activities that are most important to them. These activities should, according to the patient, be (1) relevant, (2) limiting performance, and (3) regularly performed (at least weekly). Here, the first one registered was taken as an outcome.

Lastly, the Global Perceived Effect (GPE) was also included as a measure of the perception of the patient regarding recovery and/or satisfaction with the treatment [17,32].

Measurements were considered to reflect baseline data when registered in the period from 2 weeks before the start date, as defined above, until 1 week (for treatment trajectories with a duration of 2 - 4 wk) or 2 weeks (minimum duration of 5 wk) after the start date. Similarly, measurements were considered to reflect the posttest when registered in the period from 1 to 2 weeks, depending on the duration of the treatment trajectory, before the last date of the treatment trajectory until 2 weeks after the last date. Difference scores were calculated for each outcome measure by comparing baseline and posttest values, when at least 1000 treatment trajectories were registered with that outcome measure. The threshold of 1000 treatment trajectories was set arbitrarily to ensure a substantial sample size, given the secondary use of the data over an extended study period. If grip strength of both hands was measured, the difference score on the hand with the greatest difference score was used.

Analysis

All data preparation and descriptive analyses were conducted in Stata (version 16; StataCorp LLC). Missing values were not imputed, and the significance level was set at 0.01, given the considerable size of the dataset. Descriptives of the total cohort were assessed by providing means and SDs, or frequencies and percentages, as appropriate. Statistical analyses were conducted in MLwiN (Centre for Multilevel Modelling, University of Bristol).

Patient Population Starting Physical Therapy for COVID-19 Over Time and Clinical Profile

Multilevel generalized linear mixed model analysis was conducted to examine the patient population starting physical therapy within the cohort over time (relative incidence). Defined levels were physical therapy practices, patients, and start week.

Relative incidence was determined as the proportion of treatment trajectories starting per age subgroup and gender subgroup in that start week, compared to all starting physical treatment trajectories for COVID-19 in primary health care in the Netherlands in the first 2.5 years of the pandemic. Notably, it does not give the incidence of patients starting physical therapy for COVID-19 in the Dutch population, as there is no underlying population available in the registries for persons who did not attend an associated physical therapy practice. In the multilevel analysis for relative incidence, the dichotomous incidence outcome was analyzed as logistic regression, and all start weeks (weeks 1 to 147 [March 1, 2020, to December 18, 2022]) in the final cohort were included. Gender (man and woman), age (18 - 39, 40 - 59, 60 - 79, and >80 y), and time (centered around the midpoint of the time range) as a polynomial of the

XSL•FO
RenderX

third-order (to capture the nonlinear curve over time) were combined. The combination of gender and age (8 categories in total) estimated a separate time effect for each of the 8 categories. Virus type (ie, original Wuhan, Alpha, Delta, and Omicron) was included as a fixed effect.

To examine the clinical profile of patients at the start of the treatment trajectory, the median and IQR of baseline scores on the selected outcome measures were examined.

Changes in Physical Therapy Provision for COVID-19 Over Time

Multilevel generalized linear mixed model analyses were also executed to examine the amount of physical therapy provided over time by (1) the number of consultations within closed treatment trajectories and (2) the duration of closed treatment trajectories in weeks. The defined levels were physical therapy practices, patients, and start week.

The variables for the amount of physical therapy provided, that is, the number of consultations in and the duration in weeks of a treatment trajectory, were analyzed as normal regressions on the subset of closed treatment trajectories. In these multilevel analyses, only closed treatment trajectories with start weeks 1 to 134 (March 1, 2020, to September 18, 2022) were included. In these 2 models, gender, age, and virus type were included as fixed effects. The same multilevel analyses were conducted on the subset of closed treatment trajectories wherein the number of registered comorbidities could be determined, while including the number of comorbidities (0, 1, and >1) as a fixed effect.

Use of Recommended Outcome Measures

Analyses of the use of recommended outcome measures were performed on the subset of closed treatment trajectories. To examine the use of the outcome measure in the closed treatment trajectories, the percentage of treatment trajectories, including the outcome measure as a baseline measure or as both baseline and posttest, was calculated. Furthermore, the median and IQR of the scores at baseline, at posttest, and of the difference score were calculated for each outcome measure. Difference scores less than the 25th percentile minus 1.5 times the IQR or greater than the 75th percentile plus 1.5 times the IQR were treated as outliers.

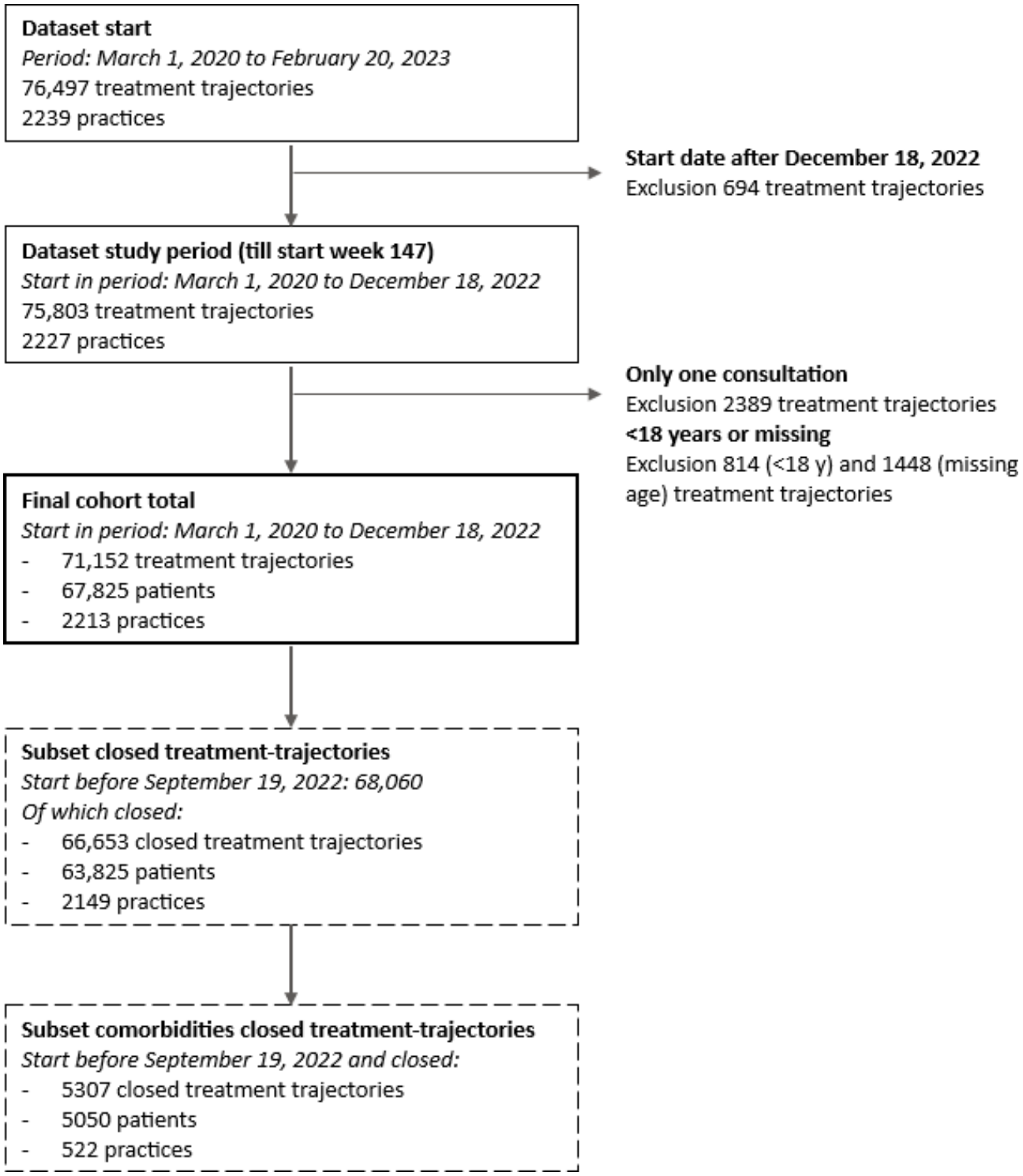
Results

Cohort Description

Figure 2 provides a flowchart of the included treatment trajectories, and Table 1 provides the descriptives of the cohort. In the final cohort, a total of 71,152 new treatment trajectories for COVID-19 were included. Most treatment trajectories started with the original variant as the dominant virus type (n=21,564, 30.3%), followed by Omicron (n=18,611, 26.2%), Delta (n=16,982, 23.9%), and Alpha (n=13,995, 19.7%) variants. In one of the treatment trajectories, the gender of the patient was missing. A total of 66,653 treatment trajectories were considered closed and thus included in the specific analyses of the number of consultations in, and the duration of, closed treatment trajectories (Figure 2). For the subset of closed treatment trajectories of patients for whom comorbidities could be determined (n=5307), 65.3% (n=3466) had no registered comorbidities, 24.0% (n=1275) had one comorbidity, and 10.7% (n=566) had multiple comorbidities.

XSL•FO
RenderX

Figure 2. Flowchart of included physical therapy treatment trajectories for COVID-19 in primary care in the Netherlands during the period 2020 to 2022 and their subsets.



XSL•FO
RenderX

Table . Descriptives of the included patients and physical therapy treatment trajectories for COVID-19 in the retrospective cohort over the years 2020 to 2022.

	Total cohort	Women	Men
Treatment trajectories, n (%)	71,152 (100)	45,478 (63.9)	25,673 (36.1)
Unique patients, n	67,825	43,253	24,571
Dominant virus type, n (%)			
Original	21,564 (30.3)	13,698 (30.1)	7866 (30.6)
Alpha	13,995 (19.7)	8322 (18.3)	5673 (22.1)
Delta	16,982 (23.9)	10,873 (23.9)	6108 (23.8)
Omicron	18,611 (26.2)	12,585 (27.7)	6026 (23.5)
Age, mean (SD)	52.2 (15.4)	49.9 (15.1)	56.2 (15.2)
Age group (y), n (%)			
20-39	15,511 (21.8)	11,684 (25.7)	3827 (14.9)
40 - 59	33,058 (46.5)	22,064 (48.5)	10,994 (42.8)
60 - 79	19,541 (27.5)	10,251 (22.5)	9290 (36.2)
>80	3041 (4.3)	1479 (3.3)	1562 (6.1)
Duration of symptoms at start of treatment (w), n (%)[a]			
0 - 2	6271 (9.7)	3787 (9.2)	2484 (10.7)
3 - 12	37,550 (58.3)	23,898 (57.9)	13,651 (58.8)
>12	20,611 (32.0)	13,526 (32.8)	7085 (30.5)
Baseline scores			
6MWT[b] (m)			
Score, median (IQR)	420 (340-490)	410 (340-480)	430 (346-504)
Count, n	27,035	17,339	9696
5TST[c] (s)			
Score, median (IQR)	13 (10-17)	13 (10-17)	13 (10-17)
Count, n	2337	1453	884
SPPB[d] (points)			
Score, median (IQR)	10 (9-12)	10 (9-12)	10 (8-12)
Count, n	4665	2961	1704
PSFS[e] (points)			
Score, median (IQR)	8 (7-10)	8 (7-10)	8 (7-10)
Count, n	55,074	35,060	20,014
FSS[f] (points)			
Score, median (IQR)	5.8 (5 - 6.3)	5.8 (5.1 - 6.3)	5.7 (4.9 - 6.3)
Count, n	995	652	343
GS[g]: left (kg)			
Score, median (IQR)	29 (22-38)	26 (20-30)	40 (32-49)
Count, n	12,306	7864	4442
GS: right (kg)			
Score, median (IQR)	30 (24-40)	27 (21-32)	42 (33-50)
Count, n	12,293	7860	4443
GS: unknown (kg)			

XSL•FO
RenderX

	Total cohort	Women	Men
Score, median (IQR)	30 (22-40)	26 (20-31)	42 (32-50)
Count, n	1845	1160	685
Subset comorbidities			
Closed treatment trajectories, n	5307	3355	1952
Assigned comorbidity category, n (%)			
0	3466 (65.3)	2353 (70.1)	1113 (57.0)
1	1275 (24.0)	720 (21.5)	555 (28.4)
>1	566 (10.7)	282 (8.4)	284 (14.5)

[a]There are missing data for this variable: total n=64,432, women n=41,211, and men n=23,220.

[b]6MWT: 6-minute walking test.

[c]5TST: 5-times sit-to-stand test.

[d]SPPB: Short Physical Performance Battery.

[e]PSFS: Patient-Specific Functional Scale.

[f]FSS: Fatigue Severity Scale.

[g]GS: grip strength.

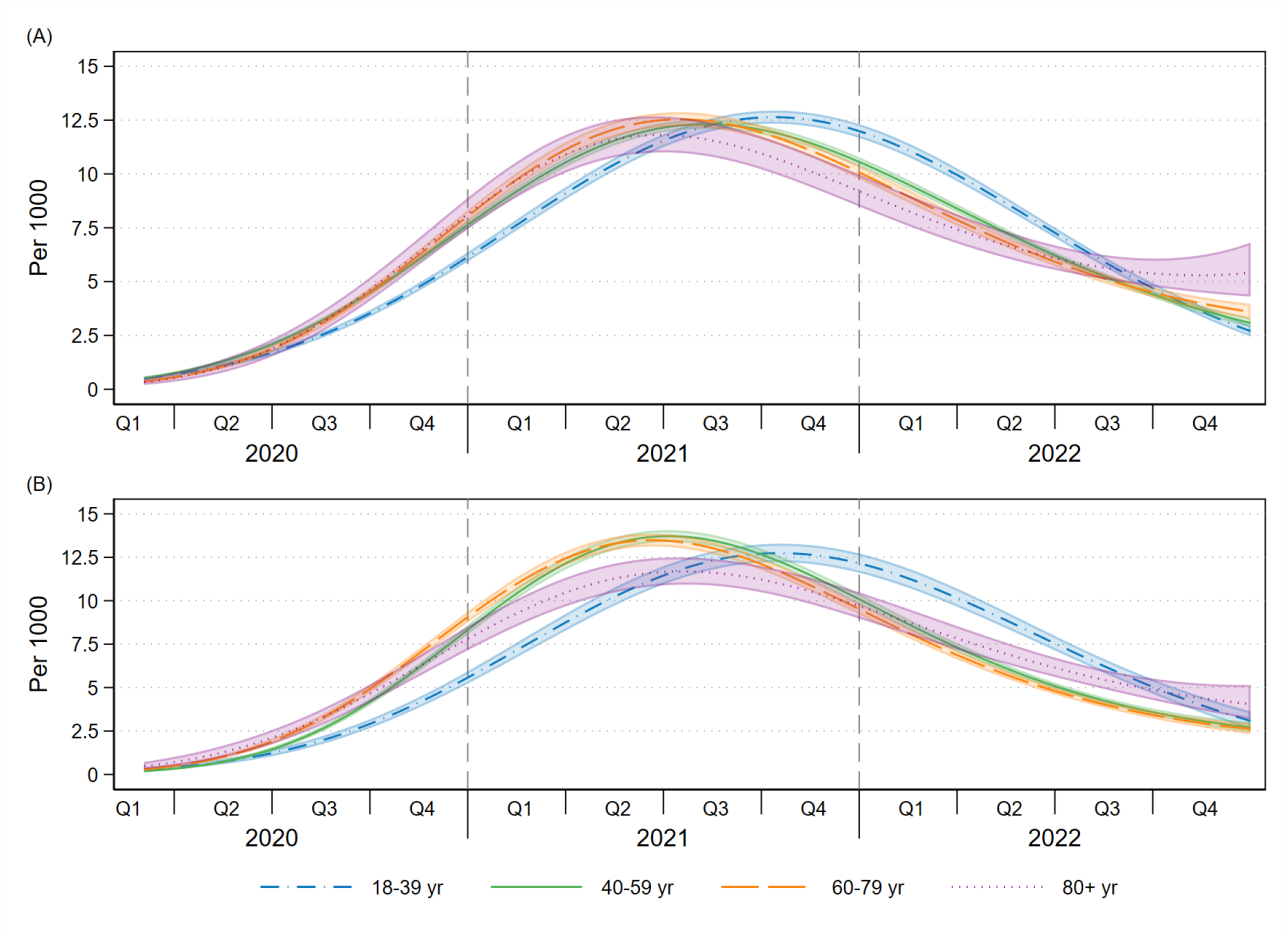
Patient Population Receiving Physical Therapy After COVID-19

Table 1 presents the baseline scores on the outcome measures registered around the start of the physical therapy treatment trajectories for COVID-19. In 83.2% (59,184/71,152) of the treatment trajectories, at least one measurement with any of the included outcome measures at baseline was registered.

Figure 3 provides a visual representation, for women (3A) and men (3B) separately, of the relative incidence and 95% CIs of the starting treatment trajectories for COVID-19 per start week, relative to all started treatment trajectories for COVID-19 in the cohort per age category, corrected for dominant virus type at the population level (see Multimedia Appendix 2 for the model). It demonstrates a difference between the subgroups in relative incidence of starting a physical therapy treatment trajectory after COVID-19 over the first 2.5 years of the pandemic. For both men and women, the 3 older age groups (40‑59, 60‑79, and >80 y) show the greatest relative incidence of starting a physical therapy treatment after COVID-19 in mid-2021. For the younger age group (18‑39 y), the peak for both women and men is in the beginning of the fourth quartile of 2021. In line with the number of treatment trajectories per subgroup (Table 1), the CIs are widest around the oldest age group (1479 treatment trajectories for women and 1562 for men) and smallest for the age groups of 40 to 59 years (22,064 treatment trajectories for women and 10,994 for men).

XSL•FO
RenderX

Figure 3. Relative incidence of the starting physical therapy treatment trajectories for COVID-19 in the Netherlands, included in the retrospective cohort over 2020 to 2022, per age category and per start week for (A) women and (B) men. Q: quartile.



Amount of Physical Therapy Provided Within Closed Treatment Trajectories

Table 2 shows the models for the duration (in weeks) and number of consultations per closed treatment trajectory. Figure 4 displays these respective models per start week, with the 95% CIs. Overall, the closed treatment trajectories had a duration of 24 weeks with 28 consultations in the first 2.5 years of the pandemic, while correcting for age, gender, and dominant virus type (Table 2). A decreasing trend in duration (in weeks) and in number of consultations within the closed treatment trajectories was shown over time of the start of the treatment trajectories, thus patients who started a physical treatment trajectory for COVID-19 in 2020 generally received more physical therapy sessions than patients who started the treatment trajectory in 2022. For the duration of the treatment trajectory, the model shows a steeper decline at the start of the pandemic (year 2020), which stabilizes over 2021, and shows some decline again in 2022 (Figure 4). The number of consultations per treatment trajectory decreases more linearly over time (Figure 4).

XSL•FO
RenderX

Figure 4. The duration (in weeks) and number of consultations in closed physical therapy treatment trajectories for COVID-19 in Dutch primary care, per start week over the period 2020 to 2022, corrected for age, gender, and dominant virus type. Q: quartile.

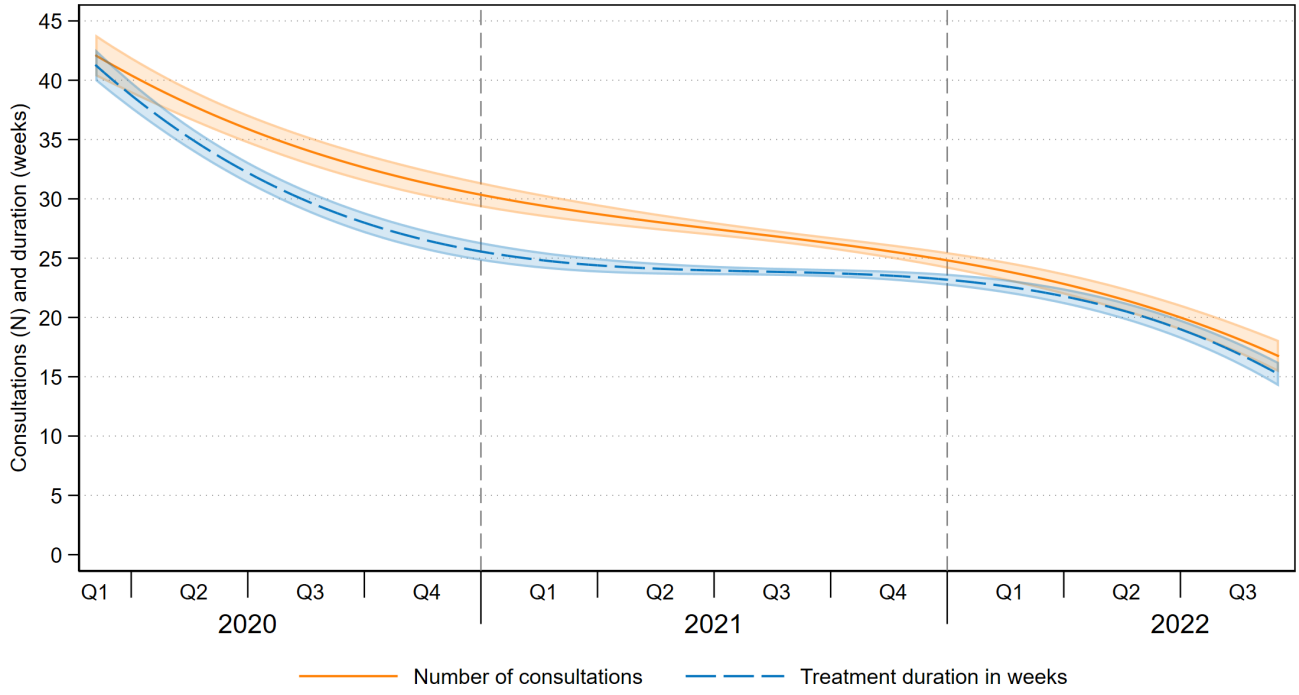


Table . Models for the duration (in weeks) and the number of consultations in closed physical therapy treatment trajectories for COVID-19 that started in the period 2020 to 2022 in Dutch primary care (N=66,652)[a].

Variable	Duration (wk)		Number of consultations	
	β (SE)	*P* value	β (SE)	*P* value
Constant	24.0 (0.22)	<.001[b]	27.7 (0.32)	<.001[b]
Time	−0.022 (0.01)	.07	−0.092 (0.02)	<.001[b]
$Time^2$	9.1×10^{-4} (1.1×10^{-4})	<.001[b]	3.5×10^{-4} (1.5×10^{-4})	.01
$Time^3$	-4.0×10^{-5} (3.0×10^{-6})	<.001[b]	-2.3×10^{-5} (3.9×10^{-6})	<.001[b]
Age group (y)[c]				
18 - 39	−1.3 (0.15)	<.001[b]	−1.9 (0.20)	<.001[b]
60 - 79	0.5 (0.14)	<.001[b]	1.4 (0.18)	<.001[b]
>80	−2.6 (0.30)	<.001[b]	−4.0 (0.39)	<.001[b]
Women[d]	1.5 (0.12)	<.001[b]	0.048 (0.16)	.76
Alpha[e]	−0.6 (0.26)	.02	−0.9 (0.34)	.008[f]
Delta[e]	−2.3 (0.53)	<.001[b]	−1.0 (0.68)	.13
Omicron[e]	−3.1 (0.68)	<.001[b]	−2.7 (0.88)	.002[f]

[a]For one patient within this selection, information on gender was missing.
[b]*P*<.001.
[c]Patients aged 40 to 59 years were the reference group for the age variable.
[d]Men was the reference group for gender.
[e]Original Wuhan virus type was the reference group for Alpha, Delta, and Omicron.
[f]*P*<.01.

Both models show a significant parabolic effect of the age categories (*P*<.001) on the amount of physical therapy provided, with a maximum among older patients aged 60 to 79 years. Overall, the duration of the treatment trajectory for patients aged between 18 to 39 years was 1.3 weeks shorter than for the reference group of patients in the 40 to 59 years age group and

XSL•FO
RenderX

1.8 weeks (1.3 plus 0.5) shorter than for the 60 to 79 years age group (Table 2). Overall, the number of consultations within a treatment trajectory was also highest for the age group of 60 to 79 years, namely 1.4 consultations higher than for the reference age group of patients (40 to 59 y) and 3.3 and 5.4 consultations higher, respectively, than for patients in the age groups 18 to 39 years and 80 years and older (Table 2). Gender was only significant in the model for the duration of the treatment trajectory, with an overall 1.5-week longer duration of the treatment trajectory for women than for men ($P<.001$).

Multimedia Appendix 3 shows the results of the regression analyses for the subset of closed treatment trajectories where it was possible to correct for the number of comorbidities (5307 treatment trajectories in total). The addition of the categories was not significant for the duration of the closed treatment trajectories (one comorbidity: β 0.649, SE 0.496; $P=.19$; more than one comorbidity: β 1.265, SE 0.707; $P=.07$), nor for the number of consultations in the closed treatment trajectories (one comorbidity: β 0.869, SE 0.665; $P=.19$; more than one comorbidity: β 1.008, SE 0.947; $P=.29$).

Registered Use of Recommended Outcome Measures

Table 3 shows the use of and the scores on the outcome measures within closed physical therapy treatment trajectories for COVID-19. Only for the PSFS, 6MWT, and grip strength were there more than 1000 baseline and posttest registrations. Overall, the PSFS was registered in 78% of the closed treatment trajectories (52,040/66,653), at baseline-only in 51% (n=33,836) and at baseline and as posttest in an additional 27% (n=18,204). A score on the 6MWT was registered as baseline in 28% (n=18,419) and as both baseline and posttest in an additional 11% (n=7316). For grip strength, this was in 15% (n=10,122) and 5% (n=3448) of the closed treatment trajectories, respectively.

XSL•FO
RenderX

Table . Usage of and scores on recommended outcome measures at baseline and at posttest in the closed physical therapy treatment trajectories for COVID-19 that started in the period 2020 to 2022 in Dutch primary care.

Test	Number of measurements		Scores on repeated measurements, median (IQR)		
	Baseline[a], n	Repeated measurement, n (outliers)	Baseline	Posttest	Difference
6MWT[b] (m)	25,735	7028 (288)	418 (349-483)	507 (440-575)	83 (40-140)
5TST[c] (s)	2281	501	—[d]	9 (7.3 - 12)	—
SPPB[e] (points)	4429	779	—	12 (10-12)	—
PSFS[f] (points)	52,040	18,096 (108)	8 (7-10)	1 (0 - 4)	-6 (–8 to –4)
FSS[g] (points)	897	191	—	3.7 (2.3 - 5)	—
GS[h]: men (kg)					
Left	4276	1139	—	—	—
Right	4265	1135	—	—	—
Unknown	646	173	—	—	—
Selected[i]	N/A[j]	1196 (120)	40 (32-47)	45 (38-53)	5 (2-9)
GS: women (kg)					
Left	7496	1839	—	—	—
Right	7493	1838	—	—	—
Unknown	1117	268	—	—	—
Selected	N/A	2013 (119)	26 (20-30)	30 (25-35)	4 (1-7)
GS: all (kg) selected	N/A	3209 (239)	30 (22-38)	34 (27-43)	4 (2-8)
GPE[k]: recovery	—	10,736	—	2 (1-2)	—
GPE: satisfaction	—	6632	—	1 (1-2)	—

[a]Numbers differ from Table 1, as the numbers in this table are given within the subset of closed treatment trajectories.
[b]6MWT: 6-minute walking test.
[c]5TSTS: 5-times sit-to-stand test.
[d]Not applicable.
[e]SPPB: Short Physical Performance Battery.
[f]PSFS: Patient-Specific Functional Scale.
[g]FSS: Fatigue Severity Scale.
[h]GS: grip strength.
[i]Scores for chosen hand.
[j]N/A: not available.
[k]GPE: global perceived effect.

On the PSFS, where a lower score reflects improvement, the median difference score was −6 (IQR −8 to −4) points over 18,096 closed treatment trajectories (Table 3). For the 6MWT and grip strength, higher scores reflect improvement. On the 6MWT, a median difference score of 83 (IQR 40-140) m was found (n=7028). On grip strength, median difference scores of 5 (IQR 2-9) kg for men and of 4 (IQR 1-7) kg for women were found over, respectively, 1196 and 2013 closed treatment trajectories.

The GPE-recovery was registered in 10,736 closed treatment trajectories, of which 81.6% (n=8761) indicated that the recovery was “much” to “very much better” since the beginning of the treatment. In 94.4% (n=6259) of the closed treatment trajectories in which GPE satisfaction was registered (n=6632), “very” to “very much satisfaction” with the treatment was indicated.

Discussion

Principal Findings

This retrospective cohort study, using real-world routinely recorded data of more than 60,000 patients from EHRs of physical therapy practices, provides insight into developments in the use of physical therapy after COVID-19 in the Dutch adult population during the first 2.5 years of the pandemic. Both the relative incidence of starting physical therapy and the amount of health care provided differed between groups of age and gender. Furthermore, more than 4 out of 5 of the treatment

XSL•FO
RenderX

trajectories included at least one measurement using an outcome measure that was recommended in Dutch treatment guidelines.

The retrospective cohort in this study seems representative of developments concerning coronavirus in the Netherlands during the first 2.5 years of the pandemic. For all subgroups of age and gender, the peak number of patients starting physical therapy after COVID-19 was in 2021, ranging from mid-2021 toward the end of 2021 for older age (>40 y) vs younger age groups (<40 y). These peaks are in accordance with the spread of coronavirus throughout the Netherlands, with peaks in the number of infections throughout 2021 and rapidly increasing infection rates among younger age groups, particularly with the Omicron variant. Furthermore, almost half of the treatment trajectories were for patients aged 40 to 59 years, and most of the patients were women. This is similar to characteristics of patients in a prospective study on allied health care in the Netherlands, with the inclusion period between March and June 2021, in which 64% of the participants were women and the mean age lay around 50 years [33]. Therefore, we expect that these data are representative of the provided physical therapy for COVID-19 in primary care in the Netherlands. The large sample size and use of real-world data increase the external validity of the findings.

The amount of physical therapy received was high for patients who started physical therapy after a COVID-19 infection in the beginning of the pandemic, but it rapidly decreased over the course of 2020. Notably, despite the reimbursement of 50 consultations over a period of 6 months for physical therapy after COVID-19 from the basic health insurance coverage [34], the number of consultations patients received was just over half of this number (28 consultations) for closed treatment trajectories. The latter indicates that the amount of physical therapy patients received was rapidly adjusted over time and that the possibility of declaring a large number of consultations for these patients has not been fully exploited by practitioners and patients. Notably, the amount of physical therapy received by patients after a COVID-19 infection showed a significant parabolic effect with age, with overall a maximum number of consultations and duration of the treatment trajectory among older patients aged 60 to 79 years. They had in general 3 consultations more per treatment trajectory than patients aged 18 to 39 years and 5 consultations more than patients aged 80 years or older. The number of consultations did not differ between men and women, but, on average, the treatment trajectory of women was 1.5 weeks longer, indicating a lower frequency of physical therapy for women compared to men.

The diagnostic code used in the primary physical therapy setting in the Netherlands does not allow for a differentiation between acute and postacute COVID-19, nor does it provide information on the type of symptom due to COVID-19 for which the patient sought help from the physical therapist. However, as, in the majority of the treatment trajectories, the patients visited the physical therapist after 3 to 12 weeks or after more than 12 weeks from the start of the symptoms for which the patient visited the physical therapist, the findings seem to apply primarily to patients with (early) postacute COVID-19. Additionally, the registered measures at baseline provide insight into the clinical profile at the start of the treatment trajectory of the patients visiting the physical therapist after a COVID-19 infection. In 83.2% (59,184/71,152) of the treatment trajectories, at least one measurement using any of the recommended outcome measures was registered at baseline. The baseline scores found in the present study are similar to the ones found in the nationwide prospective study, with a total of 992 patients, for the FSS [19], PSFS, 6MWT, 5TSTS, and grip strength [35]. Overall, these scores indicate severe fatigue [30], declined functional capacity by the 6MWT [36], and lower functional lower-limb strength [37] at baseline in this population, although there were no impairments in mobility according to the SPPB [38] nor weak grip strength for men or women [39].

Besides baseline measures, measures at posttest were also analyzed. In the majority of treatment trajectories (>80%), high perceived recovery and high satisfaction with the received physical therapy received were registered with the GPE. When interpreting these results, however, it is necessary to keep in mind that we made secondary use of routinely recorded health care data that were not registered for the goals of the present study. The outcome measures in the registries are not validated and the data might be biased to an extent as it is recorded in the EHRs by the therapists for different use than research. Furthermore, the lack of a control group in this study and knowledge on the specific content of the provided physical therapy does not allow for conclusions on causal relations or effectiveness. Therefore, caution is required when interpreting these measurement outcomes. There were 3 outcome measures for which a substantial sample size of closed treatment trajectories with both a baseline and posttest were available. On these, patients showed, on average, clinically relevant improvements from baseline to posttest in functional capacity (a >30.2 m increase on the 6MWT [40]) and patient-reported functional status (a >2.3 point decrease on the PSFS [41]) but not in grip strength (<5.0 - 6.5 kg [29]). This is in accordance with previous studies and a recent meta-analysis showing a significant increase in the 6MWT [35,42,43] and PSFS [35] compared to a control group or to baseline, but there was no clinically relevant improvement for grip strength [35,43]. Although the baseline scores and general direction of changes from baseline to posttest are similar to the ones found in the prospective study [35], the changes reported in this study are greater. As mentioned above, it is therefore important to underline that the study design in the present study only shows improvement over time at the population level within this specific population and does not allow for a distinction between natural healing over time, the effects of possible other interventions, or the effects of physical therapy.

Notably, although most of the literature supports the use of physical exercise–based rehabilitation for persons with long COVID-19 as a potentially promising and, although symptom-dependent, effective therapy [4,43-46], there are also concerns regarding the possible detrimental effect of physical exercise. A recent systematic review and meta-analysis reported that more than half of the individuals with postacute COVID-19 syndrome perceived postexertional malaise [47], and an intensive longitudinal cohort study showed that physical activity was, among other activities, associated with subsequently increased severity of various symptoms [48]. It is therefore

XSL•FO
RenderX

important that red flags, such as exertional desaturation and cardiac impairment following COVID-19, are ruled out before considering physical exercise training and that patients are assessed for postexertional symptom exacerbation, after which interventions should be modified accordingly [17,44,49]. In the present study, it is not possible to determine which patients would officially be classified as having postacute COVID-19, only that patients in the cohort were referred by a health care professional, as that was a prerequisite for reimbursement via basic health insurance.

Limitations

There are some limitations to the present study that are important to be mentioned. First of all, the observational nature of this type of study does not allow for conclusions on causal relations or effectiveness of the health care provided. Importantly, only patients who sought physical therapy for COVID-19 at a physical therapy practice participating in one of the registries are represented in the data. Therefore, no information from a control group nor from the underlying Dutch population is available (we have no information on people who do not visit a [affiliated] physical therapist). Second, as described earlier, use of outcome measures in the registries has not been validated, and for example, it is often unknown whether a patient-reported outcome such as the PSFS or GPE is indeed answered by the patient or instead by the therapist. Third, as the amount of physical therapy provided is based on treatment trajectories that are closed, naturally the proportion of shorter closed treatment trajectories is higher near the end of the data collection. However, the impact on the current results is likely modest, as we purposefully excluded the last 22 weeks in the data for the analyses of closed treatment trajectories, and the decline is already consistently visible well before the end of the data collection. Further, the addition of registered information on comorbidities was only possible for a small subgroup, and the combination of these lower numbers and the perhaps coarse outcome of “none,” “one,” or “more than one” registered comorbidities hampers the possibility to draw conclusions based on these data. Lastly, information on other possible confounders such as social economic position or ethnic origin was not available.

Conclusions

The present study shows the possibilities of using routinely collected health care data to gain insight into real-world physical therapy provision in primary care for COVID-19. The relative incidence of starting a physical therapy treatment trajectory for COVID-19 in primary care differed mainly between younger adults and middle-aged to older adults and was in accordance with the course of the pandemic in the Netherlands. Overall, patients showed severe perceived fatigue, declined functional capacity, and lower-limb strength at the start of the physical therapy and were satisfied with the treatment received. Furthermore, the amount of physical therapy provided declined over the first 2.5 years of the pandemic, differed between groups of age and between men and women, and was generally lower than the amount reimbursed by the Dutch government via the basic health insurance.

Acknowledgments

The authors would like to thank all physical therapy practices that participate in the various national registries: the Netherlands Institute for Health Services Research Primary Care Database (Nivel-PCD); the Dutch National Data Registry (LDK) of the Association for Quality in Physical Therapy (SKF); the National Registry Exercise Therapy (LDO); and the National Clinical Registry (LDF) of the Royal Dutch Society for Physical Therapy (KNGF). Generative AI was not used in any stage of preparation for this manuscript.

Funding

This project was funded by ZonMw Efficiency Studies (10390062010001) and received additional funding to set up the data-collection tool from the Royal Dutch Society for Physiotherapy, the Association for Quality in Physical Therapy, the Nivel Netherlands Institute for Health Services Research, Stichting Revalidatie en Wetenschap, and Maastricht University. The funding sources were not involved in data collection and analysis or in the publication process of the manuscript.

Data Availability

In the present study, data from the Nivel Primary Care Database (Nivel-PCD) have been used. However, access to data in Nivel-PCD is subject to the Nivel-PCD governance codes. Requests for data access can be directed to gegevensaanvragen@nivel.nl. Restrictions involve establishing a data-sharing agreement and obtaining approval from the appropriate Nivel-PCD governance bodies (privacy committee and steering committee).

Authors' Contributions

Conceptualization: WMM (lead), RAdB (equal), TJH-B (equal), RV (supporting)
Data curation: RV
Formal analysis: PS (lead), WMM (equal), RV (equal)
Funding acquisition: WMM (equal), TJH-B (equal)
Investigation: RV
Methodology: WMM (equal), PS (equal), RV (equal), ACV (supporting)

XSL•FO
RenderX

Project administration: RV (lead)
Supervision: WMM
Visualization: RV, PS
Writing – original draft: RV (lead), WM (supporting)
Writing – review & editing: RV (lead), WM (supporting), RAB (supporting), ACV (supporting), TJH-B (supporting), PS (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Definition and classification of comorbidities.
[DOCX File, 16 KB - publichealth_v12i1e88633_app1.docx]

Multimedia Appendix 2
Model of relative incidence.
[DOCX File, 18 KB - publichealth_v12i1e88633_app2.docx]

Multimedia Appendix 3
Models for the duration (in weeks) and number of consultations of closed treatment trajectories for the comorbidities subset.
[DOCX File, 16 KB - publichealth_v12i1e88633_app3.docx]

References

1. WHO COVID-19 dashboard. World Health Organization. URL: https://covid19.who.int/ [accessed 2026-08-25]
2. Grasselli G, Tonetti T, Protti A, et al. Pathophysiology of COVID-19-associated acute respiratory distress syndrome: a multicentre prospective observational study. Lancet Respir Med 2020 Dec;8(12):1201-1208. [doi: 10.1016/S2213-2600(20)30370-2] [Medline: 32861276]
3. Biehl M, Sese D. Post-intensive care syndrome and COVID-19 - implications post pandemic. Cleve Clin J Med 2020 Aug 5. [doi: 10.3949/ccjm.87a.ccc055] [Medline: 32759175]
4. Ceravolo MG, Anwar F, Andrenelli E, et al. Evidence-based position paper on physical and rehabilitation medicine professional practice for persons with COVID-19, including post COVID-19 condition: the European PRM position (UEMS PRM Section). Eur J Phys Rehabil Med 2023 Dec;59(6):789-799. [doi: 10.23736/S1973-9087.23.08315-6] [Medline: 38214046]
5. Assiri AM, Alamaa T, Elenezi F, et al. Unveiling the clinical spectrum of post-COVID-19 conditions: assessment and recommended strategies. Cureus 2024 Jan;16(1):e52827. [doi: 10.7759/cureus.52827] [Medline: 38406111]
6. A clinical case definition of post COVID-19 condition by a Delphi consensus, 6 October 2021. : World Health Organization; 2021 Oct 6 URL: https://iris.who.int/server/api/core/bitstreams/53c01712-a322-446e-ad9f-4e078b757267/content [accessed 2026-08-25]
7. Hastie CE, Lowe DJ, McAuley A, et al. True prevalence of long-COVID in a nationwide, population cohort study. Nat Commun 2023 Nov 30;14(1):7892. [doi: 10.1038/s41467-023-43661-w] [Medline: 38036541]
8. Post-covid [Article in Dutch]. RIVM (Rijksinstituut voor Volksgezondheid en Milieu). URL: https://www.rivm.nl/gezondheidsonderzoek-covid-19/kwartaalonderzoek-volwassenen/post-covid [accessed 2026-08-25]
9. Arienti C, Lazzarini SG, Andrenelli E, et al. Rehabilitation and COVID-19: systematic review by Cochrane rehabilitation. Eur J Phys Rehabil Med 2023 Dec;59(6):800-818. [doi: 10.23736/S1973-9087.23.08331-4] [Medline: 38214047]
10. Luo D, Mei B, Wang P, et al. Prevalence and risk factors for persistent symptoms after COVID-19: a systematic review and meta-analysis. Clin Microbiol Infect 2024 Mar;30(3):328-335. [doi: 10.1016/j.cmi.2023.10.016] [Medline: 37866679]
11. Chen C, Haupert SR, Zimmermann L, Shi X, Fritsche LG, Mukherjee B. Global prevalence of post-coronavirus disease 2019 (COVID-19) condition or long COVID: a meta-analysis and systematic review. J Infect Dis 2022 Nov 1;226(9):1593-1607. [doi: 10.1093/infdis/jiac136] [Medline: 35429399]
12. Michelen M, Manoharan L, Elkheir N, et al. Characterising long COVID: a living systematic review. BMJ Glob Health 2021 Sep;6(9):e005427. [doi: 10.1136/bmjgh-2021-005427] [Medline: 34580069]
13. Malik P, Patel K, Pinto C, et al. Post-acute COVID-19 syndrome (PCS) and health-related quality of life (HRQoL)—a systematic review and meta-analysis. J Med Virol 2022 Jan;94(1):253-262. [doi: 10.1002/jmv.27309] [Medline: 34463956]
14. Peter RS, Nieters A, Kräusslich HG, et al. Post-acute sequelae of covid-19 six to 12 months after infection: population based study. BMJ 2022 Oct 13;379:e071050. [doi: 10.1136/bmj-2022-071050] [Medline: 36229057]
15. Meijer WM, Veldkamp R, Verburg AC, Bie RAD, Hoogeboom TJ. Paramedische herstelzorg na COVID-19 – inzicht in verloop van de eerstelijns zorg over de tijd [Report in Dutch]. : Nivel; 2023 URL: https://www.nivel.nl/sites/default/files/bestanden/1004377.pdf [accessed 2026-08-25]

XSL•FO
RenderX

16. KNGF-standpunt Fysiotherapie bij COVID-19: Aanbevelingen voor fysiotherapeutisch handelen tijdens de ziekenhuisopname [Report in Dutch]. : Koninklijk Nederlands Genootschap voor Fysiotherapie (KNGF); 2021 URL: https://www.kennisplatformfysiotherapie.nl/app/uploads/sites/2/2024/10/kngf_standpunt_fysiotherapie_covid-19_2_tijdens-ziekenhuisopname_022.pdf [accessed 2026-08-25]
17. KNGF position statement on physical therapy during COVID-19: recommendations for physical therapy activities in the primary care setting. : Royal Dutch Society for Physical Therapy (Koninklijk Nederlands Genootschap voor Fysiotherapie – KNGF); 2022 URL: https://www.kennisplatformfysiotherapie.nl//app/uploads/sites/2/2024/10/kngf-position-statement-on-physical-therapy-during-covid-19.pdf [accessed 2026-08-25]
18. Vergoeding van paramedische herstelzorg na COVID-19: uitleg over de tijdelijke regeling versie 1 juli 2021 [Report in Dutch]. : Zorginstituut Nederland; 2021 URL: https://www.zorginstituutnederland.nl/site/binaries/site-content/collections/documents/2020/09/24/vergoeding-van-paramedische-herstelzorg-na-covid-19/Vergoeding+paramedische+herstelzorg+na+COVID+-+uitleg+over+de+tijdelijke+regeling+versie+1+juli+2021.pdf [accessed 2026-08-25]
19. De Bie RA, Verburg AC, Agasi-Idenburg C, et al. Evaluation of allied healthcare in patients recovering from COVID-19: study protocol and baseline data of s national prospective cohort study. J Rehabil Med 2022 Aug 26;54:jrm00309. [doi: 10.2340/jrm.v54.2506] [Medline: 35735900]
20. Vanhommerig JW, Verheij RA, Hek K, et al. Data Resource Profile: Nivel Primary Care Database (Nivel-PCD), the Netherlands. Int J Epidemiol 2025 Feb 16;54(2):dyaf017. [doi: 10.1093/ije/dyaf017] [Medline: 40083190]
21. Uitvoeringswet algemene verordening gegevensbescherming [Article in Dutch]. Overheid.nl (Wettenbank). URL: https://wetten.overheid.nl/BWBR0040940/2026-09-01/0 [accessed 2026-09-15]
22. PM304 declaratie paramedische hulp [Article in Dutch]. Vektis. URL: https://www.vektis.nl/standaardisatie/standaarden/PM304-3.2 [accessed 2026-08-25]
23. Varianten van het coronavirus SARS-CoV-2 [Article in Dutch]. RIVM (Rijksinstituut voor Volksgezondheid en Milieu). 2023. URL: https://www.rivm.nl/corona/actueel/virusvarianten [accessed 2023-09-28]
24. Codelijst meetinstrumenten [Report in Dutch]. : KNGF (Koninklijk Nederlands Genootschap voor Fysiotherapie); 2024 URL: https://www.kngf.nl/app/uploads/2024/07/codelijst-fy060202-meetinstrumenten.pdf [accessed 2026-08-25]
25. Bie T, de Veldkamp R, Verburg K, Meijer W. Inzet van aanbevolen meetinstrumenten in de eerstelijns fysiotherapeutische zorg bij de behandeling van COVID-19-patiënten [Report in Dutch]. : Nivel; 2023 URL: https://www.nivel.nl/sites/default/files/bestanden/1004435.pdf [accessed 2026-08-25]
26. Csuka M, McCarty DJ. Simple method for measurement of lower extremity muscle strength. Am J Med 1985 Jan;78(1):77-81. [doi: 10.1016/0002-9343(85)90465-6] [Medline: 3966492]
27. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. Am J Respir Crit Care Med 2002 Jul 1;166(1):111-117. [doi: 10.1164/ajrccm.166.1.at1102] [Medline: 12091180]
28. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994 Mar;49(2):M85-M94. [doi: 10.1093/geronj/49.2.m85] [Medline: 8126356]
29. Bohannon RW. Minimal clinically important difference for grip strength: a systematic review. J Phys Ther Sci 2019 Jan;31(1):75-78. [doi: 10.1589/jpts.31.75] [Medline: 30774209]
30. Valko PO, Bassetti CL, Bloch KE, Held U, Baumann CR. Validation of the fatigue severity scale in a Swiss cohort. Sleep 2008 Nov;31(11):1601-1607. [doi: 10.1093/sleep/31.11.1601] [Medline: 19014080]
31. Beurskens AJ, de Vet HC, Köke AJ, et al. A patient-specific approach for measuring functional status in low back pain. J Manipulative Physiol Ther 1999;22(3):144-148. [doi: 10.1016/s0161-4754(99)70127-2] [Medline: 10220712]
32. Hudak PL, Wright JG. The characteristics of patient satisfaction measures. Spine (Phila Pa 1976) 2000 Dec 15;25(24):3167-3177. [doi: 10.1097/00007632-200012150-00012] [Medline: 11124733]
33. Slotegraaf AI, Gerards MHG, Verburg AC, et al. Evaluation of primary allied health care in patients recovering from COVID-19 at 6-month follow-up: Dutch nationwide prospective cohort study. JMIR Public Health Surveill 2023 Oct 20;9:e44155. [doi: 10.2196/44155] [Medline: 37862083]
34. van Ark T. Voorwaardelijke toelating eerstelijns paramedische zorg voor patiënten na ernstige COVID-19 [Report in Dutch]. : Ministerie van Volksgezondheid, Welzijn en Sport (Ministry of Health, Welfare and Sport); 2020 URL: https://open.overheid.nl/documenten/ronl-a66315dc-b10e-4993-8ce0-8a1c6f6dbf6c/pdf [accessed 2026-08-25]
35. Gerards MHG, Verburg AC, Slotegraaf AI, et al. Primary care physical therapy and exercise therapy for people with persistent complaints after a SARS-CoV-2 infection: secondary analysis of the paraCov prospective cohort study. Phys Ther 2025 Jun 2;105(6):pzaf065. [doi: 10.1093/ptj/pzaf065] [Medline: 40344651]
36. Casanova C, Celli BR, Barria P, et al. The 6-min walk distance in healthy subjects: reference standards from seven countries. Eur Respir J 2011 Jan;37(1):150-156. [doi: 10.1183/09031936.00194909] [Medline: 20525717]
37. Bohannon RW, Bubela DJ, Magasi SR, Wang YC, Gershon RC. Sit-to-stand test: performance and determinants across the age-span. Isokinet Exerc Sci 2010;18(4):235-240. [doi: 10.3233/IES-2010-0389] [Medline: 25598584]

XSL•FO
RenderX

38. Bergland A, Strand BH. Norwegian reference values for the Short Physical Performance Battery (SPPB): the Tromsø study. BMC Geriatr 2019 Aug 8;19(1):216. [doi: 10.1186/s12877-019-1234-8] [Medline: 31395008]
39. Dodds RM, Syddall HE, Cooper R, et al. Grip strength across the life course: normative data from twelve British studies. PLoS One 2014;9(12):e113637. [doi: 10.1371/journal.pone.0113637] [Medline: 25474696]
40. Bohannon RW, Crouch R. Minimal clinically important difference for change in 6-minute walk test distance of adults with pathology: a systematic review. J Eval Clin Pract 2017 Apr;23(2):377-381. [doi: 10.1111/jep.12629] [Medline: 27592691]
41. Horn KK, Jennings S, Richardson G, Vliet DV, Hefford C, Abbott JH. The patient-specific functional scale: psychometrics, clinimetrics, and application as a clinical outcome measure. J Orthop Sports Phys Ther 2012 Jan;42(1):30-42. [doi: 10.2519/jospt.2012.3727] [Medline: 22031594]
42. Ponce-Campos SD, Díaz JM, Moreno-Agundis D, et al. A physiotherapy treatment plan for post-COVID-19 patients that improves the FEV1, FVC, and 6-min walk values, and reduces the sequelae in 12 sessions. Front Rehabil Sci 2022;3:907603. [doi: 10.3389/fresc.2022.907603] [Medline: 36188947]
43. Zheng C, Chen XK, Sit CHP, et al. Effect of physical exercise-based rehabilitation on long COVID: a systematic review and meta-analysis. Med Sci Sports Exerc 2024 Jan 1;56(1):143-154. [doi: 10.1249/MSS.0000000000003280] [Medline: 37586104]
44. Chuang HJ, Lin CW, Hsiao MY, Wang TG, Liang HW. Long COVID and rehabilitation. J Formos Med Assoc 2024 Jan;123 Suppl 1:S61-S69. [doi: 10.1016/j.jfma.2023.03.022] [Medline: 37061399]
45. Singh SJ, Daynes E, McAuley HJC, et al. Balancing the value and risk of exercise-based therapy post-COVID-19: a narrative review. Eur Respir Rev 2023 Dec 31;32(170):230110. [doi: 10.1183/16000617.0110-2023] [Medline: 38123233]
46. Antony Leo Asser P, Soundararajan K. The vital role of physiotherapy during COVID-19: a systematic review. Work 2021;70(3):687-694. [doi: 10.3233/WOR-210450] [Medline: 34719461]
47. An Y, Guo Z, Fan J, et al. Prevalence and measurement of post-exertional malaise in post-acute COVID-19 syndrome: a systematic review and meta-analysis. Gen Hosp Psychiatry 2024;91:130-142. [doi: 10.1016/j.genhosppsych.2024.10.011] [Medline: 39490027]
48. Greenwood DC, Mansoubi M, Bakerly ND, et al. Physical, cognitive, and social triggers of symptom fluctuations in people living with long COVID: an intensive longitudinal cohort study. Lancet Reg Health Eur 2024 Nov;46:101082. [doi: 10.1016/j.lanepe.2024.101082] [Medline: 39381546]
49. Clinical management of COVID-19: living guideline, June 2025. : World Health Organization; 2025 URL: https://iris.who.int/server/api/core/bitstreams/d1021eff-f570-4c22-b630-a44bf4267a6c/content [accessed 2026-08-25]

Abbreviations

5TSTS: 5-times sit-to-stand test
6MWT: 6-minute walking test
EHR: electronic health record
FSS: Fatigue Severity Scale
GDPR: General Data Protection Regulation
GPE : Global Perceived Effect
Nivel-PCD: Nivel Primary Care Database
PSFS: Patient-Specific Functional Scale
SPPB: Short Physical Performance Battery

Edited by A Mavragani, T Sanchez; submitted 28.Nov.2025; peer-reviewed by B Kocazeybek, Z Gunendi; revised version received 22.Jun.2026; accepted 12.Jul.2026; published 28.Sep.2026.

Please cite as:
Veldkamp R, Spreeuwenberg P, Verburg AC, Hoogeboom TJ, de Bie RA, Meijer WM
Physical Therapy Provision After COVID-19 in the Netherlands: Retrospective Cohort Study
JMIR Public Health Surveill 2026;12:e88633
URL: https://publichealth.jmir.org/2026/1/e88633
doi:10.2196/88633

© Renee Veldkamp, Peter Spreeuwenberg, Arie Cornelis Verburg, Thomas Johannes Hoogeboom, Rob A de Bie, Willemijn M Meijer. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 28.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic

XSL•FO
RenderX

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

XSL•FO
RenderX

Trends in Antimicrobial Consumption in Pakistan (2016 2028): Retrospective Observational Study With Forecasting

Zikria Saleem[1], PhD; Brian Godman[2], PhD; Zunaira Akbar[3], PhD; Abdul Haseeb[4], PhD; Muhammad Usman Qamar[5], PhD; Anees ur Rehman[6], PhD; Mahmoud E Elrggal[7], Pharm D

[1]Department of Pharmacy Practice, College of Pharmacy, Qassim University, Buraydah, Saudi Arabia

[2]Department of Public Health Pharmacy and Management, School of Pharmacy, Sefako Makgatho Health Sciences University, Ga-Rankuwa, South Africa

[3]Department of Pharmacy Practice, Riphah Institute of Pharmaceutical Sciences, Riphah International University, Lahore, Pakistan

[4]Department of Pharmacy Practice, Faculty of Pharmacy, University of Tabuk, Tabuk, Saudi Arabia

[5]Institute of Microbiology, Faculty of Life Sciences, Government College University, Faisalabad, Pakistan

[6]Department of Pharmacy Practice, Faculty of Pharmacy, Bahauddin Zakariya University, Multan, Pakistan

[7]College of Medicine, Al-Qunfudah, Umm Al-Qura University, Makkah, Saudi Arabia

Corresponding Author:
Zikria Saleem, PhD
Department of Pharmacy Practice, College of Pharmacy, Qassim University, Buraydah, Saudi Arabia

Abstract

Background: Antimicrobial resistance is a public health crisis exacerbated by the irrational use of antibiotics, particularly in low- and middle-income countries. Pakistan, one of the highest consumers of antibiotics globally, faces unique challenges, including unregulated sales, overuse of broad-spectrum antibiotics, and inadequate stewardship programs.

Objective: This study aimed to analyze antibiotic consumption trends in Pakistan from 2016 to 2023, project future use through 2028, and evaluate the subsequent implications for antimicrobial resistance and antimicrobial stewardship programs.

Methods: Antibiotic sales data were retrieved for Pakistan from the IQVIA MIDAS database spanning 2016 to 2023. Data were converted to defined daily doses (DDDs) and DDD per 1000 inhabitants per day (DID) using the World Health Organization Anatomical Therapeutic Chemical classification system. Data cleaning, statistical analyses, and data visualization were performed using R software (version 4.3.2) and Microsoft Excel. Trends were analyzed using linear regression, while future projections (2024 - 2028) were developed using trend-based models. Descriptive analysis was performed, and visualizations were used to illustrate findings.

Results: The total antibiotic consumption in Pakistan from 2016 to 2023 was 12.88 billion DDDs. Broad-spectrum penicillins and fluoroquinolones, each accounting for 37.7 DID, were the most consumed classes. The analysis revealed significant increases in the consumption of macrolides (+76%; rising from 2.26 to 3.99 DID) and cephalosporins (+36%; from 2.87 to 3.89 DID) from 2016 to 2023, with macrolides projected to reach 5.79 DID by 2028. Reserve antibiotics, including oxazolidinones (+354%; from 0.03-to 0.014 DID) and glycylcycline (+236%; from 0.001 to 0.0003 DID), also showed appreciable increases, reflecting greater reliance on last-line therapies. In contrast, aminoglycosides (−36%; from 0.013 to 0.14 DID) and narrow-spectrum penicillins (−30%; from 0.008 to 0.005 DID) experienced notable declines.

Conclusions: The study highlights a concerning overreliance on broad-spectrum and reserve antibiotics in Pakistan, thus underscoring the urgent need for robust antimicrobial stewardship programs and stricter regulation of over-the-counter antibiotic sales to rationalize antibiotic use. Future efforts should focus on addressing gaps in prescribing practices, improving diagnostic capacity, and monitoring stewardship program outcomes to mitigate resistance development and preserve antibiotic efficacy.

(JMIR Public Health Surveill 2026;12:e81288) doi:10.2196/81288

KEYWORDS

trends; projections; antimicrobial consumption; Pakistan; implications; antimicrobial stewardship; resistance

Introduction

Antibiotics have revolutionized modern medicine, playing a critical role in preventing and treating bacterial infections [1]. However, their widespread and often irrational use has led to an appreciable increase in antimicrobial resistance (AMR) in recent years [2]. AMR is one of the greatest public health challenges of the 21st century, threatening the effectiveness of antibiotics and the ability to treat common infections, ultimately increasing morbidity, mortality, and health care costs [2-4]. In

XSL•FO
RenderX

low- and middle-income countries (LMICs), including Pakistan, where health care systems are underresourced and antibiotic use is often unregulated, this issue becomes even more pronounced [5-8].

Pakistan is among the top consumers of antibiotics globally, a distinction that underscores the urgent need for comprehensive surveillance and targeted interventions to address current concerns with their overuse [9]. Key factors driving high antibiotic consumption in the country include easy availability of over-the-counter antibiotics, an insufficient health care infrastructure, lack of awareness among prescribers and the public regarding antibiotics and AMR, and sociocultural practices that promote self-medication [5,6,10]. The unregulated sale of antibiotics and the proliferation of generic brands further exacerbate the problem, leading to overuse and misuse of broad-spectrum antibiotics including those from the World Health Organization (WHO) Watch List [7,11-13]. The COVID-19 pandemic has further complicated the landscape of antibiotic use [14-16]. There was an unprecedented surge in the use of antibiotics during the early stages of the pandemic, driven by the misperception that antibiotics could prevent or treat COVID-19 [16-18]. These challenges, compounded by inadequate national stewardship policies and concerns with limited knowledge and antimicrobial stewardship activities, have contributed to a significant rise in AMR in Pakistan, posing a major threat to public health and economic stability [10,19].

Surveillance of antibiotic consumption is a cornerstone in combating AMR, providing insights into use patterns and identifying areas for intervention [9,20]. The number of defined daily doses (DDDs) per 1000 inhabitants per day (DID) is a standardized metric recommended by the WHO for measuring antibiotic consumption. This metric allows for meaningful comparisons over time and across countries [21]. Understanding trends in antibiotic consumption, particularly in the context of Pakistan, is crucial for informing national policies, optimizing antimicrobial stewardship programs (ASPs), and aligning with the global action plan on AMR [5]. Assessing antibiotic use patterns across sectors has gained importance following recent deliberations of the United Nations General Assembly (UNGA) on AMR, which emphasize increasing the use of antibiotics from the WHO Access List to help reduce AMR [22]. This is important in Pakistan, along with other LMICs, where we see high rates of use of antibiotics from the Watch and Reserve lists with their greater resistance potential [9,12]. Such trends must be reversed if Pakistan is to meet its National Action Plan targets and the UNGA goal of achieving 70% of antibiotic use from Access List antibiotics [22]. Previous studies have highlighted high consumption rates of antibiotics, including cephalosporins, fluoroquinolones, and macrolides, in Pakistan; however, there is currently a lack of comprehensive analyses regarding long-term trends and projections for antibiotic use in Pakistan [9,17]. This gap in knowledge hinders the country's ability to meet its National Action Plan on AMR, which is already a challenge, and contribute meaningfully to global efforts in reducing antibiotic misuse [5]. This study provides a detailed analysis of antibiotic consumption trends in Pakistan from 2016 to 2023, with projections through 2028. By analyzing consumption data and forecasting across 13 major antibiotic classes, this study establishes a counterfactual baseline of expected use and aims to highlight secular trends; identify statistically significant changes; and assess the impact of potential interventions, such as the COVID-19 pandemic, on use patterns. The findings are intended to inform policymakers, health care providers, and researchers about the current state of antibiotic use in Pakistan and the critical need for targeted stewardship efforts.

Methods

Data Source and Collection

Antibiotic consumption data for Pakistan were retrieved from the comprehensive IQVIA MIDAS database, which consolidates pharmaceutical sales data, as reported in a recent study by Klein et al [9]. The data covered the period from 2016 to 2023 and included retail pharmacy sales, capturing estimated use trends for antibiotics across the country. The sales data, recorded in estimated grams of active ingredients, were systematically converted into DDD using the Anatomical Therapeutic Chemical classification system and methodologies recommended by the WHO Collaborating Centre for Drug Statistics Methodology. The DDD conversion accounted for variations in the route of administration (oral and parenteral) for each molecule.

Data Processing

Antibiotic molecules were classified according to the Anatomical Therapeutic Chemical or DDD Index 2026 [23]. For combination drugs, components were broken into their individual active molecules, and consumption was attributed to the main antibiotic molecule based on its 1 daily dose. Antibiotic consumption rates were calculated as DID, enabling cross-year and cross-population comparison. Population data for Pakistan were sourced from the World Bank and national government statistics.

Study Period and Variables

The study analyzed 13 antibiotic classes, including macrolides, cephalosporins, fluoroquinolones, carbapenems, tetracyclines, and others. Variables included annual DIDs for each antibiotic class, percentage changes from 2016 to 2023, and projections (2024 - 2028).

Statistical Analysis

Antibiotic classes were categorized as increasing, decreasing, or stable based on percentage changes in DIDs. Classes with changes >10% were considered increasing, <−10% as decreasing, and between −10% to 10% as stable. This operational threshold was adopted by the authors as a pragmatic balance between clinical relevance and statistical sensitivity. Data cleaning, statistical analyses, and data visualization were performed using R software (version 4.3.2; R Foundation for Statistical Computing; Posit) and Microsoft Excel. For each antibiotic class, linear regression models were applied to evaluate trends from 2016 to 2023. Slope coefficients indicated the rate of annual change. Future antibiotic use (2024 - 2028) was projected using the coefficients from linear regression models. The projections assumed a consistent linear trend in use, with results expressed in DIDs. A +10% or −10% SE

XSL•FO
RenderX

margin was incorporated to reflect potential variability due to external factors such as policy changes and AMR management programs. Line graphs with distinct markers were used to show historical trends and projections. Bar charts were used to compare percentage changes, emphasizing significant trends.

Ethical Considerations

Ethics approval for the study was obtained from the Research Ethics Committee of the Department of Pharmacy Practice, Faculty of Pharmacy, Bahauddin Zakariya University, Multan (BZU-FOPDPP-2456). Participant consent was not required as the study was based on aggregated sales data and did not involve human participants or identifiable personal information. All data were aggregated and anonymized, ensuring compliance with ethical standards for research.

Results

The analysis of antibiotic use in Pakistan between 2016 and 2023, alongside projections for 2024 to 2028, revealed important trends and insights. Table 1 and Table S1 in Multimedia Appendix 1 highlight the annual consumption of antibiotics as DDDs and DID. The total antibiotic consumption in Pakistan during the study period from 2016 to 2023 was measured at an estimated 12.88 billion DDDs. Among the antibiotic classes, broad-spectrum penicillins and fluoroquinolones were the most consumed, contributing significantly to the total DDDs. These findings emphasize the dominant role of these classes in driving overall antibiotic use, highlighting their critical impact on AMR trends. Over the study period, macrolides exhibited the highest growth (+76%), increasing from 2.26 to 3.98 DID, and are projected to reach 5.79 DID by 2028. Similarly, cephalosporins increased by 35.8% from 2.86 to 3.89 DID, with projections suggesting further growth to 4.76 DID. Conversely, tetracyclines showed a slight decline (−3.3%) from 2.19 to 2.12 DID, with projections indicating a continued decrease to 2.00 DID by 2028.

Table . Consumption pattern of antibiotics as defined daily dose (DDD) per 1000 inhabitants per day (DID) in Pakistan.

Antibiotic class	Year-wise DDD per 1000 DID							
	2016	2017	2018	2019	2020	2021	2022	2023
Broad-spectrum penicillins	4.8475	4.5654	4.6502	4.7367	4.5420	4.6185	4.9834	4.7300
Fluoroquinolones	4.5628	4.4010	4.6072	4.5863	4.3664	4.8303	5.2147	5.1958
Cephalosporins	2.8666	2.8807	3.1315	3.2945	3.1519	3.6104	3.9696	3.8936
Macrolides	2.2590	2.2778	2.5770	2.6436	3.3594	3.9336	4.1285	3.9864
Tetracyclines	2.1930	2.2175	2.2884	2.2098	2.1028	2.1524	2.0863	2.1199
Sulfonamides	1.2224	1.0970	1.1353	1.2119	1.1897	1.0901	1.0633	1.0125
Aminoglycosides	0.1324	0.1164	0.1049	0.0849	0.0797	0.0876	0.0992	0.0851
Oxazolidinones	0.0313	0.0428	0.0555	0.0632	0.0707	0.1026	0.1209	0.1420
Narrow-spectrum penicillins	0.0084	0.0069	0.0098	0.0087	0.0069	0.0086	0.0083	0.0059
Carbapenems	0.0071	0.0076	0.0094	0.0101	0.0106	0.0207	0.0210	0.0218
Glycopeptide	0.0015	0.0014	0.0018	0.0021	0.0025	0.0038	0.0034	0.0045
Glycylcycline	0.0001	0.0001	0.0001	0.0002	0.0001	0.0002	0.0002	0.0003
Other antibacterials	0.0460	0.0451	0.0453	0.0470	0.0448	0.0469	0.0441	0.0377

Table 2 presents year-on-year percentage changes. Significant increases were observed for carbapenems (+207%), glycopeptides (+191%), and glycylcycline (+236%), driven largely by higher demand in clinical settings. Oxazolidinones exhibited the most dramatic increase (+354%), reflecting increased reliance on reserve antibiotics. In contrast, aminoglycosides (−36%), narrow-spectrum penicillins (−30%), and sulfonamides (−17%) experienced a notable decline, possibly due to shifts in prescribing preferences. Figure 1 illustrates trends in antibiotic use, with aminoglycosides showing a significant reduction and oxazolidinones exhibiting the largest increase in consumption over the previous 8 years (2016 - 2023).

XSL•FO
RenderX

Figure 1. Cluster changes in consumption between 2016 and 2023.

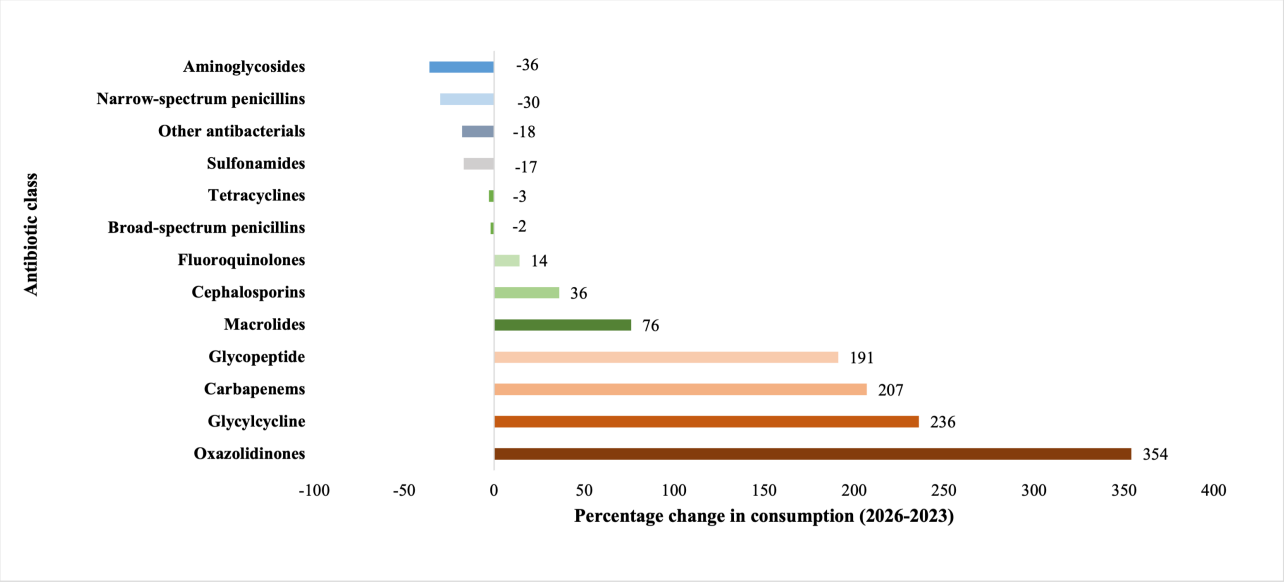


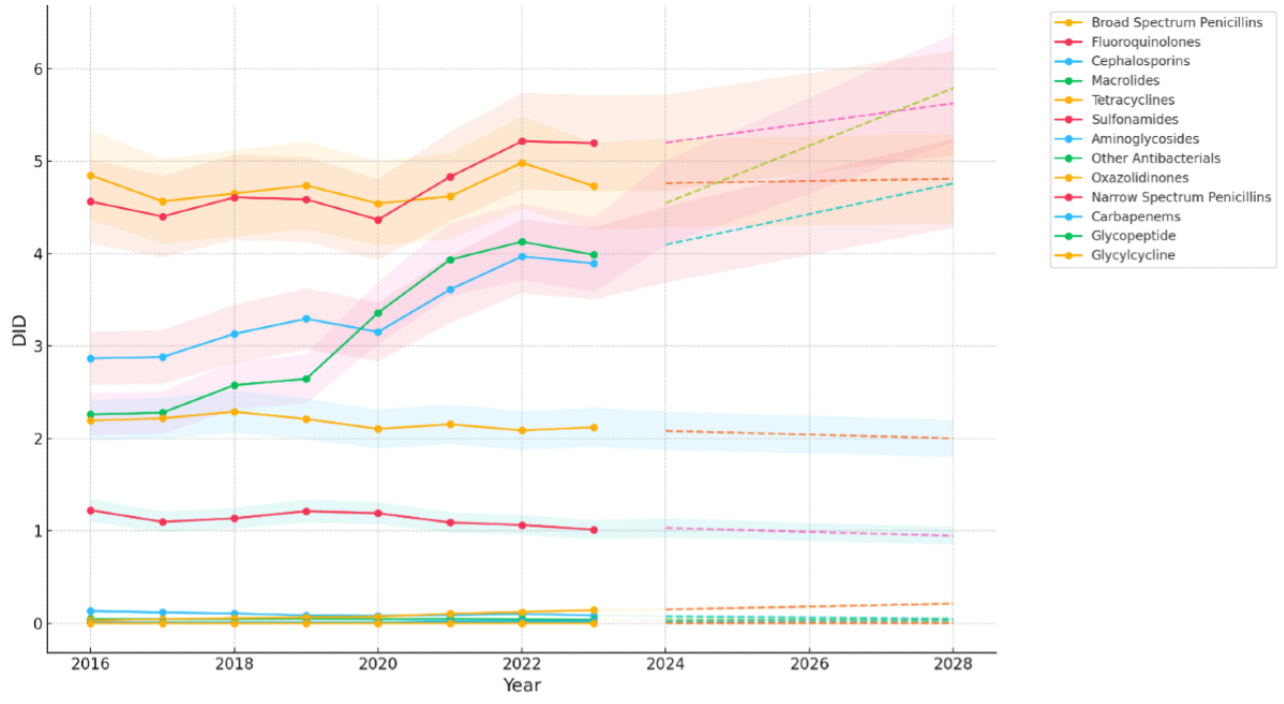
Table . Percentage change in consumption.

Antibiotic class	Year-wise change in consumption (%)							
	(2016 - 2017)	(2017 - 2018)	(2018 - 2019)	(2019 - 2020)	(2020 - 2021)	(2021 - 2022)	(2022 - 2023)	(2016 - 2023)
Broad-spectrum penicillins	−6	2	2	−4	2	8	−5	−2
Fluoroquinolones	−4	5	0	−5	11	8	0	14
Cephalosporins	0	9	5	−4	15	10	−2	36
Macrolides	1	13	3	27	17	5	−3	76
Tetracyclines	1	3	-3	−5	2	−3	2	−3
Sulfonamides	−10	3	7	−2	-8	−2	−5	−17
Aminoglycosides	−12	−10	−19	−6	10	13	−14	−36
Oxazolidinones	37	30	14	12	45	18	18	354
Narrow-spectrum penicillins	−18	41	−11	−21	26	−3	−29	−30
Carbapenems	7	24	7	5	95	1	4	207
Glycopeptide	−12	30	23	17	52	−10	30	191
Glycylcycline	−10	2	96	−9	60	10	17	236
Other antibacterials	−2	0	4	−5	5	−6	−15	−18

Figure 2 demonstrates historical trends and future projections until 2028, with projections indicated by dotted lines. Broad-spectrum antibiotics such as macrolides and cephalosporins show steep upward trajectories (2.25 DID to 5.78 DID and 2.86 DID to 4.75 DID, respectively), signaling the potential overuse of these agents, whereas fluoroquinolones, a critical group, are projected to increase modestly to 5.63 DID by 2028. The inclusion of a +10% or −10% margin of error reflects variability due to external factors such as policy changes or antimicrobial stewardship interventions.

XSL•FO
RenderX

Figure 2. Trends and projections of antibiotic use in Pakistan (2016 - 2028). DID: defined daily dose per 1000 inhabitants per day.



Discussion

Overview

This study analyzed antibiotic consumption trends in Pakistan from 2016 to 2023, with projections through 2028, highlighting critical insights into patterns of use and their implications for AMR and ASP. The findings align with global trends but also underscore the unique challenges faced by LMICs, such as Pakistan, in managing antibiotic consumption [24]. The total antibiotic consumption in Pakistan was measured at 12.88 billion DDDs, reflecting a 17% increase since 2016. This trend parallels global findings, where antibiotic consumption rose by 20.9% during a similar period, although the growth rate was slower compared to the 35.5% increase observed from 2008 to 2015 [25]. However, the COVID-19 pandemic significantly impacted global antibiotic use, leading to temporary reductions, particularly in high-income countries, where outpatient antibiotic use is dominant [9,26-28]. In Pakistan, as in other LMICs, a rebound was observed postpandemic, with consumption exceeding prepandemic levels by 2023, driven by unregulated access and increased reliance on broad-spectrum antibiotics [28,29].

Broad-spectrum penicillins and fluoroquinolones were the most consumed antibiotic classes in Pakistan, accounting for 37.7 DID each in 2023. This heavy reliance on broad-spectrum antibiotics is consistent with trends observed in other LMICs such as India and Bangladesh, where similar patterns have been linked to limited diagnostic capacity and the widespread availability of over-the-counter antibiotics [9,30,31]. In contrast, reserve antibiotics, including carbapenems, oxazolidinones, and glycylcycline, showed dramatic increases in use. These trends highlight the growing demand for last-line antibiotics in critical care settings, reflecting rising resistance to first-line therapies [13,32]. Several factors contribute to the accelerated rise in antibiotic use in developing countries, including environmental, socioeconomic, and cultural determinants, with economic growth appearing to be the most influential driver in lower-income settings [9]. During the COVID-19 pandemic, antibiotic sales also increased considerably due to empirical use in the setting of diagnostic uncertainty and concerns about bacterial coinfections [33]. The projected escalation in the use of macrolides and cephalosporins (2.25 DID to 5.78 DID and 2.86 DID to 4.75 DID, respectively) reflects an increasing dependence on broad-spectrum antibiotics. This growing dependence on Watch group antibiotics deviates from the WHO's recommended target of maintaining at least 60% of total consumption within the Access category, potentially accelerating the selection of multidrug-resistant pathogens [33].

Economic growth, urbanization, and health care infrastructure gaps are significant drivers of antibiotic overuse in Pakistan. As seen in other middle-income countries, rising gross domestic product and rapid urbanization often correlate with increased antibiotic consumption [34]. However, this growth comes at the expense of proper regulatory oversight, leading to widespread misuse. In Pakistan, the lack of robust ASPs and weak enforcement of prescription-only antibiotic policies exacerbate the problem [6,12,35]. The WHO's Access, Watch, and Reserve (AWaRe) classification system has been pivotal in promoting rational antibiotic use globally [36]. However, in Pakistan, as in many LMICs, the Watch and Reserve categories dominate consumption, reflecting poor alignment with the AWaRe recommendations [25]. In contrast, high-income countries have successfully reduced the use of high-risk antibiotics such as fluoroquinolones through stringent regulatory warnings and public awareness campaigns [9,37]. The rising consumption of fluoroquinolones and reserve antibiotics in Pakistan highlights the urgent need for similar regulatory interventions [17]. This study underscores the critical need for comprehensive ASPs tailored to the local context [38].

XSL•FO
RenderX

Strengthening regulatory frameworks, enhancing diagnostic capacity, and promoting public awareness are essential to curbing antibiotic misuse [39,40]. Investments in preventive measures, such as improved water, sanitation, and hygiene infrastructure, along with robust vaccination programs, could also significantly reduce the burden of infectious diseases and, consequently, the need for antibiotics in Pakistan [41,42].

Principal Findings

The major findings of the study are as follows:

- Total antibiotic consumption in Pakistan (2016 - 2023) reached 12.88 billion DDDs, with broad-spectrum penicillins and fluoroquinolones accounting for nearly half.
- Annual consumption of macrolides (+76%) and cephalosporins (+36%) increased sharply, with macrolides projected to reach 5.79 DID by 2028.
- Reserve antibiotics, including oxazolidinones and glycylcycline, showed marked growth, reflecting increased use of last-line therapies, while aminoglycosides and narrow-spectrum penicillins declined.

Study Limitations and Future Research

While the data in this report provide valuable insights, they have inherent limitations. First, the IQVIA MIDAS data primarily capture retail pharmacy sales and do not include hospital procurement or inpatient use, which may lead to underrepresentation of antibiotics predominantly used in hospital settings (eg, carbapenems and glycopeptides) or for severe infections. No distinction was made between prescription-only and over-the-counter antibiotics, as consumption estimates were derived from aggregated sales data. Second, the dataset lacks indication-level, health care setting–level, and prescriber-level information, preventing analysis of the clinical context or appropriateness of antibiotic use. Third, the projections for 2024 to 2028 assume linear trends, which may not account for unforeseen policy changes, health care developments, or epidemics. The study does not address regional variations within Pakistan, which could provide additional insights into localized drivers of antibiotic use. Additionally, the study focused exclusively on human antibiotic consumption, excluding veterinary and agricultural use, which are critical components of a One Health approach to AMR. Future research should integrate these dimensions to provide a more comprehensive understanding of antibiotic use trends.

Conclusions

The findings of this study highlight the urgent need for targeted interventions to optimize antibiotic use in Pakistan. The high reliance on broad-spectrum and reserve antibiotics, coupled with the rapid rebound in consumption after the pandemic, underscores the importance of strengthening ASPs and regulatory oversight. By aligning national policies with global frameworks following UNGA recommendations, Pakistan can mitigate the risks of AMR and ensure sustainable antibiotic use.

Acknowledgments

The authors would like to extend their gratitude to Eili Y Klein for their invaluable support and for allowing the use of data from their insightful study on antibiotic consumption, which significantly informed this work. The authors also extend their appreciation to Umm Al-Qura University, Saudi Arabia, for funding this research work through Grant Number 26UQU4320605GSSR01. QuillBot was used to improve the language and readability of the manuscript.

Funding

This work was funded by Umm Al-Qura University, Saudi Arabia (grant 26UQU4320605GSSR01).

Data Availability

The data presented in this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: ZS.
Data curation: ZS.
Formal analysis: ZS and MEE.
Funding acquisition: MEE.
Methodology: ZS, AH, and MUQ.
Software: ZS.
Validation: ZS, MUQ, and BG.
Investigation: ZS, AH, and AUR.
Resources: MUQ and BG.
Visualization: AH.
Project administration: ZS.
Writing—original draft preparation: ZS, AH, ZA, and MEE.
Writing—review and editing: MUQ and BG.
All authors have read and agreed to the published version of the manuscript.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Year-wise consumption of antibiotics in terms of defined daily dose and projections in terms of defined daily dose per 1000 inhabitants per day.
[DOCX File, 155 KB - publichealth_v12i1e81288_app1.docx]

References

1. Zaffiri L, Gardner J, Toledo-Pereyra LH. History of antibiotics. From salvarsan to cephalosporins. J Invest Surg 2012 Apr;25(2):67-77. [doi: 10.3109/08941939.2012.664099] [Medline: 22439833]
2. Hou J, Long X, Wang X, et al. Global trend of antimicrobial resistance in common bacterial pathogens in response to antibiotic consumption. J Hazard Mater 2023 Jan 15;442:130042. [doi: 10.1016/j.jhazmat.2022.130042] [Medline: 36182890]
3. GBD 2021 Antimicrobial Resistance Collaborators. Global burden of bacterial antimicrobial resistance 1990-2021: a systematic analysis with forecasts to 2050. Lancet 2024 Sep 28;404(10459):1199-1226. [doi: 10.1016/S0140-6736(24)01867-1] [Medline: 39299261]
4. Kumarasamy KK, Toleman MA, Walsh TR, et al. Emergence of a new antibiotic resistance mechanism in India, Pakistan, and the UK: a molecular, biological, and epidemiological study. Lancet Infect Dis 2010 Sep;10(9):597-602. [doi: 10.1016/S1473-3099(10)70143-2] [Medline: 20705517]
5. Saleem Z, Godman B, Azhar F, et al. Progress on the national action plan of Pakistan on antimicrobial resistance (AMR): a narrative review and the implications. Expert Rev Anti Infect Ther 2022 Jan;20(1):71-93. [doi: 10.1080/14787210.2021.1935238] [Medline: 34038294]
6. Saleem Z, Sono TM, Godman B. Concerns with current drug laws regarding the purchasing antibiotics without a prescription in Pakistan; ways forward to assist the national action plan. Expert Rev Anti Infect Ther 2023;21(11):1163-1165. [doi: 10.1080/14787210.2023.2260096] [Medline: 37712181]
7. Saleem Z, Saeed H, Akbar Z, et al. WHO key access antibiotics price, availability and affordability in private sector pharmacies in Pakistan. Cost Eff Resour Alloc 2021 Feb 16;19(1):10. [doi: 10.1186/s12962-021-00263-x] [Medline: 33593366]
8. Arif S, Sadeeqa S, Saleem Z. Patterns of antimicrobial use in hospitalized children: a repeated point prevalence survey from Pakistan. J Pediatric Infect Dis Soc 2021 Nov 11;10(10):970-974. [doi: 10.1093/jpids/piab026] [Medline: 34293148]
9. Klein EY, Impalli I, Poleon S, et al. Global trends in antibiotic consumption during 2016-2023 and future projections through 2030. Proc Natl Acad Sci U S A 2024 Dec 3;121(49):e2411919121. [doi: 10.1073/pnas.2411919121] [Medline: 39556760]
10. Atif M, Ihsan B, Malik I, et al. Antibiotic stewardship program in Pakistan: a multicenter qualitative study exploring medical doctors' knowledge, perception and practices. BMC Infect Dis 2021 Apr 21;21(1):374. [doi: 10.1186/s12879-021-06043-5] [Medline: 33882843]
11. Saleem Z, Faller EM, Godman B, et al. Antibiotic consumption at community pharmacies: a multicenter repeated prevalence surveillance using WHO methodology. Med Access Point Care 2021;5:23992026211064714. [doi: 10.1177/23992026211064714] [Medline: 36204499]
12. Saleem Z, Hassali MA, Godman B, et al. Sale of WHO AWaRe groups antibiotics without a prescription in Pakistan: a simulated client study. J Pharm Policy Pract 2020;13:26. [doi: 10.1186/s40545-020-00233-3] [Medline: 32774870]
13. Pauwels I, Versporten A, Drapier N, Vlieghe E, Goossens H, Global-PPS network. Hospital antibiotic prescribing patterns in adult patients according to the WHO Access, Watch and Reserve classification (AWaRe): results from a worldwide point prevalence survey in 69 countries. J Antimicrob Chemother 2021 May 12;76(6):1614-1624. [doi: 10.1093/jac/dkab050] [Medline: 33822971]
14. Sefah IA, Essah DO, Haque M, et al. COVID-19, healthcare and self-medication issues in resource limited-settings: findings and implications based on experiences in Ghana. Adv Hum Biol 2021;11(3):224-233. [doi: 10.4103/aihb.aihb_82_21]
15. Bednarčuk N, Golić Jelić A, Stoisavljević Šatara S, et al. Antibiotic utilization during COVID-19: are we over-prescribing? Antibiotics (Basel) 2023 Feb 2;12(2):308. [doi: 10.3390/antibiotics12020308] [Medline: 36830218]
16. Hackman HK, Annison L, Arhin RE, et al. Self-medication with antibiotics during the COVID-19 pandemic: a cross-sectional study among adults in Tema, Ghana. PLoS One 2024;19(6):e0305602. [doi: 10.1371/journal.pone.0305602] [Medline: 38917123]
17. Mustafa T, Niazi MR, Lakdawala Z, Mirza S. Regional and national trends in consumption of antimicrobials in Pakistan; pre and post-COVID (2019-2021). Clin Infect Dis 2023 Dec 20;77(Suppl 7):S569-S577. [doi: 10.1093/cid/ciad647] [Medline: 38118009]
18. Chowdhury K, Haque M, Nusrat N, et al. Management of children admitted to hospitals across Bangladesh with suspected or confirmed COVID-19 and the implications for the future: a nationwide cross-sectional study. Antibiotics (Basel) 2022 Jan 14;11(1):105. [doi: 10.3390/antibiotics11010105] [Medline: 35052982]

XSL•FO
RenderX

19. Saleem Z, Hassali MA, Hashmi FK, Godman B, Ahmed Z. Snapshot of antimicrobial stewardship programs in the hospitals of Pakistan: findings and implications. Heliyon 2019 Jul;5(7):e02159. [doi: 10.1016/j.heliyon.2019.e02159] [Medline: 31384689]
20. Zhang D, Cui K, Wang T, et al. Trends in and correlations between antibiotic consumption and resistance of Staphylococcus aureus at a tertiary hospital in China before and after introduction of an antimicrobial stewardship programme. Epidemiol Infect 2018 Nov 16;147:e48. [doi: 10.1017/S0950268818003059] [Medline: 30442220]
21. Guidelines for ATC classification and DDD assignment. WHO Collaborating Centre for Drug Statistics Methodology. URL: https://atcddd.fhi.no/atc_ddd_index_and_guidelines/guidelines/ [accessed 2026-04-14]
22. Mendelson M, Lewnard JA, Sharland M, et al. Ensuring progress on sustainable access to effective antibiotics at the 2024 UN General Assembly: a target-based approach. Lancet 2024 Jun 8;403(10443):2551-2564. [doi: 10.1016/S0140-6736(24)01019-5] [Medline: 38797179]
23. ATC index with DDDs. WHO Collaborating Centre for Drug Statistics Methodology. URL: https://atcddd.fhi.no/atc_ddd_index_and_guidelines/atc_ddd_index/ [accessed 2026-04-14]
24. Godman B, Haque M, McKimm J, et al. Ongoing strategies to improve the management of upper respiratory tract infections and reduce inappropriate antibiotic use particularly among lower and middle-income countries: findings and implications for the future. Curr Med Res Opin 2020 Feb;36(2):301-327. [doi: 10.1080/03007995.2019.1700947] [Medline: 31794332]
25. Klein EY, Van Boeckel TP, Martinez EM, et al. Global increase and geographic convergence in antibiotic consumption between 2000 and 2015. Proc Natl Acad Sci U S A 2018 Apr 10;115(15):E3463-E3470. [doi: 10.1073/pnas.1717295115] [Medline: 29581252]
26. Beović B, Doušak M, Ferreira-Coimbra J, et al. Antibiotic use in patients with COVID-19: a “snapshot” Infectious Diseases International Research Initiative (ID-IRI) survey. J Antimicrob Chemother 2020 Nov 1;75(11):3386-3390. [doi: 10.1093/jac/dkaa326] [Medline: 32766706]
27. Zhu B, Zhao X, Li Y, et al. Impact of the COVID-19 outbreak on the antibiotic use patterns among a rural community population in Eastern China. Antibiotics (Basel) 2022 Nov 3;11(11):1544. [doi: 10.3390/antibiotics11111544] [Medline: 36358199]
28. Nandi A, Pecetta S, Bloom DE. Global antibiotic use during the COVID-19 pandemic: analysis of pharmaceutical sales data from 71 countries, 2020-2022. EClinicalMedicine 2023 Mar;57:101848. [doi: 10.1016/j.eclinm.2023.101848] [Medline: 36776504]
29. Alshaikh FS, Godman B, Sindi ON, Seaton RA, Kurdi A. Prevalence of bacterial coinfection and patterns of antibiotics prescribing in patients with COVID-19: a systematic review and meta-analysis. PLoS One 2022;17(8):e0272375. [doi: 10.1371/journal.pone.0272375] [Medline: 35913964]
30. Saleem Z, Hassali MA, Godman B, et al. Point prevalence surveys of antimicrobial use: a systematic review and the implications. Expert Rev Anti Infect Ther 2020 Sep;18(9):897-910. [doi: 10.1080/14787210.2020.1767593] [Medline: 32394754]
31. Singh SK, Sengupta S, Antony R, et al. Variations in antibiotic use across India: multi-centre study through global point prevalence survey. J Hosp Infect 2019 Nov;103(3):280-283. [doi: 10.1016/j.jhin.2019.05.014] [Medline: 31170422]
32. Sulis G, Sayood S, Katukoori S, et al. Exposure to World Health Organization’s AWaRe antibiotics and isolation of multidrug resistant bacteria: a systematic review and meta-analysis. Clin Microbiol Infect 2022 Sep;28(9):1193-1202. [doi: 10.1016/j.cmi.2022.03.014] [Medline: 35339675]
33. Gul B, Sana M, Saleem A, et al. Antimicrobial dispensing practices during COVID-19 and the implications for Pakistan. Antibiotics (Basel) 2023 Jun 6;12(6):1018. [doi: 10.3390/antibiotics12061018] [Medline: 37370337]
34. Otaigbe II, Elikwu CJ. Drivers of inappropriate antibiotic use in low- and middle-income countries. JAC Antimicrob Resist 2023;5(3):dlad062. [doi: 10.1093/jacamr/dlad062] [Medline: 37265987]
35. Torumkuney D, Jamil B, Nizamuddin S, van Hasselt J, Pirzada U, Manenzhe R. Country data on AMR in Pakistan in the context of community-acquired respiratory tract infections: links between antibiotic susceptibility, local and international antibiotic prescribing guidelines, access to medicine and clinical outcome. J Antimicrob Chemother 2022 Sep 6;77(Suppl_1):i18-i25. [doi: 10.1093/jac/dkac213] [Medline: 36065729]
36. Sharland M, Cappello B, Ombajo LA, et al. The WHO AWaRe Antibiotic Book: providing guidance on optimal use and informing policy. Lancet Infect Dis 2022 Nov;22(11):1528-1530. [doi: 10.1016/S1473-3099(22)00683-1] [Medline: 36309019]
37. Charani E, Mendelson M, Pallett SJ, et al. An analysis of existing national action plans for antimicrobial resistance-gaps and opportunities in strategies optimising antibiotic use in human populations. Lancet Glob Health 2023 Mar;11(3):e466-e474. [doi: 10.1016/S2214-109X(23)00019-0] [Medline: 36739875]
38. Li DX, Cosgrove SE. Antimicrobial stewardship: efficacy and implementation of strategies to address antimicrobial overuse and resistance. In: Pulcini C, Ergönül Ö, Can F, Beović B, editors. Antimicrobial Stewardship: Academic Press; 2017:13-28. [doi: 10.1016/C2015-0-04271-X]
39. Thamlikitkul V, Rattanaumpawan P, Sirijatuphat R, Wangchinda W. Integrated one-day surveillance of antimicrobial use, antimicrobial consumption, antimicrobial resistance, healthcare-associated infection, and antimicrobial resistance burden

XSL•FO
RenderX

among hospitalized patients in Thailand. J Infect 2020 Jul;81(1):98-106. [doi: 10.1016/j.jinf.2020.04.040] [Medline: 32360884]
40. Agarwal J, Singh V, Das A, Nath SS, Kumar R, Sen M. Reversing the trend of antimicrobial resistance in ICU: role of antimicrobial and diagnostic stewardship. Indian J Crit Care Med 2021 Jun;25(6):635-641. [doi: 10.5005/jp-journals-10071-23861] [Medline: 34316142]
41. Lewnard JA, Lo NC, Arinaminpathy N, Frost I, Laxminarayan R. Childhood vaccines and antibiotic use in low- and middle-income countries. Nature 2020 May;581(7806):94-99. [doi: 10.1038/s41586-020-2238-4] [Medline: 32376956]
42. GBD 2017 SDG Collaborators. Measuring progress from 1990 to 2017 and projecting attainment to 2030 of the health-related Sustainable Development Goals for 195 countries and territories: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):2091-2138. [doi: 10.1016/S0140-6736(18)32281-5] [Medline: 30496107]

Abbreviations

AMR: antimicrobial resistance
ASP: antimicrobial stewardship program
AWaRe: Access, Watch, Reserve
DDD: defined daily dose
DID: defined daily dose per 1000 inhabitants per day
LMIC: low- and middle-income country
UNGA: United Nations General Assembly
WHO: World Health Organization

Edited by C Argyropoulos; submitted 25.Jul.2025; peer-reviewed by D Iryawati, G Mboowa, R Koizumi; revised version received 19.Feb.2026; accepted 21.Feb.2026; published 28.Apr.2026.

Please cite as:
Saleem Z, Godman B, Akbar Z, Haseeb A, Qamar MU, Rehman AU, Elrggal ME
Trends in Antimicrobial Consumption in Pakistan (2016 - 2028): Retrospective Observational Study With Forecasting
JMIR Public Health Surveill 2026;12:e81288
URL: https://publichealth.jmir.org/2026/1/e81288
doi:10.2196/81288

© Zikria Saleem, Brian Godman, Zunaira Akbar, Abdul Haseeb, Muhammad Usman Qamar, Anees ur Rehman, Mahmoud E Elrggal. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 28.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Implementation Effect of a Pilot Urban Primary Health Care Model on the Usage of Health Care From Medically Trained Providers for Chronic Diseases Among Slum Populations in Bangladesh: Findings From an Implementation Research Study

Md. Golam Rabbani[1], MSS; Zakir Hossain[1], MSc; Khadija Islam Tisha[1], MPH; Kamrun Nahar[1], MSS; Zillur Rahman Sakin[1], MSS; Towhida Nasrin[1], MSS; Mohammad Wahid Ahmed[1], MSS; Md. Zahid Hasan[1], PhD; Chandrasegarar Soloman[2], MPH; Margub Aref Jahangir[2], MPH; Maya Vandenent[2], MPH; Shehrin Shaila Mahmood[1], PhD

[1]Health Economics and Financing, Health Systems and Population Studies Division, icddr,b (International Centre for Diarrhoeal Disease Research, Bangladesh), Dhaka, Bangladesh

[2]UNICEF (United Nations International Children's Emergency Fund) Bangladesh, Dhaka, Bangladesh

Corresponding Author:
Shehrin Shaila Mahmood, PhD
Health Economics and Financing, Health Systems and Population Studies Division, icddr,b (International Centre for Diarrhoeal Disease Research, Bangladesh)
68, Shaheed Tajuddin Ahmed Sarani, Mohakhali
Dhaka, 1212
Bangladesh
Phone: 880 01787670999
Email: shaila@icddrb.org

Abstract

Background: Ensuring affordable and quality primary health care for urban populations remains a persistent challenge in resource-constrained countries such as Bangladesh. The country's fragmented and pluralistic urban health system exacerbates these challenges, making it particularly difficult to deliver primary health care to urban residents. In Bangladesh, the Aalo Clinic has been piloted to provide essential health services to low-income urban residents, including slum dwellers.

Objective: This study aimed to assess how this model affects health care–seeking behavior from medically trained providers (MTPs) for managing chronic diseases.

Methods: We conducted implementation research from October 2021 to August 2023 using a cross-sectional study design. To assess the effectiveness of the Aalo Clinic model, the baseline survey was conducted from January to February 2022, and the end-line surveys were carried out from April to June 2023 in the Korail, Mirpur, Shyampur, Dhalpur, and Tongi-Ershadnagar slums, which are adjacent to the clinics. In the baseline survey, we randomly selected 2000 households (with 9196 members), while 2033 households (with 8223 members) were interviewed in the end-line survey. We considered 8919 individuals in the baseline and 7994 individuals in the end-line, all aged 1 year or older. We used descriptive analysis and chi-squared tests to assess changes in health care usage from MTPs and applied logistic regression models to assess the model's influence on health care usage from MTPs while controlling for other factors.

Results: Among slum dwellers, awareness of Aalo Clinic increased by approximately 62 percentage points, from 11.46% (n=227) at baseline to 73.36% (n=1473) at end-line. The usage of health care for the management of chronic diseases from Aalo Clinics' MTPs was significantly higher at the end-line period (n=63, 11.13%) compared to the baseline period (n=6, 0.59%). Regression analysis indicated that slum dwellers were approximately 23 times more likely (95% CI 8.77-61.80) to use health care from Aalo Clinic for chronic diseases in the end-line period compared to the baseline. However, there was no substantial change in health care usage from non-MTPs, which remained stable at around 38% (baseline: n=388; end-line: n=218) in both periods, and local drug stores continue to be the primary source for managing chronic diseases.

Conclusions: Usage of Aalo Clinic MTPs for chronic disease care increased substantially compared to other MTPs among slum dwellers. However, reliance on non-MTPs, particularly local drug stores, remained stable. Ensuring the availability of essential medicines may further reduce dependence on non-MTPs among low-income urban slum residents. Future longitudinal studies are needed to evaluate the long-term impact on chronic disease management in Bangladesh.

XSL•FO
RenderX

(JMIR Public Health Surveill 2026;12:e76290) doi:10.2196/76290

KEYWORDS

urban primary health care; Aalo Clinic model; chronic disease; health care usage; medically trained provider; slum population; Bangladesh

Introduction

In low- and middle-income countries (LMICs) such as Bangladesh, ensuring affordable and quality health care remains a persistent challenge, particularly in urban areas where the provision of basic health care is limited [1]. This challenge contributes to the global setback in making significant progress toward achieving universal health coverage (UHC) by 2030: a key objective of the sustainable development goals [2]. To address this issue of providing affordable and quality health care, the World Health Organization recommends strengthening health systems with a primary health care (PHC) approach, as it can significantly contribute to delivering essential health care services [3].

Bangladesh has achieved significant success in health, leading to reduced maternal and child mortality rates [4]. However, the pressing challenge now is to improve the health of impoverished urban populations. Rapid urbanization is underway, with large numbers of rural residents migrating to cities, driven by poverty, climate change, and the hope for better economic opportunities. Currently, approximately 40% of Bangladesh's population resides in urban areas, with an estimated annual growth rate of 3% [5]. A significant portion of these urban residents live in informal settlements, such as slums, which have a concerning growth rate of 7% [6]. Cities with inadequate infrastructure are struggling to accommodate this continuous influx of people, resulting in impoverished slum dwellers facing challenges in accessing health care, among other difficulties. This situation highlights the weakness in the government's health system implementation and the basic health care for slum dwellers [7,8].

The Ministry of Health and Family Welfare has the mandate for health across the country, but urban PHC is mostly delivered by the Ministry of Local Government and Rural Development Cooperation or the Local Governance Institutions, due to a decree providing them the authority to deliver PHC in urban areas [7,9]. However, with its limited capacity, it is currently unable to provide adequate PHC. As a result, access to quality PHC for urban populations, particularly vulnerable individuals such as slum dwellers, is limited. Evidence showed that fragmented health care delivery in urban areas has been caused by unclear roles, poor planning, and a lack of coordination between the 2 authorities, among other factors [8]. Consequently, untrained informal providers, for example, drug sellers and low-quality private clinics, are relied upon by underserved populations [7,8,10-12], which worsens the health of slum dwellers and leads to catastrophic health care expenditure [13,14].

Evidence shows that nongovernmental organizations (NGOs) and development partners have tackled problems in slums with innovative strategies [7,9,15-18]. To overcome the existing barriers to accessing health care, particularly essential high-quality PHC services for urban dwellers, an urban PHC model, namely the Aalo Clinic, is being piloted in selected urban low-income areas since 2021 [19]. We aimed to assess the implementational effect of the Aalo Clinic model in prompting health care usage from medically trained providers (MTPs) for chronic diseases and to identify factors that influence the health care–seeking behavior of the patients.

Methods

Description of Aalo Clinic model

With a view to providing accessible and affordable care to the urban poor, this model has been developed based on setting up a static clinic, namely the Aalo Clinic, and outreach services through "Dorgoray Shasthya Sheba"[20]. The key features of this model are shown in Textbox 1.

XSL•FO
RenderX

Textbox 1. Key features of the Aalo Clinic model.

- Offers free medical treatment to city residents.
- Patients receive treatment for common, seasonal, and noncommunicable diseases, as well as Expanded Programme on Immunization vaccination and family planning services.
- Provides free medicines: a total of 23 different types of medications are currently being administered.
- Free 10 examinations are performed, including complete blood count, creatinine, erythrocyte count, dengue, and diabetes.
- Every step is digitized, from patient registration to prescription writing to test results.
- Once a patient has registered, clinicians can access information about their prior medical history for any future visits or follow-up care.
- The standard time for doctors to see a patient is 8-10 minutes.
- Operates 6 days a week, except Fridays and public holidays, in 2 shifts: morning shift from 8:30 AM to 2:30 PM and evening shift from 3 PM to 9 PM.
- Well-equipped (infrastructure, health workforce, or equipment) to provide primary health care services.
- Provides outreach health care services through mobile vans in the static Aalo Clinic areas, where patients can receive counseling services from paramedics, and if required, patients are referred to Aalo Clinics.
- Arranges awareness-raising activities for the respective community.
- Provides appropriate referral if required for more specialized care.
- Provides 25%-40% discounted diagnostic tests for patients referred from Aalo Clinics.
- Routinely monitors clinic performance for quality assurance.
- Records beneficiaries' feedback to cater service according to their need.

In the Aalo Clinic model, services have been provided according to the essential service package. The currently available outpatient services at the Aalo Clinics and outreach points are listed in Textbox 2.

XSL•FO
RenderX

Textbox 2. Health care services offered by the Aalo Clinic model for the urban slum populations in Bangladesh.

Maternal, neonatal, child, and adolescent health care preconception, antenatal, delivery, and postnatal

- Maternal and newborn care (antenatal care, postnatal care, and delivery referral, newborn care, etc)
- Child health and immunization (Integrated Management of Childhood Illness, acute respiratory infection, diarrhea, growth monitoring and promotion, Expanded Programme on Immunization, etc)
- Adolescent health (sexual and reproductive health, reproductive tract infection or sexually transmitted infection, dysmenorrhea, tetanus toxoid vaccine, nutrition, and counseling on mental health)
- Nutrition
- Child nutrition: assessment of nutrition status, prevention of malnutrition, and management of malnutrition
- Promote exclusive breastfeeding, complementary feeding, water, sanitation, and hygiene for children, and growth monitoring and promotion
- Support to moderate acute malnutrition children and refer severe acute malnutrition children
- Maternal nutrition (BMI screening of malnourished pregnant women, mid-upper arm circumference, and counseling for food)

Noncommunicable diseases

- Noncommunicable disease (hypertension, diabetes mellitus, chronic obstructive pulmonary disease, etc) consultation and referral
- Mental health consultation
- Sexual and gender-based violence

Management of other common conditions

- Eye care and referral
- Skin care and referral
- Geriatric (older adult) care and referral
- Minor infection and disease
- Health education and counseling

Family planning (method selection counseling, provide oral contraceptive pill, condom, injectables, and intrauterine device)

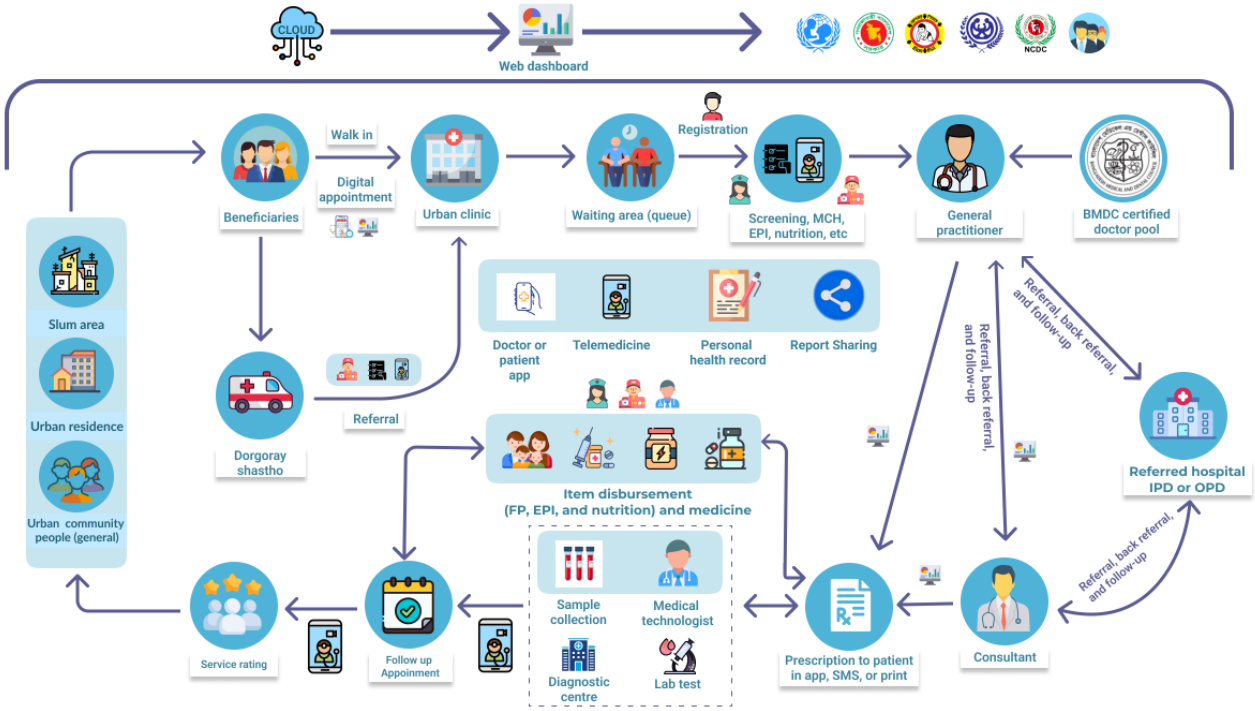
Communicable diseases (consultation on communicable diseases, referral, and consultation on sexually transmitted infection or sexually transmitted disease)

Laboratory urinalysis, Hb estimation, blood grouping and Rh typing, blood sugar, complete blood count, and electrocardiogram

The health care services of Aalo Clinic are delivered in a structured way. During the first visit to the Aalo Clinic, a patient got registered on an IT-based system using their mobile phone number. Following that, the patient receives an automated serial number to get on the queue for general practitioners (GPs). After the screening by the GPs, the patient obtains free diagnostic tests and medicines depending on the prescriptions. The doctor can refer to a higher level of health care facilities if required. Figure 1 shows the overall service flow of the Aalo Clinics.

XSL•FO
RenderX

Figure 1. Service flow of the Aalo Clinic model and its information technology-based management system for health care delivery to urban slum populations in Bangladesh (adapted and reproduced with permission from UNICEF Bangladesh). UNICEF: United Nations International Children's Emergency Fund.



Study Design

We conducted implementation research with the Aalo Clinic model between October 2021 and August 2023. With this study as a key component, we carried out repeated cross-sectional household surveys with selected slum dwellers, a primary focus group of this model, to assess the effects of the model on the usage of PHC for managing chronic diseases. The baseline survey was conducted from January to February 2022, shortly after the model's service delivery began, while the end-line survey took place from April to June 2023, one year after its implementation. While both surveys targeted the same population group, the same households were not necessarily sampled at both time points.

Study Setting and Population

The Aalo Clinic model has been implemented in 6 selected areas with Dhaka North City Corporation, Dhaka South City Corporation, Gazipur City Corporation, and Narayanganj City Corporation. In Dhaka North City Corporation, 2 Aalo Clinics were established and located adjacent to the Korail and Mirpur slums. In Dhaka South City Corporation, 2 Aalo Clinics were also established, which are located adjacent to the Shyampur and Dhalpur slums. Whereas 1 clinic was established in Gazipur City Corporation within the Tongi-Ershadnagar slum and another was established in Narayanganj City Corporation. Urban Health and Demographic Surveillance System (UHDSS) of icddr,b (International Centre for Diarrhoeal Disease Research, Bangladesh), was used as the sampling frame in this study, which includes Korail, Mirpur, Shyampur, Dhalpur, and Tongi-Ershadnagar slums. The UHDSS covers approximately 31,577 households. Details about the UHDSS and its coverage are available elsewhere [10]. However, the inclusion criteria were households covered under UHDSS in the selected slums, household members who had been living with the selected household for at least 6 months before the survey, and households that voluntarily expressed interest in participating in the current study. The exclusion criteria were households not covered by the UHDSS and those not interested in participating in this study.

Sample Size and Sampling

The household survey served the purpose of assessing the usage of the services of Aalo Clinic to assess the effectiveness of the model in increasing usage of PHC. Concerning this, we aimed to assess the awareness level of the Aalo Clinic model among slum households and expected that approximately 30% of households would be aware of this model and use health care after its implementation. Mahmood et al [12] conducted a study in the same setting and found that 32.2% of sick household members in urban slums used health care from formal health care providers. Considering this proportion, a 5% error level for a 95% CI, and a design effect of 1.2, applied conservatively to account for potential intracluster correlation within slum communities, we estimated a required sample size of 387 for each slum. Furthermore, we considered a nonresponse rate of 10% and projected that the required sample size of 430 households needs to be approached to participate in each study area (total 430×5=2150 households). Finally, we interviewed 2000 households in the baseline survey and 2033 households in the end-line survey. It is important to mention that the sample distribution did not align with the population size of each slum. An equal number of households were sampled from each slum to ensure reliable results. This uniform approach facilitated homogeneity in data collection and supported effective comparative analysis. However, we considered 8919 individuals

XSL•FO
RenderX

in the baseline and 7994 individuals in the end-line satisfying the inclusion criteria for chronic disease, with a minimum age of household members 1 year preceding the survey.

A systematic random sampling method was applied to select the households. For each slum, we established an interval between the selected households based on household size and approached them accordingly. For example, approximately 14,000 households reside in the Korail slum and 2300 in the Shyampur slum. For Korail, we used an interval of 32 (14,000/430), and for Shyampur, an interval of 5 (2300/430). If a selected household had multiple eligible members, we randomly chose one to participate in this study using the lottery method. In cases where a household had no eligible member, we proceeded to the next household until we found an eligible participant.

Data Collection Tool and Variables

We used a pretested and identical semistructured questionnaire to conduct the baseline and end-line surveys. Data were collected through face-to-face interviews with adult household members aged 18 years or older. A total of 12 trained field research assistants were recruited to carry out each survey under the supervision of a field research supervisor. To ensure data quality, this study's investigators regularly supervised the data collection team and monitored the data to promptly address any issues that arose during the data collection process in this study's areas. Information on the household members' demographics, socioeconomic characteristics, illness, health care usage, providers, and knowledge about the Aalo Clinic was collected. Household knowledge, awareness, and usage of the Aalo Clinic were assessed through respondent interviews. Awareness was defined as whether the respondent had ever heard of the Aalo Clinic, while knowledge required correctly identifying it as a local health care facility and naming at least one of its services. Usage referred to any household member receiving services from the Aalo Clinic before the survey, including consultations, disease care, diagnostics, or referrals. Respondents were also asked about their sources of information regarding the Aalo Clinic.

Usually, slum dwellers used health care from drug sellers, quacks, doctors' chambers, community-based health facilities run by government organizations and NGOs, private clinics and hospitals, government hospitals, medical colleges, and specialized health care institutes [10,12]. In this study, MTPs included doctors with at least an MBBS degree, whether GPs or specialized doctors, who provide health care services through government or nongovernment medical colleges, hospitals, private clinics, or any form of institutional health facilities. These health care organizations hire medically trained staff. In contrast, drug sellers and traditional healers who lack formal training in providing clinical services are categorized as non-MTPs in this study. A similar classification could be found elsewhere [18,21,22]. However, we divided the MTPs based on the Aalo Clinic and other health care facilities that functioned with the MTPs to assess the usage of health care from the MTPs of the Aalo Clinic model.

Evidence reported that in the context of LMICs, the income of informal workers is not stable due to seasonal variations, frequent migration, etc [23,24]. Therefore, to measure the socioeconomic status of slum households, we used asset quintiles rather than income levels. We categorized wealth status into 5 quintiles based on the available assets in the household, such as access to utility services, improved sanitation and drinking water facilities, modern technologies, furniture, etc. These asset variables were used to estimate an asset score through principal component analysis: a robust and validated method widely used in similar contexts. After adjusting for household size, the asset score was used to categorize household wealth status [25]. It is noted that the assessment of wealth quintiles is relative to the current study population and is not comparable to the wealth distribution in the general population of Bangladesh.

Data Analysis

The background information of the surveyed households, their members, and the health service usage patterns of the sick members are reported as categorical variables with frequency (n) and percentage (%). Out-of-pocket expenditure is presented as the mean and median in Bangladeshi Taka, with US $1=BDT 110. Descriptive analysis was performed to assess distributional changes among surveyed households between the baseline and end-line periods. Additionally, several inferential statistical tests were conducted to assess significant changes between the waves: the Wilcoxon rank sum test for 2 categorical variables, the Kruskal-Wallis test for more than 2 categorical variables, and the *t* test for continuous variables. These ensured the statistical rigor.

Cross-tabulation and chi-squared test were performed to assess the changes in health care usage from MTP among this study's population between the 2 waves. Further, 3 logistic regression models were applied in assessing the impact of the Aalo Clinic on health care usage from the MTP while controlling for other covariates. This analytic approach enables robust estimation of intervention impact while controlling for potential confounders. In the first model, the binary dependent variable was the "status of health care utilization from the Aalo Clinic" (1=used Aalo Clinic MTP services; 0=did not use Aalo Clinic MTP services). In the second model, the dependent variable was "MTP utilization other than Aalo Clinic" (1=used non–Aalo Clinic MTP services; 0=did not use non–Aalo Clinic MTP services). In the third model, the dependent variable was "status of health care utilization from non-MTPs" (1=used care from non-MTPs; 0=did not use services from non-MTPs). Independent predictor variables included age, sex, marital status, education, occupation, asset quintiles, distance of households, survey type, severity of illness, knowledge, and total health care expenditure. A P value of <.05 was considered statistically significant. Data analysis was conducted using Stata SE (version 16.0; StataCorp LLC).

Ethical Considerations

Ethical approval for this study was obtained from the institutional review board of the icddr,b before data collection (PR#22020). Informed written consent was taken from all the respondents before data collection, after explaining this study's objectives, the voluntary nature of participation, confidentiality assurances, and their right to withdraw at any time without

XSL•FO
RenderX

repercussions. For participants unable to read or write, the consent form was read aloud, and consent was documented via thumb impressions in the presence of an impartial witness. All data were anonymized to ensure privacy and confidentiality, with no personally identifiable information collected or disclosed. Data were stored securely with restricted access to authorized research team members only. No financial or material compensation was provided to participants for their involvement in this study. This paper does not include any images or supplementary materials that could reveal the identity of individual participants.

Results

Sociodemographic Characteristics of This Study's Participants

Both in baseline and end-line surveys, the proportion of males and females was almost equal. At the end-line survey, the proportion of members with no education (n=3047, 38.12%) and secondary and above (n=2485, 31.09%) increased slightly compared to the baseline period. Further, no substantial change was observed in terms of wealth quintiles between the periods. However, there was a sizable increase in the percentage of respondents who knew about Aalo Clinic at the end-line (n=1473, 73.36%) compared to baseline (n=227, 11.46%). Table 1 presents detailed information on the background characteristics of this study's population.

XSL•FO
RenderX

Table 1. Sociodemographic characteristics of study participants at baseline (January to February 2022) and end-line (April to June 2023) from household surveys in the urban slums of Bangladesh.

Variable	Baseline (January to February 2022)	End-line (April to June 2023)
Characteristics of household members	8919 (100)	7994 (100)
Age (years)		
Up to 19	3627 (40.67)	3142 (39.3)
20-39	3319 (37.21)	3035 (37.97)
40-59	1475 (16.54)	1383 (17.3)
60 or more	498 (5.58)	434 (5.43)
Sex		
Male	4460 (50.01)	3957 (49.5)
Female	4459 (49.99)	4037 (50.5)
Marital status		
Married	4617 (51.77)	4325 (54.1)
Unmarried	3799 (42.59)	3218 (40.26)
Others (eg, widows and divorced)	503 (5.64)	451 (5.64)
Education		
No education	3159 (35.42)	3047 (38.12)
Primary	3009 (33.74)	2462 (30.8)
Secondary and above	2751 (30.84)	2485 (31.09)
Occupation		
Currently employed	3502 (39.26)	3339 (41.77)
Currently unemployed	5310 (59.54)	4529 (56.65)
Others (eg, disabled and retired persons)	107 (1.2)	126 (1.58)
Regular earner		
Yes	3143 (35.24)	3146 (39.35)
No	5776 (64.76)	4848 (60.65)
Household-level characteristics	1981 (100)	2008 (100)
Household size		
Small (<4)	459 (23.17)	723 (36.01)
Medium (4-6)	1298 (65.52)	1164 (57.97)
Large (>6)	224 (11.31)	121 (6.03)
Household ownership status		
Owner	794 (40.08)	700 (34.86)
Tenant	1187 (59.92)	1308 (65.14)
Education of the household head		
No education	768 (38.77)	751 (37.40)
Primary	640 (32.31)	636 (31.67)
Secondary and above	573 (28.92)	621 (30.93)
Occupation of the household head		
Currently employed	1732 (87.43)	1743 (86.80)
Currently unemployed	214 (10.80)	202 (10.06)
Others (eg, disabled and retired persons)	35 (1.77)	63 (3.14)
Asset quintiles		

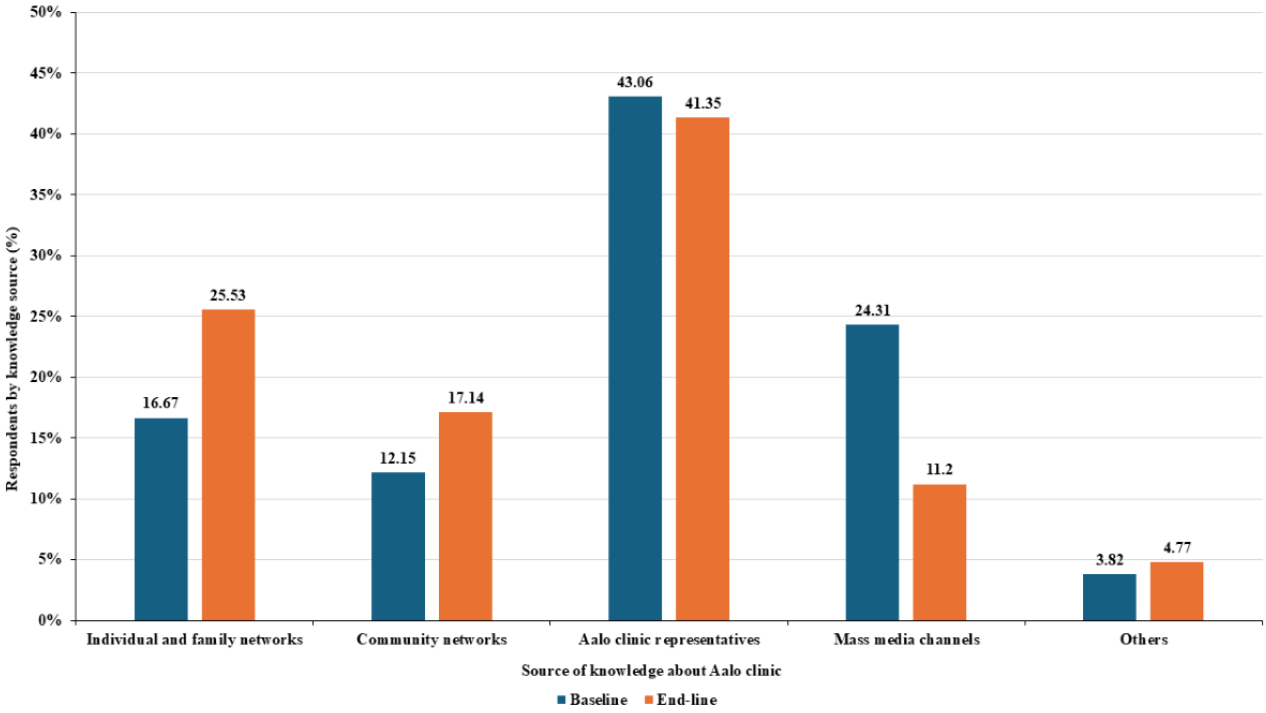
XSL•FO
RenderX

Variable	Baseline (January to February 2022)	End-line (April to June 2023)
Poorest	423 (21.35)	376 (18.73)
Poorer	382 (19.28)	416 (20.72)
Middle	389 (19.64)	408 (20.32)
Richer	410 (20.7)	388 (19.32)
Richest	377 (19.03)	420 (20.92)
City corporation		
Dhaka North City Corporation	794 (40.08)	807 (40.19)
Dhaka South City Corporation	788 (39.78)	798 (39.74)
Gazipur City Corporation	399 (20.14)	403 (20.07)
Slum location		
Korail	397 (20.04)	404 (20.12)
Mirpur	397 (20.04)	403 (20.07)
Shyampur	393 (19.84)	403 (20.07)
Dholpur	395 (19.94)	395 (19.67)
Ershadnagar	399 (20.14)	403 (20.07)
Knowledge about Aalo Clinic		
Yes	227 (11.46)	1473 (73.36)
No	1754 (88.54)	535 (26.64)

Source of Information About Aalo Clinic

Community networks were the primary source of knowledge about Aalo Clinic in both baseline (n=124, 43.06%) and end-line (n=1030, 41.35%) periods. During the baseline period, mass media channels were the second most common source (n=70, 24.31%). By the end-line, individual and family networks became the second most common source of information (n=636, 25.53%). Further details are presented in Figure 2.

Figure 2. Sources of information about the Aalo Clinic model among urban slum populations in Bangladesh during baseline (January to February 2022) and end-line (April to June 2023) household surveys.



XSL•FO
RenderX

Chronic Illness and Pattern of Health Care Usage and Expenditure

Around 19% (n=1511) of individuals had a chronic illness at the end-line period, whereas the ratio was about 16% (n=1414) in the baseline period. At the end-line, the highest number of people (n=327, 21.64%) endured chronic musculoskeletal problems, while at the baseline, cardiovascular disease was more common (n=286, 20.23%). At the baseline, around 84% (n=1188) of participants required regular treatment for chronic diseases, but only around 71% (n=1010) received treatment in the past 3 months. At the end-line, around 82% (n=1236) needed regular treatment, but only around 38% (n=566) reported receiving it in the past 3 months. Health care usage from private facilities was significantly lower at the end-line (n=202, 35.69% from n=459, 45.45%), while it was higher for Aalo Clinic (n=63, 11.13% from n=6, 0.59%). However, usage from local drug stores remained almost the same, with 36.63% (n=209) at baseline and 36.93% (n=370) at end-line. Details about the diseases, pattern of health care usage, and expenditures are shown in Table 2.

XSL•FO
RenderX

Table 2. Health care usage patterns and expenditures for chronic disease management among urban slum populations in Bangladesh at baseline and end-line surveys.

Variable	Baseline	End-line
Have a chronic illness or disease, n (%)	8919 (100)	7994 (100)
Yes	1414 (15.85)	1511 (18.9)
No	7505 (84.15)	6483 (81.1)
Endure a chronic illness for the last 12 months or more, n (%)		
Yes	1353 (95.69)	1427 (94.44)
No	61 (4.31)	84 (5.56)
Type of chronic illness or disease, n (%)		
Heart disease or cardiovascular disease	286 (20.23)	248 (16.41)
Diabetes	204 (14.43)	231 (15.29)
Chronic respiratory disease	155 (10.96)	141 (9.33)
Cancer or tumors	34 (2.40)	40 (2.65)
Hepatobiliary problems	29 (2.05)	31 (2.05)
Neurological problem or disease	107 (7.57)	97 (6.42)
Chronic eye problem	87 (6.15)	109 (7.21)
Chronic gastrointestinal problems	100 (7.07)	82 (5.43)
Skin problem	46 (3.25)	108 (7.15)
Chronic musculoskeletal problems	251 (17.75)	327 (21.64)
Chronic urinary disorder or problem	64 (4.53)	32 (2.12)
Others (eg, anemia, sexual, etc)	51 (3.61)	65 (4.30)
Daily performance decreased due to the chronic illness, n (%)		
Yes	1126 (79.63)	1151 (76.17)
No	288 (20.37)	360 (23.83)
Need regular treatment for chronic illness, n (%)		
Yes	1188 (84.02)	1236 (81.8)
No	226 (15.98)	275 (18.2)
Received treatment for the chronic illness in the last 3 months, n (%)		
Yes	1010 (71.43)	566 (37.46)
No	404 (28.57)	945 (62.54)
Type of health care provider for chronic illness, n (%)		
Aalo Clinic provider	6 (0.59)	63 (11.13)
Medically trained providers (non–Aalo Clinic)	616 (60.99)	285 (50.35)
Nonmedically trained providers	388 (38.42)	218 (38.52)
Source of treatment for chronic illness, n (%)		
Public facility	148 (14.65)	73 (12.9)
Private facility	459 (45.45)	202 (35.69)
Aalo Clinic	6 (0.59)	63 (11.13)
NGO[a] facility	9 (0.89)	10 (1.77)
Pharmacy or local drug store	370 (36.63)	209 (36.93)
Others	18 (1.78)	9 (1.59)
Designation of the service providers for chronic illness, n (%)		
Qualified doctor (allopathic)	603 (59.7)	342 (60.42)

XSL•FO
RenderX

Variable	Baseline	End-line
Qualified doctor (homeopathic, unani, or ayurvedic)	13 (1.29)	7 (1.24)
Unqualified doctor (homeopathic, unani, or ayurvedic)	10 (0.99)	10 (1.77)
Drug seller	366 (36.24)	203 (35.87)
Others (eg, traditional healer)	18 (1.78)	4 (0.71)
Distance of the facility from respondents' home (min, baseline=798, end-line=484), n (%)		
Nearby areas (within 15)	291 (36.47)	173 (35.74)
Moderately distant areas (16 to 30)	214 (26.82)	129 (26.65)
Distant areas (more than 30)	293 (36.72)	182 (37.6)
Waiting time for health care services for chronic illness (min, baseline=798, end-line=484), n (%)		
Within 10	392 (49.12)	202 (41.74)
11-20	124 (15.54)	47 (9.71)
More than 20	282 (35.34)	235 (48.55)
Total expenditure for health care seeking for chronic illness		
Mean (SD)	2608 (7330)	2365 (6654)
Median (IQR)	1200 (2190)	980 (2020)
Total expenditure of health care seeking by health care facility type for chronic illness, mean (SD)		
Public facility	3427 (8457)	2476 (3629)
Private facility	3924 (9346)	4535 (10,457)
Aalo Clinic	590 (809)	708 (601)
NGO facility	1165 (970)	362 (495)
Pharmacy or local drug store	691 (866)	850 (756)
Others	867 (903)	1336 (1896)
Total expenditure of health care seeking by health care facility type for chronic illness, median (IQR)		
Public facility	1635 (1880)	1510 (2150)
Private facility	2140 (3120)	2760 (3560)
Aalo Clinic	400 (310)	550 (570)
NGO facility	1500 (1550)	100 (546)
Pharmacy or local drug store	400 (510)	630 (500)
Others	500 (770)	800 (780)

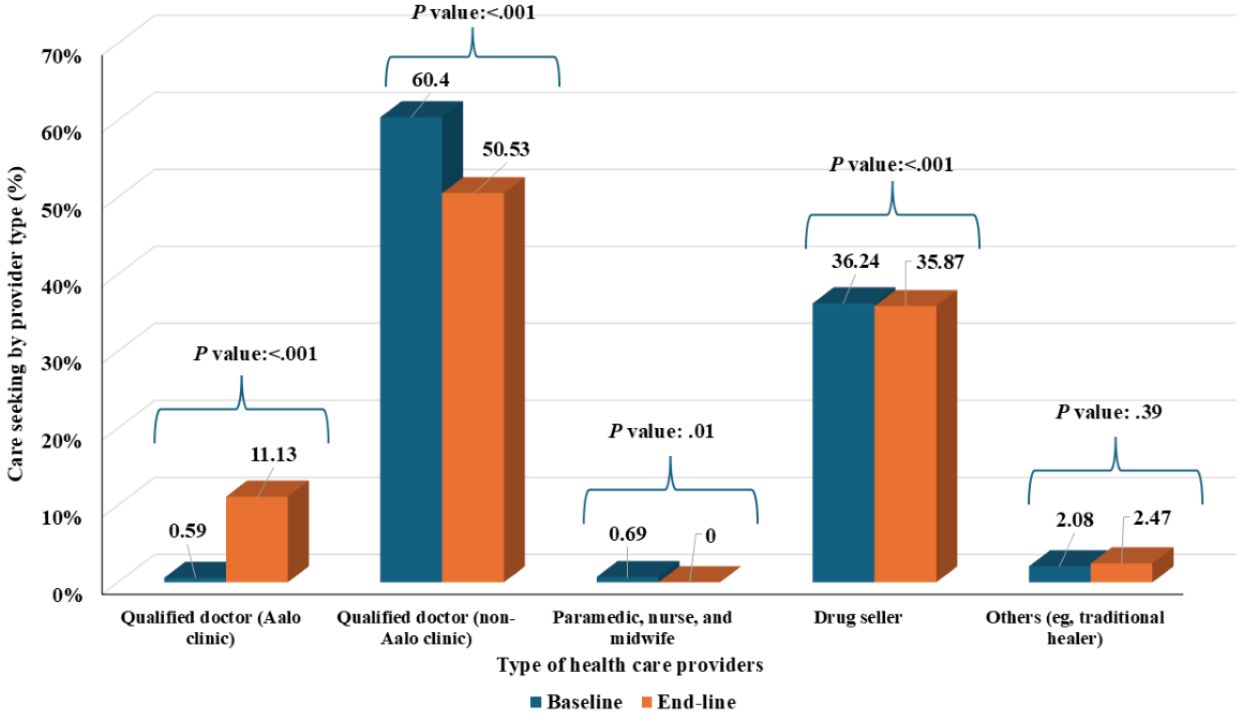
[a]NGO: nongovernmental organization.

Health Care Seeking Behavior

Figure 3 presents the preference of health care providers for chronic disease among slum dwellers. Health care usage from MTP at Aalo Clinic was significantly higher, rising from 0.59% (n=6) in the baseline period to 11.13% (n=63) in the end-line period. In both periods, non–Aalo Clinics and local drug sellers were the main health care sources, with non–Aalo Clinics at 60.40% (n=610) in the baseline and 50.53% (n=285) in the end-line, and local drug sellers at 36.24% (n=366) in the baseline and 35.87% (n=206) in the end-line. Reliance on non-MTPs declined slightly, and although this reduction was statistically significant, the magnitude of change was minimal.

XSL•FO
RenderX

Figure 3. Health care–seeking behavior for chronic diseases by type of health care providers among urban slum populations in Bangladesh during baseline and end-line surveys.



Health Care Usage of Chronic Disease Across Various Health Care Providers

At baseline, female patients (n=6, 0.98%), married individuals (n=4, 0.53%), those with no education (n=5, 0.97%), household heads with secondary education or higher (n=3, 0.97%), the unemployed (n=5, 1.33%), irregular earners (n=6, 0.97%), individuals from the richest households (n=3, 1.02%), and those who experienced a reduction in working performance due to chronic illness (n=5, 0.60%) predominantly used health care services from Aalo Clinic compared to their counterparts. A similar pattern was observed at the end-line period for the usage of Aalo Clinic MTPs. Females (n=54, 13.64%), widows or divorced individuals (n=14, 12.96%), household heads with no education (n=26, 11.82%), the unemployed (n=43, 15.69%), currently employed household heads (n=43, 9.45%), irregular earners (n=51, 15.6%), those in the richest asset quintile (n=19, 10.73%), and those who experienced a reduction in working performance due to chronic illness (n=59, 12.72%) also predominantly used health care services from Aalo Clinic compared to their counterparts. Table 3 presents a detailed distribution of patients' health care usage across various providers between 2 periods.

XSL•FO
RenderX

Table 3. Distribution of patients with chronic disease by type of health care providers used at baseline and end-line surveys in urban slums of Bangladesh.

Variable	Baseline			End-line		
	Aalo Clinic provider	Medically trained providers (non–Aalo Clinic)	Nonmedically trained providers	Aalo Clinic provider	Medically trained providers (non–Aalo Clinic)	Nonmedically trained providers
Characteristics of household members, n (%)						
Age (years)[a]						
Up to 19	0 (0)	46 (74.19)	16 (25.81)	2 (4.55)	26 (59.09)	16 (36.36)
20-39	2 (0.65)	202 (65.16)	106 (34.19)	23 (11.79)	96 (49.23)	76 (38.97)
40-59	3 (0.75)	226 (56.22)	173 (43.03)	26 (11.87)	113 (51.60)	80 (36.53)
60 or more	1 (0.42)	142 (60.17)	93 (39.41)	12 (11.11)	52 (48.15)	44 (40.74)
Sex[b]						
Male	0 (0)	255 (64.39)	141 (35.61)	9 (5.29)	88 (51.76)	73 (42.94)
Female	6 (0.98)	361 (58.79)	247 (40.23)	54 (13.64)	197 (49.75)	145 (36.62)
Marital status[c]						
Married	4 (0.53)	453 (60.48)	292 (38.99)	46 (11.19)	207 (50.36)	158 (38.44)
Unmarried	0 (0)	58 (77.33)	17 (22.67)	3 (6.38)	28 (59.57)	16 (34.04)
Others (eg, widows and divorced)	2 (1.08)	105 (56.45)	79 (42.47)	14 (12.96)	50 (46.3)	44 (40.74)
Education[d]						
No education	5 (0.97)	298 (57.98)	211 (41.05)	36 (13.38)	129 (47.96)	104 (38.66)
Primary	1 (0.38)	151 (56.77)	114 (42.86)	17 (11.26)	76 (50.33)	58 (38.41)
Secondary and above	0 (0)	167 (72.61)	63 (27.39)	10 (6.85)	80 (54.79)	56 (38.36)
Education of the household head[e]						
No education	1 (0.24)	239 (56.64)	182 (43.13)	26 (11.82)	106 (48.18)	88 (40)
Primary	2 (0.72)	161 (57.71)	116 (41.58)	22 (13.5)	79 (48.47)	62 (38.04)
Secondary and above	3 (0.97)	216 (69.9)	90 (29.13)	15 (8.2)	100 (54.64)	68 (37.16)
Occupation[f]						
Currently employed	0 (0)	236 (57.7)	173 (42.3)	13 (5.2)	121 (48.4)	116 (46.4)
Currently unemployed	5 (1.33)	220 (58.51)	151 (40.16)	43 (15.69)	137 (50)	94 (34.31)
Others (eg, disabled and retired persons)	1 (0.44)	160 (71.11)	64 (28.44)	7 (16.67)	27 (64.29)	8 (19.05)
Occupation of the household head[g]						
Currently employed	5 (0.61)	491 (60.25)	319 (39.14)	43 (9.45)	231 (50.77)	181 (39.78)
Currently unemployed	0 (0)	45 (56.96)	34 (43.04)	12 (19.67)	24 (39.34)	25 (40.98)
Others (eg, disabled and retired persons)	1 (0.86)	80 (68.97)	35 (30.17)	8 (16)	30 (60)	12 (24)
Regular earner[h]						
Yes	0 (0)	231 (58.93)	161 (41.07)	12 (5.02)	119 (49.79)	108 (45.19)
No	6 (0.97)	385 (62.3)	227 (36.73)	51 (15.6)	166 (50.76)	110 (33.64)
Household size[i]						
Small (<4)	1 (0.79)	80 (63.49)	45 (35.71)	18 (11.76)	81 (52.94)	54 (35.29)
Medium (4-6)	3 (0.43)	416 (60.29)	271 (39.28)	37 (10.76)	163 (47.38)	144 (41.86)

XSL•FO
RenderX

Variable	Baseline			End-line		
	Aalo Clinic provider	Medically trained providers (non–Aalo Clinic)	Nonmedically trained providers	Aalo Clinic provider	Medically trained providers (non–Aalo Clinic)	Nonmedically trained providers
Large (>6)	2 (1.03)	120 (61.86)	72 (37.11)	8 (11.59)	41 (59.42)	20 (28.99)
Asset quintiles[j]						
Poorest	1 (0.7)	87 (60.84)	55 (38.46)	5 (6.41)	27 (34.62)	46 (58.97)
Poorer	0 (0)	82 (56.94)	62 (43.06)	7 (8.24)	44 (51.76)	34 (40)
Middle	1 (0.49)	114 (55.61)	90 (43.9)	15 (13.27)	49 (43.36)	49 (43.36)
Richer	1 (0.44)	140 (62.22)	84 (37.33)	17 (15.04)	61 (53.98)	35 (30.97)
Richest	3 (1.02)	193 (65.87)	97 (33.11)	19 (10.73)	104 (58.76)	54 (30.51)
Ownership of a living house[k]						
Owner	5 (0.86)	363 (62.59)	212 (36.55)	37 (12.63)	154 (52.56)	102 (34.81)
Tenant	1 (0.23)	253 (58.84)	176 (40.93)	26 (9.52)	131 (47.99)	116 (42.49)
City corporation[l]						
Dhaka North City Corporation	1 (0.26)	233 (61.15)	147 (38.58)	18 (7.86)	116 (50.66)	95 (41.48)
Dhaka South City Corporation	0 (0)	243 (61.83)	150 (38.17)	20 (11.76)	92 (54.12)	58 (34.12)
Gazipur City Corporation	5 (2.12)	140 (59.32)	91 (38.56)	25 (14.97)	77 (46.11)	65 (38.92)
Study area (slums)[m]						
Korail	0 (0)	137 (68.5)	63 (31.5)	7 (5.83)	73 (60.83)	40 (33.33)
Mirpur	1 (0.55)	96 (53.04)	84 (46.41)	11 (10.09)	43 (39.45)	55 (50.46)
Shyampur	0 (0)	122 (68.16)	57 (31.84)	7 (8.64)	39 (48.15)	35 (43.21)
Dholpur	0 (0)	121 (56.54)	93 (43.46)	13 (14.61)	53 (59.55)	23 (25.84)
Ershadnagar	5 (2.12)	140 (59.32)	91 (38.56)	25 (14.97)	77 (46.11)	65 (38.92)
Type of chronic illness or disease, n (%)[n]						
Heart disease or cardiovascular disease	2 (0.9)	112 (50.68)	107 (48.42)	9 (10.59)	37 (43.53)	39 (45.88)
Diabetes	2 (1.16)	101 (58.38)	70 (40.46)	16 (14.04)	53 (46.49)	45 (39.47)
Chronic respiratory disease	0 (0)	73 (60.83)	47 (39.17)	4 (7.84)	23 (45.1)	24 (47.06)
Total expenditure of health care seeking by health care provider type for chronic illness[r]						
Mean (SD)	590 (809)	3760 (9065)	698 (867)	708 (601)	3855 (9066)	863 (802)
Median (IQR)	400 (310)	2000 (2870)	400 (540)	550 (570)	2295 (3630)	630 (500)

Econometric Analysis

Factors associated with health care usage for chronic disease from MTPs of Aalo Clinic, non–Aalo Clinic, and non-MTPs are shown in Table 4. Female household members were over 3 times more likely (odds ratio [OR] 3.68, 95% CI 1.54-8.79) to use services from MTPs at Aalo Clinic compared to their male counterparts. Furthermore, individuals residing within the nearby areas were 2 times more likely (OR 2.77, 95% CI 1.08-7.11), and individuals residing within a moderate distance from Aalo Clinic were 3 times more likely (OR 3.23, 95% CI 1.23-8.48) to use its services compared to those living at a longer distance. Compared to the baseline period, household members were 23 times more likely (OR 23.28, 95% CI 8.77-61.80) to use MTPs at Aalo Clinic during the end-line period. Patients whose daily performance decreased due to chronic illnesses were 3 times more likely (OR 3.27, 95% CI 1.16-9.23) to take services from MTPs at Aalo Clinic during the end-line period. Additionally, household members who had knowledge about Aalo Clinic were twice as likely (OR 2.13, 95% CI 1.09-4.14) to use its MTPs compared to those without such knowledge.

XSL•FO
RenderX

Table 4. Factors associated with health care usage from Aalo Clinic, non–Aalo Clinic, medically trained and non-MTPs[a] for chronic disease management among urban slum populations in Bangladesh.

Characteristics of household members	Aalo Clinic provider, OR[b] (95% CI)	Medically trained providers (non–Aalo Clinic), OR (95% CI)	Non-MTPs, OR (95% CI)
Age (years)			
Up to 19	Reference	Reference	Reference
20-39	1.44 (0.12-17.42)	1.06 (0.25-4.47)	0.77 (0.20-2.96)
40-59	1.85 (0.14-24.57)	0.83 (0.19-3.63)	0.90 (0.23-3.63)
60 or more	2.00 (0.14-28.76)	1.26 (0.27-5.81)	0.66 (0.16-2.76)
Sex			
Male	Reference	Reference	Reference
Female	3.68[c] (1.54-8.79)	0.91 (0.62-1.33)	0.85 (0.59-1.22)
Marital status			
Married	Reference	Reference	Reference
Unmarried	1.43 (0.17-12.29)	2.35 (0.60-9.21)	0.40 (0.11-1.43)
Others (eg, widows and divorced)	0.69 (0.30-1.59)	0.76 (0.47-1.23)	1.37 (0.87-2.15)
Education of the household head			
No education	0.81 (0.40-1.65)	1.25 (0.84-1.87)	0.83 (0.57-1.21)
Primary	Reference	Reference	Reference
Secondary and above	0.66 (0.30-1.45)	2.05[c] (1.33-3.16)	0.58[c] (0.38-0.87)
Occupation of the household head			
Currently employed	0.92 (0.40-2.11)	0.92 (0.57-1.49)	1.12 (0.71-1.78)
Currently unemployed	Reference	Reference	Reference
Others (eg, disabled and retired persons)	2.51 (0.62-10.13)	1.48 (0.54-4.05)	0.44 (0.16-1.19)
Distance of the facility from respondents' home			
Nearby areas	2.77[d] (1.08-7.11)	0.04[e] (0.02-0.06)	23.15[e] (13.92-38.47)
Moderately distant areas	3.23[c] (1.23-8.48)	0.09[e] (0.06-0.15)	10.23[e] (6.11-17.12)
Distant areas	Reference	Reference	Reference
Household size			
Small (<4)	0.67 (0.22-2.00)	1.55 (0.83-2.87)	0.67 (0.37-1.23)
Medium (4-6)	0.62 (0.23-1.64)	1.01 (0.62-1.65)	1.01 (0.63-1.63)
Large (>6)	Reference	Reference	Reference
Asset quintiles			
Poorest	0.56 (0.19-1.66)	1.19 (0.68-2.09)	0.99 (0.59-1.68)
Poorer	0.68 (0.24-1.96)	1.21 (0.70-2.11)	0.96 (0.57-1.60)
Middle	Reference	Reference	Reference
Richer	1.38 (0.58-3.26)	1.34 (0.82-2.21)	0.73 (0.45-1.16)
Richest	1.04 (0.44-2.49)	1.49 (0.91-2.45)	0.73 (0.46-1.17)
Study area (slums)			
Korail	Reference	Reference	Reference
Mirpur	2.47 (0.76-8.02)	0.31[e] (0.18-0.54)	2.29[c] (1.38-3.79)
Shyampur	2.44 (0.68-8.76)	0.37[c] (0.21-0.66)	2.17[c] (1.25-3.76)
Dholpur	2.47 (0.75-8.09)	0.59[d] (0.35-1.00)	1.27 (0.77-2.10)

XSL•FO
RenderX

Characteristics of household members	Aalo Clinic provider, OR[b] (95% CI)	Medically trained providers (non–Aalo Clinic), OR (95% CI)	Non-MTPs, OR (95% CI)
Ershadnagar	4.03[c] (1.36-11.91)	0.63 (0.38-1.03)	1.12 (0.70-1.79)
Survey type			
Baseline	Reference	Reference	Reference
End-line	23.28[e] (8.77-61.80)	0.33[e] (0.23-0.48)	1.22 (0.87-1.72)
Daily performance decreased due to chronic illness			
Yes	3.27[c] (1.16-9.23)	2.13[c] (1.36-3.36)	0.39[e] (0.25-0.60)
No	Reference	Reference	Reference
Knowledge about Aalo Clinic			
Yes	2.13[c] (1.09-4.14)	1.03 (0.61-1.75)	0.68 (0.41-1.10)
No	Reference	Reference	Reference
Total expenditure for the chronic illness	0.9991[e] (0.9986-0.9996)	1.0008[e] (1.0006-1.0010)	0.9993[c] (0.9992-0.9995)

[a]MTP: medically trained provider.
[b]OR: odds ratio.
[c]$P<.01$.
[d]$P<.05$.
[e]$P<.001$.

On the other hand, patients were 2 times more likely (OR 2.05, 95% CI 1.33-3.16) to use non–Aalo Clinic MTPs if their household head had a secondary or higher education level compared to those whose household head had only a primary education level. Individuals residing nearby or within a moderate distance from non–Aalo Clinic MTPs were less likely to use these services compared to those living at a longer distance. Individuals from the richest asset quintiles were about 1.5 times more likely (OR 1.49, 95% CI 0.91-2.45) to use non–Aalo Clinic MTPs compared to those in the middle quintile of this study's population. During the end-line period, household members were less likely (OR 0.33, 95% CI 0.23-0.48) to use non–Aalo Clinic MTPs compared to the baseline period. Moreover, patients whose daily performance decreased due to chronic illnesses were 2 times more likely (OR 2.13, 95% CI 1.36-3.36) to take services from MTPs at non–Aalo Clinics during the end-line period.

Further, we found that households where the household head had an education of secondary or above level were less likely (OR 0.58, 95% CI 0.38-0.87) to use non-MTPs compared to others. Individuals residing nearby and within a moderate distance from non-MTPs were about 23 times and 10 times more likely (OR 23.15, 95% CI 13.92-38.47 and OR 10.23, 95% CI 6.11-17.12) to use these services compared to those living at a longer distance, respectively. Members' richest asset quintiles were less likely (OR 0.73, 95% CI 0.45-1.16) to use non-MTPs compared to those in the middle quintile. Additionally, household members who possessed knowledge about Aalo Clinic were less likely (OR 0.68, 95% CI 0.41-1.10) to use non-MTPs compared to those without such knowledge. However, the association was not statistically significant.

Discussion

Principal Results and Comparison With Prior Work

The pilot Aalo Clinic model offered an essential package of health services, including chronic diseases, and an opportunity for outpatient PHC for the urban low-income population.

We found that health care use from Aalo Clinic MTPs for chronic diseases was significantly associated with higher use compared to other MTPs. The Aalo Clinic's unique features, such as the availability of affordable and quality health care services for chronic diseases, may significantly influence its impact on the community. The multivariate logistic regression analysis demonstrated a significant effect that individuals were over 23 times more likely to use Aalo Clinic MTP services at end-line than baseline. A similar initiative, namely the "Mohalla Clinic" in Delhi, India, has been implemented [26], and its impact aligns with the findings of this study to make primary care accessible and affordable from MTP for urban underserved populations [27,28]. However, despite these positive trends, the intervention did not substantially reduce the usage of health care from non-MTPs. Non-MTP providers remain prevalent and widely trusted within these communities. The continued reliance on non-MTPs in slum areas can be attributed to several factors, including deeply rooted health care–seeking behaviors, cultural preferences, limited medicine supply, and the convenience and accessibility of informal providers such as drug sellers and traditional healers. Shifting patient preferences away from non-MTPs is a complex and gradual process that requires sustained engagement.

Knowledgeable individuals can make informed decisions when choosing health care providers [29]. The integration of awareness measures in our analysis is a strength, offering a

XSL•FO
RenderX

deeper understanding of demand-side dynamics. Among the slum population, knowledge about Aalo Clinic increased significantly, which may be associated with greater usage of health care services from Aalo Clinic MTPs compared to other MTPs. A study conducted by Hasan et al [21] identified factors that influence health care usage from the MTP of a government-sponsored health protection scheme in Bangladesh. The scheme covers the population who are below the poverty line and live in semiurban and rural areas. However, this study reported that households with knowledge about the scheme were significantly associated with higher health care usage from the MTP of that scheme.

While evidence from LMICs often shows higher health care usage among males [30-32], our study found that females were more likely to access Aalo Clinic MTP services. Several contextual factors may explain this finding. The strategic placement of Aalo Clinics within slum areas enhances accessibility, while the provision of privacy, reduced costs, and shorter travel distances are particularly important for women in the urban poor settings of Bangladesh. These factors likely contribute to the higher uptake of services among females at Aalo Clinics compared to other MTPs. This observation is consistent with findings from studies among similar and general populations in Bangladesh [33,34], highlighting the importance of placing clinics close to communities to reduce sex-based access barriers.

This study identified a noteworthy pattern in health care usage, demonstrating a significant association between the distance from households and the type of health care provider. Our findings indicate that individuals are more motivated to seek care from providers located in closer proximity, irrespective of whether the provider is an MTP or a non-MTP. This observation aligns with previous studies, which have similarly highlighted the critical role of proximity in influencing health care–seeking behavior [35,36]. On the other hand, the significant positive association between longer distance and higher health care usage, specifically from non–Aalo Clinic MTPs, reveals an interesting trend. This could be indicative of several factors, such as specialization of health facilities for specific services, reputation, and the referral system [37]. Non–Aalo Clinic MTPs provide advanced laboratory tests, specialist consultations, and inpatient care unavailable at the more basic primary care level of Aalo Clinics. This may be associated with patients traveling from farther away or being referred to non–Aalo Clinic MTPs for services not offered locally, resulting in higher usage of these more distant providers. Hasan et al [21] also found that, though health care, medicine, and diagnostic services are free in a government-sponsored health protection scheme, a higher proportion of people still sought care from other MTPs located within a similar or greater distance.

Though there was no significant difference in health care usage from Aalo Clinic MTPs across asset quintiles, a significant difference was observed in the usage of health care services from non–Aalo Clinic MTPs and non-MTPs across asset quintiles. Patients who use health care from non–Aalo Clinic MTPs where those who belong to the richest quintiles were more likely to use health care from non–Aalo Clinic MTPs; whereas those who use health care from non-MTPs, among those who belong to the richest quintiles, were less likely to seek health care from non-MTPs. Ahmed et al [18] conducted a study in Bangladesh and showed a similar finding that individuals of the richest households are more likely to use health care than individuals of the poorest households. Moreover, evidence reported that people who belong to the richest quintiles are less likely to use health care from non-MTPs [38,39].

Nevertheless, the Aalo Clinic model has been delivering free-of-cost essential health care services for chronic diseases for the low-income slum population, a substantial portion of them opt for informal providers, such as drug sellers. This behavior may be rooted in long-standing historical and cultural patterns of health care seeking within these communities. Traditionally, many slum residents have relied on informal providers such as drug sellers, traditional healers, or unlicensed practitioners due to their easy accessibility, flexible hours, and lower immediate costs. These providers are often trusted members of the community who offer quick remedies, sometimes with payment options that formal providers cannot match. Additionally, limited awareness of the benefits of MTPs, coupled with barriers such as distance, cost, and social norms, has reinforced preferences for informal care.

Consequently, changing these entrenched behaviors will require sustained time and concerted efforts. The widespread presence of informal providers, combined with the lower travel time and out-of-pocket costs associated with accessing them, further perpetuates their use. To effectively address this challenge, it is crucial to establish and maintain satellite clinic services within the community over an extended period. These clinics can serve as accessible platforms for educating slum residents on the importance of regularly using qualified MTPs for chronic disease management, while minimizing travel and financial barriers.

Both affordability and quality of health care are critical components of UHC. Thus, it is important to emphasize these. Though Aalo Clinic has made notable advancements in enhancing access to quality PHC, concerns regarding the affordability of its services remain, primarily due to associated costs. Despite the clinic's strategic location for slum dwellers and its provision of free registration, consultations, essential medicines, and diagnostic or laboratory services, users may still incur significant health care expenses or face higher opportunity costs. These financial burdens may often arise from the limited availability of certain medicines and laboratory services at Aalo Clinics, necessitating that patients seek these resources externally, thereby increasing out-of-pocket payments. To mitigate this affordability issue, it is imperative to ensure that Aalo Clinics are adequately stocked with essential medicines and laboratory services. Close collaboration with government stakeholders responsible for urban PHC provision should be pursued to address these gaps effectively. Enhancing the operational readiness of the model could help reduce out-of-pocket expenses and improve equitable access to affordable health care services, thereby advancing progress toward UHC.

XSL•FO
RenderX

Even though the Aalo Clinic model is associated with higher health care usage, it faces a significant challenge in the form of extended waiting times for accessing services. This issue likely arises from either an influx of patients or an inadequate number of health care providers. Each Aalo Clinic center is staffed with only 1 male and 1 female GP per shift, which may prove insufficient given the patient load. Research indicates that longer waiting times are often negatively correlated with perceived quality of care [40]. To address this issue effectively, it is crucial to improve both the infrastructure and human resource capacity of the Aalo Clinic. Enhancing these aspects could help mitigate waiting times and improve overall patient satisfaction and service delivery.

Limitations

One of the key limitations is that we were unable to account for seasonal variations in health care usage because the baseline survey was conducted from January to February 2022, while the end-line survey occurred between April and June 2023. Additionally, our reliance on self-reported data introduces the possibility of recall bias. Furthermore, the repeated cross-sectional design of this study limits our ability to make causal inferences about the relationship between significant factors and health care usage. Therefore, to better understand these associations over time, future research using a longitudinal design is warranted. As we did not follow the same households over time, it is difficult to determine whether observed changes are directly attributable to the Aalo Clinic intervention or influenced by external factors, such as other NGOs, government schemes, or broader contextual changes in the urban health system. Finally, the use of equal sample sizes across slums, regardless of population size, may limit generalizability, although it allowed for better cross-site comparison.

Conclusions

The Aalo Clinic model is significantly associated with increased usage of MTPs for chronic disease care, with patients showing a shift in preference from other MTPs, particularly private providers, to Aalo Clinics. Although there was a statistically significant but modest reduction in reliance on non-MTPs, informal providers continue to serve as an important source of care in low-income, underserved slum areas. To support continuity of care for chronic diseases requiring regular treatment, Aalo Clinic authorities should maintain consistent availability of essential medications. Future longitudinal research is warranted to better understand the drivers of patient preferences for chronic diseases and the longer-term impact of the Aalo Clinic model before scale-up, providing evidence to inform strategies for strengthening urban PHC and advancing progress toward UHC for low-income slum populations in Bangladesh.

Acknowledgments

We sincerely acknowledge the funding support provided by the Swedish International Development Cooperation Agency (Sida) through UNICEF (United Nations International Children's Emergency Fund; 02150). We extend our gratitude to UNICEF, our direct donor. We also sincerely acknowledge the cooperation and support of Nari Maitree, PHD (Partners in Health and Development), and CMED (Cloud-Based Medical System) throughout this study. icddr,b (International Centre for Diarrhoeal Disease Research, Bangladesh) is also thankful to the Governments of Bangladesh and Canada for providing unrestricted or institutional support. All authors declared that they had insufficient funding to support open access publication of this paper, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee (APF) support for the publication of this paper.

Data Availability

The datasets used and/or analyzed during this study are available from the corresponding author upon reasonable request.

Funding

This study was funded by UNICEF, with support from Sida (02150). The funder, UNICEF Bangladesh, had no involvement in this study's design, data collection, analysis, and interpretation.

Authors' Contributions

MGR and SSM conceptualized the research idea. MGR, MZH, and SSM performed the data analysis and interpretation. MGR and SSM were responsible for writing the manuscript. MZH, KIT, HAP, KN, MWH, MZH, CS, MAJ, and MV contributed to writing, reviewing, and revising the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

CS, MAJ, and MV are affiliated with UNICEF.

References

1. Gitahi G, Cashin C. Universal Health Coverage. Handbook of Global Health 2021 May 12:1611-1649 [FREE Full text] [doi: 10.1007/978-3-030-05325-3_69-1] [Medline: 23882953]
2. WHO. Universal health coverage (UHC). 2025 Dec 5. URL: https://www.who.int/news-room/fact-sheets/detail/universal-health-coverage-(uhc)

XSL•FO
RenderX

3. WHO. Primary health care. 2025 Dec 5. URL: https://www.who.int/news-room/fact-sheets/detail/primary-health-care
4. Balabanova D, Mills A, Conteh L, Akkazieva B, Banteyerga H, Dash U, et al. Good health at low cost 25 years on: lessons for the future of health systems strengthening. Lancet 2013 Jun 15;381(9883):2118-2133. [doi: 10.1016/S0140-6736(12)62000-5] [Medline: 23574803]
5. World Bank Group. Urban population (% of total population) - Bangladesh. 2022. URL: https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS?locations=BD
6. Ahmed I. Building resilience of urban slums in Dhaka, Bangladesh. Procedia - Soc Behav Sci 2016;218:202-213. [doi: 10.1016/j.sbspro.2016.04.023]
7. WHO. Bangladesh health system review: health systems in transition. 2015. URL: https://apps.who.int/iris/bitstream/handle/10665/208214/9789290617051_eng.pdf?sequence=1&isAllowed=y
8. Ministry of Health and Family Welfare. National Urban Health Strategy 2020. Dhaka; 2020 Feb. URL: https://buhn.org/uploads/files/resources/reports/national/14_National%20Urban%20Health%20Strategy%202020.pdf
9. Asian Development Bank. Local government institutional assessment. 2015. URL: https://www.adb.org/sites/default/files/linked-documents/42177-013-ban-oth-03.pdf
10. icddr,b. Slum health in Bangladesh: insights for health and demographic surveillance. 2019. URL: http://dspace.icddrb.org/jspui/bitstream/123456789/9298/1/icddrb-SP154.pdf
11. Hamid SA, Begum A. Responsiveness of the urban primary health care delivery system in Bangladesh: a comparative analysis. Int J Health Plann Manage 2019;34(1):251-262. [doi: 10.1002/hpm.2626] [Medline: 30146682]
12. Mahmood SS, Hasan MZ, Hasan AMR, Rabbani MG, Begum F, Yousuf TB, et al. Health system impact of COVID-19 on urban slum population of Bangladesh: a mixed-method rapid assessment study. BMJ Open 2022;12(2):e057402 [FREE Full text] [doi: 10.1136/bmjopen-2021-057402] [Medline: 35197355]
13. Ahmed S, Ahmed MW, Hasan MZ, Mehdi GG, Islam Z, Rehnberg C, et al. Assessing the incidence of catastrophic health expenditure and impoverishment from out-of-pocket payments and their determinants in Bangladesh: evidence from the nationwide household income and expenditure survey 2016. Int Health 2022;14(1):84-96 [FREE Full text] [doi: 10.1093/inthealth/ihab015] [Medline: 33823538]
14. Health Economics Unit. Bangladesh national health accounts 1997-2020. 2022. URL: http://www.heu.gov.bd/site/notices/00b55bcc-7eb1-4ee6-adb8-dfed4e9c2eca/bd-National-Health-Accounts-1997-2020
15. UNDP. Urban social protection strategy and action plan in Bangladesh. 2020. URL: https://socialprotection.gov.bd/wp-content/uploads/2024/01/Urban-Social-Protection-Strategy-and-Action-Plan.pdf
16. Alam K, Ahmed S. Cost recovery of NGO primary health care facilities: a case study in Bangladesh. Cost Eff Resour Allocation 2010 Jun 9;8:12 [FREE Full text] [doi: 10.1186/1478-7547-8-12] [Medline: 20529379]
17. Health Economics Unit, Ministry of Health & Family Welfare. Expanding Social Protection for Health Towards Universal Coverage: Health Care Financing Strategy 2012-2032. 2012. URL: https://socialprotection.gov.bd/wp-content/uploads/2017/03/HCF-Strategy-Bd-2012-2032.pdf
18. Ahmed S, Sarker AR, Sultana M, Chakrovorty S, Ahmed MW, Dorin F, et al. The impact of community-based health insurance on the utilization of medically trained healthcare providers among informal workers in Bangladesh. PLoS One 2018;13(7):e0200265 [FREE Full text] [doi: 10.1371/journal.pone.0200265] [Medline: 29995899]
19. UNICEF. Three-month-old Sulaiman beats pneumonia: a new clinic model is offsetting the burden of health care costs in the slums of Bangladesh. 2023. URL: https://www.unicef.org/bangladesh/en/stories/three-month-old-sulaiman-beats-pneumonia
20. Bangladesh Urban Health Network (BUHN). Aalo Clinic. URL: https://buhn.org/aalo-clinic
21. Hasan MZ, Ahmed MW, Mehdi GG, Khan JAM, Islam Z, Chowdhury ME, et al. Factors affecting the healthcare utilization from Shasthyo Suroksha Karmasuchi scheme among the below-poverty-line population in one subdistrict in Bangladesh: a cross sectional study. BMC Health Serv Res 2022;22(1):885 [FREE Full text] [doi: 10.1186/s12913-022-08254-1] [Medline: 35804366]
22. Khan JAM, Ahmed S, Sultana M, Sarker AR, Chakrovorty S, Rahman MH, et al. The effect of a community-based health insurance on the out-of-pocket payments for utilizing medically trained providers in Bangladesh. Int Health 2020;12(4):287-298 [FREE Full text] [doi: 10.1093/inthealth/ihz083] [Medline: 31782795]
23. O'Donnell O, van Doorslaer E, Wagstaff A, Lindelow M. Analyzing health equity using household survey data. World Bank Institute. Washington, DC: World Bank Publications; 2008. URL: https://documents1.worldbank.org/curated/en/633931468139502235/pdf/424800ISBN978011OFFICIAL0USE0ONLY10.pdf [accessed 2026-04-02]
24. Asian Development Bank. The Informal Sector and Informal Employment in Bangladesh: Country Report 2010. 2012. URL: https://www.adb.org/sites/default/files/publication/30084/informal-sector-informal-employment-bangladesh.pdf
25. Vyas S, Kumaranayake L. Constructing socio-economic status indices: how to use principal components analysis. Health Policy Plan 2006;21(6):459-468. [doi: 10.1093/heapol/czl029] [Medline: 17030551]
26. Lahariya C. Mohalla clinics of Delhi, India: could these become platform to strengthen primary healthcare? J Family Med Prim Care 2017;6(1):1-10 [FREE Full text] [doi: 10.4103/jfmpc.jfmpc_29_17] [Medline: 29026739]
27. Yamini, Bodkhe N. Assessing the effect of mohalla clinics on healthcare system in Delhi: evidence from a primary study in west district. Indian J Hum Dev 2022;16(3):493-512. [doi: 10.1177/09737030221143220]

XSL•FO
RenderX

28. Lahariya C. Access, utilization, perceived quality, and satisfaction with health services at Mohalla (Community) clinics of Delhi, India. J Family Med Prim Care 2020;9(12):5872-5880 [FREE Full text] [doi: 10.4103/jfmpc.jfmpc_1574_20] [Medline: 33681011]
29. Lewis CL, Pignone MP. Promoting informed decision-making in a primary care practice by implementing decision aids. N C Med J 2009;70(2):136-139 [FREE Full text] [Medline: 19489371]
30. Kalin Y. Access to and utilization of health services in rural Bangladesh: a gender perspective. Lund University Publications. 2011. URL: https://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=2163324&fileOId=2163326 [accessed 2026-04-02]
31. Dupas P, Jain R. Women left behind: gender disparities in utilization of government health insurance in India. Am Econ Rev 2024;114(10):3345-3383. [doi: 10.1257/aer.20230521]
32. Saikia N, Moradhvaj, Bora JK. Gender difference in health-care expenditure: evidence from India human development survey. PLoS One 2016;11(7):e0158332 [FREE Full text] [doi: 10.1371/journal.pone.0158332] [Medline: 27391322]
33. Adams AM, Islam R, Yusuf SS, Panasci A, Crowell N. Healthcare seeking for chronic illness among adult slum dwellers in Bangladesh: a descriptive cross-sectional study in two urban settings. PLoS One 2020;15(6):e0233635 [FREE Full text] [doi: 10.1371/journal.pone.0233635] [Medline: 32542043]
34. Bertakis KD, Azari R, Helms LJ, Callahan EJ, Robbins JA. Gender differences in the utilization of health care services. J Fam Pract 2000;49(2):147-152. [Medline: 10718692]
35. Wambui WM, Kimani S, Odhiambo E. Determinants of health seeking behavior among caregivers of infants admitted with acute childhood illnesses at Kenyatta National Hospital, Nairobi, Kenya. Int J Pediatr 2018;2018:5190287 [FREE Full text] [doi: 10.1155/2018/5190287] [Medline: 30643520]
36. Anselmi L, Lagarde M, Hanson K. Health service availability and health seeking behaviour in resource poor settings: evidence from Mozambique. Health Econ Rev 2015;5(1):62 [FREE Full text] [doi: 10.1186/s13561-015-0062-6] [Medline: 26329425]
37. Abraham J, Sick B, Anderson J, Berg A, Dehmer C, Tufano A. Selecting a provider: what factors influence patients decision making? J Healthcare Manage 2011;56(2):99-116 [FREE Full text] [doi: 10.1097/00115514-201103000-00005]
38. Rodrigues R, Ilinca S, Schmidt A. Inequality and inequity in the use of long-term care services in Europe: is there reason for concern? Eurohealth (Lond) 2017;23(4):7-10 [FREE Full text]
39. Rodríguez-Sánchez B, Pascual Sáez M, Cantarero-Prieto D. Dependent, poorer, and more care-demanding? An analysis of the relationship between being dependent, household income, and formal and informal care use in Spain. Int J Environ Res Public Health 2021;18(8):4339 [FREE Full text] [doi: 10.3390/ijerph18084339] [Medline: 33921914]
40. Bleustein C, Rothschild DB, Valen A, Valatis E, Schweitzer L, Jones R. Wait times, patient satisfaction scores, and the perception of care. Am J Manag Care 2014;20(5):393-400 [FREE Full text] [Medline: 25181568]

Abbreviations

GP: general practitioner
icddr,b: International Centre for Diarrhoeal Disease Research, Bangladesh
LMIC: low- and middle-income country
MTP: medically trained provider
NGO: nongovernmental organization
OR: odds ratio
PHC: primary health care
UHC: universal health coverage
UHDSS: Urban Health and Demographic Surveillance System

Edited by A Mavragani, T Sanchez; submitted 20.Apr.2025; peer-reviewed by P Barua, E Hannah; comments to author 06.Jun.2025; accepted 25.Sep.2025; published 29.May.2026.

Please cite as:
Rabbani MG, Hossain Z, Tisha KI, Nahar K, Sakin ZR, Nasrin T, Ahmed MW, Hasan MZ, Soloman C, Jahangir MA, Vandenent M, Mahmood SS
Implementation Effect of a Pilot Urban Primary Health Care Model on the Usage of Health Care From Medically Trained Providers for Chronic Diseases Among Slum Populations in Bangladesh: Findings From an Implementation Research Study
JMIR Public Health Surveill 2026;12:e76290
URL: https://publichealth.jmir.org/2026/1/e76290
doi:10.2196/76290
PMID:

XSL•FO
RenderX

©Md. Golam Rabbani, Zakir Hossain, Khadija Islam Tisha, Kamrun Nahar, Zillur Rahman Sakin, Towhida Nasrin, Mohammad Wahid Ahmed, Md. Zahid Hasan, Chandrasegarar Soloman, Margub Aref Jahangir, Maya Vandenent, Shehrin Shaila Mahmood. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 29.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

COVID-19 Information Sources and Vaccination Status Among Californian Adults by Generation Using the 2022 California Health Interview Survey: Cross-Sectional Study

Julia Forest Zabala, BS; Melissa Sablik, MS, RN, CPN; Gina R Finical, MS, RN, CCRN; Victoria F Keeton, RN, CNS, CPNP-PC, PhD; Janice F Bell, MN, MPH, PhD, FAAN

Betty Irene Moore School of Nursing, University of California, Davis, 2570 48th Street, Sacramento, CA, United States

Corresponding Author:
Julia Forest Zabala, BS
Betty Irene Moore School of Nursing, University of California, Davis, 2570 48th Street, Sacramento, CA, United States

Abstract

Background: As communication technology advances and the digital divide grows, a deeper understanding of the influence of different information sources on vaccine uptake by generations can inform targeted public health interventions in times of future crisis. While the COVID-19 pandemic highlighted the role of media sources on the decision to receive vaccines, no studies have focused on the impact of the type and number of information sources in a population-based sample in California.

Objective: In this study, we examined associations between Californians' self-reported most relied upon COVID-19 information sources, categorized by type and measured as a count, and their COVID-19 vaccination status using data collected from the 2022 California Health Interview Survey. To address differences in information preferences and vaccine uptake by age, we also tested for potential effect modification of the relationship between relied upon COVID-19 information sources and vaccination status by generational membership (eg, Generation Z, millennials, Generation X, baby boomers, and Silent Generation).

Methods: We conducted a secondary analysis of cross-sectional data from the 2022 California Health Interview Survey. Vaccine status (any or none) was modeled as a function of information sources (or count) controlling for important sociodemographic and health confounding variables. Interaction terms of information sources (or count) by generational status were added to the models to test effect modification, and if significant, the models were stratified by generation. All analysis was survey-weighted to account for the complex survey sampling design.

Results: Compared to relying on traditional news media for COVID-19 information, relying on word of mouth (odds ratio [OR] 0.6), social media (OR 0.62), and doctors (OR 0.41) for COVID-19 information was associated with lower odds of being vaccinated for COVID-19. A dose-response relationship was identified, with each additional information source associated with 9% higher odds of being vaccinated for COVID-19. In stratified models, social media, compared to traditional news media, was associated with lower odds of vaccination for Generation X, baby boomers, and the Silent Generation.

Conclusions: Health information preferences, especially for traditional news media, are associated with COVID-19 vaccine uptake, and the information sources differ by generation. These findings provide information for stakeholders interested in vaccine hesitancy, health informatics, messaging strategies, health literacy, and future health information outreach programs during epidemics or pandemics. Dissemination of public health information should include multiple information sources to reach all individual preferences across different generations.

(JMIR Public Health Surveill 2026;12:e85904) doi:10.2196/85904

KEYWORDS

COVID-19 vaccines; COVID-19; self-report; pandemics; health literacy; social media; information sources; vaccination hesitancy; generational membership; media

Introduction

Background

Information Sources on COVID-19 Vaccination

The sources of information individuals consult and rely on have been linked to decisions on vaccination [1]. Nontraditional information sources such as social media and word of mouth vary in their reliability and verifiability and are not held to the same standards of peer review or fact-checking in the same way as information from government, news outlets, and health care [2]. Information and misinformation about COVID-19 can have dangerous consequences [3], because of the novelty of the disease and changes in information as more evidence is produced [4].

XSL•FO
RenderX

Studies in Sierra Leone and Italy have shown that the type and number of information sources used can significantly impact vaccine intention and uptake [5,6]. In an earlier study using the California Health Interview Survey (CHIS), people who reported using health care providers as their primary information source about vaccines have higher odds of human papillomavirus vaccine receipt or intention [7]. Another study found that obtaining vaccine information from the internet or relatives is associated with lower vaccination rates for measles, mumps, and rubella, and hepatitis B virus [1]. Additionally, when individuals used more than one information source to make measles, mumps, and rubella, or hepatitis B virus vaccine decisions, there was higher reported vaccine hesitancy [1]. On the contrary, a more recent study found that relying on and consulting more sources was associated with higher odds of COVID-19 vaccine uptake [8].

Vaccination Hesitancy and Uptake

Access to vaccinations and vaccination hesitancy (ie, the refusal or reluctance to receive vaccinations despite availability) are independent and joint contributors to vaccine uptake [9]. Vaccine hesitancy has been a persistent public health problem since the inception of vaccines [10], and in 2019, it was identified by the World Health Organization (WHO) as one of the top 10 global health threats [11]. Mass media, the technology used to reach the public such as radio, television, or newspapers, has historically played a significant role in influencing health beliefs and behaviors, including vaccine hesitancy and uptake [10,12].

Vaccination Uptake, Hesitancy, and Age

In 2021, vaccination uptake differed by generation, with coverage lowest among individuals aged 18 to 29 years and highest among individuals aged 65 years or older [13]. Prior literature indicates that younger adults tend to be more vaccine hesitant than members of older generations [10,12]. Few papers before the COVID-19 pandemic examined the role of information sources in vaccine uptake, perhaps because the issues were less politicized, people trusted information about vaccines from the health care and government sectors, and vaccine rates were higher overall [14]. With the COVID-19 pandemic, misinformation about vaccines became more prevalent and the question of vaccine uptake became polarized with political affiliation [14]. In one paper, social media use, often associated with younger age, is associated with higher odds of COVID-19 vaccine hesitancy [12]. There is also evidence that identified sources of information including government, science, and discussing vaccination with family and friends positively influenced COVID-19 vaccine intention [10]. However, to our knowledge, no studies have explicitly examined COVID-19 vaccine uptake by generation and information source.

Generational membership including Generation Z (Gen Z; born 1995 - 2012), millennials (born 1980 - 1994), Generation X (Gen X; born 1965 - 1979), baby boomers (born 1946 - 1964), and the Silent Generation (born 1925 - 1945) is associated with varying historical, technological, social, economic, and political circumstances, which may impact the amount and type of information sources consumed [15,16]. Media preferences and trust vary by age group with one nationally representative survey study finding that 9.5% of Gen Z participants trust their social media contacts mostly to completely (Facebook and Twitter, subsequently rebranded X) for COVID-19 information, compared to 3.8% of Silent Generation, 6.3% of millennials, 5.9% of Gen X, and 6% of baby boomer participants [15]. In the same study, 44.6% of Gen Z participants agreed or strongly agreed with the statement "COVID-19 vaccines have many known harmful side effects," compared to 44.7% of millennials, 43.2% of Gen X, 31.8% of baby boomers, and 18.7% of the Silent Generation participants [15]. However, no studies to our knowledge have examined the potential effect modification of the association between media preferences and vaccine uptake by age or generational status.

Given this context of varying trusted information sources and vaccine decision making, the WHO and Centers for Disease Control and Prevention (CDC) have monitored public communication on digital and nondigital platforms related to COVID-19 to address misinformation and promote accurate public health information [17,18]. These include social media, online platforms, news, radio, and television [17]. The WHO described an excess of both accurate and inaccurate information as an "infodemic" requiring coordinated management and has since created a framework to identify and respond to misinformation [17]. As a result, the CDC formed an Insights Unit to track and analyze information using a social listening protocol, resulting in reports that guided public health messaging [18].

Gaps in the Literature

Taken together, previous literature includes investigations of the influence of different information sources on vaccination decisions across various demographics, yet there remain gaps in understanding these dynamics across the range of information sources at the population level. Studies are also needed to focus on the number of information sources as well as the type [19] and to consider differences in these associations by generation.

Goal of This Study

To address these gaps, the current study leverages CHIS, a population-based California survey, to examine associations between the type and number of the most relied upon information sources and COVID-19 vaccination status. Most relied upon sources were a specific term created by CHIS referring to the highest ranked source. We hypothesized that the use of specific types of information sources such as "your doctor" is associated with higher rates of vaccination when compared to traditional news media (eg, television, radio, and newspaper). We also tested the effect modification of the relationship between information sources (type and number) and vaccination status by generational membership, hypothesizing that the impact of traditional news media sources on vaccination status will be greater among older adults and the impact of digital media sources (eg, social media) on vaccination status will be greater among younger adults. Additionally, we hypothesized that identifying social media as the most relied upon information source would be associated with lower odds of vaccination. The study findings can improve our understanding of media preferences as influences on health

XSL•FO
RenderX

decision behaviors and vaccine hesitancy and inform public health messaging efforts tailored to generational differences.

Methods

We conducted a secondary analysis of cross-sectional data from the 2022 CHIS.

Data Source

The CHIS is a population-based survey of noninstitutionalized Californians collected by Social Science Research Solutions, a research company, for the University of California, Los Angeles Center for Health Policy Research [20]. The 2021 - 2022 CHIS data were collected between March 18, 2021, and November 30, 2022, using a mixed-mode method surveying (phone and internet) of the representative population of Californians and geographically stratified address-based sampling. Of the adult interviews conducted, 89.5% were completed online, and 10.5% were over the telephone. CHIS strategically over-samples to account for households abstaining from participation and added a prepaid cell phone sample to reach individuals left out of general address-based sampling frameworks. Interviews were held in the following languages: Spanish, English, Chinese (Cantonese and Mandarin), Korean, Tagalog, and Vietnamese.

Sample

The study sample included all adult Californians responding to the 2021 - 2022 CHIS survey (N=21,463) [20]. No additional exclusion criteria were applied and there were no missing data.

Ethical Considerations

The University of California Davis Institutional Review Board reviewed this study and determined it was not human participants research and therefore was exempt from further review based on the use of publicly available data without identifying information (IRBNet ID 2233309). As a secondary analysis of existing data, informed consent was not obtained by the study investigators. Information about informed consent of CHIS participants is available online [20].

Variables and Measurement

Dependent Variable

Vaccine uptake, operationalized as self-reported vaccination status, is the dependent variable for both aims. Participants were asked, “Have you been fully vaccinated, partially vaccinated, or are you not vaccinated, for COVID-19?” Being fully vaccinated was defined in the CHIS as receiving 2 shots of the Pfizer or Moderna vaccine, 1 shot of the Johnson & Johnson vaccine, or 2 shots of the AstraZeneca or Sinovac vaccine (which are available in other countries, not in the United States). Partially vaccinated, not explicitly defined by the survey documentation, was interpreted to mean a vaccination series that had been initiated but not completed. Vaccinations were still being distributed to priority groups, through November 2022; therefore, to fully capture intent to vaccinate in our cohort, we categorized any level of vaccination (partial or full) as “Vaccinated” (yes=1) and “Not Vaccinated” as the reference group (no=0).

Independent Variables

Participants were asked about their most relied upon sources of COVID-19 information, categorized into 12 possible categories: newspapers, radio, television, family, friends, community leaders, religious leaders, social media, doctors, governmental agencies, places of employment, and none of the above. For the examination of the type of information source, the response options were recoded into 6 categories: traditional news media (television, radio, and newspapers), word of mouth (family, friends, religious leaders, and community leaders), social media, your doctor, your employer, or governmental agency, or none of these. The reference group was reporting traditional news media as the most reliable source. For the number of sources, the independent variable is a count of the original 12 information sources with a range of 0 to 11.

Covariates

Sociodemographic variables were included in the models as potential confounders. This included age categorized by generational membership, which was also examined as an effect modifier: Gen Z (reference group; aged 12 - 25 y), millennials (aged 26 - 39 y), Gen X (aged 40 - 54 y), baby boomers (aged 55 - 74 y), and Silent Generation (aged ≥75 y) [16]. The University of Southern California generational age ranges generally matched the CHIS age categories with some exceptions around cut-off years, for instance, the millennial category featuring young members from Gen X.

Following Joshi et al’s [21] conceptual framework on the factors that influence vaccine acceptance and hesitancy, the covariates were selected as relevant to this study and as important social and structural determinants of COVID-19 vaccination uptake [22]. They include race or ethnicity, urbanicity, self-reported sex, educational attainment, marital status, insurance status, and income. Reported race and ethnicities include White non-Hispanic (reference group), Hispanic, African American or non-Hispanic, American Indian or Alaskan Native, Asian, and Other or Two or More Races. Urbanicity was categorized as urban (reference group), mixed urbanicity, suburban, and rural. Self-reported sex was categorized as male (reference group) or female. Educational attainment was categorized as having no high school diploma (reference group), a high school diploma, some college degree, bachelor’s degree, master’s degree, or PhD. Marital status was categorized as married or living with a partner (reference group), widowed, divorced, separated, or never married. Insurance status was categorized as yes (reference group). Income was recoded into four categories: (1) less than US $39,999 (reference group); (2) US $40,000 to $79,999; (3) US $80,000 to $149,999; and (4) US $150,000 to over US $180,000.

Statistical Analysis

All analyses were conducted with Stata 18 (StataCorp LLC) and CHIS survey weights were applied to account for the complex survey sampling design. We used descriptive statistics to summarize all study variables and to compare vaccination status by generation. Multivariable logistic regression was used to model information sources first as categories, then as counts, as functions of vaccination status in models adjusted for all

XSL•FO
RenderX

covariates. We also tested the effect modification of these associations by generational membership by adding information source by generation interaction terms to the models. If any of the source by generation interaction terms were significant, we conducted and reported stratified models by generation. Statistical significance was set at $P<.05$.

Results

Sample Description

Before applying survey weights, most participants were baby boomers aged 55 to 74 years (n=8905, 41.49%), followed by Gen X aged 40 to 54 years (n=5328, 24.82%), and millennials aged 26 to 39 years (n=3598, 16.76%; Table 1). Most participants were female (n=12,238, 57.02%). Over half held a master's degree or PhD (54.5%, n=11,699) and 33.2% (n=7116) had some college, vocational education, or an associate's degree. Almost half of the participants identified as White, non-Hispanic (n=10,432, 48.6%), followed by Hispanic (n=5719, 26.7%), Asian (n=3234, 15.1%), African American (n=1072, 5%), American Indian or Alaska Native (n=129, 0.6%), and two or more races (n=877, 4.1%). Less than half lived in urban areas (n=8956, 41.7%) and 31.9% (n=6855) lived in suburban areas. More than half of the participants were married (n=12,104, 56.4%) and 33.1% (n=7107) had an annual income greater than or equal to US $150,000. Most participants had health insurance (n=20,666, 96.3%). Just over half relied on traditional news media (ie, television, radio, and newspaper) as their primary source of COVID-19 information (n=11,356, 52.9%).

XSL•FO
RenderX

Table . Sociodemographic characteristics and COVID-19 information sources among Californian adults, 2021 - 2022 (N=21,463).

Characteristics	Value, n (%)
Age by generation (y)	
Generation Z (18-25)	1106 (5.15)
Millennials (26-39)	3598 (16.76)
Generation X (40-54)	5328 (24.82)
Baby boomers (55-74)	8905 (41.49)
Silent Generation (≥75)	2526 (11.77)
Sex	
Male	9225 (42.98)
Female	12,238 (57.02)
Education	
No high school diploma	558 (2.60)
High school diploma	553 (2.58)
Some college (vocational or AS or AA included)	7116 (33.15)
BA or BS degree	1537 (7.16)
MA or MS or PhD	11,699 (54.51)
Race or ethnicity	
African-American, non-Hispanic	1072 (4.99)
American Indian or Alaska Native, non-Hispanic	129 (0.6)
Asian, non-Hispanic	3234 (15.07)
Hispanic	5719 (26.65)
White, non-Hispanic	10,432 (48.60)
Two or more races	877 (4.09)
Urbanicity or rurality	
Urban	8956 (41.73)
Mixed	2038 (9.50)
Suburban	6855 (31.94)
Rural	3614 (16.84)
Marital status	
Never married	4288 (19.98)
Married or living with partner	12,104 (56.39)
Widowed or separated or divorced	5071 (23.63)
Household income (annual, US $)	
<39,999	5556 (25.89)
40,000 - 79,999	5055 (23.55)
80,000 - 149,999	3745 (17.45)
>150,000	7107 (33.11)
Has health insurance	20,666 (96.29)
Most relied upon source of COVID-19 information	
Traditional news media (television, radio, and newspaper)	11,356 (52.91)
Word of mouth (family, friends, community, and religious leaders)	1171 (5.46)
Social media	2959 (13.79)
Your doctor	1013 (4.72)

XSL•FO
RenderX

Characteristics	Value, n (%)
Governmental agency or employer	3733 (17.39)
None of these	1231 (5.74)

Most participants were vaccinated for COVID-19 (n=19,836, 92.4%), and the vaccination rates increased slightly with age (from 88.97% among Gen Z to 96% in the Silent Generation). Among survey respondents who were vaccinated, the most relied upon information sources were their employer or government (95%) and traditional news media (94.6%). Vaccination rate was lowest (n=936, 76%) among those reporting no sources of information, whereas the highest vaccination rates (93% - 95%) were found for those reporting between 2 and 7 resources (Table 2). Among survey respondents who did not use any information sources, 76% (n=936) were vaccinated for COVID-19, whereas this rate was 80.6% (n=29) among those who used 11 resources.

Table . COVID-19 vaccination status by source count (N=21,463).

Source count	Total (N=21,463)	Vaccinated (n=19,836), n (%)
0	1231	936 (76.04)
1	5085	4665 (91.74)
2	4630	4334 (93.61)
3	4096	3890 (94.97)
4	2726	2562 (93.98)
5	1720	1618 (94.07)
6	941	880 (93.52)
7	554	517 (93.32)
8	266	243 (91.35)
9	132	122 (92.42)
10	46	41 (89.13)
11	36	29 (80.56)

Regression of Vaccination Status on Information Source Type

In the fully adjusted logistic regression model of information source type and vaccination status (Table 3), compared to respondents using traditional news media for information about COVID-19, those relying on word of mouth had 40% lower odds (P<.002); those using social media had 38% lower odds (P<.001); those relying on their doctors had 59% lower odds (P<.001); and those reporting “none of these” had 79% lower odds of being vaccinated for COVID-19 (P<.001). There were no significant differences between using employers or governmental agencies and receiving a COVID-19 vaccination compared to traditional news media.

XSL•FO
RenderX

Table . Multivariable logistic regression models of vaccination status[a].

	Information sources and vaccine status			Count of information sources and vaccine status		
	OR[b]	*P* value	95% CI	OR	*P* value	95% CI
Source count	__[c]	—	—	1.09	<.001	1.04 - 1.15
Relied upon information sources						
Traditional news media	—	—	—	—	—	—
Word of mouth	0.60	.002	0.44 - 0.82	—	—	—
Social media	0.62	<.001	0.51 - 0.77	—	—	—
Your doctor	0.41	<.001	0.29 - 0.57	—	—	—
Governmental agency or employer	1.05	.72	0.79 - 1.39	—	—	—
None of these	0.21	<.001	0.16 - 0.28	—	—	—
Age by generation (y)						
Generation Z (18-25)	—	—	—	—	—	—
Millennials (26-39)	1.01	.95	0.75 - 1.35	1.06	.70	0.78 - 1.44
Generation X (40-54)	1.34	.09	0.95 - 1.88	1.45	.04	1.03 - 2.05
Baby boomers (55-74)	2.72	<.001	1.93 - 3.83	3.20	<.001	2.27 - 4.50
Silent Generation (≥75)	3.62	<.001	2.30 - 5.71	4.35	<.001	2.76 - 6.86
Annual household income (US $)						
<39,999	—	—	—	—	—	—
40,000 - 79,000	1.36	.009	1.08 - 1.72	1.39	.006	1.10 - 1.75
80,000 - 149,000	1.69	<.001	1.29 - 2.20	1.67	<.001	1.29 - 2.18
150,000->180,000	2.37	<.001	1.90 - 2.95	2.36	<.001	1.91 - 2.92
Has health insurance	0.75	.08	0.55 - 1.04	0.75	.07	0.56 - 1.02
Race or ethnicity						
African-American, NH[d]	0.60	.004	0.43 - 0.85	0.60	.002	0.43 - 0.83
American Indian or Alaska Native, NH	0.76	.57	0.30 - 1.93	0.68	.41	0.28 - 1.70
Asian, NH	4.25	<.001	2.84 - 6.36	4.27	<.001	2.85 - 6.40
Hispanic	1.28	.008	1.07 - 1.54	1.31	.004	1.09 - 1.56
White, NH	—	—	—	—	—	—
Two or more races	1.66	.004	1.18 - 2.35	1.67	.005	1.18 - 2.38
Marital status						
Married or living with partner	—	—	—	—	—	—
Widowed or separated or divorced	0.98	.84	0.80 - 1.20	0.99	.94	0.80 - 1.22
Never married	1.35	.02	1.04 - 1.76	1.31	.04	1.01 - 1.69

XSL•FO
RenderX

	Information sources and vaccine status			Count of information sources and vaccine status		
	OR[b]	*P* value	95% CI	OR	*P* value	95% CI
Urbanicity						
Urban	—	—	—	—	—	—
Mixed	0.56	<.001	0.42 - 0.75	0.57	<.001	0.43 - 0.76
Suburban	0.84	.06	0.69 - 1.01	0.83	.049	0.69 - 0.99
Rural	0.45	<.001	0.36 - 0.57	0.44	<.001	0.35 - 0.55
Self-reported sex	1.14	.10	0.98 - 1.32	1.19	.04	1.01 - 1.39
Educational level						
No HS[e] diploma	—	—	—	—	—	—
HS diploma	0.46	.003	0.29 - 0.76	0.47	.003	0.29 - 0.77
Some college (vocational or AS or AA included)	0.66	.07	0.43 - 1.03	0.64	.045	0.41 - 0.99
BA or BS degree	0.82	.48	0.48 - 1.42	0.79	.38	0.47 - 1.33
MA or MS or PhD	1.70	.04	1.03 - 2.78	1.66	.045	1.01 - 2.72

[a]All estimates are survey weighted and generalized to a sample of N=29,560,694.

[b]OR: odds ratio.

[c]Not applicable.

[d]NH: non-Hispanic.

[e]HS: high school.

Regression of Vaccination Status on Information Source Count

In the fully adjusted model examining the count of relied upon information sources and vaccination status, each additional source was associated with 9% higher odds of being vaccinated for COVID-19 (*P*<.001; Table 3). In this model, the odds of being vaccinated were higher among Gen X, baby boomers, and Silent Generation when compared to Gen Z. Covariates in this model associated with higher odds of vaccination included an annual income of over US $40,000 (compared to <US $40,000), self-reporting race or ethnicity as Hispanic, Asian, or two or more races (compared to White non-Hispanic), never being married (compared to married or living with a partner), self-reporting male, and attaining postgraduate education (compared to no high school diploma). Covariates associated with lower odds of vaccination in the source count model included self-reporting race or ethnicity as African American (compared to White non-Hispanic), living in a rural or suburban environment (compared to urban), and attaining a high school diploma or completing some college (compared to no high school diploma).

Effect Modification

We found significant effect modification of the associations of type of information source and vaccination status by generation. Accordingly, we conducted full models stratified by generation (Table 4). For each generation, 2 models were run, one with information source type and one with information source count as the primary independent variable.

XSL•FO
RenderX

Table . Multivariable regression models of vaccination status stratified by generation[a].

Information sources	Generation Z (18-25 y; n=1106)		Millennial (26-39 y; n=3598)		Generation X (40-54 y; n=5328)		Baby boomers (55-74 y; n=8905)		Silent Generation (≥75 y; n=2526)	
	Odds ratio (95% CI)	*P* value	Odds ratio (95% CI)	*P* value	Odds ratio (95% CI)	*P* value	Odds ratio (95% CI)	*P* value	Odds ratio (95% CI)	*P* value
Traditional news media	—[b]	—	—	—	—	—	—	—	—	—
Word of mouth	0.76 (0.41 - 1.42)	.39	0.82 (0.36 - 1.86)	.64	0.41 (0.23 - 0.74)	.004	0.77 (0.44 - 1.36)	.36	1.92 (0.26 - 14.42)	.52
Social media	1.44 (0.84 - 2.46)	.18	0.69 (0.45 - 1.06)	.09	0.63 (0.40 - 0.98)	.04	0.18 (0.11 - 0.30)	<.001	0.16 (0.05 - 0.51)	.003
Your doctor	0.30 (0.07 - 1.30)	.11	1.00 (0.38 - 2.63)	.99	0.29 (0.17 - 0.50)	<.001	0.47 (0.33 - 0.67)	<.001	0.56 (0.10 - 2.97)	.49
Your employer or government agency	2.72 (0.93 - 7.96)	.07	1.02 (0.62 - 1.68)	.94	1.07 (0.65 - 1.73)	.80	0.90 (0.60 - 1.36)	.61	1.34 (0.58 - 3.10)	.49
None of these	0.39 (0.18 - 0.84)	.02	0.36 (0.16 - 0.74)	.005	0.24 (0.13 - 0.46)	<.001	0.09 (0.06 - 0.13)	<.001	0.10 (0.04 - 0.24)	<.001
Source count	1.18 (1.003 - 1.38)	.045	1.02 (0.92 - 1.13)	0.65	1.13 (1.03 - 1.24)	.01	1.11 (1.001 - 1.24)	.048	1.28 (0.99 - 1.63)	.051

[a]All estimates are survey weighted and generalized to a sample of N=29,560,694.

[b]Not applicable.

In the sample restricted to Gen Z, the only information source type associated with vaccination status was "none of these," which was associated with lower odds of being vaccinated compared to traditional news media (odds ratio [OR] 0.39, 95% CI 0.18 - 0.84; Table 4). Source count was significantly associated with higher odds of vaccination uptake (OR 1.18, 95% CI 1.003 - 1.38; Table 4). In the sample restricted to millennials, only "none of these" information sources (compared to traditional news media) was associated with lower odds of vaccination (OR 0.36, 95% CI 0.16 - 0.74). Source count was not significantly associated with vaccination uptake for millennials. In the sample restricted to Gen X, word of mouth (OR 0.41, 95% CI 0.23 - 0.74), social media (OR 0.63, 95% CI 0.40 - 0.98), your doctor (OR 0.29, 95% CI 0.17 - 0.50), and "none of these" (OR 0.24, 95% CI 0.13 - 0.46) were significantly associated with lower odds of vaccination compared to traditional news media. Source count was significantly associated with higher odds of vaccination for Gen X (OR 1.13, 95% CI 1.03 - 1.24).

In the sample restricted to baby boomers, social media (OR 0.18, 95% CI 0.11 - 0.30), your doctor (OR 0.47, 95% CI 0.33 - 0.67), and "none of these" (OR 0.09, 95% CI 0.06 - 0.13) were significantly associated with lower odds of vaccination compared to traditional news media. Source count was significantly associated with higher odds of vaccination for baby boomers (OR 1.11, 95% CI 1.001 - 1.24). In the sample restricted to the Silent Generation, social media (OR 0.16, 95% CI 0.05 - 0.51) and "none of these" (OR 0.10, 95% CI 0.04 - 0.24) were associated with lower odds of vaccination when compared to traditional news media. Source count approached but did not reach statistical significance for the Silent Generation.

Discussion

Principal Findings

In this large population-based sample of adult Californians, we find that both the identified most-relied-upon information source and the count of identified sources are robust predictors of the COVID-19 vaccination status. This supports previous evidence that information sources are associated with vaccination status [1,5]. This study contributes new knowledge to the associations between vaccine status, number of sources used, source type, and examines variation in these associations by generational membership. Thus, these are important factors to consider in the future during research and public health program development when the need arises.

In line with previous literature that found social media to be associated with negative views towards COVID-19 vaccination receipt, we found that use of social media was associated with lower odds of vaccination status compared to traditional news media and controlling for covariates. When stratified by

XSL•FO
RenderX

generation, this association held among Gen X, baby boomers, and the Silent Generation. This divergence may be explained by the personalization of social media feeds and the potential for algorithmic information silos. Similar to our findings, Moon et al [12] found higher social media use was a significant predictor of COVID-19 vaccine hesitancy. In this study, this was true across older generations such as Gen X, baby boomers, and the Silent Generation. Future research may be warranted to explore how targeted advertising or recommended pages via an algorithm may be able to amplify vaccination information to combat the spread of misinformation.

In this analysis, we relied upon traditional news media as the reference category. We found in all generational groups, with the exception of Gen Z and millennials, that relying on social media for vaccine information compared to traditional news media was associated with lower odds of vaccination uptake. Among Gen Z and millennials, relying on social media compared to traditional news media was not associated with vaccination uptake. This means that although these generational groups tend to use social media more than others, its influence is not different than that of traditional news media. These groups may view traditional news media through social media platforms, thereby minimizing the difference in the influence of these information sources. While other researchers have identified relationships between social media use and vaccine hesitancy [12], they were unable to detail what type of content was being consumed through social media. Social media may include short-form clips of traditional news media information. These inconsistencies warrant further research to better understand how different generations consume social media and how this, in turn, may impact vaccination uptake.

Our findings consistently showed that respondents reporting “none of these” information sources had lower odds of being vaccinated across all generations. The deeper meaning behind this response is unclear, though it could represent those who abstain from media or do not interact with the listed sources; this may represent a gap in the listed responses to the survey question. For instance, alternate popular sources such as podcasts were not among the multiple-choice answers provided by CHIS. It is also possible that those who did not fully understand the question or preferred not to answer selected this option.

While reporting “your doctor” as the most relied upon information source was associated with lower odds of vaccination, this was in comparison to traditional news media. At the time of this survey, information surrounding COVID-19 and vaccinations was communicated heavily through various traditional and nontraditional media sources. This finding does not indicate that doctors were not perceived as reliable sources of information for COVID-19, but rather that traditional news media sources likely played a more central role in distributing COVID-19 vaccination information and related public health messaging.

Consistent with past studies, we found that relying on a higher number of information sources was associated with higher COVID-19 vaccination rates, possibly due to accumulating greater knowledge about vaccine-related information [7,8]. Our results further identified a dose-response relationship with a greater quantity of relied-upon information sources being associated with increasing odds of being vaccinated for COVID-19. These results suggest that people who engage in more health information-seeking behaviors may be more successful in garnering accurate and updated public health advice. It is also plausible that individuals seeking multiple sources of information have a greater propensity to be vaccinated due to other personal characteristics allied with health promotion.

It should be mentioned that the vaccination rates across each generation in this sample surpassed both state and national averages [23]. This study found that Gen Z had the lowest COVID-19 vaccination rates among all generations, while the Silent Generation had the highest. In comparison, California data shows that individuals aged 40 to 69 years exhibited both higher hesitancy and higher likelihood of vaccination compared to other age groups [12]. The concept of generational imprinting can suggest that the generation that grew up during poliovirus vaccination development may view COVID-19 vaccinations in a positive light and therefore may desire to receive the vaccination as soon as they can [15]. The high rate of Silent Generation vaccine receipt in our sample may be explained by the timing of the survey, during which vaccine rollout was made available to older adults and health care workers first.

These findings on COVID-19 vaccination patterns among Californians have important public health implications about the role of information sources. They illuminate the need to tailor messaging by generation based on preferred information sources. This will support accurate and trustworthy information reaching all audiences, as the sources of information used are linked to knowledge, vaccine intention, and vaccination rates. Accessing a variety of sources may contribute to shaping knowledge and behaviors regarding vaccination. Reliance on diverse information sources could combat algorithmic information silos, with those who access multiple, reliable sources being more likely to be informed about the benefits and risks of vaccines and thereby influencing their vaccination decisions. This information contributes to our understanding of the information sources used by age generations in the context of the COVID-19 pandemic and may inform strategies for promoting health equity and widespread immunization in similar, future health contexts.

Limitations

This study had several limitations. First, due to the cross-sectional study design, we cannot establish temporal precedence or causation. Second, the data were collected during the COVID-19 pandemic, and we cannot conclude that the findings would hold before or after the pandemic. Third, because the CHIS data were collected in 2021 and 2022, there was not equal access to vaccination for all participants as vaccines were just becoming available to the public. Accordingly, we combined those who were partially vaccinated with those who were fully vaccinated into one “vaccinated” category. Partial vaccination does not necessarily indicate an intention to become fully vaccinated. We have no assurance that partially vaccinated individuals completed the vaccination series, and if they did

XSL•FO
RenderX

not, we do not know whether vaccine hesitancy was a contributing factor. Additionally, COVID-19 vaccination status is time-dependent, as vaccine availability differed across age groups and months in 2021 and 2022 when the data were collected. The CHIS dataset does not include the interview month, resulting in the inability to adjust for survey timing. Combining these categories may introduce misclassification bias and may inflate the rates of vaccination in the study and underestimate underlying vaccine hesitancy. Also, we are unable to identify the amount, quality, or precise source of information in any of the relied upon categories. For instance, social media platforms include a variety of information sources such as clinician influencers, friends and family, public health, and traditional news.

Fourth, we can only speculate about the information source choice "none of these" and what this meant to respondents. Our finding that this choice was consistently associated with lower odds of vaccination warrants follow-up in future studies, perhaps with cognitive interviews or other qualitative approaches designed to more fully understand this selection. This information would be helpful in informing public health guidance and messaging, building trust within community-academic partnerships, and efforts to improve health literacy in the community.

Fifth, because this is a secondary data analysis, our findings are subject to unmeasured confounding. Greater information about vaccination intentions could have further elucidated the findings given that these data were collected in the early phases of vaccine rollout. In addition, the CHIS categorical age ranges fit the generational age cutoffs closely with a few exceptions [16]. For instance, millennials are defined as those born between 1980 and 1994, corresponding to those aged 28 to 42 years in 2022. However, the CHIS survey had categorical age ranges that did not perfectly fit (26-29 y, 30-34 y, and 35-39 y) these generational groupings. As a result, our generational membership included or missed some members with ages close to the cut points, for example, millennials included the young members of Gen X. In addition, there may be limitations to generalizability as a result of the "healthy user effect" or positive selection bias [24]. Our sample had higher than average vaccination rates, which may relate to participants being more likely to participate in a study about health [23,24]. As a result, our sample may have different characteristics from the population of interest, as the vaccination rates in our sample were higher than the California state average [23]. Finally, because the CHIS survey is a population-based sample of California residents, findings may not be generalizable beyond California.

Today's political climate has seen a rise in antivaccine attitudes [25] alongside the perpetuation of false narratives linking vaccines to autism in children [26]. These events may risk delaying the development of new vaccines and undermining public confidence in vaccine experts. Within this context, a deeper understanding of how information sources influence vaccine decision-making may be more pertinent than ever.

Conclusions

Our findings suggest that an effective approach to disseminating accurate public health information should include multiple methods to reach all generational preferences of the public. The findings from this study on COVID-19 vaccination patterns in California provide important contributions for stakeholders interested in vaccine hesitancy, health informatics, and health literacy. They identify the importance of understanding generational differences in how people rely on media during a pandemic. These findings may also not be limited to vaccination use, and further research could expand to explore other preventative health care behaviors such as screenings, or other public health aspects that are discussed by media sources.

Acknowledgments

The authors would like to thank the anonymous reviewers for their insightful feedback and the University of California Davis Betty Irene Moore School of Nursing for its continued support.

Funding

JFZ is supported by the National Institutes of Health under grant T32 5T32NR021294-02. JFZ, MS, and GRF receive funding through scholarships at the University of California Davis Betty Irene Moore School of Nursing to support their doctoral studies. No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Data Availability

The datasets generated or analyzed during this study are publicly available in the California Health Interview Survey (CHIS) repository and can be accessed by creating a free online account [20].

Conflicts of Interest

None declared.

References

1. Charron J, Gautier A, Jestin C. Influence of information sources on vaccine hesitancy and practices. Med Mal Infect 2020 Nov;50(8):727-733. [doi: 10.1016/j.medmal.2020.01.010] [Medline: 32067795]

XSL•FO
RenderX

2. Kouzy R, Abi Jaoude J, Kraitem A, et al. Coronavirus goes viral: quantifying the COVID-19 misinformation epidemic on Twitter. Cureus 2020 Mar 13;12(3):e7255. [doi: 10.7759/cureus.7255] [Medline: 32292669]
3. Mian A, Khan S. Coronavirus: the spread of misinformation. BMC Med 2020 Mar 18;18(1):89. [doi: 10.1186/s12916-020-01556-3] [Medline: 32188445]
4. Ali SH, Foreman J, Tozan Y, Capasso A, Jones AM, DiClemente RJ. Trends and predictors of COVID-19 information sources and their relationship with knowledge and beliefs related to the pandemic: nationwide cross-sectional study. JMIR Public Health Surveill 2020 Oct 8;6(4):e21071. [doi: 10.2196/21071] [Medline: 32936775]
5. Tabacchi G, Costantino C, Cracchiolo M, et al. Information sources and knowledge on vaccination in a population from southern Italy: the ESCULAPIO project. Hum Vaccin Immunother 2017 Feb;13(2):339-345. [doi: 10.1080/21645515.2017.1264733] [Medline: 28032814]
6. Kulkarni S, Sengeh P, Eboh V, et al. Role of information sources in vaccination uptake: insights from a cross-sectional household survey in Sierra Leone, 2019. Glob Health Sci Pract 2022 Feb 28;10(1):e2100237. [doi: 10.9745/GHSP-D-21-00237] [Medline: 35294376]
7. Almeida CM, Tiro JA, Rodriguez MA, Diamant AL. Evaluating associations between sources of information, knowledge of the human papillomavirus, and human papillomavirus vaccine uptake for adult women in California. Vaccine (Auckl) 2012 Apr 19;30(19):3003-3008. [doi: 10.1016/j.vaccine.2012.01.079] [Medline: 22326776]
8. Kim Y, Kim J, Li Y. Association of sources of COVID-19 information with vaccine uptake, preventive behaviors, and perceived severity of COVID-19. J Am Med Dir Assoc 2023 Aug;24(8):1143-1150. [doi: 10.1016/j.jamda.2023.05.024] [Medline: 37400059]
9. Newcomer SR, Graham J, Irish K, et al. Identification of spatial clusters of undervaccination patterns among children aged <24 months using Immunization Information System Data, Montana, 2015-2019. Public Health Rep 2024;139(3):360-368. [doi: 10.1177/00333549231186603] [Medline: 37503702]
10. Wagner A, Juvalta S, Speranza C, Suggs LS, Drava J, COVIDisc study group. Let's talk about COVID-19 vaccination: relevance of conversations about COVID-19 vaccination and information sources on vaccination intention in Switzerland. Vaccine 2023 Aug 14;41(36):5313-5321. [doi: 10.1016/j.vaccine.2023.07.004] [Medline: 37455160]
11. Ten threats to global health in 2019. World Health Organization. 2019. URL: https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019 [accessed 2026-01-30]
12. Moon I, Han J, Kim K. Determinants of COVID-19 vaccine hesitancy: 2020 California Health Interview Survey. Prev Med Rep 2023 Jun;33:102200. [doi: 10.1016/j.pmedr.2023.102200] [Medline: 37034299]
13. Diesel J, Sterrett N, Dasgupta S, et al. COVID-19 vaccination coverage among adults—United States, December 14, 2020-May 22, 2021. MMWR Morb Mortal Wkly Rep 2021 Jun 25;70(25):922-927. [doi: 10.15585/mmwr.mm7025e1] [Medline: 34166331]
14. Bolsen T, Palm R. Politicization and COVID-19 vaccine resistance in the U.S. Prog Mol Biol Transl Sci 2022;188(1):81-100. [doi: 10.1016/bs.pmbts.2021.10.002] [Medline: 35168748]
15. Wang VHC, Silver D, Pagán JA. Generational differences in beliefs about COVID-19 vaccines. Prev Med 2022 Apr;157:107005. [doi: 10.1016/j.ypmed.2022.107005] [Medline: 35235852]
16. Research guides. Demographics. Age groups. University of Southern California Libraries. URL: https://libguides.usc.edu/busdem/age [accessed 2026-01-30]
17. Rubinelli S, Purnat TD, Wilhelm E, et al. WHO competency framework for health authorities and institutions to manage infodemics: its development and features. Hum Resour Health 2022 May 7;20(1):35. [doi: 10.1186/s12960-022-00733-0] [Medline: 35525924]
18. Kolis J, Brookmeyer K, Chuvileva Y, et al. Infodemics and vaccine confidence: protocol for social listening and insight generation to inform action. JMIR Public Health Surveill 2024 Oct 24;10:e51909. [doi: 10.2196/51909] [Medline: 39447166]
19. Venzon A. Public perceptions of COVID-19 vaccine information: quantitative analyses of survey and Twitter data in the United States [Dissertation]. : University of California; 2022 URL: https://escholarship.org/uc/item/7c93j4fx#author [accessed 2026-01-30]
20. California health interview survey (CHIS): public use files. UCLA Center for Health Policy Research. URL: https://healthpolicy.ucla.edu/our-work/public-use-files [accessed 2026-01-30]
21. Joshi A, Kaur M, Kaur R, Grover A, Nash D, El-Mohandes A. Predictors of COVID-19 vaccine acceptance, intention, and hesitancy: a scoping review. Front Public Health 2021;9:698111. [doi: 10.3389/fpubh.2021.698111] [Medline: 34485229]
22. Vardavas C, Nikitara K, Aslanoglou K, et al. Social determinants of health and vaccine uptake during the first wave of the COVID-19 pandemic: a systematic review. Prev Med Rep 2023 Oct;35:102319. [doi: 10.1016/j.pmedr.2023.102319] [Medline: 37564118]
23. Gottfried J, Park E. Americans' social media use 2025. : Pew Research Center; 2025 URL: https://www.pewresearch.org/wp-content/uploads/sites/20/2025/11/PI_2025.11.20_Social-Media-Use_REPORT.pdf [accessed 2026-01-30]
24. COVID-19 vaccination coverage and vaccine confidence among adults. Centers for Disease Control and Prevention, COVIDVaxView. 2024. URL: https://www.cdc.gov/covidvaxview/interactive/adults.html [accessed 2026-01-30]
25. Brassey J, Mahtani KR, Spencer EA, Heneghan C. Volunteer bias. Catalog of Bias. 2017. URL: http://www.catalogofbias.org/biases/volunteer-bias [accessed 2026-01-30]

XSL•FO
RenderX

26. Ortiz-Prado E, Suárez-Sangucho IA, Vasconez-Gonzalez J, et al. Pandemic paradox: how the COVID-19 crisis transformed vaccine hesitancy into a two-edged sword. Hum Vaccin Immunother 2025 Dec;21(1):2543167. [doi: 10.1080/21645515.2025.2543167] [Medline: 40791124]

Abbreviations

CDC: Centers for Disease Control and Prevention
CHIS: California Health Interview Survey
OR: odds ratio
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 14.Oct.2025; peer-reviewed by J Wu, M Rohova; accepted 22.Dec.2025; published 17.Feb.2026.

Please cite as:
Zabala JF, Sablik M, Finical GR, Keeton VF, Bell JF
COVID-19 Information Sources and Vaccination Status Among Californian Adults by Generation Using the 2022 California Health Interview Survey: Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e85904
URL: https://publichealth.jmir.org/2026/1/e85904
doi:10.2196/85904

© Julia Forest Zabala, Melissa Sablik, Gina R Finical, Victoria F Keeton, Janice F Bell. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 17.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effectiveness of Inactivated COVID-19 Vaccination Against COVID-19–Related Hospitalization and Severe Outcomes in Adults ≥80 Years During Omicron Circulation in Beijing, China: Retrospective Cohort Study

Dan Zhao[1,2,3*], MD; Ying Ma[1,2,3*], MD; Juan Li[1,2,3], PhD; Xiaomei Li[1,2,3], BSc; Zhiqiang Cao[1,2,3], PhD; Wei Yao[1], BSc; Jiang Wu[1,2], MD; Luodan Suo[1,2,3], MD

[1]Beijing Key Laboratory of Surveillance, Early Warning and Pathogen Research on Emerging Infectious Diseases, Beijing Center for Disease Prevention and Control, 16 Hepingli Middle St, Dongcheng District, Beijing, China

[2]Beijing Research Center for Respiratory Infectious Diseases, Beijing, China

[3]School of Public Health, Capital Medical University, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Luodan Suo, MD
Beijing Key Laboratory of Surveillance, Early Warning and Pathogen Research on Emerging Infectious Diseases, Beijing Center for Disease Prevention and Control, 16 Hepingli Middle St, Dongcheng District, Beijing, China

Abstract

Background: A large wave of COVID-19 caused by SARS-CoV-2 Omicron subvariants began in Beijing in early December 2022.

Objective: This study aimed to evaluate the COVID-19 vaccine effectiveness (VE) in mitigating the risk of COVID-19–related hospitalization during the epidemic.

Methods: We conducted a retrospective cohort study linking regional health care data and vaccination registry routinely collected in Beijing. All electronic medical records on COVID-19–related hospital discharges of older inpatients aged ≥80 years during November 2022 and February 2023 were included. Poisson regressions were used to estimate incidence risk ratio of COVID-19–related hospitalization, severe or critical cases, and in-hospital death compared with unvaccinated groups, adjusting for gender and age. VE was calculated as 1 minus incidence risk ratio×100%.

Results: A total of 53,789 individuals aged ≥80 years were included, 28,423 (52.84%) were male, 45,270 (84.16%) were aged 80 - 89 years, and 8519 (15.84%) were aged ≥90 years. Overall, 30,531 (56.76%) were in the vaccine group, with 4524 (8.41%) of the total participants receiving partial vaccination, 20.91% completing the primary series, and 14,761 (27.44%) receiving one booster. Additionally, 23,258 (43.24%) were in the unvaccinated group. Of the 53,789 hospitalized individuals, 17,916 (33.31%) had a COVID-19 diagnosis, 3535 (6.57%) had COVID-19–related hospitalization, 961 (1.79%) were COVID-19 severe or critical cases, and 4130 (7.68%) had in-hospital death. The analysis revealed that the VE of booster vaccination in preventing COVID-19–related hospital, severe or critical COVID-19, and in-hospital death was 63.5% (95% CI 59.8% - 66.9%), 66.9% (95% CI 60.1% - 72.6%) and 79.4% (95% CI 77% - 81.5%), the VE of primary series 56% (95% CI 51.4% - 60.2%), 66.8% (95% CI 59% - 73%) and 66.4% (95% CI 63% - 69.5%).

Conclusions: The first booster vaccination was associated with significantly reduced the risk of COVID-19–related severe outcomes in older inpatients aged ≥80 years during the Omicron-dominant period. Considering the potential selection bias and unmeasured confounders, these estimates may reflect both the VE and the better baseline health status of the vaccinated older individuals.

(JMIR Public Health Surveill 2026;12:e82915) doi:10.2196/82915

KEYWORDS

older individuals; COVID-19; vaccine effectiveness; hospitalization; booster dose; real-world evidence

XSL•FO
RenderX

Introduction

COVID-19 vaccines have saved millions of lives worldwide by providing strong protection against serious illness, hospitalization, and death [1]. As SARS-CoV-2 evolves, changes in the receptor-binding domain of the spike protein are capable of evading antibodies induced by vaccination or infection; this waning immunity and reduced vaccine effectiveness (VE) have been observed [2,3]. Since 2022, SARS-CoV-2 Omicron and its sub-lineages (including BA.1, BA.2, BA.4, BA.5, XBB, BQ.1, and JN.1) have emerged and become the dominant circulating variants worldwide [4]. A serological study demonstrated that 3 to 5 doses of the ancestral monovalent vaccine elicited neutralizing antibodies with a narrow breadth against Omicron BA.5 and BA.2.75, while failing to neutralize BQ.1.1 and XBB sublineages [5]. Although severe COVID-19 and death are currently less prevalent in the Omicron-dominant era than in previous years [5], SARS-CoV-2 continues to cause significant morbidity and mortality, especially among the older population and those with existing medical conditions, including high blood pressure, diabetes, obesity, immunosuppression, cancer, and pregnancy [1]. The World Health Organization recommends that older adults as a priority group should be included in the long-term vaccine immunization strategy. However, as the continuous evolution of SARS-CoV-2 persists, it is probable that both the currently available and newly updated variant-specific vaccines will fall behind the viral antigenic evolution [6,7].

In China, the COVID-19 vaccine was approved for emergency use in July 2020, and a territory-wide massive vaccination program was launched, first in all adults aged 18 years and older at the end of March 2021, then expanding to adolescents aged 12 - 17 years in July 2021, and to children 3 - 11 years in October 2021. The coverage of the primary COVID-19 vaccine among the entire population, individuals aged ≥60 years, and those aged ≥80 years was estimated at 90.4%, 86.6%, and 66.4%, respectively [8]. Randomized clinical trials have shown that COVID-19 vaccines provided moderate protection against SARS-CoV-2 infection [9-14]. Post-market studies have shown moderate or high levels of protection against severe illness, hospitalization, and death in Chinese adults [15,16]. However, few data are available for those aged 80 years and older, especially regarding the risk of COVID-19–related hospitalization after the first booster dose during the Omicron period. Older people are at a higher risk of severe disease outcomes from COVID-19 infection, with higher rates of hospitalization and mortality than other ages, even after receiving the booster dose [17,18]. To guide the rational use of COVID-19 vaccines, it is crucial to evaluate and update their real-world effectiveness of COVID-19 vaccines, especially in older adults.

Beijing is the capital city of China, with a population of 21.8 million in 2024. According to the Beijing Immunization Information Management System data, inactivated vaccines constituted 99.55% (63.52 million doses) of Beijing's administered COVID-19 vaccination through November 2022. China, including Beijing, adhered to a "zero-COVID-19" policy from April 2020 to early December 2022, which interrupted the transmission of SARS-CoV-2 in varying-scale COVID-19 outbreaks caused by variants [19]. Since December 2022, the zero-COVID policy has been rapidly lifted nationwide, with a major wave of Omicron sublineages. In Beijing, the cumulative infection attack rate was estimated to be 75.7% [20,21], caused by the Omicron BF.7 sub-lineage primary [22]. The roll-out of the fourth vaccine dose commenced in December 2022. In this study, we assessed the real-world VE of COVID-19 vaccines against COVID-19–related hospitalization during the Omicron wave in a well-traced cohort of older adults aged ≥80 years in Beijing.

Methods

Study Population and Design

We conducted a retrospective cohort study in Beijing. This study included Beijing residents enrolled in the Beijing Hospital Utilization Information System (BHUIS), an electronic platform managed by the Beijing Municipal Health Commission's Information Center (BMHCIC). The system aggregates standardized inpatient demographic and discharge records from all tertiary and secondary hospitals, as well as more than 50% of primary care facilities across the city, capturing 94% of all hospital admissions in Beijing between 2022 - 2023. In this study, individuals recorded in BHUIS were eligible for inclusion if he or she was aged ≥80 years (with a birth date on or before November 1, 1942), residents of Beijing, and admitted to the hospital between November 1, 2022 and February 28, 2023 (the starting and ending date of the Omicron wave). We excluded individuals who met any one criteria: individuals who had a previous COVID-19 infection before November 1, 2022 recorded in the National Notifiable Infectious Disease Surveillance System (NNIDSS); those who had any COVID-19 vaccination recorded after November 1, 2022, to avoid misclassification of exposure status due to overlapping seroconversion periods and outcome observation windows; those who received four doses of COVID-19 vaccine, as this subgroup (n=424) was undersampled for robust statistical analysis and risked exposure misclassification given their distinct vaccination timeline (eg, fourth doses administered since December in 2022).

Data Source

We obtained complete electronic medical records of older adults aged ≥80 years who were admitted to the hospital between November 1, 2022 and February 28, 2023 from the BMHCIC. BMHCIC performed data queries and abstraction in BHUIS on December 30, 2023, to account for delays in reporting. All hospital discharges in the study cohort were downloaded and included in the study. The retrieved variables included patient name, age, identification card number, hospital case number, in-hospital deaths, discharge diagnosis (principal and secondary diagnosis), date of admission, and postal code of the residing address. To avoid duplicated hospitalizations, we retained records with the earliest admission date of the same individual who had multiple hospitalizations during a 21-day period. For records occurring 21 days apart, we considered it a new event of hospitalization for the same person.

XSL•FO
RenderX

We collected the vaccination history of individuals (ie, vaccine type, dosage, and date of injection) from the Beijing Center for Disease Control and Prevention, by linking hospitalization records with the Beijing Immunization Information Management System (IIS). The IIS is a web-based system akin to other computerized population-based databases previously documented in the literature [23], in which COVID-19 vaccine recipient information can be readily queried and retrieved using the vaccinee's identification card number as an identifier. A "double match" technique, using an individual's name and identification card number as identifiers, was used to link and acquire individual's COVID-19 vaccine exposure. In our study, vaccine information recorded in the IIS can be viewed as a complete record of the study cohort, as every dose of COVID-19 vaccine administered in the administrative region was required to be entered into the IIS by nurses through client application software.

Exposures

Based on the individual's COVID-19 vaccine history before the Omicron wave started, we categorized the study cohort into four groups, (1) unvaccinated (no history of COVID-19 vaccination before November 1, 2022), (2) partially vaccinated (≥14 d after receipt of the first vaccine dose and before completing the full primary series), (3) completed primary series (≥14 d after completing the full primary series and before receipt of the first booster dose), and (4) completed the first booster vaccination (≥7 d after receipt of the booster dose). According to China's COVID-19 Vaccines Immunization Schedule, the primary series of COVID-19 vaccine entailed two doses of inactivated vaccine, one dose of adenovirus vector vaccine, or three doses of recombination vaccine, while the booster vaccination involved an additional one dose of inactivated vaccine, adenovirus vector vaccine, or recombination vaccine after the recipient had completed the primary series.

Outcomes and Covariates

The primary study outcome was hospitalization with COVID-19–related disease, defined as hospital discharge diagnosed with code U07.1 in the *International Statistical Classification of Diseases and Related Health Problems 10th Revision, 10th Revision (ICD-10)*. All hospitalizations with COVID-19 had a positive nucleic acid amplification test. Other outcomes evaluated in the study included COVID-19–related hospitalization, severe or critical illness, and in-hospital death. Considering the high infection rate of SARS-CoV-2, only inpatients with COVID-19 *ICD-10* code in "principal diagnosis" were defined as COVID-19–related hospitalization. Severe or critical illness was assessed by doctors based on the criteria provided in the Diagnosis and Treatment Protocol for COVID-19 (Version 10) [24]. Severe illness was defined as any of the following: (1) respiratory distress (ie, respiration rate ≥30 breaths per min), (2) oxygen saturation ≤93 % at rest, (3) the ratio of arterial partial pressure of oxygen and fraction of inspired oxygen ≤300 mmHg, or (4) chest imaging showing obvious lesion progression (>50% increase within 24 - 48 h). Critical COVID-19 was defined as any of the following: (1) respiratory failure requiring mechanical ventilation, (2) shock, or (3) other organ failure requiring intensive care unit admission. In-hospital death with COVID-19 was defined as the death record of inpatients with COVID-19–related disease, the use of all-cause in-hospital mortality in patients diagnosed with COVID-19 cannot represent or estimate the specific COVID-19 mortality, due to the inability to definitively attribute death causation using available health care data. We controlled for several patient-level characteristics (ie, sex and age) by stratifying the analysis of the association between vaccination and outcome within these variables.

Statistics Analysis

For comparison of baseline characteristics, we used the chi-square test or Fisher exact test for categorical variables as appropriate and Bonferroni method for their pairwise comparisons, used Kruskal-Wallis test for continuous variables and Mann-Whitney *U* test for their pairwise comparisons. We calculated the risk of COVID-19 among hospitalized patients during the study period by dividing the total number of inpatients diagnosed with COVID-19–related diseases by the size of the study cohort. The 95% CI was estimated using the normal approximation of Poisson distribution. The association (ie, incidence risk ratio) between vaccination and COVID-19–related hospitalization, severe or critical illness, and death was evaluated by comparing the risk between vaccinated and unvaccinated groups using Poisson regression models with a robust variance estimator, adjusting for sex and age. The VE was estimated as ([1-incidence risk ratio]×100%) for partially vaccinated, primary series, and booster vaccination. We conducted a sensitivity analysis restricting the outcome to hospitalizations with severe or critical COVID-19 as the principal diagnosis. Data cleaning was performed in MySQL 5.7 because of its high-performance data management capabilities. All statistical analyses were conducted using SPSS 20.0 (IBM Corp). All *P* values were 2-sided and considered significant at $P<.05$. Considering that our data consisted of almost all inpatient cases during the specified time, and 33,724 excluded records reflected cleaning and cohort eligibility criteria (Figure S1) in Multimedia Appendix 1, no survey weights were used in the analyses.

Ethical Considerations

The study was conducted in accordance with the Declaration of Helsinki and approved by the Human Research Ethics Committee of the Beijing Center for Disease Prevention and Control (approval number: 2022 - 14), the retrospective analysis without additional informed consent was approved by the review board. To safeguard participant information, all the data were deidentified after linkage. Additionally, access to the data was restricted to authorized research personnel only.

Results

Basic Characteristics of Study Population

A total of 87,513 records were registered in the study. After removing 28,551 duplicate case records (84.7% of exclusions), 1570 (4.6%) whose birthday did not meet the criteria, and 3603 (10.7%) who were vaccinated between November in 2022 and February in 2023 and ineligible for inclusion, 53,789 individuals were included in the study cohort. Of these, 28,423 (52.84%)

XSL•FO
RenderX

were male. A total of 30,531 (56.76%) belonged to the vaccination group, while 23,258 (43.24%) were unvaccinated, as detailed in Table 1. The vaccine rollout in the study cohort was shown in Figure 1.

Table . Basic characteristics of the hospitalized older patients with COVID-19 cohort by vaccination status in Beijing during the Omicron wave from November 2022 to February 2023.

Characteristics	Total, n (%)	Unvaccinated, n (%)	Partially vaccinated, n (%)	Primary vaccination, n (%)	Booster vaccination, n (%)	Chi-square (*df*)/ *H* value	*P* value
No. of persons	53,789	23,258 (43.24)	4524 (8.41)	11,246 (20.91)	14,761 (27.44)	__[e]	__[e]
Sex							
Male	28,423 (52.84)	11,723 (50.40)[a(1)]	2177 (48.12)[a(2)]	5647 (50.21)[a(1,2)]	8876 (60.13)[a(3)]	44.86[b,c]	<.001
Female	25,366 (47.16)	11,535 (49.60)[a(1)]	(51.88)[a(2)]	99 (49.79)[a(1,2)]	5885 (39.87)[a(3)]		
Age (years), median (IQR)	85 (82-88)	86 (83-89)[a(1)]	85 (82-88)[a(2)]	85 (82-88)[a(2)]	83 (81-86)[a(3)]	2515.06[d]	<.001
Age group (years)							
80-89	45,270 (84.16)	18,126 (77.93)[a(1)]	3819 (84.42)[a(2)]	9621 (85.55)[a(3)]	13,704 (92.84)[a(4)]	1526.99[b,c]	<.001
90+	8519 (15.84)	5132 (22.07)[a(1)]	705 (15.58)[a(2)]	1625 (14.45)[a(3)]	1057 (7.16)[a(4)]		

[a]indicated the result of pairwise comparisons, if two vaccination groups have the same superscript number, that means there is no significant difference between them; if they are different, there is a significant difference (*P*<.05).

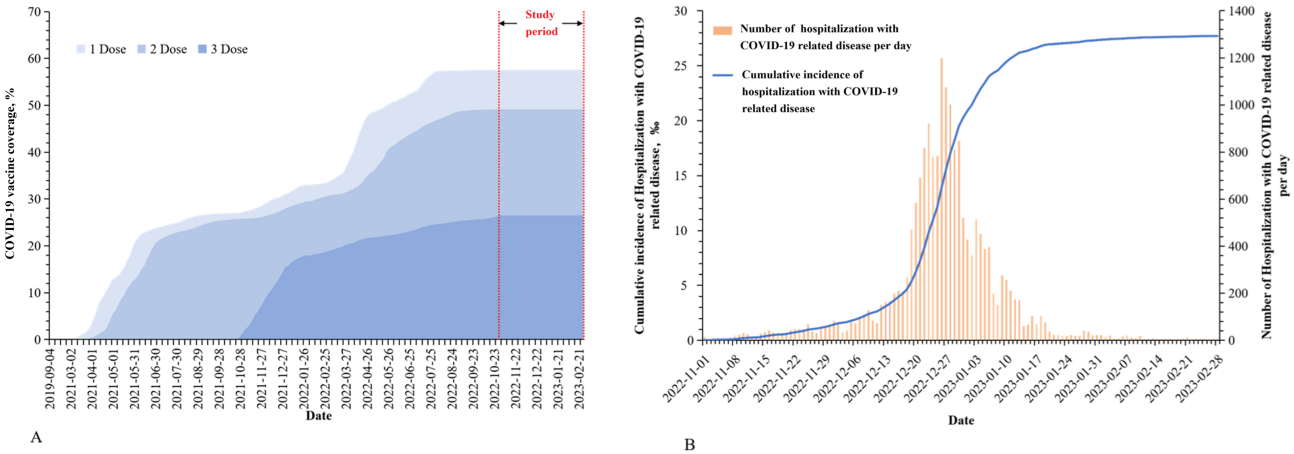
[b]Chi-square value.

[c]Degree of freedom (*df*)=3.

[d]H value.

[e]Not applicable

Figure 1. Temporal trends in COVID-19 vaccination coverage and hospitalization with COVID-19 incidence among adults aged ≥80 years from November 2022 to February 2023 in Beijing, China. (A) Cumulative daily coverage of COVID-19 vaccination, with the x-axis representing the vaccination date. This data aims to show the progression of vaccination over time. (B) The number and cumulative incidence of hospitalizations with COVID-19, where the x-axis indicates the admission date.



Hospitalization and VE

There were 17,916 (33.31%) individuals hospitalized with COVID-19, and the proportion of COVID-19 in males (36.60%, 10,403/28,423) was higher than that in females (29.62%, 7513/25,366) (*P*<.001). The proportion in inpatients aged ≥90 years (44.99%, 3833/8519) was higher than that in inpatients aged 80 - 89 years (31.11%, 14,083/45,270) (*P*<.001). Of these, 3535 (6.57%) were COVID-19–related hospitalization, 961 (1.79%) were COVID-19 severe or critical cases, and 4130 (7.68%) were in-hospital deaths in the study cohort. The incidence of hospitalization with COVID-19–related disease was estimated to be 2.66 % among 673 thousand people aged 80 years and older in Beijing. Most cases occurred in December and January of the following year (Figure 1).

The risk of COVID-19–related hospitalization in the unvaccinated group (9.76%, 2269 cases) was higher than that in the vaccinated group (partially vaccinated: 4.49%, 203/4524 cases; primary series: 4.49%, 505/11,246 cases; booster vaccination: 3.78%, 558/14,761 cases). The effectiveness of COVID-19 immunization in preventing COVID-19–related

XSL•FO
RenderX

hospitalization was 56% (95% CI 49% - 62%), 56% (95% CI 51.4% - 60.2%), 63.5% (95% CI 59.8% - 66.9%) for partial immunization, primary series, and booster vaccination, respectively (Figure 2). Figure 3 shows the VE stratified by the time interval between the last dose and hospital admission. VE of primary series was 63.7% (95% CI 58.9% - 68%) and 73.7% (95% CI 66.5% - 79.4%) during 6 - 12 months and ≥12 months, respectively. VE of booster was 43.1% (95% CI 24.4% - 57.2%) for <6 months, 44.9% (95% CI 36.5% - 52.2%) during 6 - 12 months, and 73.3% (95% CI 69.6% - 76.5%) during ≥12 months.

Figure 2. Adjusted vaccine effectiveness against severe outcomes by immunization status in adults aged ≥80 years during Omicron circulation, Beijing, China (November 2022 to February 2023) (Cohort study of hospitalized patients; vaccine effectiveness estimates via Poisson regression). (A) VE against COVID-19–related hospitalization, severe or critical disease, and in-hospital death; (B) Sensitivity analysis restricting to cases with severe or critical COVID-19 as principal diagnosis.

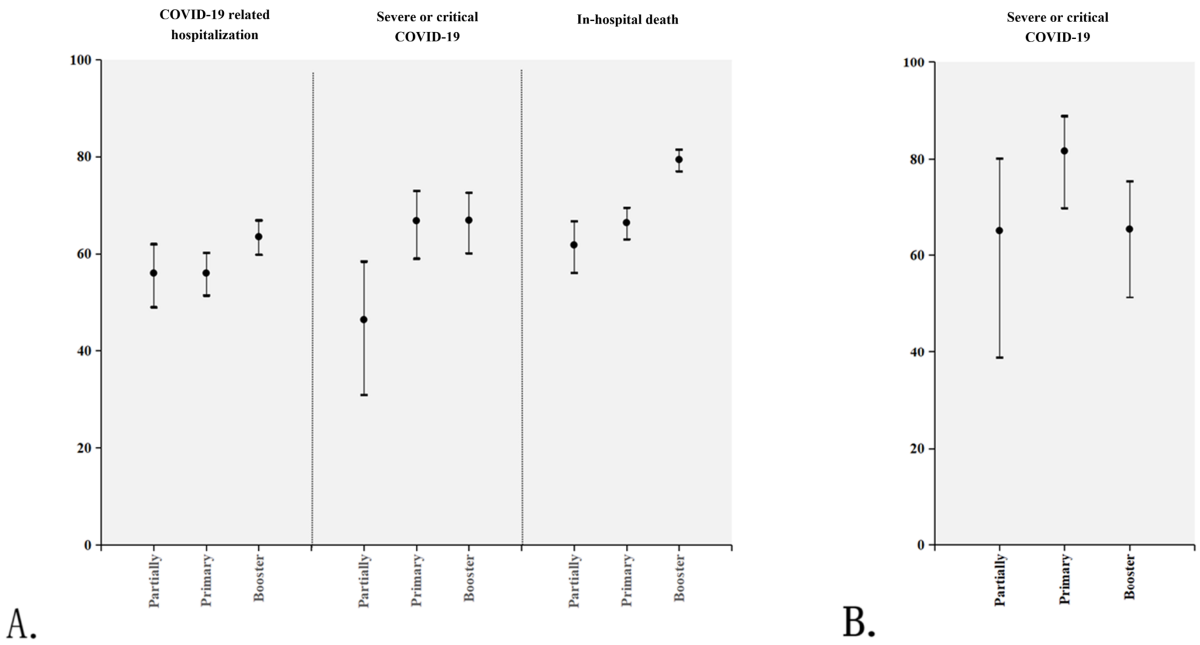
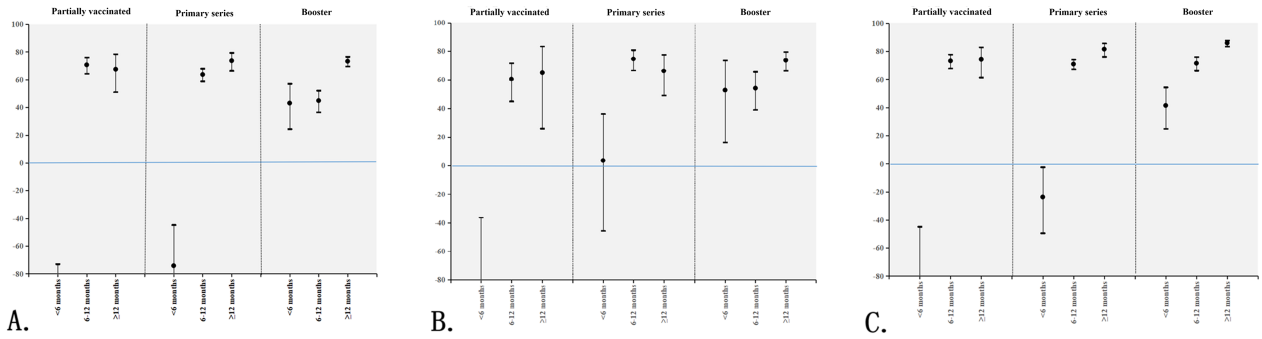


Figure 3. Stratified vaccine effectiveness with 95% CI by the time interval between the last vaccination dose and hospital admission for COVID-19 related outcomes among the older inpatients aged ≥80 years during the Omicron circulation (November 2022 to February 2023) in Beijing, China. (A) Hospitalization (principal diagnosis); (B) Severe or critical disease (principal or secondary diagnosis); (C) In-hospital death. Transverse lines that cross the vertical blue line at zero indicate no statistical significance.



The risk of severe or critical COVID-19 in the unvaccinated group (2.8%, 652 cases) was higher than that of the vaccinated group (0.92% - 1.48%). The effectiveness of COVID-19 vaccines against severe or critical COVID-19 was 46.4% (95% CI 30.9% - 58.4%), 66.8% (95% CI 59% - 73%) and 66.9% (95% CI 60.1% - 72.6%) for partial immunization, primary series, and booster vaccination, respectively (Figure 2). VE of primary series was 74.6% (95% CI 66.6% - 80.7%) during 6 - 12 months, and 66.2% (95% CI 49% - 77.6%) ≥12 months following vaccination. VE of booster was 52.9% (95% CI 16.2% - 73.5%) for <6 months, 54.2% (95% CI 39% - 65.6%) during 6 - 12 months, and 73.6% (95% CI 66.3% - 79.4%) for ≥12 months following vaccination (Figure 3).

The risk of in-hospital death with COVID-19 disease in the unvaccinated group (12.86%, 2990/23,258 cases) was higher than that of the vaccinated group (partially vaccinated: 5.11%, 231/4524 cases; primary series: 4.48%, 504/11,246 cases; booster vaccination: 2.74%, 405/14,761 cases). The effectiveness of booster vaccination against in-hospital deaths

XSL•FO
RenderX

was 79.4% (95% CI 77% - 81.5%), higher than that of primary series (66.4% [95% CI 63% - 69.5%]). The effectiveness of partial COVID-19 immunization in preventing in-hospital death was 61.8% (95% CI 56.1% - 66.7%) (Figure 2). VE of primary series was 71% (95% CI 67.3% - 74.3%) during 6 - 12 months, and 81.6% (95% CI 76.2% - 85.8%) ≥12 months following vaccination. VE of booster was 41.5% (95% CI 24.9% - 54.5%) for <6 months, 71.6% (95% CI 66.4% - 76%) during 6 - 12 months, and 86% (95% CI 83.6% - 87.9%) for ≥12 months following vaccination (Figure 3).

While VE generally declined with time since vaccination, the point estimate for ≥12 months post-booster appeared higher than the 6 - 12 months group. This anomaly might be attributable to the survivor bias, the population in this ≥12 months cohort opted for vaccination earlier (before November 2021) because they were in better health and awareness of health protection, which could lead to an overestimation of VE.

Sensitivity Analysis

If only the older individuals with severe or critical COVID-19 in "principal diagnosis" were analyzed, the VE against hospitalization decreased to 65.1% (95% CI 38.8% - 80.1%), 81.6% (95% CI 69.8% - 88.8%), 65.4% (95% CI 51.3% - 75.4%) for partial immunization, primary series, and booster vaccination, respectively (Figure 2). The VE of the booster was comparable to the findings of the main analysis, with varying precision due to reduced case counts (n=42 vs n=138 in main analysis). A similar pattern was observed in the stratified analysis based on different vaccination intervals (Table S1 and Table S2 in Multimedia Appendix 2 for time-stratified results).

Discussion

Principal Findings

During an Omicron BF.7-predominant wave of COVID-19 in Beijing city, we used electronic individual-level hospitalization records matching with record-documented COVID-19 vaccination status to calculate the characteristics and vaccination status of the inpatients aged ≥80 years and estimate the real-world effectiveness of domestic COVID-19 vaccines. This study shows that the proportion of hospitalizations diagnosed with COVID-19 disease in vaccination groups was significantly lower for those who had not been vaccinated. When given according to the recommended schedule of a 2-dose primary series followed by a booster dose 6 months later, the VE against COVID-19–related hospitalization outcomes was 63.5% - 79.4%. These results represent absolute modest VE associated with vaccine-induced protection for inpatients aged ≥80 years, uninfluenced by hybrid immunity.

COVID-19–associated hospitalization rates remain higher among older adults relative to rates among younger adults, adolescents, and children [25]. The Omicron variant of SARS-CoV-2 has been reported as more transmissible, but less severe, than previous variants. In the study of all hospitalization for population aged 80 years and older, 33.31% of inpatients had COVID-19–related diagnosis, the incidence of hospitalization with COVID-19 disease in ≥80-year-old people was estimated as 2.66% in the United Kingdom [26], the proportion of older adults aged ≥80 years who were hospitalized in all COVID-19 cases with Omicron (15.2%) was highest compared with other age groups of participants (1.47% - 5.80%). Compared with other Omicron variants, the studies in New England [27] found that 3.1% of patients infected with Omicron BA.2 required admission to hospitalization for general population (The mean age is 49.5). Up to now, older adults still were listed as the highest priority groups in the COVID-19 vaccination guidance [28]. In comparison with other countries, China initiated its vaccine roll-out campaign relatively early [29]. Despite the coverage of primary series for people aged 18 years and older in Beijing having already raised to a high level before the Omicron epidemic wave and quickly carrying out the booster program, 30% of adults aged 80 years and older remained unvaccinated or partially vaccinated, only 46.7% completing the booster shots [30]. China has the highest number of older individuals globally [31]. To prevent excessive disease severity and mortality owing to vaccination hesitancy, devoting greater effort toward vaccinating the older population is warranted given the uncertainty regarding future epidemic waves with frequent emergence of virus variants.

The effectiveness of the COVID-19 vaccines has relatively well maintained over time against the Omicron variant and its sublineages for severe disease compared to pre-Omicron variants of concern [4]. Our analysis demonstrated the primary series effectiveness against hospitalization and mortality was slightly higher than that reported in Brazilian test-negative case-control studies for those aged ≥80 years who completed primary series of CoronaVac (inactivated vaccine) during the Gamma variant predominance (38.9% [21.4% - 52.5%] and 44% [20.3% - 60.6%], respectively) [32]. This difference may be due to several factors. Circulating variants and prior infection status played significant roles. Other elements, such as the time interval between vaccination and the outcome, hospitalization criteria, population distribution, and research methods can also influence the calculated VE. However, similar to the conclusions found in most studies, our estimated VE of booster against COVID-19–related hospital admissions was higher than that of the primary series. One study in China indicated that a completed vaccination with a booster reduced the risk of asymptomatic or mild Delta or Omicron variant COVID-19 progressing to pneumonia [33]. In the study of quarantined close contacts of BA.2-infected individuals in China [34], primary series VE against pneumonia or worse infection was 66% (95% CI 57% - 73%) and 91% (95% CI 91% - 96%), booster dose VE was 74% (95% CI 69% - 79%) and 93% (95% CI 87% - 97%), respectively. Among individuals 80 years and older in the McMenamin study, VE increased from 58% to 97% when boosted [15]. In a prospective cohort study across UK nations, data on 30 million people aged ≥18 years from primary care, testing, vaccination, hospitalization, and mortality were linked. They found that compared with received primary doses, the first booster of BNT162b2 (tozinameran; Pfizer-BioNTech) or mRNA-1273 (elasomeran; Moderna) could have reduced severe outcomes (hospitalization or death) due to Omicron B.1.1.529 variant (the rate decreased from 8.8 to 7.6 events per 1000 person-y) [18]. Compared with the observed VEs in aforementioned studies conducted during the stages of Alpha,

XSL•FO
RenderX

Gamma, and initial Omicron predominance [15,35,36], our estimated result, which focuses on the period of Omicron BF.7 predominance and BA.5.2, was lower. This study also evaluated VE against in-hospital deaths with COVID-19 disease, but it was difficult to determine whether the deaths marked among inpatients with the U07.1 *ICD-10* code were actually caused by COVID-19 disease, so the direct protective effect could not be calculated due to many confounding factors (such as the reason for hospitalization, treatment measures, etc). Moreover, in line with previous research, a decline in VE was observed as age increased. During the Omicron BA.2 subvariant wave in Hong Kong, VE against severe or fatal disease was 96.3% (94.9% - 97.3%) for 2 doses of BNT162b2 and 91.7% (88.7% - 94%) for 2 doses of CoronaVac in those aged 20 - 59 years, reducing to 86.9% (80.5% - 91.3%) and 58.2% (45.1% - 68.2%) among those aged 80 years or older [15]. A parallel pattern emerged in a Shanghai Omicron BA.2 outbreak, where 3 doses of inactivated vaccines showed 96.0% (91% - 98.3%) effectiveness against severe or critical COVID-19 in the 60 - 69 age cohort, decreasing to 66.8% (−22.5% to 91%) among people aged ≥90 years [16]. Older people have a less complete immune response to antigens and a reduced ability to develop robust immunity after infections or vaccination because of the age-related decline in the functioning of the immune system [37].

For the primary and booster immunizations of the COVID-19–inactivated vaccine, the GMT against the wild-type strain declined gradually during the 12 months after the primary immunization and was lower against the Delta and Omicron variant spike proteins [38]. It is necessary to estimate the protective effect of different intervals after the latest vaccination and how long vaccinated individuals should be evaluated before future epidemic trends. As of November in 2022, the COVID-19 mass vaccination has completed and most people have gotten the latest dose for 12 months and longer. There were too few COVID-19 cases with 6 months between the last dose and exposure to estimate reliable VEs stratified by time; results are shown in Supplementary Table S2 in Multimedia Appendix 2. VE of vaccination in more than 12 months and during 6 months and 12 months after the last dose has remained 52% - 86%. The highest VE observed in the ≥12-month post-vaccination group contradicts established immune waning patterns. This paradox may reflect methodological biases, one is healthy survivor bias, the population in this ≥12-month cohort opted for vaccination earlier on because they were in better health and awareness of health protection, or the healthiest individuals survived long enough to be in the ≥12 months group, all these could lead to an overestimation of VE. Additionally, temporal confounding may have occurred, the ≥12 months group may have been infected later in the wave when hospital triage and treatment protocols improved, or the virus variant and the population's immunity status may have changed over the study period. In addition, as more than 4 doses of vaccine have been administered worldwide, it is necessary to further track and monitor the vaccination situation and VE in China.

Strengths and Limitations

One of the strengths of this study is the effective linkage between hospital admission data and vaccination records, which enables a low-cost and large-scale real-world evaluation of vaccine effects in the population. Specifically, this approach provides valuable insights into the effectiveness of prototype inactivated COVID-19 vaccines in protecting the eldest adults. The other one is our results reveal that, among inpatients aged 80 years and above who have not had a previous SARS-CoV-2 infection, there is a moderately significant level of protection. Importantly, this protection remains independent of hybrid immunity. In the context of multiple global epidemics, especially those involving Omicron variant strains, most people have experienced one or more infections. When evaluating the protective effect of vaccination against Omicron, it is difficult to exclude the confounding effects of previous infections with SARS-CoV-2 variants. However, this study still offers unique evidence for the basic protective effect of inactivated vaccines.

Our study had several limitations. First, the foremost concern is unmeasured confounding from fundamental health status differences between vaccinated and unvaccinated groups. The unvaccinated group was significantly older (86 vs 83 y) and had more females (49.60% vs 39.87%; $P<.001$) than the booster group, suggesting poorer baseline health or frailty limiting vaccine access. Critically linked is our inability to stratify by underlying diseases, a key unmeasured confounder. While we adjusted for available confounders (age and sex), unmeasured factors may persist, and the VE estimates may be influenced by unmeasured health status differences between the groups, as the unvaccinated group's inherently poorer health may lead to worse outcomes independent of vaccination status. Future studies should use targeted designs that prospectively enroll subgroups stratified by specific comorbidities, implement propensity score matching to further mitigate confounding. Second, use of all-cause in-hospital mortality among patients diagnosed with COVID-19 cannot represent or estimate the specific COVID-19 directly caused by COVID-19. This misclassification could introduce bias into the estimates of VE against such in-hospital deaths. Third, there might have been potential selection bias in our cohort design without nonhospitalized individuals. Considering the hospitalized population may have worse physical conditions than those of the nonhospitalized population, the evaluation based on hospitalized cases may underestimate the value of VE. During the establishment of the cohort, eliminating individuals who were vaccinated during the study period would introduce a selection bias to the results, which cannot represent the true situation of the population. The composition of the inpatients in this study was close to the age distribution of permanent residents in Beijing (83.09% were 80 - 89 y old, 16.91% were ≥90 y old), but the difference in gender composition (56.24% were female and 43.76% were male), although estimated VE was adjusted for age and gender would decrease this bias, our results still can only be limited to representing the hospitalized population, may not be generalizable to nonhospitalized or younger populations. Another point to consider is that, compared with experimental studies, the observational study cannot absolutely control external conditions; therefore, there could be unmeasured confounding variables.

XSL•FO
RenderX

Conclusions

In conclusion, booster vaccination was associated with significantly reduced risk of COVID-19–related severe or critical illness or in-hospital deaths in older inpatients aged ≥80 years during an Omicron-dominant period. Offering a booster dose and regularly monitoring the coverage and VE of COVID-19 vaccines remains the mainstay for mitigating impacts of upcoming epidemics. While our results demonstrate a significant association between booster vaccination and reduced severe outcomes, these estimates may reflect both VE and the better baseline health status of the vaccinated older individuals.

Funding

This research was funded by the Beijing Science and Technology Program (grant ID Z231100004823001), and Beijing Research Center for Respiratory Infectious Diseases Project (grant ID BJRID2024-006).

Data Availability

The datasets analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: JW and LS
Methodology: DZ
Software: YM
Validation: JW and LS
Investigation and formal analysis: DZ, YM, JL, XL, and ZC
Data curation: WY
Writing – original draft preparation: DZ and YM
Writing – review and editing: JL, XL, ZC, WY, JW, and LS
Visualization: DZ
Supervision and project administration: LS
Funding acquisition: JW and LS

Conflicts of Interest

None declared.

Multimedia Appendix 1
Flow chart of the study cohort.
[PNG File, 243 KB - publichealth_v12i1e82915_app1.png]

Multimedia Appendix 2
Vaccine effectiveness of COVID-19 vaccines considering different immunization status, intervals since the last dose against COVID-19–related hospitalization during the Omicron wave in Beijing, along with characteristics of hospitalized patients in the cohort.
[DOCX File, 21 KB - publichealth_v12i1e82915_app2.docx]

References

1. Coronavirus disease (COVID-19). World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/coronavirus-disease-(covid-19) [accessed 2023-12-20]
2. Cao Y, Yisimayi A, Jian F, et al. BA.2.12.1, BA.4 and BA.5 escape antibodies elicited by Omicron infection. Nature New Biol 2022 Aug;608(7923):593-602. [doi: 10.1038/s41586-022-04980-y] [Medline: 35714668]
3. Hachmann NP, Miller J, Collier ARY, et al. Neutralization escape by SARS-CoV-2 Omicron subvariants BA.2.12.1, BA.4, and BA.5. N Engl J Med 2022 Jul 7;387(1):86-88. [doi: 10.1056/NEJMc2206576] [Medline: 35731894]
4. WHO roadmap on uses of COVID-19 vaccines in the context of Omicron and high population immunity. World Health Organization. URL: https://www.who.int/publications/i/item/WHO-2019-nCoV-Vaccines-SAGE-Prioritization-2023.2 [accessed 2023-12-28]
5. Suntronwong N, Kanokudom S, Duangchinda T, et al. Neutralization of omicron subvariants and antigenic cartography following multiple COVID 19 vaccinations and repeated omicron non JN.1 or JN.1 infections. Sci Rep 2025 Jan 9;15(1):1454. [doi: 10.1038/s41598-024-84138-0] [Medline: 39789099]
6. Hogan AB, Doohan P, Wu SL, et al. Estimating long-term vaccine effectiveness against SARS-CoV-2 variants: a model-based approach. Nat Commun 2023 Jul 19;14(1):4325. [doi: 10.1038/s41467-023-39736-3] [Medline: 37468463]

XSL•FO
RenderX

7. Ledford H. How severe are Omicron infections? Nature New Biol 2021 Dec 23;600(7890):577-578. [doi: 10.1038/d41586-021-03794-8] [Medline: 34934198]
8. Press conference of the joint prevention and control mechanism of the state council [Web page in Chinese]. The Portal Website of the Central People's Government of the People's Republic of China. URL: https://www.gov.cn/xinwen/gwylflkjz221/index.htm [accessed 2023-12-28]
9. Tanriover MD, Doğanay HL, Akova M, et al. Efficacy and safety of an inactivated whole-virion SARS-CoV-2 vaccine (CoronaVac): interim results of a double-blind, randomised, placebo-controlled, phase 3 trial in Turkey. Lancet 2021 Jul 17;398(10296):213-222. [doi: 10.1016/S0140-6736(21)01429-X] [Medline: 34246358]
10. Al Kaabi N, Zhang Y, Xia S, et al. Effect of 2 inactivated SARS-CoV-2 vaccines on symptomatic COVID-19 infection in adults: a randomized clinical trial. JAMA 2021 Jul 6;326(1):35-45. [doi: 10.1001/jama.2021.8565] [Medline: 34037666]
11. Halperin SA, Ye L, MacKinnon-Cameron D, et al. Final efficacy analysis, interim safety analysis, and immunogenicity of a single dose of recombinant novel coronavirus vaccine (adenovirus type 5 vector) in adults 18 years and older: an international, multicentre, randomised, double-blinded, placebo-controlled phase 3 trial. Lancet 2022 Jan 15;399(10321):237-248. [doi: 10.1016/S0140-6736(21)02753-7] [Medline: 34953526]
12. Interim recommendations for use of the inactivated COVID-19 vaccine, CoronaVac, developed by Sinovac: interim guidance, 24 may 2021. World Health Organization (WHO) Institutional Repository for Information Sharing (IRIS). URL: https://iris.who.int/handle/10665/341454 [accessed 2023-12-28]
13. Interim recommendations for use of the inactivated COVID-19 vaccine BIBP developed by China National Biotec Group (CNBG), Sinopharm: interim guidance, first issued 7 may 2021, updated 28 october 2021, updated 15 march 2022. World Health Organization (WHO), Institutional Repository for Information Sharing (IRIS). URL: https://iris.who.int/handle/10665/352470 [accessed 2023-12-28]
14. Interim recommendations for use of the CanSinoBIO Ad5-nCoV-S [recombinant] vaccine (Convidecia™) against COVID-19: interim guidance, first issued 19 May 2022. World Health Organization (WHO), Institutional Repository for Information Sharing (IRIS). URL: https://iris.who.int/handle/10665/354409 [accessed 2023-12-28]
15. McMenamin ME, Nealon J, Lin Y, et al. Vaccine effectiveness of one, two, and three doses of BNT162b2 and CoronaVac against COVID-19 in Hong Kong: a population-based observational study. Lancet Infect Dis 2022 Oct;22(10):1435-1443. [doi: 10.1016/S1473-3099(22)00345-0] [Medline: 35850128]
16. Huang Z, Xu S, Liu J, et al. Effectiveness of inactivated COVID-19 vaccines among older adults in Shanghai: retrospective cohort study. Nat Commun 2023;14(1):2009. [doi: 10.1038/s41467-023-37673-9] [Medline: 37037803]
17. Ward IL, Robertson C, Agrawal U, et al. Risk of COVID-19 death in adults who received booster COVID-19 vaccinations in England. Nat Commun 2024 Jan 16;15(1):398. [doi: 10.1038/s41467-023-44276-x] [Medline: 38228613]
18. Agrawal U, Bedston S, McCowan C, et al. Severe COVID-19 outcomes after full vaccination of primary schedule and initial boosters: pooled analysis of national prospective cohort studies of 30 million individuals in England, Northern Ireland, Scotland, and Wales. Lancet 2022 Oct 15;400(10360):1305-1320. [doi: 10.1016/S0140-6736(22)01656-7] [Medline: 36244382]
19. Ge Y, Wu X, Zhang W, et al. Effects of public-health measures for zeroing out different SARS-CoV-2 variants. Nat Commun 2023 Aug 29;14(1):5270. [doi: 10.1038/s41467-023-40940-4] [Medline: 37644012]
20. Liu J, Ding F, Wu Y, et al. Trends of SARS-CoV-2 infection in sentinel community-based surveillance after the optimization of prevention and control measures - China, december 2022-january 2023. China CDC Wkly 2023 Feb 17;5(7):159-164. [doi: 10.46234/ccdcw2023.028] [Medline: 37009518]
21. Leung K, Lau EHY, Wong CKH, Leung GM, Wu JT. Estimating the transmission dynamics of SARS-CoV-2 Omicron BF.7 in Beijing after adjustment of the zero-COVID policy in november-december 2022. Nat Med 2023 Mar;29(3):579-582. [doi: 10.1038/s41591-023-02212-y] [Medline: 36638825]
22. Pan Y, Wang L, Feng Z, et al. Characterisation of SARS-CoV-2 variants in Beijing during 2022: an epidemiological and phylogenetic analysis. Lancet 2023 Feb 25;401(10377):664-672. [doi: 10.1016/S0140-6736(23)00129-0] [Medline: 36773619]
23. Wu W, Cao L, Zheng J, Cao L, Cui J, Xiao Q. Immunization information system status in China, 2017. Vaccine (Auckl) 2019 Oct 8;37(43):6268-6270. [doi: 10.1016/j.vaccine.2019.08.070] [Medline: 31526621]
24. Huang F, Cheng JM, Guo MN, et al. The design and implementation of the tenth revision of the International Classification of Disease clinical modification of Beijing. Chin J Hosp Adm 2011;27(11):835-838. [doi: 10.3760/cma.j.issn.1000-6672.2011.11.009]
25. Panagiotakopoulos L, Godfrey M, Moulia DL, et al. Use of an additional updated 2023-2024 COVID-19 vaccine dose for adults aged ≥65 years: recommendations of the Advisory Committee on Immunization Practices - United States, 2024. MMWR Morb Mortal Wkly Rep 2024 Apr 25;73(16):377-381. [doi: 10.15585/mmwr.mm7316a4] [Medline: 38662708]
26. Nyberg T, Ferguson NM, Nash SG, et al. Comparative analysis of the risks of hospitalisation and death associated with SARS-CoV-2 omicron (B.1.1.529) and delta (B.1.617.2) variants in England: a cohort study. Lancet 2022 Apr 2;399(10332):1303-1312. [doi: 10.1016/S0140-6736(22)00462-7] [Medline: 35305296]

XSL•FO
RenderX

27. Strasser ZH, Greifer N, Hadavand A, Murphy SN, Estiri H. Estimates of SARS-CoV-2 Omicron BA.2 subvariant severity in New England. JAMA Netw Open 2022 Oct 3;5(10):e2238354. [doi: 10.1001/jamanetworkopen.2022.38354] [Medline: 36282501]
28. SAGE updates COVID-19 vaccination guidance. World Health Organization. 2023 Mar 28. URL: https://worldhealthorganization.createsend1.com/t/d-e-zklkult-ijidulchl-y/ [accessed 2024-02-01]
29. Zhang J, Yang H, Yang M, Tan H. The role of vaccines in COVID-19 control strategies in Singapore and China. Health Policy Technol 2022 Jun;11(2):100620. [doi: 10.1016/j.hlpt.2022.100620] [Medline: 35369127]
30. Wang G, Yao Y, Wang Y, et al. Determinants of COVID-19 vaccination status and hesitancy among older adults in China. Nat Med 2023 Mar;29(3):623-631. [doi: 10.1038/s41591-023-02241-7] [Medline: 36720270]
31. Chen X, Giles J, Yao Y, et al. The path to healthy ageing in China: a Peking University-Lancet Commission. Lancet 2022 Dec 3;400(10367):1967-2006. [doi: 10.1016/S0140-6736(22)01546-X] [Medline: 36423650]
32. Ranzani OT, Hitchings MDT, Dorion M, et al. Effectiveness of the CoronaVac vaccine in older adults during a gamma variant associated epidemic of covid-19 in Brazil: test negative case-control study. BMJ 2021 Aug 20;374:n2015. [doi: 10.1136/bmj.n2015] [Medline: 34417194]
33. Li J, Song R, Yuan Z, et al. Protective effect of inactivated COVID-19 vaccines against progression of SARS-CoV-2 Omicron and Delta variant infections to pneumonia in Beijing, China, in 2022. Vaccines (Basel) 2022 Jul 29;10(8):1215. [doi: 10.3390/vaccines10081215] [Medline: 36016103]
34. Tang L, Wang FZ, Rodewald LE, et al. Real-world effectiveness of primary series and booster doses of inactivated coronavirus disease 2019 vaccine against Omicron BA.2 variant infection in China: a retrospective cohort study. J Infect Dis 2023 Aug 11;228(3):261-269. [doi: 10.1093/infdis/jiad090] [Medline: 37005365]
35. Huang Z, Xu S, Liu J, et al. Effectiveness of inactivated and Ad5-nCoV COVID-19 vaccines against SARS-CoV-2 Omicron BA. 2 variant infection, severe illness, and death. BMC Med 2022 Oct 20;20(1):400. [doi: 10.1186/s12916-022-02606-8] [Medline: 36266697]
36. Jara A, Undurraga EA, González C, et al. Effectiveness of an Inactivated SARS-CoV-2 vaccine in Chile. N Engl J Med 2021 Sep 2;385(10):875-884. [doi: 10.1056/NEJMoa2107715] [Medline: 34233097]
37. de Candia P, Prattichizzo F, Garavelli S, Matarese G. T cells: warriors of SARS-CoV-2 infection. Trends Immunol 2021 Jan;42(1):18-30. [doi: 10.1016/j.it.2020.11.002] [Medline: 33277181]
38. Li J, Xie H, Chen W, et al. Immune persistence against SARS-CoV-2 after primary and booster immunization in humans: a large-scale prospective cohort study. Vaccines (Basel) 2022 Oct 8;10(10):1677. [doi: 10.3390/vaccines10101677] [Medline: 36298542]

Abbreviations

BHUIS: Beijing Hospital Utilization Information System
BMHCIC: Beijing Municipal Health Commission's Information Center
ICD-10: *International Statistical Classification of Diseases and Related Health Problems 10th Revision*
IIS: immunization information system
VE: vaccine effectiveness

Edited by A Mavragani, T Sanchez; submitted 24.Aug.2025; peer-reviewed by A Girma, L Rodewald; revised version received 19.Jan.2026; accepted 19.Jan.2026; published 11.Mar.2026.

Please cite as:
Zhao D, Ma Y, Li J, Li X, Cao Z, Yao W, Wu J, Suo L
Effectiveness of Inactivated COVID-19 Vaccination Against COVID-19–Related Hospitalization and Severe Outcomes in Adults ≥80 Years During Omicron Circulation in Beijing, China: Retrospective Cohort Study
JMIR Public Health Surveill 2026;12:e82915
URL: https://publichealth.jmir.org/2026/1/e82915
doi:10.2196/82915

© Dan Zhao, Ying Ma, Juan Li, Xiaomei Li, Zhiqiang Cao, Wei Yao, Jiang Wu, Luodan Suo. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 11.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Cross-Sectional Study on Oral Nicotine Product Sales Trends in Scandinavia From 2018 to 2025

Marina A Murphy[1], BSc, MSc, PhD; Diane Henenberg[2], BBus; Lindsay Reese[2], BS, PhD

[1]HAYPP Limited, 33 Clarke Road, Mount Farm, Milton Keynes, United Kingdom
[2]Snusbolaget AB, Stockholm, Sweden

Corresponding Author:
Marina A Murphy, BSc, MSc, PhD
HAYPP Limited, 33 Clarke Road, Mount Farm, Milton Keynes, United Kingdom

Related Article:

This is a corrected version. See correction statement: https://publichealth.jmir.org/2026/1/e94157

Abstract

This cross-sectional analysis of more than 19 million e-commerce orders from Sweden and Norway indicates that nicotine pouches have overtaken traditional snus in market share in both countries, reinforcing the potential of nicotine pouches as a harm reduction tool.

(JMIR Public Health Surveill 2026;12:e85490) doi:10.2196/85490

KEYWORDS

Norway; Sweden; market research; sales; nicotine pouch; snus; harm reduction

Introduction

Nicotine pouches (NPs) are rapidly gaining market traction, particularly in countries with established smokeless tobacco use. They differ from snus, a widely used smokeless tobacco product in Scandinavia, in that they do not contain any tobacco leaf. Chemical analyses confirmed that NPs contain substantially fewer and lower levels of toxicants compared with snus, with toxicant profiles similar to pharmaceutical nicotine replacement therapies [1]. Nicotine is dependence forming and should not be used by vulnerable populations such as youth, pregnant women, and individuals with certain health conditions; however, for adults who choose to use nicotine, NPs are increasingly popular. The United States is the largest NP market, and emerging evidence suggests that daily NP use is most common among adults who recently quit another product and rare among people who never used nicotine [2]. This encouraging trend suggests that NPs—originally designed as a harm reduction tool [3]—are indeed displacing more harmful products. This study examines whether a similar pattern is emerging in two Nordic countries where snus has traditionally been the dominant tobacco product.

Methods

Study Design

This cross-sectional study analyzed HAYPP Group AB (Stockholm, Sweden) sales data from January 1, 2018, to September 17, 2025, on 19,528,087 purchases of snus and NPs by 1,721,752 customers from seven e-commerce websites in two countries (Sweden: snusbolaget.se, haypp.se, nettotobak.se, and snusnetto.se; Norway: snuslageret.no, snushjem.no, and snus.com). In both countries, purchases are age-verified with government-issued identification numbers that encode each customer's birth date and gender. This information is authenticated by payment service providers. After anonymization (permanently removing all identifiers), aggregated data were used to calculate volume shares by product and by gender. The number of cans of NPs or snus sold in each year was divided by the total number of cans of both products sold in each country.

Ethical Considerations

Under Swedish law (Ethical Review Act, SFS 2003:460 [4]), ethical approval is required for research involving physical intervention on humans, biological material traceable to individuals, or processing of sensitive personal data. This study analyzed fully anonymized, aggregated sales data without any personal identifiers or direct human participation. Therefore, it falls outside the scope of the Act and was not submitted for ethical review. Under the European Union (EU) General Data Protection Regulation (GDPR 2016/679 [5]), ethical approval was not required for this market research study because it used fully anonymized sales data, and no personal data or direct human participation was involved. The study complies with the International Chamber of Commerce/European Society for Opinion and Market Research (ICC/ESOMAR) International

XSL•FO
RenderX

Code on Market, Opinion and Social Research, and Data Analytics [6].

All customers consented to use of their anonymized, aggregated data when reviewing and accepting the privacy policy [7]. No compensation was provided for the use of these sales data.

Results

The sales data cover 13,995,343 and 5,532,744 individual orders from 1,253,066 and 468,686 unique customers in Sweden and Norway, respectively. NPs surpassed snus in market share in 2025 in both countries (Figure 1). The Swedish NP volume share rose from 5% in 2018 to 55% in 2025, while snus declined from 95% to 45%. In Norway, the NP share rose from 22% to 56% over the same period, with snus declining from 78% to 44%.

Women first purchased more NPs than snus in 2022 (Figure 2) due to their lower 2018 market shares of snus (~25%). Men still purchase more snus than NPs in both countries, but this could change in 2026 if trends persist.

Figure 1. Cross-sectional volume shares of the oral nicotine product markets (2018 - 2025) for nicotine pouches (NP, solid lines) and snus (dotted lines) in Sweden (left) and Norway (right).

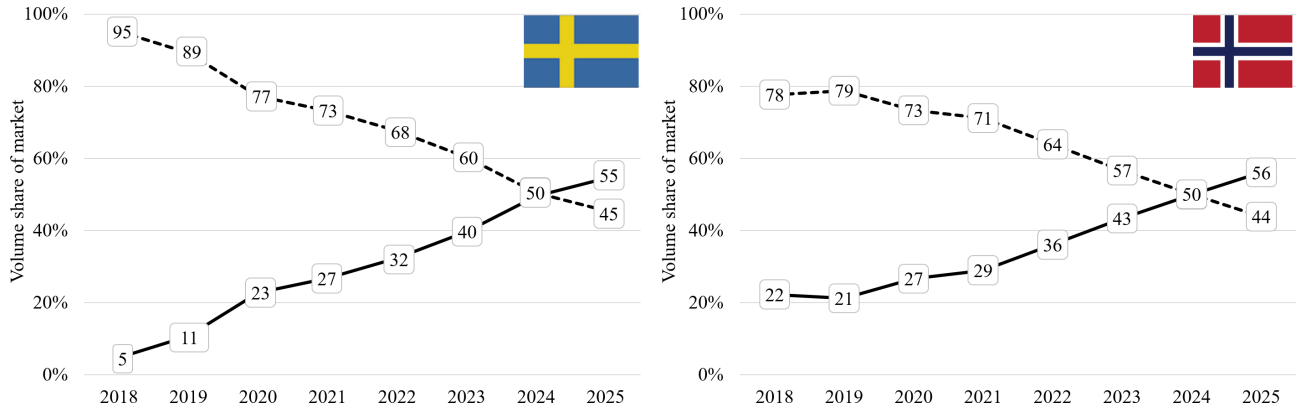
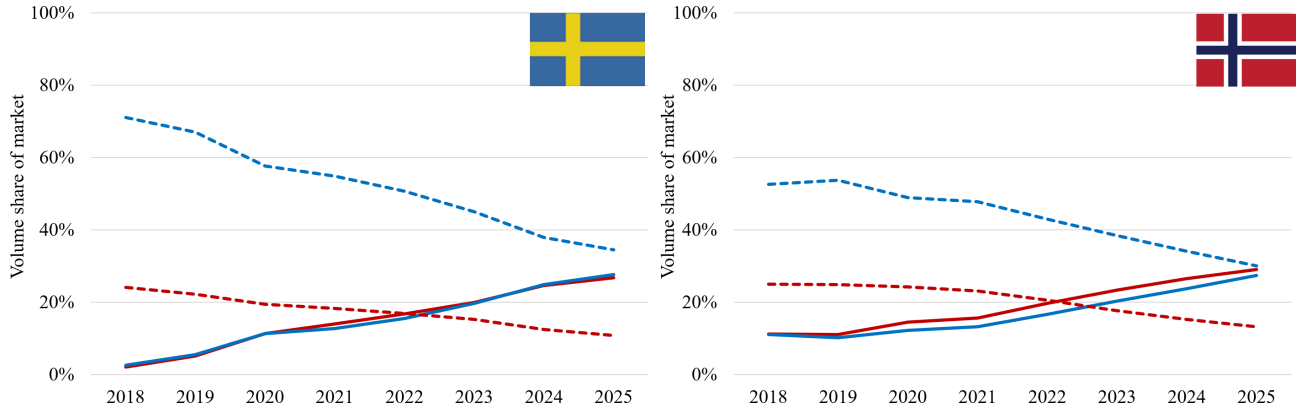


Figure 2. Cross-sectional volume shares of the oral nicotine product markets (2018 - 2025) for men (blue) and women (red) for nicotine pouches (NP, solid lines) and snus (dotted lines) in Sweden (left) and Norway (right).



Discussion

This descriptive analysis is the first cross-country comparison of NP and snus sales trends in Scandinavia. Data from more than 19 million e-commerce orders show that NPs overtook snus in the market share in Sweden and Norway in 2025, offering early evidence of market displacement in both countries. Taken together with US findings [2], these results suggest that NPs could be reshaping nicotine consumption in ways aligned with tobacco harm reduction, although ongoing monitoring and regulation will be critical to maximize benefits and minimize risks.

Although these observations do not reflect individual-level use, they highlight important shifts in consumer behavior that mirror the original "Swedish experience" where 50 years ago, men shifted from cigarettes to snus [8], and daily smoking decreased from 40% to 15% between 1976 and 2002 [9]. Over the same time frame, daily smoking rates for Swedish women dropped from 34% to 20% [9]. A similar trend materialized in Norway after low-nitrosamine snus was introduced in the 1990s, with men moving away from cigarettes before women [10,11]. The EU banned snus in 1992, but Sweden obtained an exemption. Norway is not part of the EU. Long-term epidemiological data indicate that this shift offers a substantial harm reduction benefit, as evidenced by low incidences of tobacco-related disease in Sweden compared to the rest of the EU [12].

The US Food & Drug Administration recognizes snus as a lower-risk alternative to cigarettes, authorizing eight products with the claim: "Using General Snus instead of cigarettes puts you at a lower risk of mouth cancer, heart disease, lung cancer, stroke, emphysema, and chronic bronchitis" [13]. NPs contain even fewer toxicants [1], placing them at the lowest risk end of the non-medicinal tobacco and nicotine product continuum. Maximizing public health would require encouraging switching

XSL•FO
RenderX

from higher-risk products to lower-risk products such as NPs, clear regulation, youth access prevention, and efforts to support cessation among current users who would like to stop using nicotine.

A limitation of this study is its reliance on sales data from a single e-commerce company. However, these platforms account for considerable snus and NP market shares in both countries. National surveys currently lack comparable precision. Statistik Sentrabyrå (Norway) does not distinguish between snus and NPs, and Folkhälsomyndigheten (Sweden) only began differentiating between snus and NPs in 2022. While future surveys will provide more nuanced nationally representative information, these market data likely offer a more accurate and timelier picture of an evolving Nordic nicotine landscape. Future research should explore user transitions across cigarettes, snus, NPs, and cessation, alongside long-term health and policy effects.

Acknowledgments

The sponsor HAYPP Group AB was involved in study design; the collection, analysis, and interpretation of data; the writing of the report; and the decision to submit the article for publication.

Funding

HAYPP Group AB is an online e-commerce company that sells non-combustible nicotine products and is the sponsor of this research. The Group operates independently and is not under the ownership or control of any tobacco or nicotine manufacturers.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: MAM, LR
Data curation: DH
Formal analysis: DH
Investigation: DH, MAM
Supervision: MAM
Validation: LR
Visualization: LR
Writing – original draft: LR
Writing – review & editing: LR, DH, MAM

Conflicts of Interest

MAM declares employment from HAYPP Limited, a subsidiary of HAYPP Group AB; non-financial support from UK Vaping Industry Association, for which she serves as a Director; and owns stock in HAYPP and British American Tobacco. DH and LR declare employment from Snusbolaget AB, a subsidiary of HAYPP Group AB.

References

1. Azzopardi D, Liu C, Murphy J. Chemical characterization of tobacco-free “modern” oral nicotine pouches and their position on the toxicant and risk continuums. Drug Chem Toxicol 2022 Sep;45(5):2246-2254. [doi: 10.1080/01480545.2021.1925691] [Medline: 34034614]
2. Delnevo CD, Tomaino M, Hrywna M, Bover Manderski MT. Patterns of nicotine pouch use among adults in the US, 2022-2023. JAMA Netw Open 2025 Sep 2;8(9):e2531155. [doi: 10.1001/jamanetworkopen.2025.31155] [Medline: 40920383]
3. Thornley S, McRobbie H, Lin RB, et al. A single-blind, randomized, crossover trial of the effects of a nicotine pouch on the relief of tobacco withdrawal symptoms and user satisfaction. Nicotine Tob Res 2009 Jun;11(6):715-721. [doi: 10.1093/ntr/ntp054] [Medline: 19454549]
4. Swedish Ethical Review Authority. What the Act says. URL: https://etikprovningsmyndigheten.se/en/what-the-act-says/ [accessed 2026-02-06]
5. European Union. Regulation 2016/679 (General Data Protection Regulation). URL: https://eur-lex.europa.eu/eli/reg/2016/679/oj [accessed 2026-02-02]
6. International Chamber of Commerce. ICC/ESOMAR International Code on Market, Opinion and Social Research and Data Analytics. URL: https://iccwbo.org/news-publications/business-solutions/iccesomar-international-code-market-opinion-social-research-data-analytics/ [accessed 2026-02-06]
7. HAYPP Group privacy statement. URL: https://www.haypp.com/se/privacy [accessed 2025-10-08]
8. Ramström L, Borland R, Wikmans T. Patterns of smoking and snus use in Sweden: implications for public health. Int J Environ Res Public Health 2016 Nov 9;13(11):1110. [doi: 10.3390/ijerph13111110] [Medline: 27834883]

XSL•FO
RenderX

9. Foulds J, Ramstrom L, Burke M, Fagerström K. Effect of smokeless tobacco (snus) on smoking and public health in Sweden. Tob Control 2003 Dec;12(4):349-359. [doi: 10.1136/tc.12.4.349] [Medline: 14660766]
10. Lund I, Lund KE. How has the availability of snus influenced cigarette smoking in Norway? Int J Environ Res Public Health 2014 Nov 13;11(11):11705-11717. [doi: 10.3390/ijerph111111705] [Medline: 25402565]
11. Lund M, Lund I. Smoking cessation aids and strategies: a population-based survey of former and current smokers in Norway. BMC Public Health 2022 Mar 31;22(1):631. [doi: 10.1186/s12889-022-13032-z] [Medline: 35361172]
12. Clarke E, Thompson K, Weaver S, Thompson J, O'Connell G. Snus: a compelling harm reduction alternative to cigarettes. Harm Reduct J 2019 Nov 27;16(1):62. [doi: 10.1186/s12954-019-0335-1] [Medline: 31775744]
13. FDA renews authorization for 8 general snus products to be marketed with reduced risk claim. Nov. 2024. URL: https://www.fda.gov/tobacco-products/fda-renews-authorization-8-general-snus-products-be-marketed-reduced-risk-claim [accessed 2025-10-08]

Abbreviations

EU: European Union
NP: nicotine pouch

Edited by A Mavragani, T Sanchez; submitted 08.Oct.2025; peer-reviewed by KE Lund, T Kurko; revised version received 03.Jan.2026; accepted 27.Jan.2026; published 13.Feb.2026.

Please cite as:
Murphy MA, Henenberg D, Reese L
Cross-Sectional Study on Oral Nicotine Product Sales Trends in Scandinavia From 2018 to 2025
JMIR Public Health Surveill 2026;12:e85490
URL: https://publichealth.jmir.org/2026/1/e85490
doi:10.2196/85490

© Marina A Murphy, Diane Henenberg, Lindsay Reese. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 13.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Indications for Medical Cannabis and Their Alignment With US-Wide Disease Prevalences in the US' Largest Medical Cannabis Program: Cross-Sectional Study

Sebastian Jugl, BPharm, MSc, PhD; Nicole E Smolinski, PharmD, PhD; Tansu Miriam Işılay Aydoğan, BPharm; Almut G Winterstein, RPh, PhD

Department of Pharmaceutical Outcomes and Policy, College of Pharmacy, University of Florida, 6018 Malachowsky Hall, Gainesville, FL, United States

Corresponding Author:
Almut G Winterstein, RPh, PhD
Department of Pharmaceutical Outcomes and Policy, College of Pharmacy, University of Florida, 6018 Malachowsky Hall, Gainesville, FL, United States

Abstract

This research letter compared the prevalence of qualifying conditions among patients certified to use medical cannabis in Florida to population-based prevalence estimates for the same conditions and observed marked discrepancies. These findings highlight conditions that appear to be significant reasons for patients to consider medical cannabis, providing important research priorities for future research on its effectiveness and safety.

(JMIR Public Health Surveill 2026;12:e89716) doi:10.2196/89716

KEYWORDS

medical marijuana; cannabis; qualifying condition; prevalence; cross-sectional studies; medical marijuana use registry

Introduction

The participation in medical cannabis (MC) programs has expanded rapidly in the United States—from 678,408 registered patients in 2016 to 2,974,433 in 2020 and approximately 3,641,949 in December 2024; in Florida alone, 921,698 patients were participating in the statewide MC program as of August 2025 [1-3]. Given the rapid expansion and persistent evidence gaps on therapeutic benefits and risks, rigorous studies of MC safety and effectiveness are urgently needed [4-6]. Understanding patients' and certifying providers' preferences can help identify high-priority research areas, for example when qualifying medical conditions (QMCs) with limited evidence on medical cannabis safety and effectiveness are overrepresented among certified patients relative to their prevalence in the general population. While prior studies report the most frequently certified conditions in medical cannabis programs, little is known about how conditions for which MC is used aligns with disease prevalences, and hence patient and provider preferences across different conditions to explore medical cannabis as alternative or supplemental treatment [2,7]. The objective of this study was to compare the prevalence of QMCs among medical cannabis patients in Florida to population-based prevalences of these conditions.

Methods

We analyzed data from the Florida Medical Marijuana Use Registry to identify QMC among patients with a MC certification in 2025.

The Medical Marijuana Use Registry includes data from all Florida patients who were certified by a physician to receive MC for one or more legally QMCs (the condition that is intended to be treated). Data is generated via standardized online templates from physicians authorized to order medical cannabis for patients. We excluded records with a missing patient identifier or recorded sex, non-resident status or implausible certification dates (closure date preceding the start date).

We extracted US-specific disease prevalences from peer-reviewed literature and scientific reports, focusing on Florida or, if unavailable, the southern United States, or the United States. We prioritized reports that estimated prevalence across the entire population over reports for selected subpopulation, eg, only those with private health insurance.

Allowing multiple QMCs per patient, we calculated condition-specific prevalences by dividing the number of patients with a specific qualifying medical condition on their MC certification in 2025 by the total number of patients with MC certification in 2025. Patients with multiple QMCs were only counted once in the denominator. We then computed prevalence ratios by dividing the Medical Marijuana Use

XSL•FO
RenderX

Registry data-derived prevalence for each qualifying medical condition by the population-based prevalence of this condition in Florida.

We supplemented reported prevalence with the level of evidence supporting the effectiveness of medical cannabis for each condition based on the National Academy of Sciences, Engineering, and Medicine report from 2017 that comprehensively investigated the evidence for health effects in cannabis and cannabinoids and an update from a recent published comprehensive mapping review [4,5]. Data analysis was conducted with SAS software (version 9.4).

Ethical Considerations

This study was approved by the University of Florida Institutional Review Board (IRB202002648) and Florida Department of Health IRB [2020‐101-UFL]. Informed consent was waived. Data in this research did not involve participant compensation. Confidentiality of the data was ensured through anonymity and stringent security measures.

Results

Of 779,673 qualifying patients in the Medical Marijuana Use Registry in 2025, 8493 (1.09%) were excluded: 7304 (0.94%) nonresidents, 1186 (0.15%) with implausible certification dates, and 3 (<0.01%) with missing recorded sex—yielding 771,180 qualifying patients. The median age was 46.0 years (IQR 34.0‐61.0); 342,258 (44.38%) were female and 428,922 (55.62%) were male. The most prevalent QMCs per 100,000 patients certified to use medical cannabis in Florida in 2025 (Table 1) were posttraumatic stress disorder (PTSD) (81,519), chronic pain (24,081), and multiple sclerosis (12,938). The three QMCs with the lowest prevalence per 100,000 certified patients were acquired immunodeficiency syndrome (AIDS) (174), amyotrophic lateral sclerosis (ALS) (302), and human immunodeficiency virus (HIV) (860). When compared to population-based prevalences, most conditions were more common in the Medical Marijuana Use Registry.

We observed pronounced variations in the ratio of Medical Marijuana Use Registry-based and general population prevalences across QMCs. Notably, ALS demonstrated the highest ratio (77.4), indicating a substantially higher representation among medical cannabis users. This was followed by MS (47.5), and Crohn’s disease (10.4). In contrast, chronic nonmalignant pain (1.0) had similar, and conditions like AIDS (0.6) had lower representation in the Medical Marijuana Use Registry than in the general population. The level of evidence supporting effectiveness of medical cannabis as treatment was substantial for epilepsy, multiple sclerosis, and chronic non-malignant pain, and limited or insufficient for the remainder.

XSL•FO
RenderX

Table . Qualifying medical condition prevalence in the 2025 Medical Marijuana Use Registry and ratio of Medical Marijuana Use Registry-based prevalence estimates to population-based prevalence estimates of qualifying medical condition, and level of evidence. Cross-sectional descriptive analysis of qualifying medical conditions recorded in the Florida Medical Marijuana Use Registry among Florida residents in 2025. For each qualifying condition, the table reports the estimated US-wide prevalence, the registry-based prevalence among Floridians, the ratio of registry-based to population-based prevalence, and the level of evidence supporting therapeutic effects of medical cannabis for that condition.

Qualifying condition	Estimated US-wide prevalence of the condition (per 100,000)	Medical Marijuana Use Registry based prevalence of Floridians with the condition in 2025 (per 100,000)	Ratio of Medical Marijuana Use Registry based prevalence to population-based prevalences	Level of evidence of therapeutic effect
Cancer	591[a]	5917	10.0	None or insufficient[b]
Epilepsy	1700[c]	1441	0.9	Substantial[d]
Glaucoma	1620[e]	1301	0.8	Limited evidence
HIV[f]	704[g]	860	1.2	Limited evidence
AIDS[h]	308[g]	174	0.6	Limited evidence
PTSD[i]	4700[j]	81,189	17.3	Limited evidence
ALS[k]	3.9[l]	302	77.4	None or insufficient
Crohn's disease	258[m]	2683	10.4	None or insufficient
Parkinson disease	412[n]	579	1.4	None or insufficient
Multiple sclerosis	273[o]	12,938	47.5	Substantial[p]
Chronic nonmalignant pain	24,300[q]	24,081	1.0	Substantial

[a]International Agency for Research on Cancer. Cancer Today: GLOBOCAN 2022, version 1.1. Global Cancer Observatory. February 8, 2024 [8]. [May 13, 2026].

[b]Based on effectiveness for disease modification.

[c]Florida Department of Health. Behavioral Risk Factor Surveillance System (BRFSS) data viewer. FLHealthCHARTS [9].

[d]Evidence primarily based on Dravet and Lennox-Gastaut syndrome.

[e]Ehrlich JR et al. [10].

[f]HIV: human immunodeficiency virus.

[g]Centers for Disease Control and Prevention [11].

[h]AIDS: acquired immunodeficiency syndrome.

[i]PTSD: post-traumatic stress disorder.

[j]Goldstein RB et al [12]

[k]ALS: amyotrophic lateral sclerosis.

[l]Mehta P et al [13].

[m]Lewis JD et al [14].

[n]The Lewin Group, Inc. [15]

[o]Wallin MT, et al [16].

[p]Evidence based on patient-reported spasticity outcomes.

[q]Lucas JW, Sohi I [17].

Discussion

In this cross-sectional study, we found that the prevalence of QMCs among certified MC users in Florida differed from general population disease prevalences suggesting that patients' and physicians' perceptions of MC's role as alternative treatment differs across conditions [18]. Indications with limited or substantial evidence for medical cannabis's effectiveness were evenly distributed among under- and overrepresented conditions, indicating that the level of evidence for medical cannabis's effectiveness is not the dominant factor in driving certifications for MC. This finding aligns with previous research suggesting that the evidence base for a qualifying condition appears not to be the main aspect facilitating certifications for medical cannabis [7].

Our study's limitations include its geographic confinement to Florida, inability to ascertain patients' primary reason for MC use as this might be different from certifiable conditions under Florida law, and variation in physician certification practices of registry recorded QMCs. We also relied on prevalence estimates from different sources that were not always specific to Florida and used different approaches. Future studies should test reproducibility in other states, develop state-specific general-population prevalence estimates for QMCs to improve comparisons, and examine the relationship between patients'

XSL•FO
RenderX

QMCs in the Medical Marijuana Use Registry and patients' documented medical histories.

To conclude, our findings define important research priorities regarding the evaluation of both effectiveness and safety of MC for conditions that appear to be important reasons for patients to consider MC. Future studies to investigate the effectiveness and safety of cannabis should consider these research priorities.

Acknowledgments

The authors thank Arthur Rosales for assistance with data analysis. The authors used ChatGPT (GPT-5 Thinking, from OpenAi, 2025) to support grammar, spelling, and language-clarity edits. All AI-assisted edits were reviewed by the authors, who take full responsibility for the final content of the manuscript.

Funding

The study was funded by the Florida state-funded Consortium for Medical Marijuana Clinical Outcomes Research. SJ, NES, and AGW received salary support from the Florida state–funded Consortium for Medical Marijuana Clinical Outcomes Research while the study was conducted. The funder had no involvement in the study design; data collection, analysis, or interpretation; manuscript preparation; or decision to submit the manuscript for publication. No grant or award number was assigned.

Data Availability

A de-identified version of the Florida Medical Marijuana Use Registry can be provided upon request to the Consortium for Medical Marijuana Clinical Outcomes Research only to Consortium affiliated researchers. To request access to this data, please reach out to mmj.outcomes@cop.ufl.edu to complete a Data Request Application (DRA) and a Confidentiality and Data Usage Agreement (CDA).

Authors' Contributions

Conceptualization: AGW, SJ
Data curation: NES, SJ, TA
Formal analysis: SJ
Methodology: AGW, NES
Supervision: AGW
Writing - Original draft: SJ
Writing - review & editing: AGW, NES, SJ, TA

Conflicts of Interest

SJ, NES, and AGW received salary support from the Florida state–funded Consortium for Medical Marijuana Clinical Outcomes Research while this study was conducted. AGW is the director of the Consortium for Medical Marijuana Clinical Outcomes Research. AGW has received funding for research unrelated to this work from Merck. All other authors declare no competing interests.

References

1. Trends of registered patients, authorized physicians, and dispensing of medical marijuana. Consortium for Medical Marijuana Clinical Outcomes Research. 2021. URL: https://mmjoutcomes.org/evidence/medical-marijuana-utilization/ [accessed 2024-11-15]
2. Boehnke KF, Dean O, Haffajee RL, Hosanagar A. U.S. Trends in registration for medical cannabis and reasons for use from 2016 to 2020: an observational study. Ann Intern Med 2022 Jul;175(7):945-951. [doi: 10.7326/M22-0217] [Medline: 35696691]
3. Project MP. Medical cannabis patient numbers. MPP. URL: https://www.mpp.org/issues/medical-marijuana/state-by-state-medical-marijuana-laws/medical-marijuana-patient-numbers/ [accessed 2025-11-07]
4. The Health Effects of Cannabis and Cannabinoids: The Current State of Evidence and Recommendations for Research: National Academies Press; 2017. [doi: 10.17226/24625]
5. Jugl S, Okpeku A, Costales B, et al. A mapping literature review of medical cannabis clinical outcomes and quality of evidence in approved conditions in the USA from 2016 to 2019. Med Cannabis Cannabinoids 2021 Jun;4(1):21-42. [doi: 10.1159/000515069] [Medline: 34676348]
6. Cannabis policy impacts public health and health equity. National Academies of Sciences, Engineering, and Medicine.: The National Academies Press; 2024. URL: https://www.nationalacademies.org/read/27766/chapter/1#ii [accessed 2026-08-12]
7. Boehnke KF, Sinclair R, Gordon F, et al. Trends in U.S. medical cannabis registrations, authorizing clinicians, and reasons for use from 2020 to 2022. Ann Intern Med 2024 Apr;177(4):458-466. [doi: 10.7326/M23-2811] [Medline: 38588545]

XSL•FO
RenderX

8. Tools for visualizing the GLOBOCAN cancer estimates in 2024. International Agency for Research on Cancer. URL: https://gco.iarc.who.int/today/ [accessed 2026-05-13]
9. Behavioral risk factor surveillance system (BRFSS) dashboard. FLHealthCharts. URL: https://www.flhealthcharts.gov/ChartsDashboards/rdPage.aspx?rdReport=BRFSS.Dataviewer&bid=0148 [accessed 2026-05-20]
10. Ehrlich JR, Burke-Conte Z, Wittenborn JS, et al. Prevalence of glaucoma among US Adults in 2022. JAMA Ophthalmol 2024 Nov 1;142(11):1046-1053. [doi: 10.1001/jamaophthalmol.2024.3884] [Medline: 39418040]
11. NCHHSTP atlasplus: tables. Centers for Disease Control and Prevention. URL: https://gis.cdc.gov/grasp/nchhstpatlas/tables.html [accessed 2026-05-13]
12. Goldstein RB, Smith SM, Chou SP, et al. The epidemiology of DSM-5 posttraumatic stress disorder in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions-III. Soc Psychiatry Psychiatr Epidemiol 2016 Aug;51(8):1137-1148. [doi: 10.1007/s00127-016-1208-5] [Medline: 27106853]
13. Mehta P, Raymond J, Nair T, et al. Prevalence of ALS in all 50 states in the United States, data from the National ALS Registry, 2011-2018. Amyotroph Lateral Scler Frontotemporal Degener 2024 Nov;25(7-8):687-693. [doi: 10.1080/21678421.2024.2358786] [Medline: 38826088]
14. Lewis JD, Parlett LE, Jonsson Funk ML, et al. Incidence, prevalence, and racial and ethnic distribution of inflammatory bowel disease in the United States. Gastroenterology 2023 Nov;165(5):1197-1205. [doi: 10.1053/j.gastro.2023.07.003] [Medline: 37481117]
15. The Lewin Group. Economic burden of Parkinson's and atypical parkinsonism in the United States. URL: https://www.michaeljfox.org/sites/default/files/media/document/Economic%20Burden%20of%20Parkinson%E2%80%99s%20and%20Atypical%20Parkinsonism%20in%20the%20United%20States_0.pdf [accessed 2026-05-13]
16. Wallin MT, Culpepper WJ, Campbell JD, et al. The prevalence of MS in the United States. Neurology 2019 Mar 5;92(10). [doi: 10.1212/WNL.0000000000007035]
17. Lucas JW, Sohi I. Chronic Pain and High-impact Chronic Pain in U.S. Adults, 2023. NCHS Data Brief 2024 Oct(518):CS355235. [doi: 10.15620/cdc/169630] [Medline: 39751180]
18. Astin JA. Why patients use alternative medicine: results of a national study. JAMA 1998 May 20;279(19):1548-1553. [doi: 10.1001/jama.279.19.1548] [Medline: 9605899]

Abbreviations

AIDS: acquired immunodeficiency syndrome
ALS: amyotrophic lateral sclerosis
HIV: human immunodeficiency virus
MC: medical cannabis
PTSD: posttraumatic stress disorder
QMC: qualifying medical condition

Edited by A Mavragani, T Sanchez; submitted 16.Dec.2025; peer-reviewed by S Van den Bulck, W Qin; revised version received 02.Jul.2026; accepted 12.Jul.2026; published 08.Sep.2026.

Please cite as:
Jugl S, Smolinski NE, Aydoğan TMI, Winterstein AG
Indications for Medical Cannabis and Their Alignment With US-Wide Disease Prevalences in the US' Largest Medical Cannabis Program: Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e89716
URL: https://publichealth.jmir.org/2026/1/e89716
doi:10.2196/89716

© Sebastian Jugl, Nicole E Smolinski, Tansu Miriam Işılay Aydoğan, Almut G Winterstein. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 8.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Research Letter

Intersecting Sex Work and Substance Use Risk Among Sexual and Gender Minoritized Individuals Recruited Online in San Francisco, California: Survey Results

Sean Arayasirikul[1], PhD; Jarett Maycott[2], MPH; Angela Olivares[1]

[1]The Legacy Center, Joe C Wen School of Population & Public Health, University of California, Irvine, Irvine, CA, United States
[2]San Francisco Department of Public Health, San Francisco, CA, United States

Corresponding Author:
Sean Arayasirikul, PhD
The Legacy Center
Joe C Wen School of Population & Public Health
University of California, Irvine
856 Health Sciences Quad
Irvine, CA, 92697-3957
United States
Phone: 1 9498242680
Email: sean.arayasirikul@uci.edu

Abstract

This research letter examines sex work and substance use associations in a sample of sexual and gender minoritized individuals recruited online in San Francisco, California. This study found that a history of sex work was prevalent and that people with a history of sex work were more likely to recently report using controlled substances and experience domestic violence.

(JMIR Public Health Surveill 2026;12:e83555) doi:10.2196/83555

KEYWORDS

sex work; substance use; domestic violence; intimate partner violence; online recruitment; sexual and gender minoritized community health; social media

Introduction

Sexual and gender minoritized (SGM) communities experience victimization because of structural oppression [1]. The confluence of social exclusion, economic hardship, and the internet supports the formation of a sex work economy [2,3]. People in sex work experience health inequities due to their unique occupational health exposures [4]. Yet there is little research on online help-seeking SGM populations [5]. We address this gap and characterize the prevalence of sex work history among SGM individuals seeking help online and their sexual and substance use–related prevention and treatment needs.

Methods

Ethical Considerations

The study was approved by the University of California, San Francisco institutional review board (20-33169). Participants provided informed consent, could opt out anytime, and received US $30. Data were deidentified to protect participants' privacy and confidentiality.

Study Design and Recruitment

This is a cross-sectional analysis of 409 people recruited online in San Francisco by using social medial advertisements on Facebook, Instagram, and Grindr in 2022-2024. Advertisements sought out potential participants seeking help for substance use, mental health, and HIV. Once they clicked on the advertisement and were directed to a study interest website, they were contacted by staff to screen for eligibility, informed consent was obtained electronically, and they completed a baseline assessment. Eligibility criteria included aged 18 years or older, seeking help for substance use prevention/treatment or related health topics (eg, HIV, mental health), live in San Francisco, identify as a man who had sex with men or a trans woman, and had smartphone access.

Measures and Analysis

We analyzed the following demographics: age, race/ethnicity, gender identity, sexual orientation, HIV status, housing stability,

XSL•FO
RenderX

and socioeconomic status. Recent substance use was measured by asking how many days in the last 30 days, if any, did they use each of the following substances: tobacco, vaping products, binge drinking (having 5 or more alcoholic beverages at the same time or within a couple of hours of each other for people assigned male sex at birth), marijuana, prescription opioids, nonprescription opioids, other prescription drugs, illicit drugs (crack/cocaine, amphetamine/methamphetamine, hallucinogens, inhalants), and injection drug use [6]. Responses were recoded dichotomously for any recent use (yes/no). We measured participants' history of sex work if they had ever had sex with someone in exchange for money, drugs, or shelter (yes/no) [6]. We assessed participants' risk perception of harm when engaging in condomless sex and sex while using drugs or alcohol. Responses were dichotomized into low risk (no risk, slight risk, unknown risk) and high risk (moderate risk, great risk). Self-efficacy to refuse condomless sex was assessed by participants' agreement with the statement, "I could refuse if someone wanted to have sex without a condom or dental dam" [6]. Responses were dichotomized into agree (strongly agree, agree) and disagree (disagree, strongly disagree). Domestic violence was measured by asking participants if anyone with whom they had an intimate relationship had emotionally, physically, or sexually abused them in the last 3 months (yes/no) [6]. STATA version 17 was used to create logistic regression models to test associations between history of sex work and recent substance use and sexual risk outcomes, adjusted for age, gender identity, race/ethnicity, socioeconomic status, and housing stability.

Results

Many participants (168/409, 41.08%) reported a history of sex work. The most frequently reported in the last 30 days was illicit drugs (227/409, 55.50%) followed by marijuana (220/409, 53.79%), binge drinking (160/409, 39.12%), and tobacco (155/409, 37.90%). A majority of participants perceived condomless sex (287/409, 70.17%) and sex while using drugs or alcohol (349/409, 85.33%) as high risk and could refuse condomless sex (361/409, 88.26%). Almost 22.98% (94/409) had experienced domestic violence in the last 3 months (Table 1).

XSL•FO
RenderX

Table 1. Sample characteristics, recent substance use, sexual risk perception, and self-efficacy to refuse condomless sex and behaviors among sexual and gender minoritized people recruited online in San Francisco, California, 2022-2024 (N=409).

	Overall (N=409), n (%)	Sex work status	
		Yes (n=168), n (%)	No (n=241), n (%)
Demographics			
Age (y)			
18-29	103 (25.18)	31 (18.45)	72 (29.88)
30-39	112 (27.38)	54 (32.14)	58 (24.07)
40-49	77 (18.83)	34 (20.24)	43 (17.84)
50+	117 (28.61)	49 (29.17)	68 (28.22)
Race/ethnicity			
White	169 (41.32)	71 (42.26)	98 (40.66)
Latino/a/x/e	113 (27.63)	51 (30.36)	62 (25.73)
Asian, Pacific Islander, and Native Hawaiian	50 (12.22)	13 (7.74)	37 (15.35)
Black	38 (9.29)	16 (9.52)	22 (9.13)
More than one or other	39 (9.54)	17 (10.12)	22 (9.13)
Gender identity			
Cisgender man	327 (79.95)	123 (73.21)	204 (84.65)
Transgender woman or gender expansive	82 (20.05)	45 (26.79)	37 (15.35)
Sexual orientation			
Bisexual	47 (11.49)	22 (13.10)	25 (10.37)
Gay/lesbian	282 (68.95)	105 (62.50)	177 (73.44)
Other	48 (11.74)	22 (13.10)	26 (10.79)
Straight/heterosexual	32 (7.82)	19 (11.31)	13 (5.39)
HIV status			
Not a person with HIV	278 (67.97)	96 (57.14)	182 (75.52)
Person with HIV	131 (32.03)	72 (42.86)	59 (24.48)
Housing stability			
Stable	337 (82.40)	123 (73.21)	214 (88.80)
Unstable	72 (17.60)	45 (26.79)	27 (11.20)
Socioeconomic status			
Above federal poverty line	338 (82.64)	124 (73.81)	214 (88.80)
Below federal poverty line	71 (17.36)	44 (26.19)	27 (11.20)
Recent substance use (last 30 days)			
Used tobacco			
No	254 (62.10)	81 (48.21)	173 (71.78)
Yes	155 (37.90)	87 (51.79)	68 (28.22)
Used vaping products			
No	288 (70.42)	109 (64.88)	179 (74.27)
Yes	121 (29.58)	59 (35.12)	62 (25.73)
Binge drinking			
No	249 (60.88)	116 (69.05)	133 (55.19)
Yes	160 (39.12)	52 (30.95)	108 (44.81)
Used marijuana			

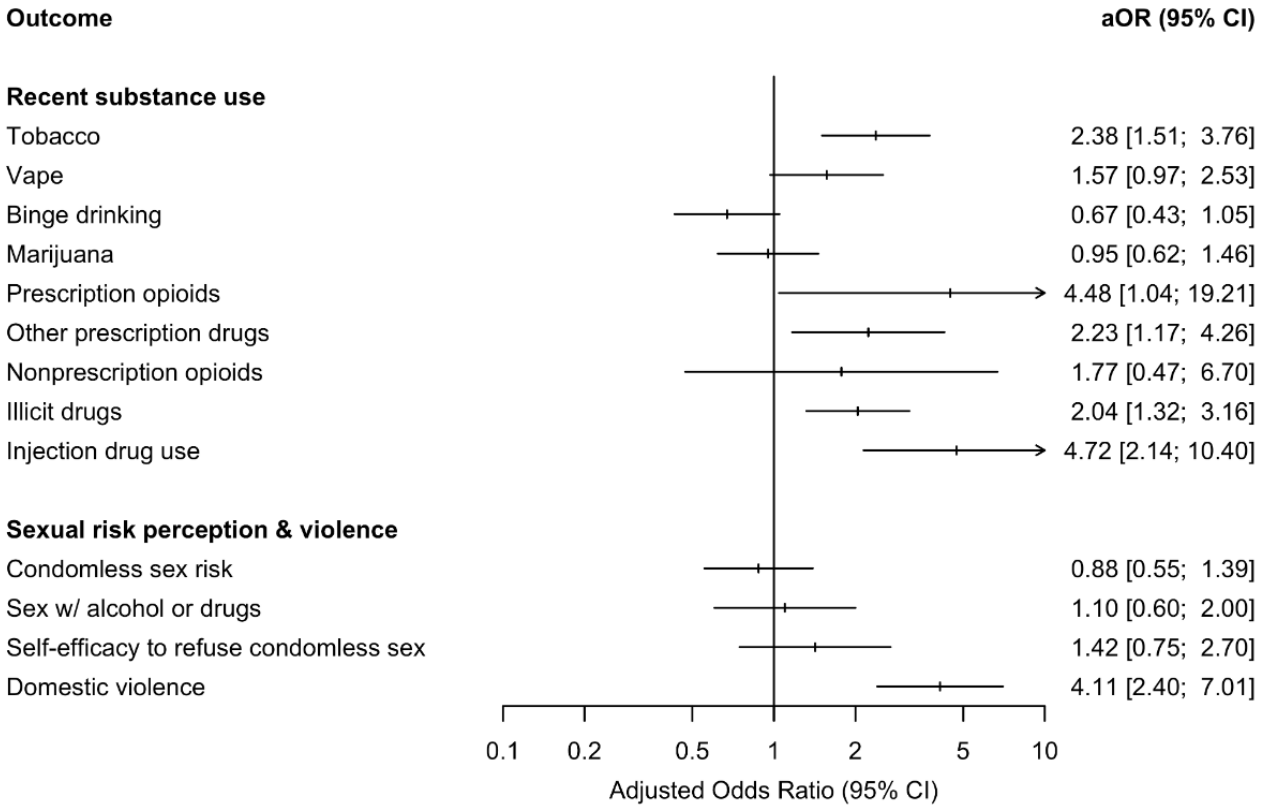
XSL•FO
RenderX

	Overall (N=409), n (%)	Sex work status	
		Yes (n=168), n (%)	No (n=241), n (%)
No	189 (46.21)	77 (45.83)	112 (46.47)
Yes	220 (53.79)	91 (54.17)	129 (53.53)
Used prescription opioid drugs			
No	398 (97.31)	160 (95.24)	238 (98.76)
Yes	11 (2.69)	8 (4.76)	3 (1.24)
Used nonprescription opioid drugs			
No	396 (96.82)	159 (94.64)	237 (98.34)
Yes	13 (3.18)	9 (5.36)	4 (1.66)
Used other prescription drugs			
No	361 (88.26)	141 (83.93)	220 (91.29)
Yes	48 (11.74)	27 (16.07)	21 (8.71)
Used illicit drugs			
No	182 (44.50)	56 (33.33)	126 (52.28)
Yes	227 (55.50)	112 (66.67)	115 (47.72)
Injected drugs			
No	366 (89.49)	135 (80.36)	231 (95.85)
Yes	43 (10.51)	33 (19.64)	10 (4.15)
Sexual risk perception, efficacy and behaviors			
Condomless sex			
High risk	287 (70.17)	122 (72.62)	165 (68.46)
Low risk	122 (29.83)	46 (27.38)	76 (31.54)
Sex while using drugs or alcohol			
High risk	349 (85.33)	142 (84.52)	207 (85.89)
Low risk	60 (14.67)	26 (15.48)	34 (14.11)
Self-efficacy to refuse condomless sex			
Disagree	48 (11.74)	25 (14.88)	23 (9.54)
Agree	361 (88.26)	143 (8.12)	218 (90.46)
Domestic violence (past 3 months)			
No	315 (77.02)	102 (60.71)	213 (88.38)
Yes	94 (22.98)	66 (39.29)	28 (11.62)

Adjusting for potential confounders, people with a history of sex work (Figure 1) were more likely to report using tobacco (adjusted odds ratio [aOR] 2.38, 95% CI 1.51-3.76), prescription opioid drugs (aOR 4.48, 95% CI 1.04-19.21), other prescription drugs (aOR 2.23, 95% CI 1.17-4.26), illicit drugs (aOR 2.04, 95% CI 1.32-3.16), and have injected drugs (aOR 4.72, 95% CI 2.14-10.40) compared to people with no sex work history. People with a history of sex work had 4.11 times the odds of experiencing domestic violence recently compared to counterparts with no sex work history (95% CI 2.40-7.01). No statistically significant association was observed between history of sex work and self-efficacy to refuse condomless sex.

XSL•FO
RenderX

Figure 1. Associations between history of sex work and recent substance use and between sexual risk perception and violence among sexual and gender minoritized people recruited online in San Francisco, California, 2022-2024 (N=409). aOR: adjusted odds ratio.



Discussion

We found that SGM participants with a history of sex work were more likely to report recent use of controlled substances. Although no statistically significant associations exist between history of sex work and sexual risk perception level and self-efficacy to refuse condom use, our findings indicate that domestic violence may be heightened for SGM people with a history of sex work. SGM populations report similar or greater rates of domestic or intimate partner violence compared to their non-SGM counterparts [7], and intimate partner violence is a common experience among people who engage in sex work [8]. Relationship dynamics among SGM partnerships can vary in composition with relationship to sexual and gender roles, power, and how violence is enacted and experienced. A recent systematic literature review found that bidirectional violence was the most common among SGM intimate partners compared to non-SGM partners [9]. The interplay between substance use, violence, and trauma are complex and must be explored in future [10]. This study has limited generalizability because of its design and possible sampling bias toward people seeking help online who may be experiencing heightened risk. Despite this, these findings reinforce the need for substance use prevention efforts to serve people with a history of sex work and address domestic violence as a unique violence exposure. Although stigma remains a significant barrier for people engaged in sex work and those using substances, addressing domestic violence may offer a whole-person alternative to intervene on the material impacts of violence and substance use behaviors that may be both enabling and coping mechanisms of violence. Public health interventions that cross-train substance use providers about domestic violence and sex work literacy, and conversely, are needed to better facilitate screening, referral, and treatment.

Acknowledgments

The authors would like to thank all the participants in this study.

Funding

This work was funded by the Substance Abuse and Mental Health Service Administration (SAMHSA; award H79SP082077). The funder was not involved in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

The datasets generated or analyzed during this study are not publicly available as they contain details that could be used to identify participants but are available from the corresponding author on reasonable request.

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. Hoskin RA. "Femininity? it's the aesthetic of subordination": examining femmephobia, the gender binary, and experiences of oppression among sexual and gender minorities. Arch Sex Behav 2020 Oct;49(7):2319-2339. [doi: 10.1007/s10508-020-01641-x] [Medline: 32198553]
2. Aggarwal NK, Consavage KE, Dhanuka I, Clement KW, Bouey JH. Health and health care access barriers among transgender women engaged in sex work: a synthesis of U.S.-based studies published 2005-2019. LGBT Health 2021 Jan;8(1):11-25. [doi: 10.1089/lgbt.2019.0243] [Medline: 33297834]
3. Jackson KJ, Santos G. Advertising patterns of internet-based male sex workers who have sex with men (MSMSW): the association between LGBTQIA+ events and advertising for work during the 2022 pride season. Am J Mens Health 2023;17(5):15579883231205984 [FREE Full text] [doi: 10.1177/15579883231205984] [Medline: 37822094]
4. Bernier T, Shah A, Ross LE, Logie CH, Seto E. The use of information and communication technologies by sex workers to manage occupational health and safety: scoping review. J Med Internet Res 2021 Jun 24;23(6):e26085 [FREE Full text] [doi: 10.2196/26085] [Medline: 34185001]
5. McDermott E. Asking for help online: lesbian, gay, bisexual and trans youth, self-harm and articulating the 'failed' self. Health (London) 2015 Nov;19(6):561-577. [doi: 10.1177/1363459314557967] [Medline: 25413341]
6. SAMHSA. National Survey on Drug Use and Health: Summary of Methodological Studies, 1971-2014. Rockville, MD: Substance Use and Mental Health Services Administration; Nov 2014.
7. Kurbatfinski S, Letourneau N, Marshall S, McBride D, Novick J, Griggs K, et al. Myths and misconceptions of intimate partner violence among sexual and gender minorities: a qualitative exploration. Front Sociol 2024;9:1466984 [FREE Full text] [doi: 10.3389/fsoc.2024.1466984] [Medline: 39758189]
8. Zemlak JL, Alexander KA, Wilson D, Singer R, Williams JS, Sherman SG. Sex workers' experiences of screening for intimate partner violence. J Obstet Gynecol Neonatal Nurs 2024 Jul;53(4):397-405. [doi: 10.1016/j.jogn.2024.02.005] [Medline: 38490264]
9. Machado A, Sousa C, Cunha O. Bidirectional violence in intimate relationships: a systematic review. Trauma Violence Abuse 2024 Apr;25(2):1680-1694. [doi: 10.1177/15248380231193440] [Medline: 37594220]
10. Morton S, Curran M, Barry O'Gorman M. Adverse childhood experiences, domestic violence and substance misuse: an action research study on routine enquiry and practice responses. Front Psychiatry 2022;13:892849 [FREE Full text] [doi: 10.3389/fpsyt.2022.892849] [Medline: 35903635]

Abbreviations

aOR: adjusted odds ratio
SGM: sexual and gender minoritized

Edited by A Mavragani, T Sanchez; submitted 04.Sep.2025; peer-reviewed by MG Curtis, RQ De Torres; comments to author 03.Dec.2025; revised version received 30.Dec.2025; accepted 02.Jan.2026; published 11.Feb.2026.

Please cite as:
Arayasirikul S, Maycott J, Olivares A
Intersecting Sex Work and Substance Use Risk Among Sexual and Gender Minoritized Individuals Recruited Online in San Francisco, California: Survey Results
JMIR Public Health Surveill 2026;12:e83555
URL: https://publichealth.jmir.org/2026/1/e83555
doi:10.2196/83555
PMID:41672477

©Sean Arayasirikul, Jarett Maycott, Angela Olivares. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 11.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Investigating the Impact of Online Abortion Myths on Health Care Providers and Advocates: Interview Study

Rachel E Moran[1], PhD; Julia Simoes[2]; Taylor Agajanian[3]; Izzi Grasso[1]; Anna Beers[4], PhD; Amanda Lock Swarr[5], PhD; Emma Spiro[1], PhD; Anna Swan[1], PhD

[1]Center for an Informed Public, Information School, University of Washington, 4000 15th Ave NE, Seattle, United States
[2]Department of Communication, University of Washington, Seattle, United States
[3]School of Communication, Northwestern University, Chicago, United States
[4]University of North Carolina at Chapel Hill, Chapel Hill, United States
[5]Department of Gender, Women & Sexuality Studies, University of Washington, Seattle, United States

Corresponding Author:
Rachel E Moran, PhD
Center for an Informed Public, Information School, University of Washington, 4000 15th Ave NE, Seattle, United States

Abstract

Background: Abortion is a common and safe medical intervention with a long history of practice in the United States. Despite this, inaccurate and misleading information persists, including falsehoods about the accessibility, legality, safety, and lived experience of abortion-related health care. Further, the prevalence and circulation of abortion myths online and offline have intensified following the overturning of Roe vs Wade in May 2022. While myths surrounding abortion have been widely documented, limited research has examined how these myths shape the everyday work of abortion health care providers and advocates, particularly in relation to patient–practitioner shared decision-making (SDM).

Objective: This study examines how abortion myths affect abortion health care professionals—including practitioners in states where abortion access is legal, and advocates connecting patients with care in states with limited abortion health care. Through interviews with advocates and professionals, we document the nature of circulating myths, the common information sources where myths are being spread, and the routes professionals take to debunk myths when they arise, or, similarly, the barriers that limit their ability or desire to debunk abortion myths.

Methods: We conducted in-depth qualitative interviews with abortion health care professionals across the WWAMI (Washington, Wyoming, Alaska, Montana, and Idaho) medical region. Interviews explored the types of abortion myths encountered in clinical settings, how providers respond to misinformation during patient interactions, and the perceived impact of myths on care delivery and SDM processes. Data were analyzed using an iterative constant comparative thematic analysis of the interview transcripts.

Results: Participants reported encountering a wide range of abortion-related myths from predominantly online (social media) and some offline sources, including misinformation about safety, legality, fertility impacts, and procedural experiences. Providers described dedicating substantial time to correcting misinformation, addressing fear and confusion, and rebuilding trust within clinical encounters. Myths were found to complicate SDM by shaping patients' expectations, emotional responses, and perceived options for care, in addition to taking up time and resources for providers of care to counter damaging myths.

Conclusions: Abortion myths have tangible effects on clinical practice and SDM, placing additional communicative and emotional labor on providers and potentially undermining patient autonomy. Addressing misinformation is therefore critical not only for public understanding but also for supporting equitable, patient-centered abortion care in a post-Roe landscape.

(JMIR Public Health Surveill 2026;12:e77050) doi:10.2196/77050

KEYWORDS

abortion; health myths; health information seeking; social media; health misinformation

Introduction

Background

Abortion is a common and safe medical intervention with a long history of practice within the United States; yet, misleading information around abortion (and pregnancy viability more broadly) persists [1,2]. Inaccurate information about abortion pertains to the realities of abortion health care (ie, information about medical interventions, medication, and aftercare) [3], falsehoods about the long-term impacts of having an abortion, and misleading information around the legality of, and access to, abortion health care [4]. The overturning of abortion protections upheld by Roe vs Wade in the 2022 Dobbs vs

XSL•FO
RenderX

Jackson Women's Health Organization ("Dobbs" for short) has exacerbated the conditions of uncertainty under which rumors and misinformation thrive [5], leading health researchers to call abortion misinformation "the next infodemic" [6]. Accordingly, given the advent of digital information-seeking for health, individuals seeking abortion information online may be met with a polluted information environment in which it is difficult to parse between falsehoods and scientific truth. Difficulties in information-seeking undermine an individual's ability to make informed health decisions and have additional negative ramifications for their relationships with health care providers.

To better understand the ramifications of online (and offline) health myths on patient-practitioner relationships, this paper examines the impact of abortion-related myths post Dobbs from the perspective of professionals involved in abortion health care and related advocacy, including abortion practitioners and advocates. Accurate knowledge is vital for shared decision-making (SDM) between patients and practitioners to occur, with persistent myths undermining the ability of practitioners to establish relationships of trust with patients and base decisions on a mutually agreed-upon knowledge foundation. Through in-depth interviews, we explore the types of myths health care professionals encounter and how they navigate abortion-related myths when they arise, and document their impacts. Our analysis offers insights into the interconnection among legal confusion, uncertainty, and the spread of misinformation, and highlights productive routes health care professionals use to counter the harmful effects of health myths and provide a foundation for SDM.

Extant Research

Health Misinformation

The spread of misleading health information is a longstanding concern for public health professionals and for science communication more broadly [7]. Further, the politicization of health-related topics has exacerbated the spread of misleading information [8], reducing the effectiveness of science communication around health and leading to significant public health impacts [9,10]. This is particularly acute in relation to abortion health care [11,12]. The overturning of historical protections of abortion access under Roe vs Wade marks a "new legal quagmire" for abortion access [13] and renews public conversation and contestation about abortion health care. Moreover, the Dobbs decision came at a time of renewed political and legislative attacks on bodily autonomy that leverage scientific misinformation as core evidence [14].

Extant research into the spread of misleading health claims primarily uses the term "misinformation" to capture the unintentional spread of nonfactual health information and "disinformation" to capture nonfactual information intentionally spread to mislead [15]. Similarly, public health scholars have used the term "myth" to describe the spread of health-related statements that are widely disseminated but have either no supportive scientific evidence or strong evidence to the contrary [16]. Given the politicization of the term "misinformation" and the nuances of scientific information that cannot necessarily be binned into the dichotomy of "true" or "false" [11], we use the term "abortion myths" throughout this work. Further, "myth" captures information that can be seen as misleading without attaching intentionality to its spread. Within the context of abortion-related information, there have historically been clear instances in which institutions have weaponized false information to intentionally discourage pregnant people from seeking out abortions. However, the complexities of the online space mean information provenance and the intentions of every poster and/or sharer of information are difficult to track. Accordingly, while we (and our interviewees) often refer to information as "misleading" given its nonfactual nature, the term "myth" affords space to avoid specific claims on intention.

Abortion Myths

The scope of abortion-related myths and their spread in online and offline environments has been well-documented. Recent research, for example, highlights the prevalence of misleading information about abortion and herbal "at-home abortion remedies" on video-sharing site TikTok (ByteDance) [17]. Earlier work explored the spread of inaccurate information related to abortion and breast cancer, infertility, and mental health problems by pregnancy crisis centers [18,19]. Further research has highlighted the role of abortion myths in pro- and antiabortion advocacy. The work by Gantt-Shafer [20] examines the impact of myths on the overall climate around abortion, finding that media representations or misrepresentations of abortion can normalize rage as an acceptable response to the existence of abortion and further the demonization of people who seek out this form of health care.

Documenting the perspectives of health care providers is particularly necessary post Dobbs as abortion health care professionals navigate a new legal climate around abortion health care [21]. Myths that arise in conjunction with new laws exacerbate confusion around legality. Location-specific research points to these complexities. For instance, a report from the Idaho Coalition for Safe Healthcare found that Idaho—a state that outlawed abortion following the Dobbs decision—lost 22% of its obstetrics and gynecology (OB-GYNs) in the year following the decision. Doctors in the state have argued that the scope of Idaho's abortion ban is unclear, creating confusion on "what care they can and can't provide patients" [22]. As such, it is important to capture how legal confusions and the rumors that arise out of this confusion impact provider well-being.

Accordingly, in the wake of the Dobbs decision, a wave of perspectives papers in public health and medical journals have called for the active participation of health care professionals in reproductive justice advocacy work [23,24]. Within this has been a call for practitioners to consider illuminating abortion myths as part of the professional obligations of health care providers [25]. However, little research exists documenting the specific impacts of abortion myths these health care providers encounter and how they attempt to debunk misleading information about abortion when it arises.

Theories of SDM

While SDM has become the norm for patient-practitioner relationships within the United States [26], there is no singular definition describing the process. According to Resnicow et al

XSL•FO
RenderX

[27], most models of SDM consider it a "collaborative process by which both the patient and health care provider contribute to medical decisions, with patients being full partners in decisions about their care..." In defining the underpinning theoretical mechanisms supporting SDM as a relational care route, Resnicow et al, present two relevant models—self-determination theory (SDT) and the difficulty × motivation matrix. SDT [28] proposes a continuum of motivational regulation—ranging from controlled motivation, that is, behavioral change emanating from external pressure and guidance, to autonomous motivation—self-determined change through internal motivation [28]. Further, SDT advances three fundamental human needs—competence, relatedness, and autonomy—that, once met, enhance behavior change by affording more autonomous motivation. Building on this, Resnicow et al [27] point toward their previously established difficulty × motivation matrix [29] to explain how the effectiveness and style of SDM needed may also be impacted by the degree of difficulty the behavioral change requires. More complex, repeated, and long-term health behaviors, for example, smoking cessation, dietary change, require more patient autonomy than one-off or less complex health decisions. Taken together, the theoretical underpinnings of SDM highlight the utility and necessity of patient autonomy in health decisions and foreground the role of information access (foundational for patient competency) in effective SDM.

Empirical studies of abortion health care examine the potential role of SDM in ensuring effective health outcomes and patient satisfaction [30]. Fagot [30] argues that SDM is an effective care strategy for patients seeking medication abortion, but one that is "multifaceted and difficult to implement in health care settings." Difficulty arises due to systemic barriers to accessing abortion care, such as cost, logistical barriers, and information overwhelm. Studies outside the US context confirm similar issues. Research by Frederico et al [31] in Mozambique highlighted how gender and power inequalities and a lack of knowledge at the national level surrounding the legality of abortion limit women's autonomy in decision-making around abortion health care. Shared-decision making thus requires health practitioners (and broader advocates) to uncover and attend to barriers that undermine the ability of patients to autonomously engage with health information and thus effectively uphold their role within the shared patient-practitioner decision-making relationship. One such barrier is increasing distrust in medical professionals [32], which is further connected with the rise of medical misinformation as individuals look elsewhere (increasingly online) for answers to their health queries, and the persistence (or perception) of health misinformation causes further distrust in health care providers [12]. However, there exists a dearth of research examining (1) the spread of abortion-related myths in the US-context post Dobbs, and (2) exploring how such myths impact practitioners in their ability to foreground SDM in their patient relationships.

To address current gaps in research, this project seeks to answer the following research questions (RQs):

- RQ1: what types of myths do providers and advocates encounter?
- RQ2: how do they navigate myths when they arise?
- RQ3: what impacts do they see myths having on abortion health care?

Methods

Sampling and Recruitment Methods

The research team conducted 14 semistructured, in-depth interviews with abortion health care providers and advocates between October 2023 and January 2024. Interviewees were recruited via email through snowball sampling [33] using email listservs and personal connections with researchers and practitioners in the WWAMI (Washington, Wyoming, Alaska, Montana, and Idaho) medical region, plus the neighboring state of Oregon. The inclusion criteria for interviews defined appropriate interviewees as individuals working or who have worked professionally within the US abortion space—either as licensed medical or adjacent health care professionals or reproductive rights advocates affiliated with professional organizations—who attest to having been impacted in some way by the spread of false or misleading information about abortion. Meeting the latter criteria was determined within the recruitment process through initial email advertisements (asking for participants who had encountered misinformation about abortion within their professional work) and in confirmation through sharing (and confirmation of awareness) of the research objectives within informed consent. We faced difficulties in recruiting health care providers in states, notably Idaho, where abortion access is banned. To expand our potential interview sample, we recruited advocates working within abortion and health care organizations to advance abortion access, in addition to health care practitioners who either practiced abortion lawfully under their home state's laws or provided general health care (usually OB-GYN and family-planning-related care) in banned states. Snowball sampling began with abortion health care providers within our university-affiliated hospital system, who then referred us to email listservs for abortion-related professionals and directly connected us with practitioners from other states within their professional networks. The email listservs used were aimed at either members of professional organizations associated with abortion and related health care or broader organizations and collectives engaged in work to advance abortion rights. Our resultant sample (Table 1) includes health care practitioners—in the main OB-GYNs and 2 primary care physicians and 2 abortion advocates—both working in abortion-specific organizations connecting individuals in banned states to health care and engaging in legislative advocacy.

XSL•FO
RenderX

Table . List of participants with state location and professional role.

Participant ID	State	Professional role
001	MT	Health care practitioner
002	WA	Health care practitioner
003	WA	Health care practitioner
004	WA	Health care practitioner
005	OR	Health care practitioner
006	MT	Health care practitioner
007	MT	Health care practitioner
008	WA	Health care practitioner
009	WY	Abortion advocate
010	MT	Health care practitioner
011	ID	Health care practitioner
012	WY	Health care practitioner
013	ID	Abortion advocate
014	AK	Health care practitioner

Ethical Considerations

This project was reviewed by the Institutional Review Board at the University of Washington and deemed exempt from federal human participants review (STUDY00018372). The University of Washington's Human Subjects Division determined that this interview study qualified for exempt status (category 2). Accordingly, formal written informed consent was not obtained. Instead, interviewees were provided with a statement blurb about the purpose of the research project, how data (interview transcripts) would be obtained, anonymized, and analyzed, and contact details for this study's team (Multimedia Appendix 1). Study details were emailed to potential participants before scheduling an interview time, and we repeated this at the beginning of interviews (before recording), at which time this study's team repeated that participants could remove themselves and their data from this study at any time (Multimedia Appendix 1). Given the sensitive nature of this study's subject matter, interviewee privacy was of the utmost importance. Interviewees were given a number identifier that was not attached to their name and/or contact information, and only two biographical details—their state and a generalized professional title (Table 1)—were collected to ensure confidentiality. All participants were compensated for their time with a US $50 gift card.

Research Team Positionality

All interviews were conducted by the lead author, with a secondary interviewer from the broader research team. The research team consists of social scientists from different disciplines researching issues related to health information-seeking, reproductive health and advocacy, and social media. Researchers were all PhDs trained in qualitative research methods, including interviews and thematic analysis. All of the research team were employed by the University of Washington either as full-time academic researchers or as doctoral students and research assistants. Researchers within the team identify as either female or nonbinary.

Interview Protocol

Interviews were semistructured, following a general agenda with intentional space for participants to direct conversation based on what they wanted to share. A protocol was developed by the research team that focused on capturing the professional experiences of interviewees in relation to the informational climate surrounding abortion to answer the core RQs related to common myths (RQ1), how they navigate these myths in their professional work (RQ2), and the broader impacts they see abortion myths as having on abortion health care (RQ3). Specific questions in the agenda pertained to experiences of abortion myths within patient interactions, considerations of public knowledge around abortion health care and prominent myths and misconceptions, experiences of patient information-seeking about abortion, and strategies interviewees deploy when they encounter abortion myths in their patient interactions (see Multimedia Appendix 2 for interview protocol). Interviews were conducted via Zoom (Zoom Communications, Inc) and lasted, on average, 45 minutes.

Audio recordings were taken for transcripts to be obtained and then destroyed. The research team met to debrief after each interview and discuss emergent themes before formal coding. We continued interviews until the research team agreed that we had achieved data saturation [34], that is, we reached a point of "informational redundancy" [35], where new interviewees provided similar answers to several prior interviews.

Data Analysis

Researchers conducted a grounded thematic analysis of the interview transcripts [36], starting from the interview dataset to ascertain common themes, then codifying themes into a specific codebook applied to interview transcripts to gauge prevalence, capture any contradictions or nuances between

XSL•FO
RenderX

perspectives, and surface appropriate descriptive quotes from interviews. An initial team of 4 researchers used qualitative data analysis software Atlas.TI to open-code the full set of transcripts to identify initial themes and points of interest. In the open-coding phase, researchers read through transcripts and labeled data (usually single sentences or brief paragraphs) of interest to the core RQs. Labels consisted of in vivo or descriptive codes. The full research team then met and used visual collaboration software Miro to undertake a collaborative clustering activity in which researchers colocated similar codes, identified salient themes from the clusters using a constant comparative method, and discussed connections between clusters. A smaller group of researchers then used the results of the clustering activity to form the basis of a codebook (Multimedia Appendix 3), which was used by the full research team to code all 14 transcripts. The findings section highlights the most pertinent themes that emerged from this analysis.

Results

RQ1: What Types of Myths Do Providers and Advocates Encounter?

Throughout the interview process, providers and advocates highlighted several central myths they encounter in their work that are predominantly procedural (around reproductive health and abortion as a medical intervention) or legal. Table 2 highlights the core thematic codes emergent from interview analysis regarding the types, sources, and routes to combating the central myths encountered by providers and advocates.

Table . Core themes emergent from interview analysis regarding major myths, sources of myths, and routes to combating myths.

Core theme	Subthemes	Description
Sources of misinformation	Offline	Descriptions of offline sources of misinformation, for example, friends and family, church, legal framework, etc
Sources of misinformation	Online	Description of online sources of misinformation (not social media), for example, Google, websites
Sources of misinformation	Social media	Description of social media sources of information, for example, TikTok, Facebook (Meta), X (formerly known as Twitter), etc
Sources of misinformation	Crisis pregnancy centers	Explicit mentions of the role of crisis pregnancy centers in “care” giving and information-seeking
Types of misinformation	Procedural misinformation	Any misinformation related to abortion procedures and the risks and pain involved
Types of myths	Future risks and impacts	Misinformation related to potential future impacts of having an abortion, for example, breast cancer, mental health, and infertility
Types of myths	Legal misinformation	Misinformation related to the legality of getting, delivering, or assisting with an abortion
Types of myths	Misinformation in options counseling	Misinformation given during counseling, for example, around options available to patients
Types of myths	Misinformation regarding the reputation of the reproductive health field	Mentions of misinformation related to the reputation of health care providers (providers are “bad people”) and institutions such as Planned Parenthood
Strategies for combating myths	Barriers to addressing misconceptions	Mentions of the barriers that exist around addressing abortion misconceptions (ie, limited time with patients)
Strategies for combating myths	Communication strategies	Mentions of the ways providers communicate with patients about abortion, their misconceptions about abortion, and how they build interpersonal trust with patients
Strategies for combating myths	Evidence-based information sharing	Mentions of how providers share information with patients as a way to combat misconceptions
Strategies for combating myths	Using social media	Mentions of the use or potential of social media to combat misconceptions among a broader audience
Strategies for combating myths	Seeking out different perspectives	Intentionally seeking out differing perspectives about abortion to build understanding about what distinct ideas or misconceptions exist around abortion

XSL•FO
RenderX

Procedural myths and misconceptions emerge mainly around the safety of abortion, both medication and procedural. Providers stressed that while much of the discourse around abortion presents it as a painful, dangerous, or risky procedure, it is exceedingly safe, especially compared to pregnancy, which can be very high-risk. One provider noted:

> *I think also there's a lot of disinformation about the safety of abortion in general. Where we know that it is much, much safer than continuing a pregnancy, and there's generally no significant risk of infertility or infection, or significant bleeding that is associated with having an abortion.* [D004]

Additionally, multiple providers noted that there are large gaps in public knowledge about medication abortion—mifepristone and misoprostol—which can be safely self-administered. D008 argued:

> *I think there's a lot of belief about abortion needing to be medicalized. And I think, like there are like organizations and then also people in the family planning community providing this care, who are trying to do more work and supporting people who choose to have abortions on their own as far as self-sourcing medication abortion, or self-managing. But I think there's—I guess I worry about patients who choose that even coming into the system with complications, or, you know, bleeding, or whatever it might be and then providers being like, Oh, you're so irresponsible for doing that. Whereas, like actually, abortion is really safe and you know, there are studies showing people do know how far along they are in pregnancy. They can take a questionnaire to self-assess. You know their risk for ectopic and like, safely have self-managed abortion.*

While procedural myths and misconceptions most commonly come from patients, myths around legality are also pervasive among the practitioners and advocates themselves—including myths about what was or was not legal and where. D003 noted:

> *Yeah, I mean, I think that it makes, it may impact, not just patients, but colleagues, staff. You know, misunderstandings about the legality of abortion like, persist in my own department about like what Washington State's laws mean, in fact. And so like, if we can't understand it as physicians and clinicians like, how are we gonna expect, you know, like other folks to understand it, if you know, if it's like, literally part of our job.*

However, while mounting legislation in some banned and restricted rural states often leads to confusion and/or misinformed assumptions about access, D009 also noted some of these assumptions predated the Dobbs decision and were rooted in preexisting rural community attitudes about abortion:

> *It's understandable that people don't quite know if it's legal or not. But even before that a lot of people just had an assumption. Oh, I didn't think it was legal in Wyoming, you know.*

Interviewees were also asked about the source of the myths and misconceptions they encountered in their industry and in their encounters with patients. Offline sources were identified as common sources of myths around abortion by our participants, with crisis pregnancy centers (CPCs) noted most frequently, followed by interpersonal relationships and other offline sources. On CPCs, 1 provider noted:

> *I mean, it's extremely angering whenever I even talk about them, because I do agree like not being a legal person myself, I don't completely understand the legal reasons they're allowed to exist and provide this false counseling and information to patients regarding like patients own health. Like people, are literally like pretending to put an ultrasound on, don't even know how to operate an ultrasound machine like lying to patients about how far along they are in pregnancy like trying to convert them to religions.*

Online sources were also highlighted by interviewees, including media such as informational websites (sometimes run by CPCs), news media, and social media. Many of the providers interviewed discussed actively avoiding social media (and even online news media in some cases) for several reasons, including a lack of free time but predominantly due to the volatility of social media and an impression that the amount of online misinformation would be overwhelming. One provider noted:

> *I haven't been on social media, for you know, many, many years, and so, in fact, that probably means that I'm not exposed to, you know, sources of kind of misinformation that other people are, and certainly patients might be exposed to.* [D006]

The aversion to online media expressed by participants could help explain why offline sources emerged as the leading source of abortion myths across interviews, despite extant research documenting the importance of social media in information-seeking. Moreover, it highlights a potential gap in knowledge for practitioners, as protecting their own mental well-being through staying away from social media also results in their (known) lack of awareness of the full extent of myths circulating on online platforms. Despite their personal avoidance of social media, interviewees did highlight the emergence of social media as a source of reproductive health myths from interactions with their patients, with specific mention of TikTok and Reddit.

In talking through the types of myths that were pervasive in their practice and/or geographic communities, interviewees reasoned why certain myths retained salience within the community or within the US more broadly, such as religion and language. The importance of religion in the framing of life and the sanctity of life led to a belief in myths around gestational limits, the availability of certain procedures, future fertility, and health, in addition to stigma and shame. One provider argued:

> *I think fear tactics are the most impactful where there is so much misinformation about the safety of abortion, and especially in religious worlds, the idea that abortion is a sin, or that someone will go to hell if they have an abortion, I think that is extremely impactful in people of faith, and in making them either*

XSL•FO
RenderX

> *feel extremely confused or, or making the decision very difficult for them, or actually making them change their decision. And I think that that is very unfortunate, and very negatively impactful both for the patient as well as for the providers taking care of them.* [D004]

Interviewees also discussed the weaponization of language around abortion, especially the use of sensationalized nontechnical terms which they saw as contributing to misleading narratives of abortion health care. For example, D014 discussed the use of the term "partial-birth abortion," which they implored is "not even a thing." According to The Guttmacher Institute, the term "partial-birth abortion" is a medically inaccurate term coined by the National Right to Life Committee [37]. The Cornell Legal Information Institute defines the practice as "an abortion in which the person performing the abortion— (A) deliberately and intentionally vaginally delivers a living fetus...for the purpose of performing an overt act that the person knows will kill the partially delivered living fetus; and (B) performs the overt act, other than completion of delivery, that kills the partially delivered living fetus" [38].

RQ2: How Do They Navigate Myths When They Arise?

Interviewees expressed hesitancy at debunking abortion-related myths through outright fact-checking. D005 argued that they "don't like to try to fact check them [patients]. I just try to kind of nudge them towards openness."

Part of this hesitancy can be attributed to the intersection of the salience of abortion myths with growing distrust in medical institutions. Several interviewees highlighted the parallels between COVID-19 and reproductive health myths in the postpandemic climate, emphasizing an antigovernment, antiregulatory stance that contributes to myths about public health—propagated through nonmedical, sometimes adversarial sources. D005 discussed changing their entire approach when talking to patients under this newfound climate of medical distrust, while D008 noted their difficulty gaining back trust despite being an expert in their field:

> *And so there's the unknown. And then individuals hear things, and they think that's the truth. And then it's just misinformation. And then I think you also have a community of people that might be anti-medicine, anti-physician ever since Covid. And if anything does come out that has some type of credibility behind it, it is immediately dismissed because they just can't even go there because of lack of trust. And that's a whole other thing is trying to gain back trust of people to hear us to say, this is, this is how it works, this is how it is.*

Moreover, D008 argued institutional distrust may be particularly salient for marginalized patients seeking care, noting:

> *I think like fear is a much stronger emotion than others. And so I think it's easy if you hear something that instills fear in you like I think it's easy to like get stuck on that idea or that thought and focus on that especially if you are somebody who might not even trust the healthcare system like, who can you trust, you know, like, I think, patients for good reasons, like because of historical coercion, you know, forced sterilization, other things the healthcare system has perpetuated on marginalized populations.*

Such findings echo long-standing concerns around medical mistrust and its impact on maternal medicine, particularly within Black communities [39], in addition to rising concerns about how the overturn of Roe vs Wade will impact patient-provider trust [40].

Instead of directly debunking myths, interviewees advocated that a core part of their role (in health care provision and advocacy) was to understand the motivations underpinning the sharing of and belief in myths and to redirect patients toward better information. Interviewees saw this approach as a route to "non-judgmental care" (D006). To understand patients' information-seeking and reorient them away from misleading information, interviewees engage in several communications strategies, including providing evidence-based information from a variety of vetted online and offline sources and engaging in motivational interviewing (MI) techniques. Several interviewees highlighted the utility of providing patients with handouts or directing them to online resources in countering myths and allowing patients to access accurate information in their own time. Interviewees named the ACOG (American College of Obstetricians and Gynecologists), the National Abortion Federation, Planned Parenthood, and a host of locally specific abortion organizations as useful sites they directed patients toward. Several interviewees highlighted how overwhelming information-seeking and decision-making around abortion health care can be and how this necessitated that practitioners provide additional resources beyond the clinical appointment. As one interviewee argued:

> *I usually honestly print out an ACOG bulletin for them [patients]. I feel like that's helpful for me to be like, "hey this information is overwhelming when you're going through a conversation like this on a clinic visit"...I'll give them information from ACOG and then usually try to have them come back in a few days.* [D007]

In addition, MI, that is, asking open questions to explore a patient's perspectives and underlying rationale, emerged as a core practice across interviews. Interviewees summarized MI as follows:

> *I'll ask a little more about that belief, how they came to that belief, what their source was, how they feel about that particular piece of information.* [D004]

MI laid a foundation upon which practitioners felt more comfortable to dispel abortion-related myths. Interviewees saw the practice as less dismissive than simply fact-checking and afforded them the ability to build trust with their patients. Moreover, MI allowed abortion health care professionals to identify which patients are likely to change their minds, and when. D001 argued:

> *I basically do motivational interviewing like: "how willing and ready are you to think differently on this? Zero? Okay, moving on, not going to waste my breath.*

XSL•FO
RenderX

> *It's not worth my energy...[I have] limited time and resources with a patient that you have to choose to spend them wisely."*

However, interviewees also named several complicating factors when navigating abortion-related myths when they arise in patient interactions. As highlighted previously, limited time with patients curtailed their ability to debunk myths both in terms of logistics—short consultation periods meant that debunking myths is not always a priority in caregiving—and in limiting the ability of health care professionals to build the trust necessary for debunking to be impactful. D003 summarized the impact of this limited interaction, explaining why it was unsurprising that patients trusted health care professionals less than those in their everyday circle:

> *I might be somebody with some fancy letters behind my name, but they're just meeting me and so if their auntie, their religious leader, their sibling, their best friend, you know some super trusted person. Like that carries a ton of weight with people and so it might just be that they may or may not believe me.*

Further, health care practitioners spoke about not always being aware of what information patients encountered, mostly due to them intentionally staying away from social media. The (nonmedical) abortion advocates interviewed were more aware of information spreading on social media, whereas common across the (medical) practitioners interviewed was a desire to stay away from social media and avoid abortion-related news coverage due to the mental health toll of being immersed in the subject matter in both their professional and personal lives. Practitioners knew from patient interactions that social media is playing a large part in information-seeking but had limited direct knowledge of the kinds of information that exist online. In these conversations, several interviewees named TikTok as a source of patient information-seeking. D014 argued:

> *TikTok is now the new doctor, [it] used to be Dr. Google. Now it's Dr. TikTok. I don't know how to dance! How am I supposed to compete with TikTok now that's where people are getting their information?*

As such, interviews demonstrated uncertainty among practitioners in engaging directly in fact-checking of abortion-related myths, predominantly because they do not always know the extent of misleading information a patient may have been exposed to and have not had the chance to sufficiently build trust and rapport with patients for fact-checking to be effective. D008 encapsulated such difficulties:

> *I think as a provider...I only know the surface level of what patients are even wondering about, because I wonder if patients even feel comfortable sometimes asking the questions they're really thinking.*

RQ3: What Impacts Do They See Myths Having on Abortion Health Care?

Several interviewees noted how their perspectives were limited to the patients they encountered and that a hidden population of pregnant people who choose not to seek out health care providers or abortion advocates may be impacted by misleading information. D004 characterized antiabortion messaging as misinformation intentionally targeted to dissuade people from having abortions:

> *I think that there's a very intentional conscious output of misinformation that's driven by people who are against abortion access and do not want people to be having abortions.*

Further, interviewees highlighted how, despite sharing misinformed views on the safety of abortion, the potential of lasting consequences, the legality, and even the pain associated with procedures, patients still chose to go forward with terminations.

> *But patients have those fears because of that disinformation and present for care anyway, because they do not want to be, and cannot be, pregnant at that time.* [D003]

For those who do access care, abortion myths may act to amplify fear, cement stigma, and increase the emotional impacts of abortion health care [41]. In addition to the impacts of negative sentiments associated with abortion on patients, interviews highlighted how abortion myths have similarly led to negative implications for providers. Providers are subject to attacks on their credibility, on a professional and individual level, that emerge from misunderstandings about the nature of their abortion-related work. In addition to negatively impacting their well-being, such misinformed attacks have also led health care professionals to avoid abortion-related work. D001 argued that they would love to do abortions, but it "would be professional suicide where I'm at."

Interviewees spoke extensively about their physical safety and ensuring the physical safety of fellow abortion health care providers and advocates. One interviewee described a recent incident in which someone used a "shotgun to blast two holes in our [the clinic's] front door" (D006), another provider's clinic was burned down by an arsonist just before it was due to open (D009), and yet another recounted how their apartment had been broken into, trashed, and abortion-related messages were left on a whiteboard (D010). Interviewees saw such acts of violence and intimidation as associated with the sustained spread of myths within the antiabortion movement. Long-cemented myths about the nature of abortion procedures and about fetal development and viability have radicalizing effects for some, driving more extreme opponents to target abortion health care providers with violence and intimidation. Resultantly, across interviews, professionals highlighted the mental and physical toll of working within abortion health care and fears for their physical safety. In addition to concerns about physical safety, interviewees feared that the potential for violence could drive health care professionals out of abortion work. D006 argued:

> *That's [violence] the kind of thing that really has a potential for causing disruption. Right? If my colleagues quit because they're fearful for their own safety, then we can't see patients.*

At the extreme, abortion myths contribute to the escalation of antiabortion violence, and in the everyday, they reinforce long-standing stigma around abortion, abortion-seekers, and

XSL•FO
RenderX

abortion providers [42]. D001 described the community shaming they and their colleagues experienced, despite not performing abortions, because of misled beliefs about abortion health care:

> *I literally have patients who won't see me because they think I actually do abortions every day. My mentor here had someone stand up in church and yell at her and start screaming that she did abortions because she did D&Cs for miscarriages.*

Myths about abortion health care encompass a wide array of reproductive health topics and, notably, result in stigma around lifesaving, often emergency, care necessary for miscarriages. D001 described an interaction where they explained care options to a patient experiencing a stillbirth at 31 weeks, resulting in the patient sobbing that they did not "want to have an abortion" because they were equating the dilation and curettage procedure needed to safely complete a miscarriage with an elective abortion.

In contemplating the impacts of abortion myths and their interaction with societal stigma, interviewees also highlighted how these myths act as barriers to care because of the time needed to debunk myths and settle associated feelings of stigmatization.

> *If you spend more time sort of working to dispel this [stigma], this might delay someone getting an abortion.* [D002]

Delaying an abortion is not necessarily negative if patients are early enough in their pregnancy and live in a state that affords them options. However, recent and ongoing legal changes around gestational time limits complicate options for pregnant people; as D007 summarized, "if it turns into a second trimester, then they [may not] be able to find a place that does that."

The impact of abortion myths thus requires consideration of the legal framework for abortion, particularly post Dobbs. In addition to novel legal myths around abortion arising in the complicated landscape following the overturning of Roe vs Wade, interviews highlighted an uptick in abortion-related conversation, resulting in the potential for further amplification of abortion myths. D004 posited:

> *I don't know if ... there's necessarily more accurate information out there than there used to be, but there's definitely just more information in general, [more] talking about abortion.*

Discussion

Principal Findings

Interviews captured abortion professionals' perspectives on the types of myths circulating about abortion health care and the impact they are having on their work. The post-Dobbs context has brought to the fore a novel arena for abortion-related myths focused on legality. Interviewees expressed concern over the rapidly altering legal landscape around their work and how this leaves both the public and abortion professionals unclear on the legalities of specific abortion-related (and broader) procedures. Given this dynamic uncertainty, an undeniable information void has emerged for those seeking up-to-date information about the legality of abortion in their specific location.

An exploration of the nature of abortion myths in information-seeking, attempts to mitigate misleading information in the clinic, and the perceived impacts of abortion myths holds significant insights for models of SDM. Previous examinations of SDM in clinical settings explore the patient-side of SDM dynamics, in addition to the role of the practitioner in allowing for patient autonomy [43,44]. Interview findings cement the difficulties of engaging in SDM, and in particular, highlight how misleading and false information can undermine the foundations of patient competence, relatedness, and autonomy [29] central to SDT and thus at the core of effective SDM. In an information environment flooded with abortion-related myths, the ability of patients to engage autonomously in care-related discussions is further eroded, and the role of practitioners in debunking myths may complicate a patient's sense of autonomy. Moreover, uncertainty around abortion myths (specifically about legality) among practitioners highlights a need to attend further to the practitioner-side of SDM dynamics and not take for granted the competence and sense of autonomy felt by practitioners in their navigation of patient care plans.

Conversations also looked to understand practitioners' and advocates' perspectives on the information-seeking environment for abortion health care. Two key takeaways emerge from the analysis findings. First, online sources—including new social media sites such as TikTok—are primary sites for abortion-related information. However, professionals express concern over the credibility of information patients are accessing on these platforms, given the prevalence and diversity of myths they encounter in their work. Coupled with the fact that many interviewees acknowledged that they do not use social media, often purposefully to avoid personal overexposure to stigma and the negative emotional toll of encountering abortion myths, means they can only be reactive within patient interactions as they are not personally aware of the social media landscape. Emergent academic research is instructive in learning about the kinds of myths—for example, around safety, how to access and use oral abortive medications, legality, pain, etc [2,6]—but more work needs to be done to provide abortion health care professionals with knowledge about viral myths that does not require them to take on the individual burden of being active on social media. Further, as social media becomes a central authority for health information for patients, attending to the nature of health information accessible (and prominent) on social media platforms is crucial to building effective SDM.

Second, abortion health care professionals have developed useful communication strategies that enable productive conversations around myths. Instead of looking to directly fact-check or debunk myths presented to them, most interviewees described a communication process akin to MI—that is, asking open questions about the nature of the information presented and its utility or value to the patient. MI allows professionals to build a rapport with patients in addition to getting to the heart of why a specific abortion myth is particularly salient to that person. Accordingly, MI appears to be a core tool in enabling SDM to occur within a polluted information environment. Future

XSL•FO
RenderX

research should explore how MI techniques can be strengthened through proactive knowledge of what types of myths are spreading in the digital information environment and how MI impacts a patient's sense of autonomy in SDM relationships.

Finally, interviews capture a pernicious cycle of distrust. Distrust of abortion providers—often underpinned by stigma and/or misled views about abortion and providers—may lead information-seekers to look to online and alternative, mostly online, sources of information about abortion health care. In doing so, information-seekers encounter further myths about abortion procedures, the consequences or impacts of abortion, and the legality of abortion. These myths further fracture trust in abortion health care providers, and in the extreme, they can lead to the harassment of practitioners, as highlighted by several interviewees. Moreover, distrust at any level undermines the ability for SDM to occur, as patients are unwilling to engage in decision-making models that afford practitioners agency, and vice versa for practitioners concerned that patients cannot give informed consent without accessible, factual information. Effectively addressing distrust most often requires individual interactions [45], yet this is impossible to achieve if patients are too distrustful to talk to health care professionals about abortion, or even access health care at all.

Limitations and Future Work

This research is necessarily limited by its focus on health care providers and advocates. Interviews capture the perspectives of abortion health care professionals, meaning that discussions of the impact of abortion myths are filtered through the experiences of these professionals, rather than directly capturing the lived experiences of those seeking abortion health care. Further, we recruited professionals with a preexisting awareness of the issue of misinformation in their field. While we anticipate that this is a common awareness within abortion health care, there may be practitioners who do not feel as acutely the impact of abortion myths on their work. As such, future research should supplement this understanding through interviews and/or survey work designed to document the spread and impact of abortion myths within patient and broader populations. In addition, interview research and subsequent analysis are grounded within the perspectives of the research team conducting this project. We recognize that our positionality, both as academic researchers and associated with our individual identities, may impact how interviewees responded to answers within the interview context and similarly biased our thematic analysis. Our research methodology is designed to reduce the impact of researcher bias, as is transparency in our reporting; however, research from outside of academia may identify additional or conflicting perspectives on abortion myths that we did not capture.

Further, as highlighted by practitioner concerns, this work only captures the perspectives of patients who access health care, excluding a large population of information-seekers who do not access abortion-related health care. Accessing this missing population is difficult, but future work should look to address this gap to gain further understanding of the impact of misleading abortion information. Finally, this work took place shortly after the Dobbs decision, wherein 13 states activated trigger laws that banned or heavily restricted abortion [46]. This immediate period can be characterized by heightened confusion, as states figure out how to enact said bans and as challenges have emerged around specific cases and restrictions more broadly. Longitudinal research should look to examine how abortion myths shift (in prevalence and nature), as the aftermath of the Dobbs decision becomes more embedded in the legal frameworks of low and no-access states.

Conclusions

Misinformation, myths, and rumors thrive in conditions of uncertainty [47], and the legal complexity and uncertainty post Dobbs (Cohen et al, unpublished data, 2022) have created an environment ripe for the spread of novel and long-standing abortion myths. Moreover, the costs of myths within abortion health care are substantial. Interviews capture the strain abortion myths create in the ability of practitioners to pursue effective SDM with patients, including undermining trust in patient-practitioner relationships, delaying desired and/or essential health care to negative outcomes, and even causing violence and harassment directed at health care providers. The future of informed consent and effective SDM for abortion health care relies upon access to authoritative information. The current online information environment concerning abortion is muddled, with added complexity because of the idiosyncrasies of access according to geography and legality. Abortion myths are long-standing, underpinned by active disinformation spread by CPCs about the nature and impact of abortion procedures and by antiabortion activists. The resultant information environment is highly vulnerable to the spread of abortion myths that undermine patient autonomy. Accordingly, abortion health care exists as an important case study for understanding models of SDM in an era of information disorder. This project highlights how health myths undermine the conditions necessary for autonomy and self-determination within health decision-making, suggesting a need to update theories underpinning SDM to account for the realities of patient information seeking.

Acknowledgments

The authors would like to thank the health care practitioners and advocates who gave their time and expertise to this research. Generative AI was not used in the preparation of this paper or broader project activities.

Funding

This study was made possible through funding from Population Health at the University of Washington, University of Washington's Center for an Informed Public, the John S. and James L. Knight Foundation, and the National Science Foundation's Graduate Research Fellowship program. Population Health funding was used to pay for the time of a postdoctoral fellow to conduct analysis

XSL•FO
RenderX

and for participant incentives. Members of the research team were, at the time of research, funded by fellowships from the Knight Foundation and the National Science Foundation. This study was made possible through funding from Population Health at the University of Washington, University of Washington's Center for an Informed Public, the John S. and James L. Knight Foundation, and the National Science Foundation's Graduate Research Fellowship program. Population Health funding was used to pay for the time of a postdoctoral fellow to conduct analysis and for participant incentives. Members of the research team were, at the time of research, funded by fellowships from the Knight Foundation and the National Science Foundation.

Data Availability

The datasets generated and analyzed during this study are not publicly available due to the sensitive nature of the subject matter and the anonymity given to participants. However, anonymized raw data and analysis are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: RM, TA, IG, AB, ALS, ES, AS
Data curation: RM, AS
Formal analysis: RM, JS, TA, IG, AB, AS
Funding acquisition: RM, AS, ES
Investigation: RM, AS
Methodology: RM, ALS, ES, AS, TA
Project administration: RM, JS, AS
Supervision: ALS, ES
Writing – original draft: RM, JS, TA, IG, AS
Writing – review & editing: RM

Conflicts of Interest

None declared.

Multimedia Appendix 1
Consent (short form).
[DOCX File, 21 KB - publichealth_v12i1e77050_app1.docx]

Multimedia Appendix 2
Interview protocol.
[DOCX File, 1676 KB - publichealth_v12i1e77050_app2.docx]

Multimedia Appendix 3
Codebook.
[XLSX File, 9 KB - publichealth_v12i1e77050_app3.xlsx]

References

1. Facts are important: abortion is healthcare. American College of Obstetricians & Gynecologists. URL: https://www.acog.org/advocacy/facts-are-important/abortion-is-healthcare [accessed 2026-07-14]
2. Patev AJ, Hood KB. Towards a better understanding of abortion misinformation in the USA: a review of the literature. Cult Health Sex 2021 Mar;23(3):285-300. [doi: 10.1080/13691058.2019.1706001] [Medline: 32202213]
3. Heymann O, Odum T, Norris AH, Bessett D. Selecting an abortion clinic: the role of social myths and risk perception in seeking abortion care. J Health Soc Behav 2022 Mar;63(1):90-104. [doi: 10.1177/00221465211044413] [Medline: 34605701]
4. Pleasants E. Challenging abortion disinformation online. UWIRE Text. 2021 Feb 9. URL: https://link.gale.com/apps/doc/A651561104/AONE?u=wash_main&sid=bookmark-AONE&xid=7f024776 [accessed 2026-07-14]
5. Spiro E, Starbird K. Rumors have rules. Issues Sci Technol 2023 Apr 1;29(3):47-49. [doi: 10.58875/CXGL5395]
6. Pagoto SL, Palmer L, Horwitz-Willis N. The next infodemic: abortion misinformation. J Med Internet Res 2023 May 4;25:e42582. [doi: 10.2196/42582] [Medline: 37140975]
7. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. J Med Internet Res 2021 Jan 20;23(1):e17187. [doi: 10.2196/17187] [Medline: 33470931]
8. Druckman JN. Threats to science: politicization, misinformation, and inequalities. Ann Am Acad Pol Soc Sci 2022 Mar;700(1):8-24. [doi: 10.1177/00027162221095431]
9. Beers A, Nguy n S, Starbird K, West JD, Spiro ES. Selective and deceptive citation in the construction of dueling consensuses. Sci Adv 2023 Sep 22;9(38):eadh1933. [doi: 10.1126/sciadv.adh1933] [Medline: 37738338]

XSL•FO
RenderX

10. Goldstein CM, Murray EJ, Beard J, Schnoes AM, Wang ML. Science communication in the age of misinformation. Ann Behav Med 2020 Dec 1;54(12):985-990. [doi: 10.1093/abm/kaaa088] [Medline: 33416836]
11. Godoy M. Doctors and advocates tackle a spike of abortion misinformation – in Spanish. NPR. 2022 Nov 3. URL: https://www.npr.org/sections/health- shots/2022/11/03/1133743997/doctors-and-advocates-tackle-a-spike-of-abortion-misinformation-in-spanish [accessed 2026-08-06]
12. Stimpson JP, Park S, Adhikari EH, Nelson DB, Ortega AN. Perceived health misinformation on social media and public trust in health care. Med Care 2025 Sep 1;63(9):686-693. [doi: 10.1097/MLR.0000000000002180] [Medline: 40793916]
13. Kulczycki A. Dobbs: navigating the new quagmire and its impacts on abortion and reproductive health care. Health Educ Behav 2022 Dec;49(6):924-928. [doi: 10.1177/10901981221125430] [Medline: 36173003]
14. Lockmiller C. Decoding the misinformation-legislation pipeline: an analysis of Florida Medicaid and the current state of transgender healthcare. J Med Libr Assoc 2023 Oct 2;111(4):750-761. [doi: 10.5195/jmla.2023.1724]
15. Krishna A, Thompson TL. Misinformation about health: a review of health communication and misinformation scholarship. Am Behav Sci 2021 Feb;65(2):316-332. [doi: 10.1177/0002764219878223]
16. Kessler SH, Bachmann E. Debunking health myths on the internet: the persuasive effect of (visual) online communication. J Public Health (Berl) 2022 Aug;30(8):1823-1835. [doi: 10.1007/s10389-022-01694-3]
17. Sharevski F, Loop JV, Jachim P, Devine A, Pieroni E. Abortion misinformation on TikTok: rampant content, lax moderation, and vivid user experiences. arXiv. Preprint posted online on Jan 12, 2023 URL: https://arxiv.org/pdf/2301.05128 [accessed 2026-08-06]
18. Bryant AG, Narasimhan S, Bryant-Comstock K, Levi EE. Crisis pregnancy center websites: information, misinformation and disinformation. Contraception 2014 Dec;90(6):601-605. [doi: 10.1016/j.contraception.2014.07.003] [Medline: 25091391]
19. Bryant AG, Levi EE. Abortion misinformation from crisis pregnancy centers in North Carolina. Contraception 2012 Dec;86(6):752-756. [doi: 10.1016/j.contraception.2012.06.001] [Medline: 22770790]
20. Gantt-Shafer J. “They just went after us:” reproductive justice advocacy at an abortion fund. Front Commun 2020;5:501276. [doi: 10.3389/fcomm.2020.501276]
21. Sabbath EL, McKetchnie SM, Arora KS, Buchbinder M. US obstetrician-gynecologists’ perceived impacts of post-Dobbs v Jackson state abortion bans. JAMA Netw Open 2024 Jan 2;7(1):e2352109. [doi: 10.1001/jamanetworkopen.2023.52109] [Medline: 38231510]
22. Luchetta J. Report shows dramatic exodus of Idaho obgyns since repeal of Roe v. Wade. Oregon Public Broadcasting. 2024 Feb 23. URL: https://www.opb.org/article/2024/02/22/report-shows-fewer-idaho-obgyns-since-repeal-of-roe-v-wade/ [accessed 2026-08-06]
23. Kozhimannil KB, Hassan A, Hardeman RR. Abortion access as a racial justice issue. N Engl J Med 2022 Oct 27;387(17):1537-1539. [doi: 10.1056/NEJMp2209737] [Medline: 36069823]
24. Wilkinson B, Onwuzurike C, Bartz D. Restrictive state abortion bans - a reproductive injustice. N Engl J Med 2022 Mar 31;386(13):1197-1199. [doi: 10.1056/NEJMp2119364] [Medline: 35333482]
25. Carroll S, Joshi D, Espey E. Physicians and healthcare professionals as advocates for abortion care and reproductive choice. Curr Opin Obstet Gynecol 2022 Dec 1;34(6):367-372. [doi: 10.1097/GCO.0000000000000833] [Medline: 36342009]
26. Légaré F, Thompson-Leduc P. Twelve myths about shared decision making. Patient Educ Couns 2014 Sep;96(3):281-286. [doi: 10.1016/j.pec.2014.06.014] [Medline: 25034637]
27. Resnicow K, Catley D, Goggin K, Hawley S, Williams GC. Shared decision making in health care: theoretical perspectives for why it works and for whom. Med Decis Making 2022 Aug;42(6):755-764. [doi: 10.1177/0272989X211058068] [Medline: 34784805]
28. Deci EL, Ryan RM. Intrinsic Motivation and Self-Determination in Human Behavior: Springer New York; 1985. [doi: 10.1007/978-1-4899-2271-7]
29. Resnicow K, Teixeira PJ, Williams GC. Efficient allocation of public health and behavior change resources: the “difficulty by motivation” matrix. Am J Public Health 2017 Jan;107(1):55-57. [doi: 10.2105/AJPH.2016.303526] [Medline: 27925833]
30. Fagot E. Closing the gap: actualising shared decision-making through effective medication abortion patient follow-up care. BMJ Open Qual 2020 Mar;9(1):e000740. [doi: 10.1136/bmjoq-2019-000740] [Medline: 32209594]
31. Frederico M, Michielsen K, Arnaldo C, Decat P. Factors influencing abortion decision-making processes among young women. Int J Environ Res Public Health 2018 Feb 13;15(2):329. [doi: 10.3390/ijerph15020329] [Medline: 29438335]
32. Hooper MW. From distrust to confidence: can science and health care gain what’s missing? Pew Research. 2024 Oct 17. URL: https://www.pew.org/en/trend/archive/fall-2024/from-distrust-to-confidence-can-science-and-health-care-gain-whats-missing [accessed 2026-08-06]
33. Noy C. Sampling knowledge: the hermeneutics of snowball sampling in qualitative research. Int J Soc Res Methodol 2008 Oct;11(4):327-344. [doi: 10.1080/13645570701401305]
34. Fusch PI, Ness LR. Are we there yet? Data saturation in qualitative research. Qual Rep 2015;20(9):1408-1416. [doi: 10.46743/2160-3715/2015.2281]
35. Sandelowski M. Theoretical saturation. In: Given LM, editor. The SAGE Encyclopedia of Qualitative Research Methods: Sage Publications, Inc; 2008:875-876. [doi: 10.4135/9781412963909.n456]

XSL•FO
RenderX

36. Braun V, Clarke V. Toward good practice in thematic analysis: avoiding common problems and be(com)ing a knowing researcher. Int J Transgend Health 2023;24(1):1-6. [doi: 10.1080/26895269.2022.2129597] [Medline: 36713144]
37. Bans on specific abortion methods used after the first trimester. Guttmacher Institute. URL: https://www.guttmacher.org/state-policy/explore/bans-specific-abortion-methods-used-after-first-trimester [accessed 2026-07-14]
38. Partial-birth abortion. Legal Information Institute. URL: https://www.law.cornell.edu/definitions/uscode.php?width=840&height=800&iframe=true&def_id=18-USC-1932217061-836166418&term_occur=999 [accessed 2026-08-06]
39. Smith KL, Shipchandler F, Kudumu M, Davies-Balch S, Leonard SA. “Ignored and invisible”: perspectives from Black women, clinicians, and community-based organizations for reducing preterm birth. Matern Child Health J 2022 Apr;26(4):726-735. [doi: 10.1007/s10995-021-03367-1] [Medline: 35072869]
40. Mitchell E. Black women’s reproductive health and legacies of distrust. Black Perspectives. 2022 Aug 3. URL: https://www.aaihs.org/black-womens-reproductive-health-and-legacies-of-distrust/ [accessed 2026-08-06]
41. Nguyen M, Cartwright AF, Upadhyay UD. Fear of procedure and pain in individuals considering abortion: a qualitative study. Patient Educ Couns 2023 Mar;108:107611. [doi: 10.1016/j.pec.2022.107611] [Medline: 36603469]
42. Steinberg JR, Tschann JM, Furgerson D, Harper CC. Psychosocial factors and pre-abortion psychological health: the significance of stigma. Soc Sci Med 2016 Feb;150:67-75. [doi: 10.1016/j.socscimed.2015.12.007] [Medline: 26735332]
43. Brown SL, Salmon P. Reconciling the theory and reality of shared decision-making: a “matching” approach to practitioner leadership. Health Expectations 2019 Jun;22(3):275-283. [doi: 10.1111/hex.12853] [Medline: 30478979]
44. Charles C, Gafni A, Whelan T. Shared decision-making in the medical encounter: what does it mean? (Or it takes at least two to tango). Soc Sci Med 1997 Mar;44(5):681-692. [doi: 10.1016/S0277-9536(96)00221-3]
45. Chandra S, Mohammadnezhad M, Ward P. Trust and communication in a doctor- patient relationship: a literature review. J Healthcare Commun 2018;03(3):36. [doi: 10.4172/2472-1654.100146]
46. Winny A. How abortion trigger laws impact mental health. Johns Hopkins Bloomberg School of Public Health. 2024 Feb 6. URL: https://publichealth.jhu.edu/2024/abortion-bans-impact-on-mental-health#:~:text=Immediately%20following%20the%20Dobbs%20decision,the%20event%20Roe%20was%20overturned [accessed 2026-07-14]
47. Krafft PM, Spiro ES. Keeping rumors in proportion: managing uncertainty in rumor systems. Presented at: CHI ’19: Proceedings of CHI Conference on Human Factors in Computing Systems; May 4-9, 2019; Glasgow, Scotland p. 1-11. [doi: 10.1145/3290605.3300876]

Abbreviations

ACOG: American College of Obstetricians and Gynecologists
CPC: crisis pregnancy center
MI: motivational interviewing
OB-GYN: obstetrics and gynecology
RQ: research question
SDM: shared decision-making
SDT: self-determination theory
WWAMI : Washington, Wyoming, Alaska, Montana, and Idaho

Edited by A Mavragani; submitted 06.May.2025; peer-reviewed by H McMahon, TA Weitz, W Liedtke; revised version received 05.May.2026; accepted 19.May.2026; published 16.Sep.2026.

Please cite as:
Moran RE, Simoes J, Agajanian T, Grasso I, Beers A, Swarr AL, Spiro E, Swan A
Investigating the Impact of Online Abortion Myths on Health Care Providers and Advocates: Interview Study
JMIR Public Health Surveill 2026;12:e77050
URL: https://publichealth.jmir.org/2026/1/e77050
doi:10.2196/77050

© Rachel E Moran, Julia Simoes, Taylor Agajanian, Izzi Grasso, Anna Beers, Amanda Lock Swarr, Emma Spiro, Anna Swan. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 16.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Perceived Difficulty Accessing Primary Health Care and Its Associations With Sociodemographic Factors and Exposure to Violence in a Violent Urban Community in Rio de Janeiro, Brazil: Cross-Sectional Study

Ana Paula da Cunha[1,2*], PhD; Jurema Corrêa da Mota[3*], PhD; Francisco Inacio Bastos[1*], PhD

[1]Institute of Scientific and Technological Communication and Information in Health, Oswaldo Cruz Foundation, Avenida Brasil, 4365, Rio de Janeiro, Brazil

[2]Joaquim Venâncio Polytechnic School of Health, Oswaldo Cruz Foundation, Rio de Janeiro, Brazil

[3]Viral Hepatitis Laboratory, Oswaldo Cruz Institute, Oswaldo Cruz Foundation, Rio de Janeiro, Brazil

[*]all authors contributed equally

Corresponding Author:
Ana Paula da Cunha, PhD
Institute of Scientific and Technological Communication and Information in Health, Oswaldo Cruz Foundation, Avenida Brasil, 4365, Rio de Janeiro, Brazil

Abstract

Background: Urban communities affected by chronic violence face persistent challenges in accessing primary health care (PHC). Although violence is widely recognized as a determinant of health inequities, there exists limited evidence on how direct exposure to violent events is related to individuals' perceptions of difficulty in accessing PHC services in highly vulnerable urban environments.

Objective: The study aims to examine how sociodemographic characteristics and exposure to violent events are associated with perceived difficulty in accessing PHC in standardized hypothetical mental health scenarios among residents of a community affected by chronic urban violence in Rio de Janeiro.

Methods: A cross-sectional survey was conducted between October 8 and November 1, 2024, with 1006 adult residents who completed 4 PHC vignettes and reported past-year exposure to violent events. Participants completed a structured questionnaire that included sociodemographic information, past-year exposure to interpersonal violence, and evaluations of 4 hypothetical vignettes describing different clinical or social situations. Vignette responses were recorded using an ordinal Likert-type scale ranging from "very easy" to "extremely difficult." Descriptive analyses were performed for all variables, and associations between sociodemographic characteristics and violence indicators were examined using Pearson chi-square tests, with effect sizes estimated using the ϕ coefficient or Cramér *V*, as appropriate, and their respective 95% CIs. Differences in vignette ratings across sociodemographic groups were assessed using Kruskal-Wallis tests. To explore the underlying relationships among violence indicators, vignette responses, and sociodemographic variables, exploratory factor analysis and multiple correspondence analysis were conducted.

Results: Violence exposure was frequent. Threats of harm were reported by 27.7% (157/567) of Black participants and 30.5% (72/236) of non-Black participants. Slapping was reported by 20.8% (118/567) of Black individuals and 28% (66/236) of non-Black individuals, while being pushed or restrained affected 21.2% (50/236) and 26.7% (63/236), respectively. Significant sex differences emerged for being punched or kicked (female 17.2% [75/435] vs male 23.9% [88/368]; *P*=.02). Kruskal-Wallis tests indicated few differences in vignette-based perceptions across groups, except for race in the alcohol-dependence vignette (χ^2_1=7.384, *P*=.006). Factor analysis identified 2 dimensions: violence variables loaded strongly on one factor (0.83 - 0.96) and vignette responses on another (0.66 - 0.83). Multiple correspondence analysis results showed that forms of violence contributed mainly to Dimension 1 (9% - 15%) and vignette categories to Dimension 2 (7% - 12%).

Conclusions: Exposure to violence was widespread, with differences in some forms of violence observed between men and women. Violence-related variables and vignette-based perceptions of difficulty accessing PHC were predominantly represented by distinct dimensions in the exploratory multivariate analyses. Strengthening PHC in violence-affected urban settings requires addressing structural barriers that shape residents' access to health care.

(JMIR Public Health Surveill 2026;12:e94922) doi:10.2196/94922

XSL•FO
RenderX

KEYWORDS

urban violence; primary health care access; sociodemographic characteristics; health care barriers; interpersonal violence

Introduction

Urban violence constitutes a major public health challenge and is closely linked to structural inequalities, social exclusion, and territorial segregation in large urban centers. Foundational analyses emphasize that violence functions both as a determinant and an outcome of broader social processes, shaping patterns of illness, emotional distress, and access to essential services [1-7]. These processes disproportionately affect populations living in deprived territories, where state presence is limited and residents routinely navigate environments marked by chronic insecurity. More recent evidence from Rio de Janeiro reinforces the complex relationship between armed violence and adverse living conditions, showing that violence is embedded in broader territorial and social dynamics rather than reflecting a simple linear association with socioeconomic deprivation [8].

These dynamics have historically been particularly evident in favelas and other socioeconomically vulnerable urban territories. Long-term research on the city's favelas has documented profound transformations in living conditions alongside the persistence of social marginalization, inequality, and violence, demonstrating how improvements in some dimensions of urban life may coexist with enduring processes of social exclusion [9].

In Rio de Janeiro, neighborhoods such as the Manguinhos hub of favelas exemplify how structural inequities intersect with urban violence. Historical accounts and community-based documentation describe Manguinhos as a territory characterized by fragmented urban development, precarious infrastructure, and recurrent episodes of armed conflict, all of which affect the daily lives of residents [10]. The historical formation of Manguinhos has involved heterogeneous processes of occupation, housing interventions, urbanization policies, and disputes over urban space, reflecting broader processes that have shaped the development of favelas in Rio de Janeiro [11]. In this territory, residents experience restricted mobility, environmental hazards, and constant exposure to situations of risk, illustrating how violence shapes daily trajectories, emotional well-being, and opportunities for social participation [12].

In Brazil, PHC is organized within the Unified Health System (Sistema Único de Saúde), primarily through the Family Health Strategy (FHS), a community- and territory-based model in which multidisciplinary teams are responsible for defined populations. The expansion of the FHS has substantially increased access to PHC; however, important barriers persist, including insufficient funding and service capacity, large populations assigned to individual teams, shortages and unequal distribution of health professionals, organizational constraints, and difficulties in ensuring timely and continuous care [13]. These challenges may be particularly relevant in socioeconomically vulnerable territories, where structural and territorial conditions can interact with the organization of health services and further constrain access to care.

The consequences of armed violence in vulnerable territories may also extend to the organization and provision of health care. Health services operating in violence-affected communities may experience disruptions in their activities, while health professionals and residents may face constraints that interfere with the provision and use of care. Recent evidence from vulnerable communities in Rio de Janeiro has shown that armed violence may interact with the functioning of primary health care (PHC) services. In particular, the presence or absence of physicians in PHC teams partially mediated associations between armed violence and selected self-reported health conditions, highlighting the interconnected relationships among territorial violence, health outcomes, and health service organization [14].

International research reinforces the evidence that living in contexts of concentrated poverty and community violence is associated with multiple health consequences, including chronic stress, reduced social cohesion, and behavioral adaptations to perceived danger [15-17]. Studies in diverse global settings demonstrate that residents of high-violence neighborhoods tend to alter their routines, avoid moving through certain areas, and develop heightened perceptions of threat, which may influence their engagement with public services and their interpretations of hypothetical or real situations involving risk or conflict [17]. In the context of PHC, these territorial and social constraints are particularly relevant because access depends not only on the availability of services but also on individuals' ability to reach, engage with, and use them within the conditions of their everyday lives [18].

Understanding how individuals perceive situations of violence in such contexts is essential for public health. Perception-based data capture dimensions of insecurity, emotional burden, and daily constraints that are not fully captured by official indicators or epidemiological records. Examining how sociodemographic characteristics intersect with reported experiences of violence helps illuminate patterns of vulnerability within territories marked by persistent inequality. Understanding perceptions of access to health care is also relevant in these settings, as perceived barriers may capture dimensions of accessibility that are not fully reflected by indicators of service availability or usage [8,19]. Examining exposure to violence alongside sociodemographic characteristics may therefore help clarify whether direct experiences of violent events are associated with how residents perceive potential difficulties in obtaining PHC.

Despite growing recognition of violence as an important determinant of health and emerging evidence concerning the effects of armed violence on health services in vulnerable urban territories [14], there remains limited evidence on whether individual exposure to interpersonal violence is associated with perceived difficulty in accessing PHC. This distinction is important because the effects of violence on territories and on the organization and functioning of health services do not

XSL•FO
RenderX

necessarily imply that individuals who have personally experienced violence perceive greater difficulty in obtaining care. Examining this relationship may help distinguish individual experiences of violence from broader structural, organizational, and territorial barriers to health care access.

The objective of this study is to examine how sociodemographic characteristics and exposure to violent events are associated with perceived difficulty in accessing PHC in standardized hypothetical mental health scenarios, among residents of a community affected by chronic urban violence in Rio de Janeiro.

Methods

Study Design and Participants

This is a cross-sectional analytical study including 1006 adults residing in Manguinhos, a community affected by armed violence in Rio de Janeiro, Brazil. The reporting of this study followed the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines (Checklist 1). Participants were recruited between October 8 and November 1, 2024, using a quota sampling strategy based on sex and age group to approximate the demographic distribution of the local adult population. A formal a priori power analysis was not performed because the study used a nonprobability quota sampling design nested within a broader population survey. A target of 1000 interviews was established to allow proportional allocation of participants according to sex and age group based on secondary population data provided by the Pereira Passos Institute. The final analytical sample comprised 1006 participants. Recruitment was conducted in public spaces throughout the community by trained interviewers. Individuals were eligible if they were 18 years or older, resided in the study area, and agreed to participate by providing written informed consent. All participants completed a structured questionnaire covering sociodemographic characteristics, exposure to interpersonal violence, and evaluations of 4 vignettes addressing clinical and social conditions.

Quota sampling has limitations compared with probability sampling, particularly regarding representativeness; however, it was chosen to facilitate the recruitment of participants living in an impoverished community with a large proportion of residents in unstable housing (eg, shacks and provisional shelters built with fragile materials) and a highly mobile population. Potential participants were screened for eligibility before the questionnaire was administered. Individuals who did not meet the eligibility criteria were not enrolled, and the electronic questionnaire was programmed to terminate automatically for ineligible respondents. Consequently, all completed questionnaires included in the analyses were obtained from eligible participants.

Study Variables

Sociodemographic characteristics included sex, age group, and race/color. Exposure to interpersonal violence was assessed with dichotomous questions that asked whether participants had experienced threats, various forms of physical aggression, or attacks involving objects such as knives or firearms in the past year.

Race/color was self-reported according to the Brazilian census classification. For analytical purposes, participants who self-identified as Black or Brown were categorized as Black, whereas those who self-identified as White, Asian, or Indigenous were categorized as non-Black.

Participants also evaluated 4 standardized narrative vignettes describing a hypothetical case of depression and anxiety, a hypothetical case of alcohol dependence, a person with a disadvantaged background but without mental health problems, and a neutral (control) vignette. The complete wording of the vignettes is provided in Multimedia Appendix 1. After reading each vignette, participants answered the following question: “If you had a relative or acquaintance living in the situation described above, how would you rate the level of difficulty they would have in receiving care at the PHC unit serving your neighborhood?” Responses were recorded using a 5-point Likert scale ranging from “very easy” to “extremely difficult,” with an additional “don’t know/no answer” option. The vignette responses were analyzed descriptively and subsequently incorporated into exploratory factor analysis and multiple correspondence analysis (MCA) to investigate the underlying structure of perceived difficulty in accessing PHC.

Analytical Procedures

Descriptive analyses included absolute and relative frequencies for all variables. Associations between sociodemographic characteristics and violence indicators were examined using Pearson chi-square tests, and the effect size was calculated using the ϕ coefficient or Cramér V, with their respective 95% CIs [20].

Differences in vignette ratings according to sex, age, and race/color were analyzed using Kruskal-Wallis tests, given the ordinal distribution of the responses. Although the Kruskal-Wallis test is rank-based and distinct from the chi-square test of independence, its test statistic (H) asymptotically follows a chi-square distribution. Therefore, the results are reported using the chi-square statistic (χ^2) and corresponding P values.

Correlation Matrix and Factor Analysis

A mixed correlation matrix was computed to explore the relationships among vignette responses, violence indicators, and sociodemographic variables. Polychoric correlations were used for ordinal items, tetrachoric correlations for dichotomous variables, and Pearson correlations for numerically recoded sociodemographic variables. This matrix was used as input for the exploratory factor analysis, conducted using principal axis factoring with oblimin rotation [21]. The number of factors was determined based on eigenvalues and parallel analysis. Factor loadings, communalities, and explained variance were examined to support interpretation [22].

MCA

Given the predominantly categorical nature of the dataset, MCA was subsequently carried out to identify multivariate patterns of association between categories of vignette responses, interpersonal violence indicators, and sociodemographic variables. The analysis produced eigenvalues, explained inertia,

XSL•FO
RenderX

category coordinates, squared cosine (cos^2) values, and contribution percentages for the main dimensions, which guided the interpretation of the underlying structure.

All analyses were performed in R software (version 4.4.2; R Foundation for Statistical Computing), using the packages *tidyverse* [23], *psych* [24], *FactoMineR* [25], *factoextra* [26], and *MASS* [27].

Ethical Considerations

The study was approved by the Research Ethics Committee of the Oswaldo Cruz Foundation (FIOCRUZ), under CAAE (Certificado de Apresentação para Apreciação Ética [Certificate of Presentation for Ethical Consideration]) number 78822424.9.0000.5241 (approval number 6.873.690). All participants were informed about the study objectives, the voluntary nature of participation, and guarantees of confidentiality. They signed the informed consent form, in accordance with resolution number 466/2012 of the Brazilian National Health Committee on Ethics.

All data were deidentified prior to analysis, and no personally identifiable information was included in the analytical dataset. Participant confidentiality was maintained throughout the study.

Participants did not receive financial compensation for their participation.

Results

Table 1 presents the distribution of sociodemographic characteristics according to different situations of violence reported in the previous 12 months. For the question “Has someone threatened to harm or injure someone important to you?,” the proportions were similar across race categories: among non-Black participants, 69.5% (164/236) responded “no” and 30.5% (72/236) “yes,” while among Black participants, 72.3% (410/567) responded “no” and 27.7% (157/567) “yes” (P=.47, ϕ=0.028, 95% CI 0.000 - 0.097). Regarding sex, 72.4% (315/435) of women and 70.4% (259/368) of men reported no such threat (P=.58, ϕ=0.022, 95% CI 0.000 - 0.103). Across age groups, the distribution also showed similar percentages, with “no” responses ranging from 69.7% (140/201) to 72.7% (208/286) (P=.76, Cramer V=0.044, 95% CI 0.000 - 0.103).

XSL•FO
RenderX

Table . Distribution of absolute and relative frequencies of sociodemographic variables according to situations of violence, 2024[a].

Variable and category	Participants, n (%)		*P* value[b]	φ/Cramer *V* (95% CI)
	No	Yes		
In the past 12 months, has someone threatened to harm or injure someone important to you?				
Race			.47	0.028 (0.000-0.097)
Non-Black	164 (69.5)	72 (30.5)		
Black	410 (72.3)	157 (27.7)		
Sex			.58	0.022 (0.000-0.103)
Female	315 (72.4)	120 (27.6)		
Male	259 (70.4)	109 (29.6)		
Age group (y)			.76	0.044 (0.000-0.103)
18-29	140 (69.7)	61 (30.3)		
30-44	226 (71.5)	90 (28.5)		
45 or older	208 (72.7)	78 (27.3)		
In the past 12 months, has someone slapped you?				
Race			.04	0.078 (0.008-0.146)
Non-Black	170 (72)	66 (28)		
Black	449 (79.2)	118 (20.8)		
Sex			.06	0.069 (0.000-0.138)
Female	347 (79.8)	88 (20.2)		
Male	272 (73.9)	96 (26.1)		
Age group (y)			.02	0.097 (0.008-0.161)
18-29	147 (73.1)	54 (26.9)		
30-44	236 (74.7)	80 (25.3)		
45 or older	236 (82.5)	50 (17.5)		
In the past 12 months, has someone pushed you, grabbed you forcefully, or thrown an object at you with the intention of causing harm?				
Race			.11	0.060 (0.000-0.129)
Non-Black	173 (73.3)	63 (26.7)		
Black	447 (78.8)	120 (21.2)		
Sex			.02	0.084 (0.015-0.153)
Female	350 (80.5)	85 (19.5)		
Male	270 (73.4)	98 (26.6)		
Age group (y)			.36	0.051 (0.000-0.111)
18-29	152 (75.6)	49 (24.4)		
30-44	239 (75.6)	77 (24.4)		
45 or older	229 (80.1)	57 (19.9)		
In the past 12 months, has someone punched, kicked, or pulled you by the hair?				
Race			.14	0.055 (0.000-0.124)
Non-Black	180 (76.3)	56 (23.7)		
Black	460 (81.1)	107 (18.9)		
Sex			.02	0.083 (0.014-0.151)
Female	360 (82.8)	75 (17.2)		
Male	280 (76.1)	88 (23.9)		

XSL•FO
RenderX

Variable and category	Participants, n (%)		*P* value[b]	φ/Cramer *V* (95% CI)
	No	Yes		
Age group (y)			.09	0.078 (0.000-0.141)
18-29	155 (77.1)	46 (22.9)		
30-44	245 (77.5)	71 (22.5)		
45 or older	240 (83.9)	46 (16.1)		
In the past 12 months, has someone attempted or actually strangled, suffocated, or intentionally burned you?				
Race			.12	0.059 (0.000-0.128)
Non-Black	204 (86.4)	32 (13.6)		
Black	513 (90.5)	54 (9.5)		
Sex			.48	0.029 (0.000-0.098)
Female	392 (90.1)	43 (9.9)		
Male	325 (88.3)	43 (11.7)		
Age group (y)			.12	0.073 (0.000-0.136)
18-29	177 (88.1)	24 (11.9)		
30-44	276 (87.3)	40 (12.7)		
45 or older	264 (92.3)	22 (7.7)		
In the past 12 months, has someone threatened or injured you with a knife, firearm, or any other weapon or object?				
Race			.22	0.047 (0.000-0.116)
Non-Black	206 (87.3)	30 (12.7)		
Black	513 (90.5)	54 (9.5)		
Sex			.35	0.037 (0.000-0.106)
Female	394 (90.6)	41 (9.4)		
Male	325 (88.3)	43 (11.7)		
Age group (y)			.10	0.076 (0.000-0.139)
18-29	176 (87.6)	25 (12.4)		
30-44	278 (88)	38 (12)		
45 or older	265 (92.7)	21 (7.3)		

[a]The ϕ coefficient was calculated for the race and sex variables, whereas Cramér *V* was calculated for age group.
[b]Chi-square test.

A high proportion of both Black and non-Black participants reported not having been slapped: 72.0% (170/236) of non-Black individuals and 79.2% (449/567) of Black individuals answered "no" (*P*=.04, ϕ=0.078, 95% CI 0.008 - 0.146). Among women, 79.8% (347/435) reported not having been slapped, compared to 73.9% (272/368) of men (*P*=.06, ϕ=0.069, 95% CI 0.000 - 0.138). Across age groups, the proportion of "no" responses varied from 73.1% (147/201) to 82.5% (236/286; *P*=.02, Cramer *V*=0.097, 95% CI 0.008 - 0.161; Table 1).

For being pushed or grabbed forcefully, or having an object thrown with the intention of causing harm, 73.3% (173/236) of non-Black individuals and 78.8% (447/567) of Black participants reported "no" (*P*=.11, ϕ=0.060, 95% CI 0.000 - 0.129). Women reported not experiencing this event in 80.5% (350/435) of cases compared with 73.4% (270/368) of men (*P*=.02, ϕ=0.084, 95% CI 0.015 - 0.153). Across age groups, "no" responses ranged from 75.6% (152/201) to 80.1% (229/286; *P*=.36, Cramer*V*=0.051, 95% CI 0.000 - 0.111; Table 1).

Physical aggression involving punching, kicking, or pulling hair was not reported by 76.3% (180/236) of non-Black participants and 81.1% (460/567) of Black participants (*P*=.14, ϕ=0.055, 95% CI 0.000 - 0.124). A difference between sexes was observed in the proportions of "yes" responses: 17.2% (75/435) among women and 23.9% (88/368) among men (*P*=.02, ϕ=0.083, 95% CI 0.014 - 0.151). Across age groups, "yes" responses ranged from 16.1% (46/286) to 22.9% (46/201) across groups (*P*=.09, Cramer *V*=0.078, 95% CI 0.000 - 0.141; Table 1).

Most participants reported not experiencing attempted or actual strangulation, suffocation, or burning in the previous 12 months. This proportion was slightly higher among Black participants (513/567, 90.5%) than among non-Black participants (204/236, 86.4%), although the difference was not statistically significant

XSL•FO
RenderX

(P=.12, ϕ=0.059, 95% CI 0.000 - 0.128). A similar pattern was observed by sex: 90.1% (392/435) of women and 88.3% (325/368) of men reported not experiencing these acts (P=.48, ϕ=0.029, 95% CI 0.000 - 0.098). Across age groups, the proportion reporting no such violence ranged from 87.3% (276/316) to 92.3% (264/286; P=.12, Cramer V=0.073, 95% CI 0.000 - 0.136; Table 1).

For threats or injuries involving a knife, firearm, or another weapon, 87.3% (206/236) of non-Black participants and 90.5% (513/567) of Black participants responded "no" (P=.22, ϕ=0.047, 95% CI 0.000 - 0.116). Women reported not experiencing this type of violence in 90.6% (394/435) of cases, compared with 88.3% (325/368; P=.35, ϕ=0.037, 95% CI 0.000 - 0.106). Across age groups, "no" responses ranged from 87.6% (176/201) to 92.7% (265/286; P=.10, Cramer V=0.076, 95% CI 0.000 - 0.139; Table 1).

Sex comparisons revealed similar vignette response scores for the depression vignette (χ^2_1=0.51, P=.48), the alcohol-dependence vignette (χ^2_1=0.04, P=.84), the vignette without mental health problems (χ^2_1=0.15, P=.70), and the control vignette (χ^2_1=0.01, P=.91; Table 2).

Table . Comparison of vignette response scores by sex, race/color, and age group.

Grouping variable and vignette	Chi-square (*df*) statistic	*P* value
Sex		
Depression vignette	0.508 (1)	.48
Alcohol-dependence vignette	0.041 (1)	.84
No mental health problems vignette	0.146 (1)	.70
Control vignette	0.013 (1)	.91
Race/color		
Depression vignette	0.009 (1)	.92
Alcohol-dependence vignette	7.384 (1)	.006
No mental health problems vignette	3.140 (1)	.08
Control vignette	0.429 (1)	.51
Age group		
Depression vignette	1.565 (2)	.46
Alcohol-dependence vignette	1.561 (2)	.46
No mental health problems vignette	4.340 (2)	.11
Control vignette	0.125 (2)	.94

Race and color comparisons revealed similar vignette response scores for the depression vignette (χ^2_1=0.01, P=.92), the vignette without mental health problems (χ^2_1=3.14, P=.08), and the control vignette (χ^2_1=0.43, P=.51). A statistically significant difference was observed only for the alcohol-dependence vignette (χ^2_1=7.38, P=.006; Table 2).

Age-group comparisons revealed similar vignette response scores for the depression vignette (χ^2_2=1.57, P=.46), the alcohol-dependence vignette (χ^2_2=1.56, P=.46), the vignette without mental health problems (χ^2_2=4.34, P=.11), and the control vignette (χ^2_2=0.13, P=.94; Table 2).

Table 3 presents the factor loadings for the 2-factor solution derived from principal axis factoring with oblimin rotation. The 4 vignette items loaded more strongly on Factor 2, with values of 0.73 for the depression vignette, 0.83 for the alcohol-dependence vignette, 0.80 for the vignette without mental health problems, and 0.66 for the control vignette.

XSL•FO
RenderX

Table . Factor loadings[a] for a 2-factor solution.

Variable	Factor 1 (PA1[b])	Factor 2 (PA2)	h^2
Depression vignette	0.05	0.73	0.54
Alcohol-dependence vignette	0.00	0.83	0.70
No mental health problems vignette	−0.02	0.80	0.63
Control vignette	−0.04	0.66	0.43
Threatened to hurt or injure someone important	0.83	0.02	0.69
Slapped (past 12 months)	0.94	0.05	0.89
Pushed/held down (past 12 months)	0.95	−0.02	0.90
Punched/kicked (past 12 months)	0.96	0.00	0.92
Strangled/suffocated/burned (past 12 months)	0.92	−0.05	0.85
Threatened/attacked with a weapon	0.84	−0.01	0.71
Sex	0.10	0.11	0.02
Age group	−0.10	−0.12	0.03
Race/color	−0.08	−0.02	0.01

[a]Factor loadings were obtained using principal axis factoring with oblimin rotation.

[b]PA: principal axis.

All 6 violence-related items loaded more strongly on Factor 1, with loadings ranging from 0.83 (threatened to harm or injure someone important) to 0.96 (punched or kicked in the past 12 months). The variable “threatened/attacked with a weapon” showed a loading of 0.84 on Factor 1, while “strangled/suffocated/burned” loaded 0.92 on the same factor (Table 3).

Sociodemographic variables showed low loadings on both factors, with values of 0.10 and 0.11 for sex, –0.10 and –0.12 for age group, and –0.08 and –0.02 for race/color. Communality values (h^2) ranged from 0.01 (race/color) to 0.92 (punched/kicked in the past 12 months; Table 3).

Table 4 displays the squared cosine ($\cos^2$) values and the percentage contributions of each category to Dimensions 1 and 2 of the MCA. Across the violence-related variables, high $\cos^2$ values on Dimension 1 were observed for the “yes” and “no” categories of punched/kicked (0.745), slapped (0.717), pushed/restrained (0.699), strangled/suffocated/burned (0.563), threatened or injured with a weapon (0.520), and threatened to hurt someone important (0.507). These categories also made the largest contributions to Dimension 1, with values ranging from 9.01% to 14.75% for the “yes” categories.

XSL•FO
RenderX

Table . Squared cosine (cos²) and dimension contributions (%) for each variable category in the multiple correspondence analysis (MCA).

Variable and category	Cos²		Contribution	
	Dimension 1	Dimension 2	Dimension 1 (%)	Dimension 2 (%)
Punched/kicked				
Yes	0.745	0.016	14.75	0.45
No	0.745	0.016	3.76	0.11
Slapped				
Yes	0.717	0.021	13.73	0.59
No	0.717	0.021	4.08	0.18
Pushed/restrained				
Yes	0.699	0.038	13.41	1.07
No	0.699	0.038	3.96	0.32
Strangled/suffocated/burned				
Yes	0.563	0.028	12.48	0.92
No	0.563	0.028	1.50	0.11
Threatened/injured with a weapon				
Yes	0.520	0.026	11.56	0.85
No	0.520	0.026	1.35	0.10
Threatened to hurt someone important				
Yes	0.507	0.023	9.01	0.60
No	0.507	0.023	3.59	0.24
Depression vignette				
Very easy	0.017	0.335	0.39	11.43
Easy	0.014	0.251	0.31	8.18
Moderate	0.009	0.004	0.18	0.12
Very difficult	0.022	0.129	0.35	2.96
Extremely difficult	0.007	0.041	0.13	1.15
Alcohol-dependence vignette				
Very easy	0.018	0.353	0.43	12.27
Easy	0.031	0.289	0.67	9.25
Moderate	0.000	0.003	0.00	0.08
Very difficult	0.030	0.151	0.46	3.33
Extremely difficult	0.000	0.040	0.00	1.18
No mental health problems vignette				
Very easy	0.019	0.343	0.44	11.87
Easy	0.026	0.216	0.57	6.96
Moderate	0.000	0.000	0.01	0.00
Very difficult	0.036	0.149	0.58	3.49
Extremely difficult	0.000	0.050	0.00	1.49
Control vignette				
Very easy	0.023	0.316	0.54	10.76
Easy	0.022	0.135	0.47	4.23
Moderate	0.003	0.001	0.06	0.01
Very difficult	0.020	0.155	0.33	3.64

XSL•FO
RenderX

Variable and category	Cos²		Contribution	
	Dimension 1	Dimension 2	Dimension 1 (%)	Dimension 2 (%)
Extremely difficult	0.001	0.032	0.01	1.00
Sex				
Female	0.011	0.013	0.12	0.21
Male	0.011	0.013	0.14	0.25
Age group (y)				
18 - 29	0.004	0.012	0.08	0.33
30 - 44	0.004	0.000	0.06	0.01
≥45	0.015	0.007	0.25	0.15
Race/color				
Non-Black	0.011	0.004	0.19	0.10
Black	0.011	0.004	0.08	0.04

For the vignette response categories, the highest cos² values on Dimension 2 were observed for the "very easy" and "easy" categories across all 4 vignettes. In the depression vignette, cos² values were 0.335 and 0.251; in the alcohol-dependence vignette, 0.353 and 0.289; in the vignette without mental health problems, 0.343 and 0.216; and in the control vignette, 0.316 and 0.135. These categories also showed the highest contributions to Dimension 2, ranging from 6.96% to 12.27% (Table 4).

Sociodemographic variables showed low cos² values in both dimensions. For sex, cos² values were 0.011 for Dimension 1 and 0.013 for Dimension 2, with contributions below 0.25%. Age groups showed cos² values ranging from 0.004 to 0.015 for Dimension 1 and from 0.000 to 0.012 for Dimension 2. Race/color categories showed cos² values of 0.011 on Dimension 1 and 0.004 on Dimension 2, with contributions below 0.20% (Table 4).

Discussion

The findings of this study indicate that exposure to interpersonal violence was frequent in the study population, while perceptions regarding the difficulty of the clinical and social situations described in the vignettes formed a distinct analytical dimension from violence exposure. In addition, perceived difficulty in accessing PHC showed limited variation across sociodemographic groups. This separation was consistently observed across exploratory multivariate analyses, including factor analysis and MCA, which revealed 2 clearly differentiated dimensions: one associated with violence-related experiences and another associated with perceived difficulty accessing PHC as assessed through the vignette scenarios. Differences between experiences and perceptions under stressful conditions have been documented in the psychological and public health literature. As discussed by Krieger [28], such challenging "crossroads" call for further theoretical development and innovative empirical studies.

Discrepancies between actual experiences of violence and individual or community perceptions of violence have been documented across diverse settings, ranging from local conflicts, such as turf wars [29], to large-scale and protracted conflicts [30]. Such differences in perceptions and responses may contribute to variation in psychological responses to adversity, including posttraumatic stress and resilience. However, resilience may be challenged by prolonged or cumulative exposure to adversity, as illustrated by recent analyses of the combined effects of war and environmental stressors [31].

The prevalence levels observed in the sample are consistent with global patterns of physical violence and threats, as documented in international epidemiological studies. Evidence from large multicountry surveys has repeatedly shown that exposure to physical or psychological aggression is widespread across different populations and settings [32-36].

Although most comparisons did not indicate statistically significant sociodemographic differences, the distribution of violence types within groups was broadly aligned with the international literature. Men reported higher frequencies of some forms of aggression, consistent with studies showing that men experience substantial levels of nonpartner physical violence, especially in community or nondomestic contexts [37-40].

At the same time, women reported higher exposure in certain categories in other international surveys, particularly those involving coercive control or severe violence [36,41]. Such patterns reinforce the notion that interpersonal violence is multifaceted and extends beyond intimate-partner contexts.

Limited variation was observed in vignette-response scores by sex, race/color, or age group, with only one statistically significant result (race/color in the alcohol-dependence vignette). International studies using vignette-based approaches have similarly documented variation in how individuals evaluate standardized health-related scenarios across sociodemographic groups [3,42].

Factor analysis results demonstrated a clear 2-factor structure: violence-related items loaded strongly on the first factor, and vignette-evaluation items loaded on the second, with minimal loadings of sociodemographic variables. This separation suggests that, in this sample, previous exposure to interpersonal violence and perceived difficulty accessing PHC through the

XSL•FO
RenderX

vignette scenarios represented distinct analytical dimensions [36].

The MCA results were consistent with this distinction, with "yes" categories for violence defining Dimension 1 and vignette categories defining Dimension 2. Similar patterns of violence clustering have been reported in multicountry studies examining violence and psychosocial variables [37,43,44].

Taken together, the results suggest that violence exposure and perceptions of difficulty accessing PHC appear to represent distinct dimensions in this sample. International literature indicates that interpersonal violence is a strong predictor of mental-health outcomes such as depression, anxiety, posttraumatic stress disorder, and suicidality [33,36,38]. These findings highlight the importance of distinguishing previous exposure to interpersonal violence from perceptions of difficulty accessing PHC, as both may be shaped by different individual, organizational, and contextual factors [45-48].

The limited differentiation in vignette responses across sociodemographic groups also suggests that perceived difficulty accessing PHC cannot be attributed to demographic characteristics alone. Broader structural, organizational, and territorial factors should therefore be considered when developing strategies to improve access to PHC in violence-affected communities [36,37,49,50].

Overall, the results indicate that, although exposure to violence was common and patterned in ways broadly consistent with international evidence, violence-related variables and vignette-based perceptions of difficulty accessing PHC were predominantly represented by distinct dimensions in the exploratory multivariate analyses. This distinction underscores the importance of considering perceived access as a multidimensional construct that may be influenced by factors beyond individual experiences of violence.

These findings also have important implications for the organization of PHC in violence-affected territories. The distinction observed between violence-related variables and vignette-based perceptions of difficulty accessing PHC in the exploratory multivariate analyses highlights the need to consider broader structural, organizational, and social factors that may shape perceived access. Consequently, strategies aimed at improving access should extend beyond the identification of violence exposure and include actions to strengthen service availability, continuity of care, welcoming practices, and coordination of care within PHC.

In the Brazilian context, these implications should be considered in relation to the organization of PHC and the structural conditions affecting access to care. Although the expansion of the FHS has contributed substantially to increased PHC coverage and access, persistent organizational and structural barriers may limit the ability of services to provide timely and continuous care, particularly in socially vulnerable territories [13]. In settings affected by chronic armed violence, these barriers may be compounded by disruptions in service functioning, restrictions on the mobility of residents and health professionals, and difficulties in maintaining stable health care teams. Recent evidence from impoverished communities in Rio de Janeiro has shown that armed violence can affect PHC functioning, with the presence or absence of physicians in primary care teams partially mediating the associations between exposure to armed violence and selected health conditions [14]. Thus, perceptions of difficulty in accessing PHC in Manguinhos should be interpreted not solely in relation to individual characteristics, but also within the broader context of territorial violence and the structural and organizational conditions that may shape access to health services.

However, the observed distinction between violence-related variables and vignette-based perceptions of difficulty accessing PHC should be interpreted with caution. It should not be interpreted as evidence that violence has no influence on perceived access to PHC. Rather, it may reflect the multidimensional nature of access, which is influenced by organizational characteristics of health services, previous experiences with care, social support, and contextual factors that were not directly assessed in this study.

This study has some limitations. First, its cross-sectional design precludes causal inference. Second, the outcome was based on responses to standardized hypothetical vignettes, which reflect perceived difficulty in accessing PHC rather than actual health care usage or experienced barriers to access. Although the vignettes were designed to assess perceived difficulty accessing PHC rather than mental health literacy, responses to hypothetical mental health scenarios may also be influenced by participants' beliefs and understanding of mental health conditions, which have been shown to vary within community populations. Third, although the analyses included key sociodemographic characteristics, other potentially relevant contextual and socioeconomic determinants of access, such as education, income, employment status, previous experiences with health services, and geographic accessibility, were not evaluated. These factors may influence perceived difficulty accessing PHC and should be considered in future studies using multivariable analytical approaches.

Furthermore, the distinction observed between violence-related variables and vignette-based perceptions of difficulty accessing PHC in the exploratory multivariate analyses should be interpreted with caution. This pattern should not be interpreted as evidence of the absence of an association between violence exposure and perceived difficulty accessing PHC. Future studies using complementary measures of access and multivariable analytical approaches may help clarify this relationship.

Understanding the factors that shape perceptions of access to PHC is essential for strengthening health systems in communities affected by chronic urban violence. Although violence-related variables and vignette-based perceptions of difficulty accessing PHC were predominantly represented by distinct dimensions in the exploratory multivariate analyses, the findings suggest that broader structural, organizational, and social determinants deserve greater attention in policies and interventions aimed at improving access. Future studies conducted in different contexts and using complementary measures of access are needed to further clarify these relationships and support the development of more equitable PHC services.

XSL•FO
RenderX

Acknowledgments

The authors kindly acknowledge the Rio de Janeiro State Research Foundation (FAPERJ) for its support. During the preparation of this manuscript, the authors used ChatGPT (GPT-5.6 Sol, OpenAI) to assist with language editing, improving readability, and refining the organization of the text. The AI tool was not used to generate, analyze, or interpret study data, nor to draw scientific conclusions. All outputs generated by the tool were carefully reviewed, verified, and edited by the authors, who take full responsibility for the content of the published manuscript.

Funding

This study was supported by the Rio de Janeiro State Research Foundation (FAPERJ) through the Health Network Program (grant E-26/010.002428/2019), coordinated by FIB. FIB also received career grant awards from FAPERJ and CNPq. The funding agencies had no role in the study design; data collection, analysis, or interpretation; manuscript preparation; or the decision to submit the manuscript for publication. The opinions expressed herein are the authors' exclusive responsibility and do not necessarily reflect those of their institutions or funders.

Data Availability

The datasets generated and/or analyzed during the current study are not publicly available because they contain information that could compromise participant confidentiality and are subject to restrictions imposed by the Research Ethics Committee. Deidentified data may be made available from the corresponding author upon reasonable request, subject to approval by the Research Ethics Committee and applicable institutional and ethical requirements.

Authors' Contributions

Conceptualization: APDC, FIPMB
Data curation: APDC
Formal analysis: APDC, JCdM
Investigation: APDC
Methodology: APDC, JCdM, FIPMB
Project administration: APDC
Supervision: FIPMB
Validation: APDC, JCdM, FIPMB
Visualization: APDC
Writing – original draft: APDC
Writing – review and editing: APDC, JCdM, FIPMB

Conflicts of Interest

None declared.

Multimedia Appendix 1
Full wording of the standardized vignettes used in the study.
[PDF File, 128 KB - publichealth_v12i1e94922_app1.pdf]

Checklist 1
STROBE checklist for cross-sectional studies.
[PDF File, 163 KB - publichealth_v12i1e94922_app2.pdf]

References

1. Minayo MCS. Social violence from a public health perspective. Cad Saude Publica 1994;10(Suppl 1). [doi: 10.1590/S0102-311X1994000500002]
2. Hollie BD, Coolhart D. "A larger system is placing people in this predicament": a qualitative exploration of living amongst urban violence and the impact on mental health and relationships in the Black community. Contemp Fam Ther 2020;42:319-334. [doi: 10.1007/s10591-020-09546-6]
3. Borg BA, Rabinak CA, Marusak HA. Violence exposure and mental health consequences among urban youth. Curr Psychol 2023;42(10):8176-8185. [doi: 10.1007/s12144-021-02141-4] [Medline: 38107483]
4. Coid J, Zhang Y, Ullrich S, et al. Interpersonal violence in a deprived Scottish urban area with aggregations of physical health risks and psychiatric morbidity: an ecological study. BMC Public Health 2021 Jun 12;21(1):1121. [doi: 10.1186/s12889-021-11167-z] [Medline: 34118918]

XSL•FO
RenderX

5. Green G, Swartz A, Tembo D, et al. A scoping review of how exposure to urban violence impacts youth access to sexual, reproductive and trauma health care in LMICs. Glob Public Health 2023 Jan;18(1):2103581. [doi: 10.1080/17441692.2022.2103581] [Medline: 35938416]
6. Pérez-Sastré MA, García-Peña C, Ramos-Lira L, Ortiz-Hernández L. Beyond direct exposure to violence: effects of living in disordered and violent communities on psychological distress in young Mexican people. Cad Saude Publica 2024;40(1):e00058123. [doi: 10.1590/0102-311XEN058123] [Medline: 38324861]
7. Vieira-Meyer A, Vieira NF, Machado FA, et al. Urban violence and its effect on primary health care. Eur J Public Health 2025 Oct;35(Supplement_4). [doi: 10.1093/eurpub/ckaf161.1420]
8. da Cunha AP, Corrêa da Mota J, Bastos FI. A city under siege: time-series analysis of shootouts and the association with health and living conditions in Rio de Janeiro, Brazil, 2017 to 2022. J Interpers Violence 2026 Aug;41(15-16):5504-5525. [doi: 10.1177/08862605251356234] [Medline: 40708161]
9. Perlman J. Favela: Four Decades of Living on the Edge in Rio de Janeiro: Oxford University Press; 2010.
10. Dicionário de Favelas Marielle Franco. Favela de manguinhos [Article in Portuguese]. Wikifavelas. 2024. URL: https://wikifavelas.com.br/index.php/Dicion%C3%A1rio_de_Favelas_Marielle_Franco [accessed 2026-07-10]
11. Fernandes TM, Costa RGR. As comunidades de Manguinhos na história das favelas no Rio de Janeiro. Tempo 2013;19(34):117-133. [doi: 10.5533/TEM-1980-542X-2013173410]
12. Lyra D, editor. Infância e trânsito em manguinhos [Report in Portuguese]. : CECIP (Centro de Criação de Imagem Popular); 2012 URL: https://www.cecip.org.br/site/wp-content/uploads/2016/08/Manguinhos-e-seus-Caminhos-relatorioCPFoco.pdf [accessed 2026-07-10]
13. Tesser CD, Norman AH, Vidal TB. Acesso ao cuidado na Atenção Primária à Saúde brasileira: situação, problemas e estratégias de superação [Article in Portuguese]. Saúde debate 2018;42(spe1):361-378. [doi: 10.1590/0103-11042018s125]
14. Mota JCD, Cunha APD, Bastos FI. Armed violence and shootouts in Rio de Janeiro: relevance to impoverished communities and primary care health services. Cien Saude Colet 2025 Oct;30(10):e12982025. [doi: 10.1590/1413-812320253010.12982025] [Medline: 41221872]
15. Sampson RJ. Great American City: Chicago and the Enduring Neighborhood Effect: University of Chicago Press; 2012.
16. Aiyer SM, Zimmerman MA, Morrel-Samuels S, Reischl TM. From broken windows to busy streets: a community empowerment perspective. Health Educ Behav 2015 Apr;42(2):137-147. [doi: 10.1177/1090198114558590] [Medline: 25512073]
17. Krug EG, Dahlberg LL, Mercy JA, Zwi AB, Lozano R, editors. World report on violence and health. : World Health Organization; 2002 URL: https://iris.who.int/server/api/core/bitstreams/a25476ed-8585-47f3-986e-9d0e7f5e9f1b/content [accessed 2026-09-01]
18. Levesque JF, Harris MF, Russell G. Patient-centred access to health care: conceptualising access at the interface of health systems and populations. Int J Equity Health 2013 Mar 11;12:18. [doi: 10.1186/1475-9276-12-18] [Medline: 23496984]
19. Topbas M, Dilaver I, Karakullukcu S, et al. What influences the perceived access to healthcare services? An investigation using structural equation modeling. Front Public Health 2026;14:1844139. [doi: 10.3389/fpubh.2026.1844139] [Medline: 42326952]
20. Ferguson CJ. An effect size primer: a guide for clinicians and researchers. Prof Psychol Res Pract 2009;40(5):532-538. [doi: 10.1037/a0015808]
21. Ekström J. A generalized definition of the polychoric correlation coefficient. : Department of Statistics, University of California, Los Angeles (UCLA); 2011 URL: https://escholarship.org/content/qt583610fv/qt583610fv.pdf?v=lg [accessed 2026-09-01]
22. Jackson JE. Oblimin rotation. In: Armitage P, Colton T, editors. Encyclopedia of Biostatistics: John Wiley & Sons; 2005. [doi: 10.1002/0470011815]
23. Wickham H, Averick M, Bryan J, et al. Welcome to the tidyverse. J Open Source Softw 2019;4(43):1686. [doi: 10.21105/joss.01686]
24. Revelle W. Psych: procedures for psychological, psychometric, and personality research. R package version 2.6.1. CRAN (Comprehensive R Archive Network). 2026. URL: https://CRAN.R-project.org/package=psych [accessed 2026-09-01]
25. Lê S, Josse J, Husson F. FactoMineR: an R package for multivariate analysis. J Stat Soft 2008;25(1):1-18. [doi: 10.18637/jss.v025.i01]
26. Kassambara A, Mundt F, Erdey L. Factoextra: extract and visualize the results of multivariate data analyses. CRAN (Comprehensive R Archive Network). 2026. URL: https://CRAN.R-project.org/package=factoextra [accessed 2026-09-01]
27. Venables WN, Ripley BD. Modern Applied Statistics with S, 4th edition: Springer; 2002.
28. Krieger N. Living and dying at the crossroads: racism, embodiment, and why theory is essential for a public health of consequence. Am J Public Health 2016 May;106(5):832-833. [doi: 10.2105/AJPH.2016.303100] [Medline: 27049420]
29. Modan GG. Turf Wars: Discourse, Diversity, and the Politics of Place: Blackwell Publishing; 2007. [doi: 10.1002/9780470773970]
30. Litz BT. Resilience in the aftermath of war trauma: a critical review and commentary. Interface Focus 2014 Oct 6;4(5):20140008. [doi: 10.1098/rsfs.2014.0008] [Medline: 25285196]

XSL•FO
RenderX

31. Döring S, Hecker T, Hall J. Drought amplifies the psychological burden of war. Nat Sustain 2026. [doi: 10.1038/s41893-026-01882-z]
32. Garcia-Moreno C, Jansen HA, Ellsberg M, Heise L, Watts CH. Prevalence of intimate partner violence: findings from the WHO multi-country study on women's health and domestic violence. Lancet 2006 Oct 7;368(9543):1260-1269. [doi: 10.1016/S0140-6736(06)69523-8] [Medline: 17027732]
33. Ellsberg M, Jansen H, Heise L, Watts CH, Garcia-Moreno C. Intimate partner violence and women's physical and mental health in the WHO multi-country study on women's health and domestic violence: an observational study. Lancet 2008 Apr 5;371(9619):1165-1172. [doi: 10.1016/S0140-6736(08)60522-X] [Medline: 18395577]
34. Sardinha L, Maheu-Giroux M, Stöckl H, Meyer SR, García-Moreno C. Global, regional, and national prevalence estimates of physical or sexual, or both, intimate partner violence against women in 2018. Lancet 2022 Feb 26;399(10327):803-813. [doi: 10.1016/S0140-6736(21)02664-7] [Medline: 35182472]
35. Clark CJ, Bergenfeld I, Cheong YF, et al. Patterns of women's exposure to psychological violence: a global examination of low- and middle-income countries. SSM Popul Health 2023;24:101500. [doi: 10.1016/j.ssmph.2023.101500] [Medline: 37727254]
36. White SJ, Sin J, Sweeney A, et al. Global prevalence and mental health outcomes of intimate partner violence among women: a systematic review and meta-analysis. Trauma Violence Abuse 2024 Jan;25(1):494-511. [doi: 10.1177/15248380231155529] [Medline: 36825800]
37. Abramsky T, Watts CH, Garcia-Moreno C, et al. What factors are associated with recent intimate partner violence? Findings from the WHO multi-country study on women's health and domestic violence. BMC Public Health 2011 Feb 16;11:109. [doi: 10.1186/1471-2458-11-109] [Medline: 21324186]
38. Fanslow JL, Malihi Z, Hashemi L, Gulliver P, McIntosh T. Prevalence of interpersonal violence against women and men in New Zealand: results of a cross-sectional study. Aust N Z J Public Health 2022 Apr;46(2):117-126. [doi: 10.1111/1753-6405.13206] [Medline: 34978353]
39. Scott-Storey K, O'Donnell S, Ford-Gilboe M, et al. What about the men? A critical review of men's experiences of intimate partner violence. Trauma Violence Abuse 2023 Apr;24(2):858-872. [doi: 10.1177/15248380211043827] [Medline: 35094633]
40. Fanslow JL, Gulliver PJ, Mellar BM, Selak V, Mcintosh TKD. Men's lifetime interpersonal violence exposure and associations with hospitalization at the national level: results from a New Zealand population-based study. Scand J Public Health 2025 Aug;53(6):638-649. [doi: 10.1177/14034948251355754]
41. Kifle ME, Aychiluhm SB, Anbesu EW. Global prevalence of intimate partner violence during the COVID-19 pandemic among women: systematic review and meta-analysis. BMC Womens Health 2024 Feb 17;24(1):127. [doi: 10.1186/s12905-023-02845-8] [Medline: 38368323]
42. Ciciurkaite G, Pescosolido BA. Mental health literacy and public stigma: examining the link in 17 countries. Med Res Arch 2024 Jul;12(7). [doi: 10.18103/mra.v12i7.5471] [Medline: 39238545]
43. Jewkes R, Jordaan E, Myrttinen H, Gibbs A. Masculinities and violence: using latent class analysis to investigate the origins and correlates of differences between men in the cross-sectional UN Multi-country Study on Men and Violence in Asia and the Pacific. J Glob Health 2020 Dec;10(2):020439. [doi: 10.7189/jogh.10.020439] [Medline: 33437463]
44. Pettersson R, Lucas S, Strandh M. Patterns of violence exposure in a life-course perspective and associations to mental and physical health problems and health-related risk behaviors among women and men in Sweden: a latent class analysis. SSM Popul Health 2025 Dec;32:101874. [doi: 10.1016/j.ssmph.2025.101874] [Medline: 41281886]
45. Ogbe E, Harmon S, Van den Bergh R, Degomme O. A systematic review of intimate partner violence interventions focused on improving social support and/ mental health outcomes of survivors. PLoS One 2020;15(6):e0235177. [doi: 10.1371/journal.pone.0235177] [Medline: 32584910]
46. Paphitis SA, Bentley A, Asher L, Osrin D, Oram S. Improving the mental health of women intimate partner violence survivors: findings from a realist review of psychosocial interventions. PLoS One 2022;17(3):e0264845. [doi: 10.1371/journal.pone.0264845] [Medline: 35299229]
47. Mellen EJ, Hatzenbuehler ML. Sexual violence-related stigma, mental health, and treatment-seeking: a multimodal assessment in a population-based study of young adults. J Interpers Violence 2023 Oct;38(19-20):11243-11271. [doi: 10.1177/08862605231179715] [Medline: 37491905]
48. Mellen EJ, Kim DY, Edenbaum ER, Cellini J. The psychosocial consequences of sexual violence stigma: a scoping review. Trauma Violence Abuse 2025 Jul;26(3):389-402. [doi: 10.1177/15248380241279860] [Medline: 39377179]
49. Bhavsar V, Bhugra D, Persaud A. Public policy, mental health and a public health approach to violence. Int J Soc Psychiatry 2020 Nov;66(7):629-632. [doi: 10.1177/0020764020915236] [Medline: 32223492]
50. Oram S, Fisher HL, Minnis H, et al. The Lancet Psychiatry Commission on intimate partner violence and mental health: advancing mental health services, research, and policy. Lancet Psychiatry 2022 Jun;9(6):487-524. [doi: 10.1016/S2215-0366(22)00008-6] [Medline: 35569504]

Abbreviations

CAAE: Certificado de Apresentação para Apreciação Ética (Certificate of Presentation for Ethical Consideration)

XSL•FO
RenderX

FHS: Family Health Strategy
MCA: multiple correspondence analysis
PHC: primary health care
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Mavragani, T Sanchez; submitted 08.Mar.2026; peer-reviewed by B Chakalov, F Carrouel; revised version received 21.Aug.2026; accepted 24.Aug.2026; published 28.Sep.2026.

Please cite as:
Cunha APD, da Mota JC, Bastos FI
Perceived Difficulty Accessing Primary Health Care and Its Associations With Sociodemographic Factors and Exposure to Violence in a Violent Urban Community in Rio de Janeiro, Brazil: Cross-Sectional Study
JMIR Public Health Surveill 2026;12:e94922
URL: https://publichealth.jmir.org/2026/1/e94922
doi:10.2196/94922

© Ana Paula da Cunha, Jurema Corrêa da Mota, Francisco Inacio Bastos. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 28.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Factors Influencing Response Rates in National Internet-Based Questionnaires: Analysis of a Large National Survey

Søren Birkeland[1,2], LLM, PhD; Anders Haakonsson[1], MSc, PhD; Sören Möller[3], MSc, PhD

[1]Odense University Hospital, Open Patient data Exploratory Network (OPEN), J. B. Winsløws Vej 21, 3rd floor, 5000 Odense, Denmark

[2]Forensic Mental Health Research Unit Middelfart (RFM), Department of Regional Health Research, University of Southern Denmark, Middelfart, Region of Southern Denmark, Denmark

[3]Department of Public Health, University of Southern Denmark, Odense, South Denmark, Denmark

Corresponding Author:
Søren Birkeland, LLM, PhD
Odense University Hospital, Open Patient data Exploratory Network (OPEN), J. B. Winsløws Vej 21, 3rd floor, 5000 Odense, Denmark

Abstract

Background: Digital mailbox platforms established by authorities for secure communication with the general populace are increasingly available for distribution of surveys. While such solutions constitute an attractive way of easily distributing surveys, nonresponse remains a critical problem, warranting measures to increase participation.

Objective: We aimed to investigate questionnaire characteristics predicting participation and the effect of reminders among men in a survey distributed through a national digital mailbox solution.

Methods: This study used an internet-based survey among men aged 45 to 70 years; the survey was sent out in 2 waves through a digital mailbox used by Danish authorities. Predictors for response were analyzed using logistic regression.

Results: Of 24,000 invitations, 22,288 participants accepted, with 6756 (30.3%) completing the survey. We found some evidence suggesting a rule of thumb that reminders may add another half to the response rate achieved in a preceding invitation, with 21% and 10% of responses obtained before and after the reminder, respectively (in the first wave). Longer questionnaire variants (odds ratio [OR] 0.77, 95% CI 0.66 - 0.88; $P<.001$ [longest vs shortest variants]), questionnaires with attached informational material (OR 0.92, 95% CI 0.85 - 0.98; $P=.02$[with vs without material]), or with potentially distressful emotional content (OR 0.91, 95% CI 0.85 - 0.98; $P=.008$ [most vs least distressful content]) tended to yield lower response rates. Furthermore, older age was associated with a higher response rate when comparing the oldest (65 - 70 years) to the youngest (45 - 55 years) age groups (OR 2.49, 95% CI 2.31 - 2.70; $P<.001$).

Conclusions: We confirmed some previous findings obtained with other modes of survey distribution on the association between response rates and questionnaire characteristics, showing that it is important to pay attention to specific characteristics when designing internet-based surveys distributed through digital mailboxes. Nonresponse may constitute a considerable and increasing challenge, and reminders may help substantially increase participation.

(JMIR Public Health Surveill 2026;12:e89761) doi:10.2196/89761

KEYWORDS

multidisciplinary research; public health; social research; survey methods; survey participation

Introduction

In many countries, electronic survey distribution through internet-based services and web platforms is a relatively new opportunity for research, and it is being increasingly used for gathering data to inform policy and practice [1]. The Danish digital mailbox (e-Boks) is one example [1], with similar digital mailbox solutions available in a rising number of countries (Norway, Sweden, Greenland, and Ireland) [2]. Survey delivery through this channel has provided many new advantages and opportunities for easy survey distribution and straightforward subsequent analysis of the gathered electronic data. Furthermore, a digital mailbox with web access constitutes an easy platform for potential respondents to access and fill out surveys without requiring, for example, manually filling out and posting paper material; therefore, it could be hypothesized that such platforms would promote research participation. Despite these advantages, nonparticipation remains a critical challenge for internet-based surveys, including those distributed via digital mailboxes, with potential implications for practice, policymaking, and public opinion [1]. Moreover, there is significant evidence from multiple studies that participation rates in public health surveys have been generally decreasing for the last several decades [3,4]. Correspondingly, there is emerging evidence that participation in web-based health-related surveys is declining [5-7]. Various

XSL•FO
RenderX

measures have been proposed to mitigate this, such as use of monetary incentives to increase response rates [6]. Another traditional approach is to use reminders [7,8].

In parallel, previous research has demonstrated that participation varies with geography, with participation usually being higher in less-populated areas [9-13]; participation also usually increases with older age, although this tendency can change when older people are less familiar with, for example, digital platforms [5-7,13-17].

Participation may, however, also be dependent on features of the survey [18]. Most of our knowledge on this point derives from surveys from the predigital era. Hence, limited research exists on internet-based solutions and, particularly, on surveys distributed through national-level web platforms. This paper reports on experiences from a research project using a national authority's digital mailbox solution, focusing on survey characteristics that predict response rates and the effect of reminders.

Methods

Overview

This is a secondary analysis of data obtained from a large, cross-sectional, national-level web survey about health care user involvement in decision-making regarding prostate-specific antigen (PSA) testing; the survey was distributed among adult men aged 45 to 70 years; this age range was chosen based on US Preventive Services Task Force and European recommendations for PSA screening for early detection of prostate cancer [19,20]. The survey was developed with public and patient involvement and has been previously described in detail [21]. We previously reported an overall good representation of the general population of men aged 45 to 70 years in our sample; however, response rates tended to increase with older age and a location of residence that was rural and had a lower tax base, fewer people with higher education, and fewer people with non-Western backgrounds [13]. In agreement with prior research, our analysis also suggested that the response rate may have varied across different versions of the questionnaire [13,18,22-24].

The questionnaire contained items covering sociodemographic information, personality (using the validated Big Five Inventory-10 [25]), preferences for control regarding health care decision-making (using the validated Control Preferences Scale [26]), purpose-designed health-related questions, and hypothetical case vignettes concerning PSA screening for prostate cancer in men. Each scenario had an identical core structure but differed regarding the degree of patient involvement, particularly information delivery (5 levels, from "no information" to "substantiated information," ie, fully shared decision-making with a decision aid provided as an attached PDF information sheet); the decision to have a PSA test or not; and 3 different outcomes (no prostate cancer, diagnosis of a treatable prostate cancer, and diagnosis of an eventually lethal prostate cancer) potentially posing 3 different levels of distress. Because of differences in scenario content, questionnaire variants inevitably had various lengths (ie, word count), ranging from short to long variants requiring participants to read through extensive survey material and ranging from variants with low potential for emotional distress to variants with high potential for distress. To ensure a quasi-random assignment of participants, they were sorted by date of birth and then consecutively allocated into 1 of the 30 different case vignette scenarios driving the key "exposure" variable.

In Denmark, all residents are registered in a civil registration system with a unique personal identification number. The sample for the study was identified through use of these civil registration numbers, which were obtained from the Danish Health Data Authority. Civil registration numbers permit matching to national-level statistical data on age and region and to municipal-level data from the Danish municipal statistical database [27] on the residential location of individuals, including data on rurality (ie, population density), socioeconomic surroundings (ie, municipal tax base), municipal educational level, and municipal proportion of residents with a non-Western background. The latter data are commonly used as a standard measure of state and municipality in Denmark, as well as for research purposes [28,29].

We used REDCap and distributed the survey through the digital mailbox e-Boks, which was introduced by the Danish authorities in 2001; from 2014 onward it has been a mandatory communication platform used by Danish authorities for secure communication with residents of Denmark. Residents receive an email and/or text message when they receive new mail in their private e-Boks account, which can be accessed directly with multifactor authentication to safeguard data security. e-Boks has a very high coverage of the Danish adult population, with only a small proportion of residents having actively deregistered (in 2017 this was 9.9%) [30].

To ensure a sufficient sample size, we obtained permission to invite 2 independent random samples of men aged 45 to 70 years in 2 waves, each receiving an initial invitation; a 14-day reminder sent to nonrespondents. The 2 waves were unconnected in so far as they were 2 different nonoverlapping random samples of the population. We sent out the first (winter) wave of invitations on January 24, 2019, with the 14-day reminders sent on February 7, 2019. We launched the second (spring) wave on March 7, 2019, with the 14-day reminders sent to nonrespondents on March 21, 2019.

Statistics and Power Analysis

Power calculations and the required sample size for conducting the main study have been previously reported [31]. Categorical characteristics were reported as counts and proportions, while numerical characteristics (eg, response time) were reported as medians with IQRs or means with SDs. Odds ratios (ORs) and *P* values for predictors of response to the questionnaire were estimated by logistic regression, both for each predictor on its own and in a multivariate model adjusting for all predictors mutually. As we previously found that response rates varied with age, rurality, tax base, education, and background, we included these variables in our analysis [13]. We did not use survey weights, as our invited sample was a random sample of the total population; hence, it represented a balanced sample. The participating individuals were not perfectly representative

XSL•FO
RenderX

of the population, but as our outcome was completion of the survey and not the answers provided to the survey, applying weights would have counteracted the purpose of the study, that is, to investigate patterns in who responds to a survey.

Ethical Considerations

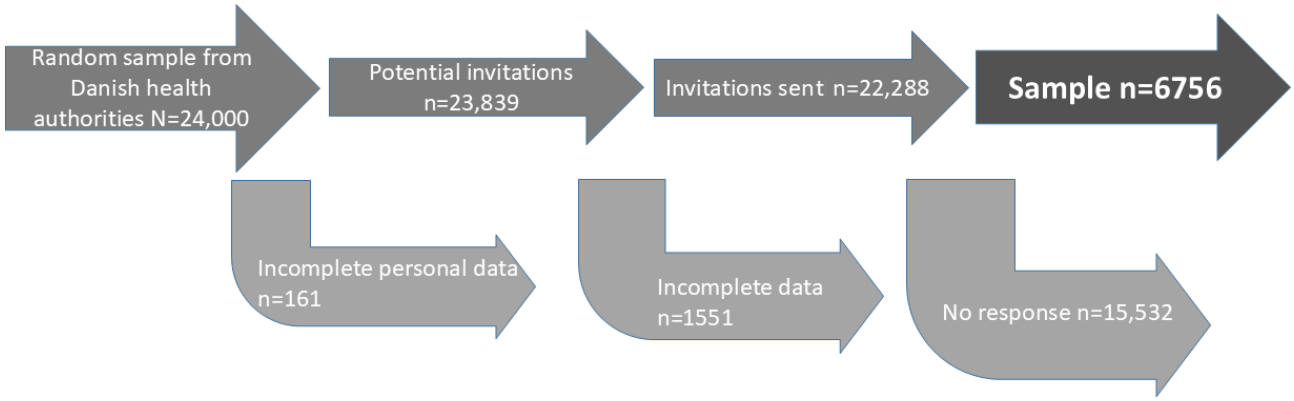
According to Regional Committees on Health Research Ethics for Southern Denmark assessment 2018 - 1906, the project was in accordance with Danish research ethics legislation. Furthermore, use of personal data was approved by the Danish Data Protection Agency (assessment 18/31191). Informed consent for survey participation was obtained through the email invitation delivered to participants in the digital mailbox and, following pseudonymization, the data were stored in a secured environment. Participants received no compensation.

Results

Of 24,000 potential invitees, we excluded 161 (0.67%) from the study, as at the time of the survey, they were no longer aged between 45 and 70 years. Of the 23,839 invitations sent out, 1551 (6.5%) were undeliverable because no digital mailbox was available and were marked incomplete. This number is in accordance with the expected compliance with the Danish national digital mailbox infrastructure [32]. Among the 22,288 people successfully invited through the digital mailbox, 15,532 (69.7%) provided partial or no response, and 6756 (30.3%) completed the questionnaire (Figure 1).

Following the first wave, 30.6% (3395/11,111) of respondents provided a complete response, including 20.6% (2293/11,111) before the 14-day reminder and an additional 9.9% (1102/11,111) after the reminder. Following the second wave, the complete response rate was 30.1% (3361/11,177), including 21.0% (2352/11,177) before the reminder and 9.0% (1009/11,177) after the reminder.

Figure 1. Flow chart showing retrieved sample from authorities, invitations sent out, and responses achieved constituting the sample.

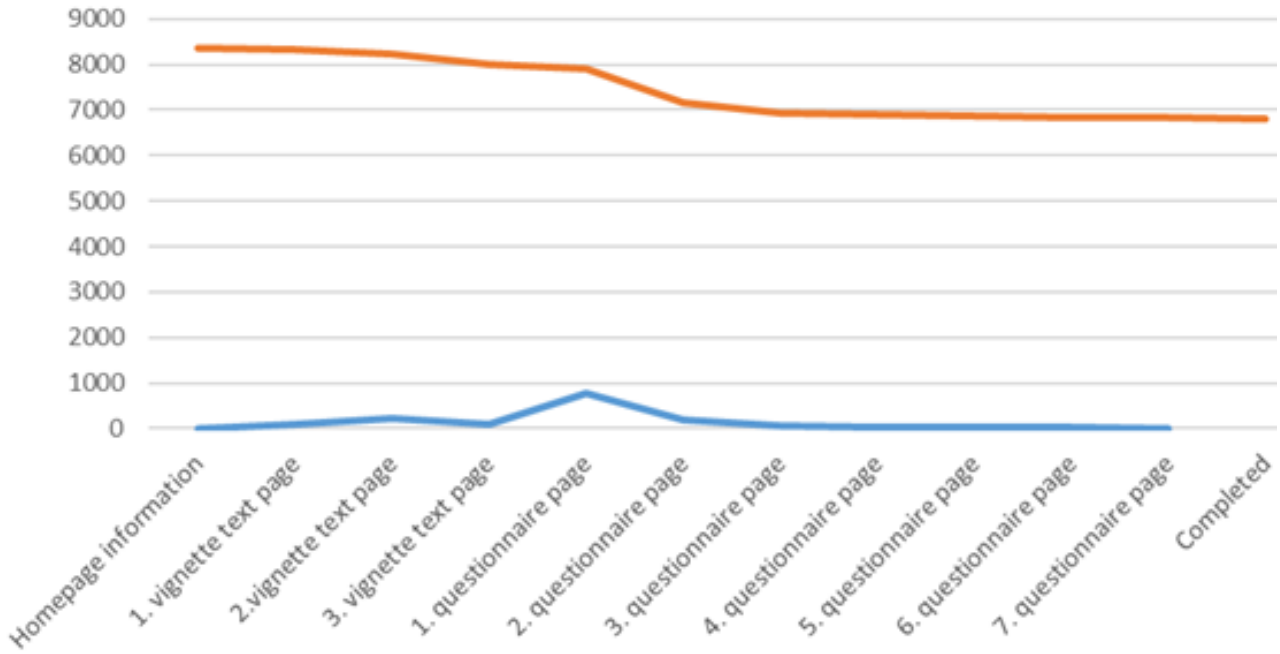


Among respondents, 30% (2022/6756) responded on the day the survey was launched (Thursday), 28% (1883/6756) on the second day (Friday), 14% (948/6756) on the third day (Saturday), 10% (648/6756) on the fourth day (Sunday), 8% (567/6756) on the fifth day (Monday), 6% (387/6756) on the sixth day (Tuesday), and 4% (301/6756) on the seventh day (Wednesday), showing a gradual decrease. Most participants responded in the afternoon and early evening.

Among 1579 people who started the questionnaire but did not complete it (a number that might include people who started more than once), 1564 finished the survey homepage describing the purpose of the survey, participants' rights, the handling of research data, and the instructions, which included pilot-testing information that the survey would take approximately 10 minutes to fill out. After this, 1477 finished the first case vignette text page, 1251 the second page, and 1149 the third page. The majority of dropouts occurred after the first questionnaire page, with only 378 remaining. Following the second questionnaire page (including variants with extra text), 177 remained; 120 remained after the third questionnaire page, 97 after the fourth questionnaire page, 79 after the fifth questionnaire page, 61 after the sixth questionnaire page; finally, 51 people remained after the seventh questionnaire page but then dropped out before finalizing the questionnaire (Figure 2).

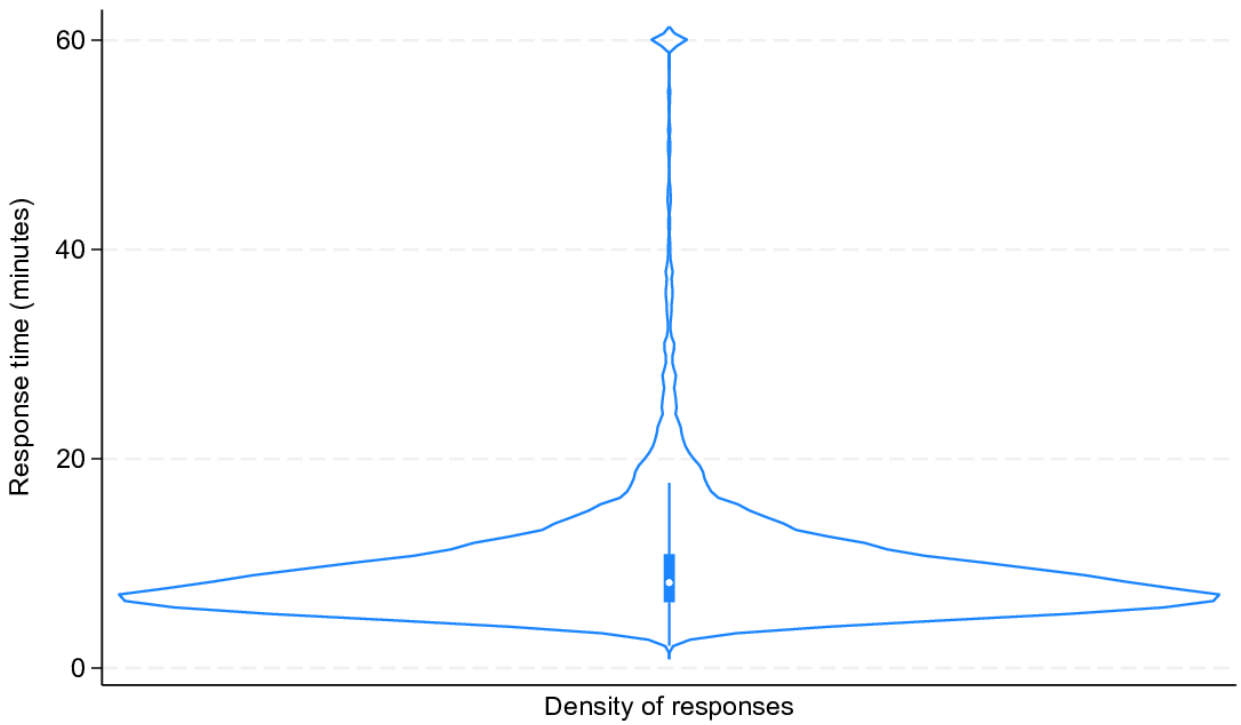
XSL•FO
RenderX

Figure 2. Remaining participants and number of dropouts through the questionnaire, from the homepage and to the pages of the questionnaire itself. The orange curve shows remaining participants at each stage and the blue curve shows the number of dropouts.



The average response time was 11.1 (SD 46.7) minutes and the median response time was 8.2 (IQR 6.3-10.9) minutes. The large difference between the average and median response times was because some responders needed several hours to complete the survey (Figure 3). There was some correlation between the time used to respond to the questionnaire and the number of middle-category responses, but this only reached statistical significance for questionnaire items 2 ($P<.001$) and 9 ($P=.02$). The median age of respondents was 59 (IQR 53-65) years. Other survey respondent characteristics are presented in Table 1.

Figure 3. Violin plot of response times in minutes, with 33 responses of more than 60 minutes winsorized to 60 minutes to improve readability.



XSL•FO
RenderX

Table . Survey respondent characteristics and number of respondents compared to number of nonrespondents.

Respondent characteristics	Respondents (n=6756), n (%)	Nonrespondents (n=15,532), n (%)
Age groups (years)		
Youngest (45 - 55)	2160 (32)	7384 (48)
Middle (55 - 65)	2774 (41)	5662 (36)
Oldest (65 - 70)	1822 (27)	2486 (16)
Municipal tax base per person (DKK[a])		
Lower half (157,077 - 180,414)	3516 (52)	7669 (49)
Upper half (180,437 - 377,409)	3240 (48)	7863 (51)
Urban or rural area (dichotomized according to population density in persons/km2)		
Rural (15 - 145)	3553 (53)	7755 (50)
Urban (147 - 11,949)	3203 (47)	7777 (50)
Proportion of people in a municipality with a non-Western background		
Least half (0.8% - 3.9%)	3587 (53)	7864 (51)
Highest half (3.9% - 14.4%)	3169 (47)	7668 (49)
Proportion of people aged 25 - 64 years with higher education		
Lower half (15.3% - 28.1%)	3592 (53)	7884 (51)
Upper half (28.2% - 59.8%)	3164 (47)	7648 (49)
Location (5 regions)		
Capital Region	1740 (26)	4660 (30)
Zealand	1104 (16)	2456 (16)
Southern Denmark	1611 (24)	3350 (22)
Central Region	1614 (24)	3423 (22)
Northern Jutland	687 (10)	1643 (11)

[a]1 DKK=US $0.156.

Table 2 presents the associations between response rates and different survey characteristics. In addition to the target group characteristics, several features of the survey were statistically significantly associated with the survey response rate, including length, attachment of PDF information material with the survey (as a decision aid, as described above; unadjusted analysis), and emotional content (adjusted analysis).

XSL•FO
RenderX

Table . Association of survey and respondent parameters with response rates in a logistic regression (raw and adjusted analyses).

			Response rate, n (%)	Unadjusted OR[a] (95% CI)	*P* value	Adjusted OR (95% CI)	*P* value
Survey characteristics							
	Questionnaire length (words)						
		900 - 1000 (level 1)	33 (727/2223)	1 (reference)		1 (reference)	
		1000 - 1100 (level 2)	30 (1799/5963)	0.89 (0.80 - 0.99)	.03	0.90 (0.81-1.00)	.048
		1100 - 1200 (level 3)	31 (1614/5173)	0.93 (0.84 - 1.04)	.20	0.96 (0.85 - 1.09)	.53
		1200 - 1300 (level 4)	30 (1320/4438)	0.87 (0.78 - 0.97)	.01	0.99 (0.84 - 1.18)	.95
		2000 - 2100 (level 5)	30 (892/3001)	0.87 (0.77 - 0.98)	.02	0.87 (0.77 - 0.98)	.02
		2100 - 2200 (level 6)	27 (404/1490)	0.77 (0.66 - 0.88)	<.001	0.87 (0.72 - 1.06)	.17
	Material attached or not						
		No attachment	31 (5460/17,797)	1 (reference)	—[b]	N/A[c]	N/A
		With attachment	29 (1296/4491)	0.92 (0.85 - 0.98)	.02	N/A	N/A
	Content (3 levels of emotional distress)						
		No	31 (2308/7467)	1 (reference)	—	1 (reference)	—
		Little	31 (2304/7406)	1.01 (0.94 - 1.08)	.79	0.98 (0.89 - 1.08)	.70
		Much	29 (2144/7415)	0.91 (0.85 - 0.98)	.008	0.86 (0.75 - 0.99)	.03
Survey distribution characteristics							
	Timing of survey invitation						
		Winter wave (January 2019)	31 (3395/11,111)	1 (reference)	—	1 (reference)	—
		Spring wave (March 2019)	30 (3361/11,177)	0.98 (0.92 - 1.03)	0.43	0.98 (0.93 - 1.04)	.58
Characteristics of survey target group							
		Age (years)					
			Youngest (45-55)	23 (2160/9544)	1 (reference)	—	1 (reference)	—
			Middle (55-65)	33 (2774/8436)	1.67 (1.57 - 1.79)	<.001	1.67 (1.56 - 1.78)	<.001
			Oldest (65-70)	42 (1822/4308)	2.51 (2.32 - 2.71)	<.001	2.49 (2.31 - 2.70)	<.001
		Municipal tax base per resident					
			Lower half	31 (3516/11,185)	1 (reference)	—	1 (reference)	—
			Upper half	29 (3240/11,103)	0.90 (0.85 - 0.95)	<.001	1.01 (0.92 - 1.11)	.83
		Urban or rural area (dichotomized according to population density)					
			Rural (lower populated half)	31 (3553/11,308)	1 (reference)	—	1 (reference)	—

XSL•FO
RenderX

		Response rate, n (%)	Unadjusted OR[a] (95% CI)	*P* value	Adjusted OR (95% CI)	*P* value
	Urban (upper populated half)	29 (3203/10,980)	0.90 (0.85 - 0.95)	<.001	1.08 (0.99 - 1.19)	.10
Proportion of residents with a non-Western background in municipality						
	Lower half	31 (3587/11,451)	1 (reference)	—	1 (reference)	—
	Upper half	29 (3169/10,837)	0.91 (0.86 - 0.96)	.001	0.95 (0.87 - 1.03)	.20
Proportion of residents aged 25 - 64 y with higher education						
	Lower half	31 (3592/11,476)	1 (reference)	—	1 (reference)	—
	Upper half	29 (3164/10,812)	0.91 (0.86 - 0.96)	.001	0.95 (0.88 - 1.03)	.24
Geographical location						
	Capital Region	27 (1740/6400)	1 (reference)	—	1 (reference)	—
	Zealand	31 (1104/3560)	1.20 (1.10 - 1.32)	<.001	1.05 (0.92 - 1.20)	.48
	Southern Denmark	32 (1611/4961)	1.29 (1.19 - 1.40)	<.001	1.24 (1.13 - 1.37)	<.001
	Central Region	32 (1614/5037)	1.26 (1.16 - 1.37)	<.001	1.29 (1.16 - 1.43)	<.001
	Northern Jutland	29 (687/2330)	1.12 (1.01 - 1.24)	.03	1.16 (1.04 - 1.29)	.007

[a]OR: odds ratio.

[b]Not applicable.

[c]Not included in adjusted analysis due to full collinearity with questionnaire length.

Discussion

Summary of Main Study Findings

This study examined questionnaire characteristics that predicted participation in an internet-based survey that used a national digital mailbox; the results showed that reminders added another half to the response rate, while longer questionnaire variants, questionnaires with attached informational material, and questionnaires with potentially distressful emotional content tended to yield lower response rates. Below, we discuss these findings in the context of the existing literature.

Study Findings and Comparison to Existing Literature

Questionnaire Length and Content

We found that the response rate was inversely associated with the length of the questionnaire. Among those who started completing the questionnaire but gave up, the majority stopped immediately after switching from the first introductory text page to the first questionnaire item requiring a response. After this point, the dropout rates were greatest for the earlier items. Many dropouts, however, also occurred during the part of the questionnaire with vignette variants that had more text material. Corresponding with this, various previous studies have found a relationship between increasing questionnaire length and decreasing response rates in both mail and internet surveys [18,22-24], a relationship that was also established in a Cochrane review of postal and electronic surveys [8]. The relationship between questionnaire length and response rate seems to hold whether looking at the number of survey pages, screens, or items or at the time to complete the questionnaire [18]. In our survey, the average response time was 11 minutes, while some previous research has suggested that a completion time of no more than 13 minutes is necessary to obtain a good response rate [18]. Likewise, our finding that questionnaire variants with more potential to cause distress had statistically significantly lower response rates confirms findings from the Cochrane review [8] and a widely held belief that questions causing emotional discomfort decrease participation [18,33].

Geographical and Target Group Factors

We found that older men had higher response rates than younger men. On the other hand, people residing in areas with a higher tax base, areas with a higher proportion of residents with higher education, and areas with more residents with a non-Western background had lower response rates than people living in lower tax-base areas, areas with fewer people with higher education, and areas with fewer people with a non-Western background. In addition, response rates were lower in urban areas than in rural areas. There has been a paucity of research on the role of geographical factors in survey response rates [18]. However,

XSL•FO
RenderX

in a Norwegian study using data from a national survey of patient experiences with maternity care, Sjetne et al [16] found that the likelihood of responding to surveys was strongly influenced by geographical location. In our study, we found that response rates were lowest in the Danish Capital Region. Corresponding with this, in addition to an overall decline in survey participation rates, Tolonen and colleagues [12] found an overall lower participation rate in the capital region. Regarding target group characteristics affecting response rates, Sjetne et al [16] found that the likelihood of responding to a survey was strongly influenced by background variables, including age. Hence, older age was associated with higher response probability. Likewise, in a previous national study conducted with women in the United Kingdom, respondents were more likely to be older, married, living in the least-deprived areas, and to be born in the United Kingdom [17]. Hence, our findings suggest that older people may not necessarily be less inclined to participate in research making use of web surveys and digital platforms. Findings of higher participation among older people thereby contradict previous digital health technology evaluation studies highlighting nonparticipation by older people and those with less experience with digital platforms [34].

Effect of Reminders

Previous studies have found large variation in the effect of reminders [18,35]. Some variation may, however, be explained by differences between paper-based and internet surveys. For example, Cantuaria and Blanes-Vidal [35] compared survey responses collected with mail and internet methods and found that reminders had a stronger effect on increasing internet survey response rates than mail response rates. Their study was relatively small; they initially invited all participants through mail, finding that 2 reminders added a little more than half to the response rate in the mail survey and almost doubled the number of internet responses [35]. We made no such comparison, as we invited all participants via the Danish digital mailbox and only used 1 reminder; moreover, our reminder did not achieve the same effectiveness as the one sent to internet participants by Cantuaria and Blanes-Vidal [35]. In our study, there was a 31% response rate following the first wave, including 21% before the 14-day reminder and an additional 10% after the reminder. Likewise, in the second wave, the reminder added roughly 50% to the response rate, suggesting a rule of thumb that reminders may add another half to the response rate achieved by an initial invitation. Previous survey studies have suggested comparable effects of reminders. For example, a Danish national self-administered health survey found that cumulative response rates after first, second, and third mailings were 36.7%, 50.3% and 59.5%, respectively [14]. A recent Danish national self-administered health survey found that the cumulative response proportions were 19.1%, 28.9%, and 37.2%, simultaneously pointing to a tendency toward declining participation in public health surveys [7]. Even if it did not investigate authorities' digital mailbox solutions as a mode of survey distribution, a Cochrane review of methods to increase the response to postal and electronic questionnaires found a similar pattern, and findings from previous studies suggest that our rule of thumb may extend to additional reminders, a way of increasing response rates often used in survey research [8]. In parallel, ethical considerations related to respect for people's right to prioritize their time and resources might speak in favor of limiting the use of reminders.

The Cochrane review found that, for postal questionnaire studies, the odds of response significantly increased when using a special (for example registered or certified) delivery service [8]. Thus, using modes of invitations signaling authority and quality may be beneficial in terms of promoting participation; however, this again gives rise to ethical considerations. For example, exposing large numbers of people to surveys causing emotional distress or exposing people to low-quality questionnaire studies that take advantage of platforms established by authorities to communicate important information may undermine people's trust in the platform. The "quality stamp" only works if its credibility is not watered down through overuse by researchers and by people starting to perceive the platform as irrelevant. An overall response rate just exceeding 30% to a public health survey benefiting from the web and platform opportunities described in our study is not high. Our findings should be seen in the context of emerging evidence that participation in web-based, health-related surveys is decreasing [5-7]. Among the possible causes of decreasing participation are, again, survey fatigue, with the increasing number of surveys being sent out making people less likely to respond to any single request [3-7].

Limitations

A limitation of this study was that it only included men aged 45 to 70. Response rates have previously been found to be greater among women than men [5,36]. Another limitation was that the study was conducted using data from 2019; it could be claimed that much has changed in the digital world since then. Still, electronic surveys are widely used for research, and digital mailboxes are an increasingly used channel for distributing electronic surveys in countries with such platforms. Moreover, we did not test whether people who responded after the reminder differed from those who responded before in terms of, for example, response patterns.

It can be questioned whether findings on participation and response rates to surveys distributed via a mandatory, high-population coverage, authority-provided digital mailbox can be generalized, or to what extent, to voluntary or low-coverage mailboxes. High-coverage and mandatory systems may tend to decrease coverage bias, while voluntary or low-coverage systems may suffer to a higher extent from nonresponse bias and self-selection errors. Additionally, in mandatory systems, individuals are legally required to register and check their mailboxes. As has been discussed above, when a survey is distributed through a mandatory system, recipients may associate it with official or legal duties, driving up response rates. Contrarily, voluntary platforms lack this legal obligation and authoritative weight. Moreover, in mandatory systems with almost complete population coverage, the sampling frame matches the target population, while in voluntary or low-coverage systems, certain demographics (eg, older people, low-income individuals, and those with lower digital literacy) may be structurally excluded or less likely to opt in, thereby

XSL•FO
RenderX

skewing results. Finally, in voluntary systems, respondents who choose to engage may have stronger preexisting opinions on the survey topic than nonrespondents. Mandatory systems may thus capture a broader spectrum of the public, including indifferent individuals. Thus, the findings of the present study may be generalizable to other mandatory or high-coverage systems, whereas estimating the general population's behavior in areas with voluntary or lower-coverage digital mailbox solutions may to some extent require statistical correction. The latter can be done, for example, through use of propensity score matching (thereby modeling the probability of an individual responding through use of administrative data, followed by reweighting of the sample to match nonparticipants) or through adjustment of data using known population benchmarks (eg, age, education, and income level) to correct for misrepresented groups.

As indicated in the Methods section, the assignment procedure was based on date-of-birth sorting rather than truly random sorting, representing a potential limitation.

Finally, the study only investigated a limited number of factors that can affect survey participation and did not examine factors such as question wording (eg, keeping questions simple), display type (eg, graphics and progress indicators), use of incentives (eg, gifts), or whether the survey is specifically aimed at the invited group (eg, a survey about an illness sent to people with that illness) [6,8,18].

Conclusion

Online surveys are now routinely used for gathering data in social research, public health, and related disciplines, and in some countries, surveys increasingly use web platforms provided by national authorities. However, nonparticipation remains a problem that can bias findings. In our study, survey characteristics such as length, file attachment, and emotionally distressful content were associated with lower response rates. Our study therefore largely confirms what is already known from predigital survey research and online research using other modes of survey distribution. Existing patterns seemingly hold in a national digital mailbox setting, thereby emphasizing that traditional principles regarding survey design and factors affecting survey participation also apply to newer technologies. This underscores the importance of considering these factors when conducting internet-based surveys distributed through digital mailboxes. Hence, questionnaires should be kept as short as possible, and the estimated completion time should be communicated honestly. Response rates can also still be considerably increased through reminders. Nonetheless, participation in public health surveys continues to decline, and despite easier procedures, it is therefore more important than ever that researchers, before launching a new survey, evaluate if the required information is already available through, for example, existing datasets, or whether it can be obtained through other means.

Acknowledgments

No AI was used in any portion of the manuscript generation.

Funding

This study was funded by the Danish Health Insurance Foundation (grant 17-B-0038) and the Lilly and Herbert Hansen Foundation (grant 100063).

Data Availability

Data are available from the corresponding author upon reasonable request.

Authors' Contributions

SB collected, analyzed, and interpreted the data used in this study and was a major contributor in writing the manuscript. AH and SM assisted in analyzing the data, interpreting the data, and writing the manuscript. All authors read, commented on, and approved the final manuscript.

Conflicts of Interest

None declared.

Checklist 1
STROBE checklist.
[PDF File, 254 KB - publichealth_v12i1e89761_app1.pdf]

References

1. Ebert JF, Huibers L, Christensen B, Christensen MB. Paper- or web-based questionnaire invitations as a method for data collection: cross-sectional comparative study of differences in response rate, completeness of data, and financial cost. J Med Internet Res 2018 Jan 23;20(1):e24. [doi: 10.2196/jmir.8353] [Medline: 29362206]
2. Modernising Ireland through digital postboxes. e-Boks. 2021. URL: https://blog.e-boks.com/modernising-ireland-through-digital-postboxes [accessed 2026-09-03]

XSL•FO
RenderX

3. Harrison S, Alderdice F, Henderson J, Redshaw M, Quigley MA. Trends in response rates and respondent characteristics in five National Maternity Surveys in England during 1995-2018. Arch Public Health 2020;78(1):46. [doi: 10.1186/s13690-020-00427-w] [Medline: 32509303]
4. Galea S, Tracy M. Participation rates in epidemiologic studies. Ann Epidemiol 2007 Sep;17(9):643-653. [doi: 10.1016/j.annepidem.2007.03.013] [Medline: 17553702]
5. Christensen AI, Lau CJ, Kristensen PL, et al. The Danish National Health Survey: study design, response rate and respondent characteristics in 2010, 2013 and 2017. Scand J Public Health 2022 Mar;50(2):180-188. [doi: 10.1177/1403494820966534] [Medline: 33161874]
6. Atchison CJ, Gilby N, Pantelidou G, et al. Strategies to increase response rate and reduce nonresponse bias in population health research: analysis of a series of randomized controlled experiments during a large COVID-19 study. JMIR Public Health Surveill 2025 Jan 9;11:e60022. [doi: 10.2196/60022] [Medline: 39791251]
7. Jezek AH, Ekholm O, Thygesen LC, Christensen AI. The impact of reminders on representativeness and survey estimates among web-mode invited in the Danish National Health Survey. Eur J Public Health 2025 Apr 1;35(2):256-262. [doi: 10.1093/eurpub/ckae176] [Medline: 39531250]
8. Edwards PJ, Roberts I, Clarke MJ, DiGuiseppi C, Woolf B, Perkins C. Methods to increase response to postal and electronic questionnaires. Cochrane Database Syst Rev 2023 Nov 30;11(11):MR000008. [doi: 10.1002/14651858.MR000008.pub5] [Medline: 38032037]
9. Kjøller M, Thoning H. Characteristics of non-response in the Danish Health Interview Surveys, 1987-1994. Eur J Public Health 2005 Oct;15(5):528-535. [doi: 10.1093/eurpub/cki023] [Medline: 16051660]
10. Demarest S, Van der Heyden J, Charafeddine R, Tafforeau J, Van Oyen H, Van Hal G. Socio-economic differences in participation of households in a Belgian national health survey. Eur J Public Health 2013 Dec;23(6):981-985. [doi: 10.1093/eurpub/cks158] [Medline: 23183496]
11. Torvik FA, Rognmo K, Tambs K. Alcohol use and mental distress as predictors of non-response in a general population health survey: the HUNT study. Soc Psychiatry Psychiatr Epidemiol 2012 May;47(5):805-816. [doi: 10.1007/s00127-011-0387-3] [Medline: 21544604]
12. Tolonen H, Koponen P, Borodulin K, Männistö S, Peltonen M, Vartiainen E. Differences in participation rates between urban and rural areas are diminishing in Finland. Scand J Public Health 2018 Nov;46(7):755-757. [doi: 10.1177/1403494817748737] [Medline: 29307266]
13. Birkeland SF, Haakonsson AK, Pedersen SS, Rottmann N, Barry MJ, Möller S. Sociodemographic representativeness in a nationwide web-based survey of the view of men on involvement in health care decision-making: cross-sectional questionnaire study. J Med Internet Res 2020 Sep 2;22(9):e19517. [doi: 10.2196/19517] [Medline: 32663149]
14. Christensen AI, Ekholm O, Kristensen PL, et al. The effect of multiple reminders on response patterns in a Danish health survey. Eur J Public Health 2015 Feb;25(1):156-161. [doi: 10.1093/eurpub/cku057] [Medline: 24855288]
15. Udby CL, Riis A, Thomsen JL, Rolving N. Does the use of telephone reminders to increase survey response rates affect outcome estimates? An ancillary analysis of a prospective cohort study of patients with low back pain. BMC Musculoskelet Disord 2021 Oct 20;22(1):893. [doi: 10.1186/s12891-021-04787-4] [Medline: 34670521]
16. Sjetne IS, Iversen HH, Holmboe O, Helgeland J. Response probabilities and response-mode preferences in a self-administered survey. BMC Res Notes 2019 May 27;12(1):289. [doi: 10.1186/s13104-019-4328-7] [Medline: 31133060]
17. Redshaw M, Heikkilä K. Delivered With Care: A National Survey of Women's Experience of Maternity Care. : National Perinatal Epidemiology Unit, University of Oxford; 2010 URL: https://researchonline.lshtm.ac.uk/id/eprint/2548656/ [accessed 2026-09-03]
18. Fan W, Yan Z. Factors affecting response rates of the web survey: a systematic review. Comput Human Behav 2010 Mar;26(2):132-139. [doi: 10.1016/j.chb.2009.10.015]
19. US Preventive Services Task Force, Grossman DC, Curry SJ, et al. Screening for prostate cancer: US Preventive Services Task Force recommendation statement. JAMA 2018 May 8;319(18):1901-1913. [doi: 10.1001/jama.2018.3710] [Medline: 29801017]
20. Mottet N, Bellmunt J, Bolla M, et al. EAU-ESTRO-SIOG guidelines on prostate cancer. Part 1: screening, diagnosis, and local treatment with curative intent. Eur Urol 2017 Apr;71(4):618-629. [doi: 10.1016/j.eururo.2016.08.003] [Medline: 27568654]
21. Birkeland S, Pedersen SS, Haakonsson AK, Barry MJ, Rottmann N. Men's view on participation in decisions about prostate-specific antigen (PSA) screening: patient and public involvement in development of a survey. BMC Med Inform Decis Mak 2020 Apr 6;20(1):65. [doi: 10.1186/s12911-020-1077-4] [Medline: 32252729]
22. Yammarino FJ, Skinner SJ, Childers TL. Understanding mail survey response behavior: a meta-analysis. Public Opin Q 1991;55(4):613. [doi: 10.1086/269284]
23. Dillman DA, Sinclair MD, Clark JR. Effects of questionnaire length, respondent-friendly design, and a difficult question on response rates for occupant-addressed census mail surveys. Public Opin Q 1993;57(3):289. [doi: 10.1086/269376]
24. Burkhart Q, Orr N, Brown JA, et al. Associations of mail survey length and layout with response rates. Med Care Res Rev 2021 Aug;78(4):441-448. [doi: 10.1177/1077558719888407] [Medline: 31747849]

XSL•FO
RenderX

25. Rammstedt B, John OP. Measuring personality in one minute or less: a 10-item short version of the Big Five Inventory in English and German. J Res Pers 2007 Feb;41(1):203-212. [doi: 10.1016/j.jrp.2006.02.001]
26. Degner LF, Sloan JA. Decision making during serious illness: what role do patients really want to play? J Clin Epidemiol 1992 Sep;45(9):941-950. [doi: 10.1016/0895-4356(92)90110-9] [Medline: 1432023]
27. De kommunale nøgletal. Økonomi og Indenrigsministeriet. URL: https://www.noegletal.dk/noegletal/ntStart.html [accessed 2026-09-03]
28. Olsen KR, Sørensen TH, Gyrd-Hansen D. Association between productivity, list size, patient and practice characteristics in general practice. Ugeskr Laeger 2010 Apr 19;172(16):1192-1196. [Medline: 20423660]
29. Birkeland S, Christensen RD, Damsbo N, Kragstrup J. Patient complaint cases in primary health care: what are the characteristics of general practitioners involved? Biomed Res Int 2013;2013:807204. [doi: 10.1155/2013/807204] [Medline: 24027764]
30. Jensen HAR, Ekholm O, Davidsen M, Christensen AI. The Danish health and morbidity surveys: study design and participant characteristics. BMC Med Res Methodol 2019 May 3;19(1):91. [doi: 10.1186/s12874-019-0733-9] [Medline: 31053088]
31. Birkeland S, Bismark M, Barry MJ, Möller S. Is greater patient involvement associated with higher satisfaction? Experimental evidence from a vignette survey. BMJ Qual Saf 2022 Feb;31(2):86-93. [doi: 10.1136/bmjqs-2020-012786] [Medline: 33888595]
32. Statistik om digital post. Digitaliseringsstyrelsen. 2019. URL: https://digst.dk/it-loesninger/digital-post/om-loesningen/tal-og-statistik/#accordion-2019 [accessed 2026-09-07]
33. Mavletova A, Couper MP. Sensitive topics in PC web and mobile web surveys: Is there a difference? Surv Res Methods 2013;7(3):191-205. [doi: 10.18148/srm/2013.v7i3.5458]
34. Poli A, Kelfve S, Motel-Klingebiel A. A research tool for measuring non-participation of older people in research on digital health. BMC Public Health 2019 Nov 8;19(1):1487. [doi: 10.1186/s12889-019-7830-x] [Medline: 31703655]
35. Cantuaria ML, Blanes-Vidal V. Self-reported data in environmental health studies: mail vs. web-based surveys. BMC Med Res Methodol 2019 Dec 12;19(1):238. [doi: 10.1186/s12874-019-0882-x] [Medline: 31830906]
36. Smith T. Sex and the Gss: Nonresponse Differences – an Update. : NORC at the University of Chicago; 1979.

Abbreviations

OR: odds ratio
PSA: prostate-specific antigen

Edited by A Mavragani, T Sanchez; submitted 16.Jan.2026; peer-reviewed by DG Ross, K Jermsittiparsert, N Sakdapat, Y Deng; revised version received 06.Jul.2026; accepted 12.Jul.2026; published 30.Sep.2026.

Please cite as:
Birkeland S, Haakonsson A, Möller S
Factors Influencing Response Rates in National Internet-Based Questionnaires: Analysis of a Large National Survey
JMIR Public Health Surveill 2026;12:e89761
URL: https://publichealth.jmir.org/2026/1/e89761
doi:10.2196/89761

© Søren Birkeland, Anders Haakonsson, Sören Möller. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 30.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Examining Artificial Intelligence Chatbots' Responses in Providing Human Papillomavirus Vaccine Information for Young Adults: Qualitative Content Analysis

Alfu Laily[1], PhD; Laura M Schwab-Reese[1], PhD; Megan Davish[1], MPH; Emily Cahue[1], MPH; Kathryn J LaRoche[1], PhD; Natalia M Rodriguez[1,2], PhD; Robert J Duncan[3], PhD; Randolph D Hubach[1], PhD; Monica L Kasting[1,2], PhD

[1]Department of Public Health, Purdue University, West Lafayette, IN, United States

[2]Cancer Prevention and Control Program, Indiana University Simon Comprehensive Cancer Center, Indianapolis, IN, United States

[3]Human Development and Family Studies, Colorado State University, Fort Collins, CO, United States

Corresponding Author:
Alfu Laily, PhD
Department of Public Health
Purdue University
812 Mitch Daniels Blvd
West Lafayette, IN, 47907-2060
United States
Phone: 1 (765) 496 1826
Email: alaily@purdue.edu

Abstract

Background: The growing use of artificial intelligence (AI) chatbots for seeking health-related information is concerning, as they were not originally developed for delivering medical guidance. The quality of AI chatbots' responses relies heavily on their training data and is often limited in medical contexts due to their lack of specific training data in medical literature. Findings on the quality of AI chatbot responses related to health are mixed. Some studies showed the quality surpassed physicians' responses, while others revealed occasional major errors and low readability. This study addresses a critical gap by examining the performance of various AI chatbots in a complex, misinformation-rich environment.

Objective: This study examined AI chatbots' responses to human papillomavirus (HPV)–related questions by analyzing structure, linguistic features, information accuracy and currency, and vaccination stance.

Methods: We conducted a qualitative content analysis following the approach outlined by Schreier to examine 4 selected AI chatbots' (ChatGPT 4, Claude 3.7 Sonnet, DeepSeek V3, and Docus [General AI Doctor]) responses to HPV vaccine questions. These questions, simulated by young adults, were adapted from items on the Vaccine Conspiracy Beliefs Scale and Google Trends. The selection criteria for AI chatbots included popularity, accessibility, countries of origin, response update methods, and intended use. Two researchers, simulating a 22-year-old man or woman, collected 8 conversations between February 22 and 28, 2025. We used a deductive approach to develop initial code groups, then an inductive approach to generate codes. The responses were analyzed based on a comprehensive codebook, with codes examining response structure, linguistic features, information accuracy and currency, and vaccination stance. We also assessed readability using the Flesch-Kincaid Grade Level and Reading Ease Score.

Results: All AI chatbots cited evidence-based sources from reputable health organizations. We found no fabricated information or inaccuracies in numerical data. For complex questions, all AI chatbots appropriately deferred to health care professionals' suggestions. All AI chatbots maintained a neutral or provaccine stance, corresponding with scientific consensus. The mean and range of response lengths varied [word count; ChatGPT: 436.4 (218-954); Claude: 188.0 (138-255); DeepSeek: 510.0 (325-735); and Docus: 159.4 (61-200)], as did readability [Flesch-Kincaid Grade Level; ChatGPT: 10.7 (6.0-14.9); Claude: 13.2 (7.7-17.8); DeepSeek: 11.3 (7.0-14.7); and Docus: 12.2 (8.9-15.5); and Flesch-Kincaid Reading Ease Score; ChatGPT: 46.8 (25.4-72.2); Claude: 32.5 (6.3-67.3); DeepSeek: 43.7 (22.8-67.4); and Docus: 40.5 (19.6-58.2)]. ChatGPT and Claude offered personalized responses, while DeepSeek and Docus lacked this. Occasionally, some responses included broken or irrelevant links and medical jargon.

Conclusions: Amidst an online environment saturated with misinformation, AI chatbots have the potential to serve as an alternative source of accurate HPV-related information to conventional online platforms (websites and social media). Improvements

XSL•FO
RenderX

in readability, personalization, and link accuracy are still needed. Furthermore, we recommend that users treat AI chatbots as complements, not replacements, to health care professionals' guidance on clinical settings.

(JMIR Public Health Surveill 2026;12:e79720) doi:10.2196/79720

KEYWORDS

artificial intelligence; health communication; papillomavirus vaccines; large language models; qualitative research

Introduction

Globally, human papillomavirus (HPV) was responsible for an estimated 620,000 new cancer cases in women and 70,000 in men in 2019 [1]. In the United States, HPV causes roughly 48,000 new cancer cases annually and is the most common sexually transmitted infection [2,3]. Although effective vaccines are available [4], uptake among eligible adults remains suboptimal. In 2022, only 57.7% of eligible women and 34.8% of eligible men (ages 19-26 years) had received more than one dose of the HPV vaccine [5]. Adequate knowledge about HPV vaccination plays a critical role in promoting vaccine uptake among eligible adults [6].

The internet has become a widely used resource for people seeking information on health-related topics, including the HPV vaccine [7,8]. Information about the HPV vaccine is a popular online search topic in the United States, with Google Trends indicating an annual search traffic growth of 8.6% from 2010 to 2021 [8]. Research has also demonstrated that HPV vaccine communication uses specific search terms (eg, #HPV and #HPVVaccine) that are prevalent on online social media platforms such as Instagram [9,10], YouTube [11], and Twitter (now known as X) [12]. While these online platforms serve as a powerful tool for disseminating health information, they also provide a potential space for inaccurate information to spread, including inadvertently false information (misinformation) and deliberate falsehoods (disinformation) [13,14]. On Instagram, a primarily visual platform, approximately 55.8% of postings regarding HPV were provaccine, whereas 42.2% had an antivaccine perspective [9]. Additionally, the hashtag *#Gardasil* on Instagram has been used to circulate conspiracy theories and unsupported and false claims of HPV vaccine-related injuries [10]. YouTube, a popular video-centric social media space, users-generated content regarding the HPV vaccine showed a greater number of likes in videos with a negative tone compared to those with a positive tone [11]. On X, a platform for brief text-based posts, content containing erroneous information regarding the HPV vaccine were 5.44 times more likely to be shared than educative content [12], indicating that inaccurate posts received higher audience engagement. Given these trends, the World Health Organization (WHO) has identified the "uncontrolled dissemination of misinformation," particularly on the topic of vaccination, as one of the most pressing health challenges for the coming decade [15].

The quality of online information related to the HPV vaccine could potentially influence vaccine decision-making [16]. Individuals who seek health information online showed lower acceptance of the HPV vaccination, which may be associated with discouraging information found online [17]. Furthermore, the overwhelming amount of information available online can lead to doubt and anxiety, making it challenging for individuals to make informed health decisions [18]. The abundance of information, much of which contains inaccurate information, may also contribute to low levels of HPV knowledge, as literature indicated that only 36.1% of young adults in the United States were aware that HPV causes more cancers than just cervical cancer [19]. This is concerning because knowledge of HPV correlates positively with HPV vaccination intention and uptake [20-22].

In recent years, there has been a growing inclination toward using internet-based artificial intelligence (AI) chatbots for accessing health-related information, including topics related to sexual health [23-26]. AI chatbots are digital systems created to mimic conversations with humans through text or voice, and their application in health care is rapidly expanding, ranging from answering general patient questions about sexual health to supporting medication adherence [25,27]. Large language models (LLMs) are AI tools built on multilayer recurrent neural networks and trained on extensive datasets of natural language to produce text that resembles human language [28]. AI chatbots are powered by LLMs, which are trained to generate human-like responses [29]. This generative ability relies on deep learning techniques, systems that mimic how the human brain handles information using neural networks to find patterns in data [30-32]. Examples of currently available AI chatbots based on deep learning include ChatGPT and DeepSeek [30,31,33], which analyze user input and then offer intelligent, contextually relevant responses for general purposes. In contrast, Docus is specifically designed to analyze health information, offer tailored health management, and provide laboratory test interpretations to facilitate health-related decision-making [34].

AI chatbots like ChatGPT are inevitably being used in support of health care settings, even though they were not originally intended for such purposes [35]. This is partly attributed to the user-friendly nature of LLMs used in the AI chatbot, making it convenient for individuals seeking health information [36]. Additionally, young adults aged <25 years and those identifying as technologically savvy had higher levels of AI chatbot acceptability [27,37]. However, it is important to note that the accuracy of AI chatbots' responses, which partly depends on the quality of their training data, is compromised in medical contexts due to the lack of specific training data in medical literature [35,38]. The inability of AI chatbots to clarify their decision-making process presents challenges in identifying and rectifying potential biases or errors [35,39,40].

Several studies have revealed complex and often contradictory findings regarding the quality of AI chatbots' responses on health-related topics. A study demonstrated that ChatGPT produced higher-quality responses than those provided by physicians from patients' questions pooled from a social media

XSL•FO
RenderX

forum [41]. A specific knowledge-based chatbot trained exclusively on HPV vaccine information outperformed general-purpose generative pretrained transformer models in both accuracy and relevance [42]. In tasks involving physician-generated medical questions, responses generated from ChatGPT 3.5 and ChatGPT 4 were rated as "nearly all correct" or "completely correct" 50% of the time [43]. However, for these, the median accuracy scores exceeded mean scores, suggesting the presence of occasional but substantial wrong answers [43]. Bing Copilot's responses to queries about prescription medications in the United States were found to be difficult to read, as indicated by low Flesch Reading Ease Scores, raising concerns about its readability and accessibility [44]. Although prior studies have evaluated AI chatbot responses, they typically focused on a single system, using physician-generated questions or purpose-built chatbots, which may not capture the full range of real-world interactions of various AI chatbots from the perspective of general public users.

While AI chatbots stand out as a remarkable technological feat, their application in providing health information raises concerns regarding the perpetuation of misinformation or inaccurate recommendations and therefore requires attention. To address this, the primary aim of this study is to investigate how 4 selected AI chatbots respond to HPV vaccine-related questions, with wording modified to reflect the language commonly used by young adults (22 years). The secondary aim of the study was to examine the content's structure and patterns, linguistic features, information accuracy, information currency, and the AI chatbots' stance on HPV vaccination, whether supportive, neutral, or contrary. This study distinguishes itself by conducting a qualitative content analysis of a diverse sample of both general-purpose and health-specific AI chatbots on HPV-related topics. By using questions tailored to the language of young adults, it provides a more ecologically valid assessment of AI chatbots' performance and their utility in the real world.

Methods

Question Set

We created semistructured questions adapted from the Vaccine Conspiracy Beliefs Scale (VCBS) items [45] and Google Trends query [46] as prompts asked in AI chatbots. The original VCBS consists of seven statements: (1) vaccine safety data are often fabricated, (2) immunizing children is harmful and this fact is covered up, (3) pharmaceutical companies cover up the dangers of vaccines, (4) people are deceived about vaccine efficacy, (5) vaccine efficacy data are often fabricated, (6) people are deceived about vaccine safety, and (7) the government is trying to cover up the link between vaccines and autism. We modified the original VCBS scripts to reflect language commonly used by young adults and tailored the content to focus specifically on the HPV vaccine (Multimedia Appendix 1). We also used Google Trends [46], a publicly accessible web service, to capture search-term volume for the Google search engine over time. In order to investigate potential questions regarding the HPV vaccine (search term: HPV vaccine) during the past 12 months (from January 1, 2024, to January 1, 2025). Google Trends recorded both top queries (the most popular or "evergreen" inquiries within the specified search parameters, maintaining relative consistency over time) and rising queries (the inquiries with the most substantial increase in search frequency compared to the previous time period, signifying a surge in relative interest) [46]. We identified HPV-related search terms from both top and rising queries and rephrased them into questions (Multimedia Appendix 1). Finally, 2 researchers (EC and MD) within the young adult age group reviewed the question set and provided feedback to ensure clarity and relevance to their communication style.

Sample and Data Collection

We collected 8 conversations (2 each from ChatGPT 4, Claude 3.7 Sonnet, DeepSeek V3, and Docus [General AI Doctor]) of AI chatbots' responses to an HPV-related question set. The versions used for ChatGPT, Claude, and DeepSeek were the latest available without a paid subscription at the time of testing, while Docus was tested using the latest version available with a paid subscription. Our researchers started the "chat" by introducing users' assigned sex at birth (male or female) and age (22 years old) in order to simulate a young adult population. Supervised by female (LMS-R, KJL, NMR, and MLK) and male (RJD and RDH) faculty members, all of whom have PhDs, 3 trained female public health graduate students (AL, MD, and EC) gathered 2 conversations from each AI chatbot, using different IP addresses, on different days, and with different sex enactments, for a total of 8 conversations from all AI chatbots. We copied and pasted those responses into a word processor and then uploaded them into HyperRESEARCH 4.5.4 (Researchware, Inc), a qualitative data analysis software. Since no human participants were involved, we did not establish a relationship prior to the commencement of the study nor created field notes. Furthermore, we only followed our predefined questions set, without introducing any off-topic or unrelated prompts into the AI chatbots during the "chat." No one else was present during data collection, except for the research team members. The duration of each collection ranged from 14 minutes to 27 hours and 5 minutes; one AI chatbot's unpaid subscription enabled a wait time after inputting several prompts, and there were server errors that prolonged the data collection process.

Rationale for Selecting the Reviewed AI Chatbots

The selected AI chatbots represented a diverse range of popularity, accessibility, country of origin, response update method, and intended use. ChatGPT had the highest global daily visits, while DeepSeek and Claude had lower visits [47]. ChatGPT was chosen primarily because it has the highest user traffic, making it likely that a significant number of young adults also engage with the platform. Due to the sudden surge in popularity of DeepSeek during the study period, as well as its origin in China, DeepSeek was selected for inclusion. Other than Docus (General AI Doctor), users could access reviewed AI chatbots without a subscription. ChatGPT and Claude were based in the United States, DeepSeek originated from China, and Docus originated from Armenia. We distinguished AI chatbots' update methods between dynamic and static by examining whether each AI chatbot accessed real-time updated online information (dynamic) or relied on fixed, pretrained

XSL•FO
RenderX

knowledge (static). We assessed the update methods by asking each AI chatbot, "What is the most recent information you can access?" ChatGPT responded with the current date (February 5, 2025), indicating dynamic updates, while Claude (October 2024), DeepSeek (June 2024), and Docus (October 2023) provided fixed cutoff dates, showing they rely on static knowledge. Other than Docus, which was developed specifically to provide health-related consultations, other AI chatbots were developed for general purposes (Table 1). Therefore, to add nuances to our study, Docus was included in this study because of its specific model.

Table 1. Artificial intelligence chatbots' characteristics for sample selection.

	ChatGPT 4	Claude 3.7 Sonnet	DeepSeek V3	Docus (General AI Doctor)
Global daily visit [47]	100-150 million	<15 million	<50 million	__[a]
Access	Free	Free	Free	Paid subscription
Originated from	United States	United States	China	Armenia
Information update methods	Dynamic	Static	Static	Static
Purpose	General	General	General	Health consultation

[a]Not available.

Data Analysis

We conducted a qualitative content analysis following the approach outlined by Schreier [48]. Initially, 3 researchers with qualitative methods training (AL, EC, and MD) thoroughly reviewed all AI chatbot responses to gain a general understanding of the data. During this phase, we identified preliminary ideas that could serve as potential code groups and codes. We adopted a deductive, concept-driven approach to develop initial code groups based on our research objectives and question set. Within these groups, we applied an inductive, data-driven approach to generate codes based on all AI chatbot responses. The code-building process involved the use of subsumption and progressive summarizing strategies. As we reviewed the responses, we identified a distinct idea and checked whether an existing code captured this idea. If the idea was already represented, we ensured the code was grouped under the appropriate code group. If not, we created a new code to reflect the previously uncaptured idea. This iterative process continued until no new ideas emerged, thereby automatically meeting the criteria of saturation.

After generating initial code drafts, 3 researchers (AL, EC, and MD) collaborated to compile code groups and codes into a single comprehensive codebook (Multimedia Appendix 2). We reviewed the codebook to ensure that each code was mutually exclusive, and the codebook was collectively exhaustive to represent our data. A faculty member (LMS-R) provided feedback on the codebook before its application. To enhance consistency in code application, researchers (AL, EC, and MD) met and pilot-coded a subset of the data. Due to the straightforward and descriptive nature of the codebook, no coding disagreements arose. During code application, 2 independent researchers double-coded each AI chatbot response. Any discrepancies in coding were discussed and resolved.

We analyzed AI chatbot responses based on structure and patterns, linguistic features, accuracy, information currency, and stance on the HPV vaccination. Structure and pattern captured the structural and content-related elements of responses, including organization format (eg, paragraphs, bullet points, and numbered lists), logical reasoning and relevance of information to the question, details and length, and how the response concluded (eg, with a summary, follow-up question, or suggested next steps). Linguistic features included the use specific terminology, use of user-centric (referencing users' age or sex) and emotional validation sentences (acknowledging user concerns and feeling or providing reassurance), use of language to convey statistics (eg, numeric vs relative terms), and assessed readability using the Flesch-Kincaid Grade Level and Reading Ease Score to determine suitability for lay audiences [49]. The Flesch-Kincaid Grade Level indicates the educational level required to understand a text passage, ranging from fifth grade (level 5) to college graduate (level 12+) [50,51]. The Flesch-Kincaid Reading Ease Score measures how easy it is to read a text passage, with scores ranging from 0 to 100 and higher scores indicating greater ease [52]. We adapted cutoff scores of ≤8 for the Flesch-Kincaid Grade Level and ≥50 for the Reading Ease Score to determine the readability for lay audiences [49]. Accuracy was assessed based on factual claims and numerical data in the AI chatbot responses, considering their alignment with current scientific evidence and HPV vaccine guidelines, including source credibility from reputable organizations (eg, the Centers for Disease Control and Prevention [CDC] and WHO) or peer-reviewed journals. Sources cited or links included in responses were manually checked to verify their credibility and to ensure they directed users to the information referenced in the AI chatbot's response. Our research team consists of experts in sexual health (AL [53,54], KJL [55,56], RDH [57,58], and MLK [59]), health communication (AL [60,61], LMS-R [62,63], RDH [64], and MLK [65]), LLM and AI (LMS-R [66,67]), and HPV vaccination (AL [54], NMR [68], and MLK [69]), ensuring content accuracy. Content was considered current if information or fact sheets were published within 12 months or studies published within 5 years from data collection. HPV vaccination stance was assessed from the response to each question and categorized as pro, neutral, or contra.

The Flesch-Kincaid scores, how the response concluded, and vaccination stance were also assessed from each AI chatbot's response to a question. However, we did not assess Flesch-Kincaid scores for responses containing links to avoid

XSL•FO
RenderX

potential skewed results. Since each response set included 21 to 24 questions, this resulted in a corresponding range of 21 to 24 data points per set for conclusion style and vaccination stance, while Flesch-Kincaid scores had a range of 9 to 21 data points for Grade Level and Reading Ease scores.

Ethical Considerations

As no human subjects were involved, ethics board reviews were not required. Furthermore, this study adheres to the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist. Specifically, our Methods section includes information from domain 1 (research team and reflexivity), domain 2 (study design), and parts of domain 3 (analysis and findings). The remainder of domain 3 is reported in our Results section.

Results

We collected the 8 conversations between February 22 and 28, 2025, using 21-24 questions modified from VCBS items [45] and Google Trends queries [46]. Numbers following AI chatbot names (eg, Claude_1 and Claude_2) indicate dataset identifiers used to distinguish different responses from the same AI chatbots.

Structure and Pattern

In structure and pattern of the responses, we observed organization format, logical reasoning, information relevance, details and length, and conclusion styles of each response.

Organization

In terms of the sentence organization format when answering one question, all AI chatbots responded more often using bulleted or numbered lists, or a combination of both, instead of paragraph-style responses. ChatGPT more frequently used bulleted points nested under a numbered list than other formatting styles. Claude more often used bulleted points but sometimes responded using only numbered lists or bullet points under numbered lists. DeepSeek typically used an organization of bulleted points under a numbered list, and at times used bulleted points nested within other bulleted points under a numbered list:

> *1. For Individuals Aged 9-14 Years*
>
> - Number of Doses: 2 doses
> - Schedule:
> - First Dose: Given at the start of the series. Second Dose: Given 6-12 months after the first dose.
>
> *This age group has a stronger immune response to the vaccine, so only two doses are needed for full protection.* [DeepSeek_2]

Docus more frequently presented responses with bullets nested under other bullets or under numbers, but it also occasionally provided responses in paragraph style:

> *Individuals who experience a life-threatening allergic reaction, such as anaphylaxis, after receiving an HPV vaccine should not continue with the vaccine series. It is crucial to consult a healthcare provider for personalized guidance and to discuss alternative options for protection against HPV-related diseases. Your healthcare provider can assess the situation and provide the best advice tailored to your specific medical needs.* [Docus_1]

Logic and Relevance of Information

All AI chatbots provided logical reasoning to support their answers, often using phrases such as “Here’s why:” [Claude_1, ChatGPT_1, DeepSeek_2, and Docus_1], “Why trust this information?” [ChatGPT_2], and “Why the difference in doses?” [DeepSeek_1]. These explanations were followed with relevant, evidence-based information to justify their responses. For example, when a user asked, “Is the info about the HPV vaccine’s safety even real, or is it just fake?,” AI chatbots offered a clear rationale grounded in scientific research:

> *7. Why Trust This Information?*
>
> *The evidence supporting the HPV vaccine’s safety comes from scientific consensus, not just anecdotal reports or single studies. The vaccine has been in use since 2006, and the data from over a decade of global use demonstrates its positive impact on public health.* [ChatGPT_2]
>
> *The HPV vaccine underwent rigorous clinical trials before approval, involving thousands of participants to ensure its safety and efficacy.* [Docus_2]

Details and Length

Some AI chatbots included additional details that, while related, were not directly relevant to the question asked, making their responses unnecessarily long. For instance, when DeepSeek was asked, “Why is the HPV vaccine not recommended after 26 years old?,” it not only explained the main reason but also introduced a new topic: “What Does ‘Shared Clinical Decision-Making’ Mean?” [DeepSeek_2]. Similarly, when asked, “Does the HPV vaccine also protect against other STDs like HIV or chlamydia?,” Docus veered into broader prevention advice, stating: “To reduce the risk of other STDs, including HIV and chlamydia, it’s important to practice safe sexual behaviors, such as...” [Docus_2].

On the other hand, we identified instances where AI chatbot responses did not provide sufficient detail. For example, ChatGPT stated, “Early safety reviews identified rare instances of anaphylaxis, approximately 1.7 cases per million doses, and episodes of syncope (fainting), often related to anxiety” [ChatGPT_1], without clarifying what was meant by “anxiety,” such as fear of needles or an underlying anxiety disorder which could be important for user understanding. Similarly, Claude mentioned, “There’s no evidence of systematic deception about HPV vaccine effectiveness” [Claude_1] but did not provide an explanation of what systematic deception means.

In our sample of 4 AI chatbots, we found that ChatGPT and DeepSeek tended to provide longer and more superfluous responses to questions than Claude and Docus. In this context, superfluous refers to answers that included excessive or unnecessary details and lacked straightforwardness. DeepSeek (510.0 words) had the highest mean word count for responses to each question, while Docus (159.4 words) had the lowest.

XSL•FO
RenderX

Conclusion Styles

To conclude their responses to a question, AI chatbots included a summary, follow-up question, further direction, or a combination of these. All AI chatbots frequently provided summaries and often used starting phrases like "The bottom line:" [ChatGPT_1], "In summary" [ChatGPT_2], or "Key takeaways" [DeepSeek_2]. We did not find any responses from DeepSeek or Docus that concluded only with a follow-up question, while ChatGPT and Claude occasionally ended their responses with follow-up questions, such as:

> *Would you like to dive deeper into any specific studies or concerns about the vaccine?* [ChatGPT_1]
>
> *Would you like more specific information about any of these side effects?* [Claude_2]

AI chatbots often provided further directions that typically encouraged users to consult a "healthcare provider" [ChatGPT_2, Claude_2, DeepSeek_1, and Docus_2] or refer to trusted sources like the "CDC or WHO" [DeepSeek_1 and Docus_1] to confirm users' understanding or if users would like to look for further information.

Linguistic Features

The AI chatbots exhibited diverse linguistic features, from the use of jargon or complex medical terms to providing sentences to foster users' emotional validation. Not all AI chatbots provided responses tailored to individual users, and emotional validation varied. Some AI chatbots had more understandable responses than others, as evidenced by Flesch-Kincaid scores.

Terminology

All AI chatbots occasionally used jargon or medical vocabulary that was infrequently used in lay audience's everyday speech. Lay users would need to seek out the definitions in order to fully comprehend the provided sentences. Uncommon terminology used included "cervical epithelial neoplasia" and "primary ovarian insufficiency" [ChatGPT_1], "immunosuppressive medications" and "post-market surveillance" [ChatGPT_2], "precancerous lesion" [Claude_1], "vertical transmission" [Claude_2], "salicylic acid or cryotherapy kits" [DeepSeek_1], "fomite transmission" [DeepSeek_2], "anaphylaxis" [Docus_1], and "oropharynx" [Docus_2]. Furthermore, all 4 AI chatbots used the term "gender" when addressing questions on the recipients of the HPV vaccine, despite "sex" being the more precise term in this context, as it pertains to biological distinctions. For instance, Docus remarked:

> *HPV also affects males and can lead to other types of cancers and genital warts. Therefore, the vaccination recommendations have expanded to include both genders.* [Docus_2]

User Centric and Emotional Validation

DeepSeek and Docus did not provide a user-centric experience, indicated by not incorporating essential users' personal information in their responses such as age or sex when applicable. Conversely, ChatGPT and Claude incorporated this contextual information. In response to the question, "How many shots do I need to get for the HPV vaccine?," DeepSeek and Docus offered a broad explanation in correspondence to the age of HPV vaccine initiation and did not provide a personalized reference. ChatGPT provided a specialized segment, "Is it too late to receive the HPV vaccine at 22?" [ChatGPT_1], and Claude stated, "Since you're 22, you would follow the 3-dose schedule" [Claude_2]. Related to emotional validation, all AI chatbots offered a degree of emotional affirmation, particularly when addressing apprehension around vaccine safety and efficacy. Docus exhibited the least affirming tone for emotional validation, providing only the short "Certainly!" [Docus_1]. Conversely, ChatGPT, Claude, and DeepSeek used more detail and affirming language, likely fostering a sense of acknowledgment among users, for instance:

> *That's a great question, and it's totally understandable to want to be sure about the safety of any vaccine.* [ChatGPT_2]
>
> *It's normal to have questions about vaccine safety.* [Claude_2]
>
> *It's understandable why some people might feel like they've been lied to about HPV vaccine safety.* [DeepSeek_2]

Additional quotes related to emotional validation can be found in Multimedia Appendix 3.

Relative vs Numeric

All AI chatbots used either relative or exact numerical statistics depending on whether precision was necessary. For instance, when citing specific studies, ChatGPT presented comprehensive numerical data:

> *Research published in The BMJ highlighted that women vaccinated between ages 12 and 13 showed an 83.9% decrease in cervical cancer rates and a 94.3% reduction in cervical intraepithelial neoplasia grade 3.* [ChatGPT_1]

In a similar vein, when addressing HPV vaccine recommendations, AI chatbots used precise age ranges and dosage quantities instead of ambiguous phrases such as "teens" or "some doses." For example:

> *Two doses: for those starting the vaccine series at age 9-14.* [ChatGPT_2]
>
> *Ages 9-14: Recommended for routine vaccination (2-dose series).* [Claude_1]
>
> *For those aged 9-14: Two doses are recommended, with the second dose given 6-12 months after the first.* [DeepSeek_2]
>
> *For those starting the vaccine series before age 15, two doses are recommended.* [Docus_1]

Conversely, when precise numbers were unnecessary to understand the context, AI chatbots used relative terminology, such as "nearly," "some," or "hundreds," exemplified by statements like:

> *In clinical trials, the vaccine was nearly 100% effective at preventing precancerous cervical lesions.* [DeepSeek_2]

XSL•FO
RenderX

The HPV vaccine has been the subject of hundreds of independent studies from universities, research hospitals, and international health organization. [ChatGPT_1]

Flesch-Kincaid Scores

Overall, ChatGPT provided responses with the lowest mean Grade Level and the highest Reading Ease Score among the 4 AI chatbots. In contrast, Claude's scores had the highest mean Grade Level and lowest mean Reading Ease Score, suggesting that its responses may be more complex than those of the other AI chatbots and potentially less accessible to a lay audience (Figures 1 and 2).

Figure 1. Artificial intelligence chatbots' Flesch-Kincaid Grade Level (range and mean).

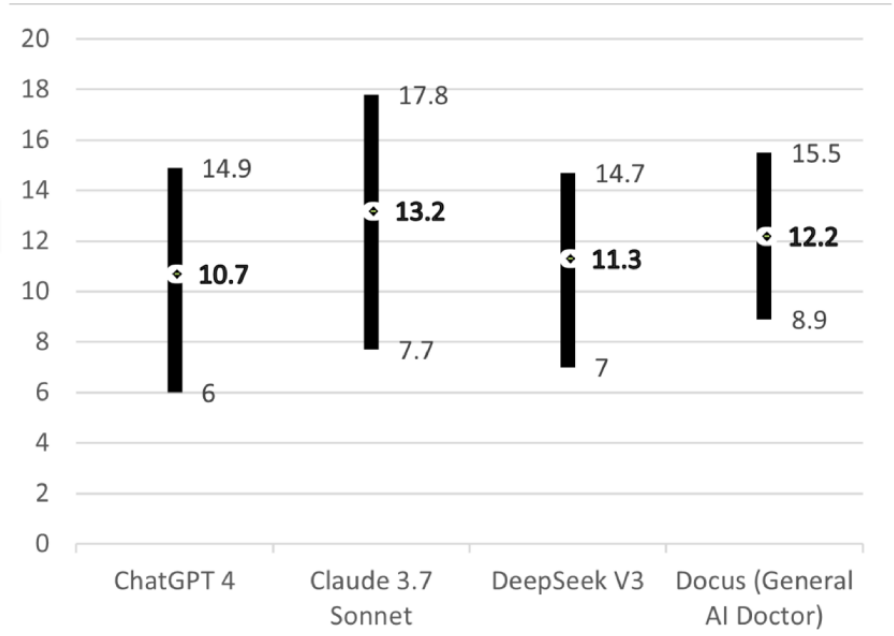
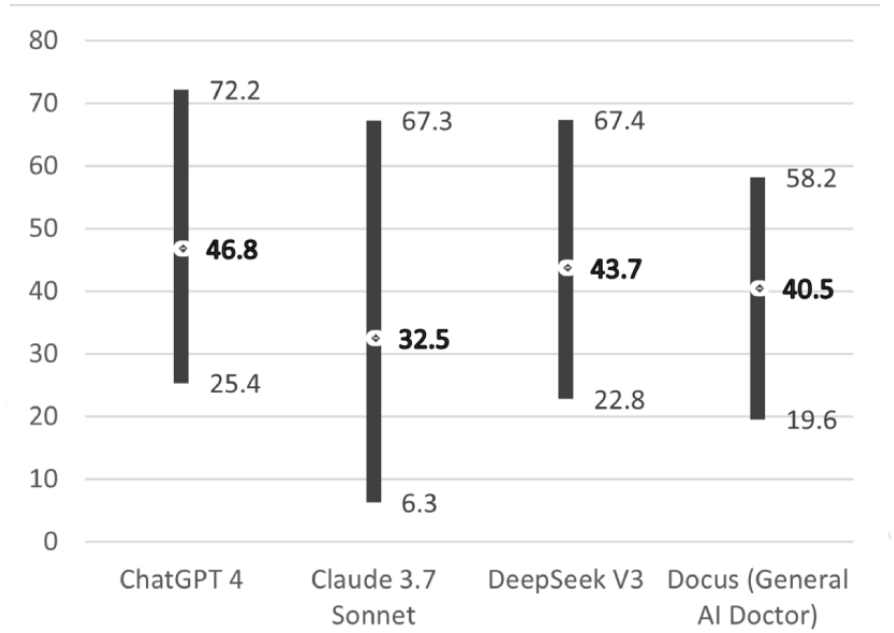


Figure 2. Artificial intelligence chatbots' Flesch-Kincaid Reading Ease Score (range and mean).



Accuracy

To examine accuracy, we observed the credibility of links provided in the responses, data accuracy, and intersentence coherence.

Information Credibility

When citing links, all AI chatbots included functional links to real, existing websites from reputable organizations, such as the WHO [ChatGPT_1], the American College of Obstetricians and Gynecologists [Claude_1], the European Medicines Agency [DeepSeek_2], and the Centers for Disease Control and

XSL•FO
RenderX

Prevention (CDC) [Docus_2]. We found no fabricated or nonexistent links, with no fake studies, fact sheets, or infographics generated. All linked studies were from peer-reviewed academic journals, including the *BMJ* [ChatGPT_1] and the *New England Journal of Medicine* [Docus_1]. However, some of the links provided, particularly those from CDC, were no longer functional at the time of data analysis and led to broken pages displaying "Page Not Found" errors. Furthermore, in a few instances, the AI chatbot responses included links unrelated to the topic asked. For example, when a user requested information about HPV vaccine safety, ChatGPT provided a link to a page on feminizing hormone therapy [70], which contained no information about the HPV vaccine. Similarly, when asked for a study on HPV vaccine safety, Docus linked to a measles vaccine study [71], which also had no relevance to the HPV vaccine (Table 2).

Table 2. Comparison of 4 artificial intelligence (AI) chatbot responses based on selected measurements.

Chatbots	ChatGPT 4	Claude 3.7 Sonnet	DeepSeek V3	Docus (General AI Doctor)
Structure and pattern				
Organization format[a]	Bulleted points nested under numbered list	Bulleted points	Bulleted points nested under numbered list	Bulleted points nested under bulleted points
Word counts per question, mean (range)	436.4 (218-954)	188.0 (138-255)	510.0 (325-735)	159.4 (61-200)
Conclusion styles[b]	Summary and further direction	Summary and further direction	Summary and further direction	Further direction
Linguistic features				
Jargon/medical terminology[c]	Immunosuppressive medications, postmarket surveillance	Precancerous lesion prevention, vertical transmission	Cryotherapy kits, fomite transmission	Anaphylaxis, oropharynx
Personalization[d]	Yes	Yes	No	No
Emotional validation[e]	"That's a great question, and it's totally understandable to want to be sure about the safety of any vaccine."	"It's normal to have questions about vaccine safety."	"It's understandable why some people might feel like they've been lied to about HPV vaccine safety."	"Certainly!"
Accuracy				
Links credibility[f]	Yes	Yes	Yes	Yes
Fabricated sources[g]	No	No	No	No
Incorrect links[h]	Yes	No	No	Yes
Statistics accuracy[i]	Yes	Yes	Yes	Yes
Currency[j]	No	Yes	Yes	Yes
Vaccination stance[k]	Pro	Neutral	Pro	Pro

[a]Most frequently used response structure to a question (eg, using bullet points, numbered lists, paragraphs, or combinations).
[b]Most frequent way responses ended to each question (eg, with summary, follow-up questions, further suggestion, or combinations).
[c]Example of jargon/medical terminology used.
[d]Referenced user's sex and ages when necessary.
[e]Example of affirming language used to acknowledge user's feeling.
[f]Always provided real link from reputable organization or peer-reviewed journal.
[g]Ever provided link/information that are made up or nonexistent.
[h]Ever provided links that are not related to the discussed topic.
[i]Always provided/cited accurate statistic numbers based on scientific evidence.
[j]Always provided information or fact sheets that were not older than 12 months or studies not older than 5 years.
[k]Most frequent stance on HPV vaccination per question (eg, contra, neutral, or pro).

Data Accuracy and Coherence

We found no inaccuracies in the numerical data or statistics provided by the AI chatbots in our sample. Additionally, when responses included links to studies, the statistical values cited in the responses matched those in the referenced sources. However, we did identify an instance of internal inconsistency in Claude's and DeepSeek's responses. Claude initially stated, "Based on extensive safety monitoring and research, there aren't specific demographic groups who systematically experience

XSL•FO
RenderX

significantly more problems with HPV vaccines" [Claude_2]. Yet, it immediately followed with, "However, there are a few considerations worth noting:" and listed groups who should avoid the vaccine, such as individuals with yeast allergies. Similarly, DeepSeek mentioned, "There are currently three vaccines available: 1. Gardasil 9: Protects against 9 types of HPV (6, 11, 16, 18, 31, 33, 45, 52, and 58). It is the most commonly used vaccine in many countries" [DeepSeek_2], but only mentioned Gardasil 9 and failed to elaborate on the other two vaccines referenced in the previous sentence or specify that they are not all available worldwide. For example, only Gardasil 9 is currently available in the United States.

Currency

Some of the studies cited by ChatGPT were more than 5 years old when addressing whether the HPV vaccine affects fertility. While these studies were not outdated, we applied a 5-year cutoff as a point of comparison to assess the currency. For example, ChatGPT referenced, "A 2018 study in Pediatrics" and "a 2019 Australian study" [ChatGPT_1]. However, we did not find any information or fact sheets that were older than 12 months or studies older than 5 years in responses from the other AI chatbots.

HPV Vaccination Stance

Across all responses, none of the AI chatbots expressed opposition to HPV vaccination, as the tone remained either neutral or supportive. Even when asked whether individuals with a history of life-threatening allergic reactions to the HPV vaccine should get vaccinated, none of the AI chatbots directly advised against it. Instead, they consistently recommended consulting health care professionals. Examples include:

> *Talk to an allergist or immunologist before getting vaccinated.* [ChatGPT_1]
>
> *Your healthcare provider can help assess the specific situation and determine whether it was truly a severe allergic reaction or another type of side effect that might not prevent further vaccination.* [Claude_1]
>
> *...the decision to receive the HPV vaccine requires careful consideration and consultation with a healthcare provider.* [DeepSeek_1]
>
> *Your healthcare provider can assess the situation and provide the best advice tailored to your specific medical needs.* [Docus_1]

Discussion

Principal Findings

Our study examined the response of 4 AI chatbots (ChatGPT 4, Claude 3.7 Sonnet, DeepSeek V3, and Docus [General AI Doctor]) in answering HPV vaccine–related questions, with an emphasis on structure and patterns, linguistic features, accuracy, information currency, and stance on HPV vaccination. Our results showed that while all AI chatbots responded based on scientific research and reputable websites, there were some variations in response length, details, readability, and personalization. All of the AI chatbots presented valid scientific references without fabrication of sources, but there were instances of AI chatbots providing broken or irrelevant links. While statistical data were accurate for all AI chatbots, several showed inconsistencies across sentences and lack of informational completeness. Although the AI chatbots did not disseminate incorrect information, their utility in public health communication is still mixed in effectively providing information about HPV vaccination in a young adult population.

Our AI chatbots sample provided responses based on scientific data, cited websites of reputable organizations, and displayed no evidence of falsification of sources. These findings imply that AI chatbots can offer factually accurate information on the HPV vaccine, similar to recent studies that discussed AI chatbots' inclusion of accurate medical information in their responses [43,72]. A systematic review also highlighted that AI chatbots have primarily been used to deliver factual information in response to users' vaccine-related questions [73]. However, when we verified the links provided in the responses, we occasionally discovered broken or irrelevant links. The majority of the broken links came from the CDC website, which may have been restructured in accordance with directives from the current administration [74,75]. Consequently, certain links that were operational at the time of data collection may have become unreachable during later analysis. As public health information is often shaped by political leadership, static AI chatbot content may quickly become outdated if it is not updated rapidly. Despite finding fewer than 5 irrelevant links, this instance may increase apprehension regarding the credibility of these AI chatbots as a medium to disseminate health information. Although these irrelevant links were not fabricated, as is often a concern in studies examining AI chatbot "hallucination" where responses contain made-up references that do not correspond to any existing studies [76,77], they still pose challenges. Users who attempt to verify information through these links may experience reduced trust or increased frustration when attempting to make informed health decisions.

The readability of AI chatbot responses in our sample, assessed using Flesch-Kincaid's Grade Level and Ease Score, would be challenging for a lay audience to comprehend. The mean Grade Level scores for each AI chatbot ranged from 10.7 to 13.2 (ChatGPT: 10.7; Claude: 13.2; DeepSeek: 11.3; and Docus: 12.2). The Reading Ease Score was 32.5 to 46.8 (ChatGPT: 46.8; Claude: 32.5; DeepSeek: 43.7; and Docus: 40.5). Therefore, ChatGPT was the most comprehensible and Claude the least comprehensible in each Flesch-Kincaid measure. Nevertheless, all AI chatbot means of Grade Level were beyond 8 and the Reading Ease Scores were below 50, indicating that all AI chatbots in our sample were deemed difficult for a lay audience to comprehend [49]. A recent study on AI chatbots also reported a high mean of Grade level score and a low mean of Reading Ease score [51]. This is concerning, as health information that exceeded suggested readability levels may create barriers to comprehension for persons with lower health literacy [78]. Additionally, the prevalent use of medical jargon among all AI chatbots in our sample may exacerbate the readability issue. Many lay audiences without medical backgrounds may not be familiar with medical terminologies such as "cervical epithelial neoplasia" and "immunosuppressive medications." Without providing definitions or further

XSL•FO
RenderX

explanation of these terminologies, it could further impede understanding. Moreover, the excessive length of responses from ChatGPT and DeepSeek may overwhelm users. Recent studies indicated that information overload, described as a condition in which an individual is unable to comprehend or respond to stimuli due to excessive information [79], may counterproductively reduce the ability to process information [80] and eventually affect HPV vaccination decision-making. In light of these challenges revealed by our study, AI developers should explore AI chatbot optimization approaches to improve readability for lay audiences, by implementing features like real-time readability adjustment, automatic jargon detection, and using summary-first formats that provide essential points up front, with optional details available on demand.

The degree of personalized responses varied across AI chatbots in our sample, with ChatGPT and Claude often referencing user characteristics such as age and sex in their responses, while DeepSeek and Docus did not, despite these user characteristics being provided at the beginning of the conversation. Personalization in health communication has been shown to enhance user experience by increasing perceived benefits and self-efficacy [81]. When AI chatbots did not include personalization, they may have missed opportunities to deliver tailored information that could have enhanced the uptake of HPV vaccination. In addition to personalization, using affirming language that acknowledges user concerns or feelings can further improve user experience and eventually improve engagement, particularly when discussing stigmatized topics like sexually transmitted infections caused by HPV [82]. However, it is important to note that implementing personalization and affirming language may require AI chatbots to retain user information, sometimes without the user's explicit consent, which in turn raises concerns about data privacy and security [83,84]. A study found that 47% of the respondents were concerned about the security of information entered into AI chatbots, which may reduce AI chatbots' acceptability in health care services [27]. To minimize these risks, robust data protection protocols are essential to safeguard user privacy when interacting with AI chatbots [83].

The uniformly neutral or provaccine stance exhibited by all AI chatbots in our sample corresponds with the scientific consensus, even when asked questions rooted in vaccination conspiracy beliefs. We found that AI chatbots maintained evidence-based stances without embracing antivaccine rhetoric. If AI chatbots take an antivaccine stance, it can fuel misinformation, such as antivaccine statements are scientifically unconfirmed and often rooted in personal stories [10,85]. While neutrality may not directly challenge misinformation, it offers an opportunity to present information in a nonjudgmental manner without using strongly persuasive messaging, helping to build trust among vaccine-hesitant individuals [86,87]. In challenging questions where AI chatbots were questioned about obtaining a vaccine despite potential life-threatening reactions, AI chatbots typically deferred to health care providers for evaluation rather than becoming the decision maker. By guiding users to health care professionals, AI chatbots effectively position themselves as supportive tools rather than substitutes for health care providers [88]. Chatbots can enhance online interactions between patients and health care professionals, especially when navigating complex medical questions [89].

Limitations

This study has multiple limitations. First, our data collection was performed during a defined period (February 22-28, 2025), documenting AI chatbot responses at a singular timeframe of their evolution. Due to the rapid advancement of AI technology, these findings may not accurately represent the current capabilities of these platforms. Second, our analysis focused on text-based responses and did not assess multimedia components that could support information dissemination on certain AI chatbots. Third, there was a time lapse of several weeks between data collection and data analysis, during which some of the links provided in AI chatbot responses were no longer available due to shifts in political leadership. Fourth, while we simulated young adult users, actual interactions may differ based on user's specific characteristics and differences in query formulation. Furthermore, it may not fully capture the diversity and natural language patterns seen in actual queries from young adults. Our question set may also not fully capture the breadth of real-world HPV-related concerns among young adults, although its development based on Google Trends data may reflect what users are commonly searching for in an online space. Furthermore, we did not measure actual user comprehension or behavioral outcomes following interactions with these AI chatbots, limiting our ability to assess their real-world relationship on HPV vaccination decisions. Fifth, we used AI chatbots with different update methods (static and dynamic) that may partially explain why some AI chatbots provided less relevant information. Finally, the generalizability of our findings may be limited as our sample focused on a subset of AI chatbots mimicking young adult users in the United States. We analyzed 2 queries per AI chatbot, which limits the ability to fully assess response variability and consistency within each platform. We focused on a select group of AI chatbots based on some sets of criteria, not comprehensively including more AI chatbots. Therefore, results may not extend to other populations, age groups, or other AI chatbots not assessed in this study.

Conclusions and Recommendations

Our study highlights the potential of AI chatbots as sources of generally accurate HPV vaccine information for young adults, potentially serving as alternatives to the misinformation-rich health content often found online. However, challenges remain in verifying the sources or links provided, ensuring readability, and maintaining personalization. These challenges underscore the urgent need for the development and adoption of standardized evaluation tools to ensure the efficacy of AI chatbots when they are used to supply health information to lay audiences. Such standardized tools would enable consistent benchmarking of AI chatbots' performance and facilitate meaningful comparisons in their ability to deliver reliable health information to lay audiences. In addition, standardized tools could help identify gaps in AI chatbot knowledge and functionality, guiding developers in refining these systems. To ensure information remains current and reliable, AI chatbots should implement regular knowledge updates and systematic checks for source accuracy. To improve readability, AI chatbots

XSL•FO
RenderX

should explore optimization approaches such as real-time readability adjustment. To enhance personalization, developers should design AI chatbot responses that are tailored to users while safeguarding their data privacy and security.

Our findings also signal the emergence of fostering "AI health literacy," which refers to the ability of individuals to understand and verify health information generated by AI tools. Moving forward, public health interventions must focus not only on the supply side by improving AI chatbots technology, but also on the demand side by equipping users with skills to verify AI-generated content to mitigate the overreliance on misinformation. We also recommend that users appropriately position AI chatbots as complements to, rather than replacements for, health care professionals' guidance. Given that AI chatbots sometimes adopt a neutral stance when discussing HPV vaccination, understanding the implications of this neutrality for combating misinformation and promoting vaccine confidence remains an important direction for future research.

Finally, moving forward, health policies should focus on partnering with technology firms to establish a systematic process for AI chatbots to help users distinguish high-quality and low-quality AI chatbots used in health care services. By addressing these considerations, AI chatbots could become accessible and valuable tools for improving HPV vaccine knowledge and potentially increasing vaccination rates among young adults.

Acknowledgments

No generative artificial intelligence (AI) was used to create the content of this manuscript. AI tools were only used to enhance grammar and readability before submission.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability

The datasets used and analyzed during the study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: AL, LMS-R, KJL, NMR, RJD, RDH, MLK
Formal analysis: AL, EC, MD
Investigation: EC, MD, KJL, NMR, RDH, MLK
Methodology: AL, KJL, NMR, RJD, RDH, MLK
Supervision: LMS-R, KJL, NMR, RJD, RDH, MLK
Validation: LMS-R
Writing – original draft: AL, EC, MD
Writing – review & editing: AL, LMS-R, KJL, NMR, RJD, RDH, MLK

Conflicts of Interest

MLK has received investigator-initiated research funding from Merck, administered through Purdue University, and has served as a consultant to Merck.

Multimedia Appendix 1
Questions set.
[PDF File (Adobe PDF File), 109 KB - publichealth_v12i1e79720_app1.pdf]

Multimedia Appendix 2
Study codebook.
[PDF File (Adobe PDF File), 118 KB - publichealth_v12i1e79720_app2.pdf]

Multimedia Appendix 3
Additional quotes.
[DOCX File , 16 KB - publichealth_v12i1e79720_app3.docx]

References

1. Human papillomavirus and cancer. World Health Organization. 2024 Mar 5. URL: https://www.who.int/news-room/fact-sheets/detail/human-papilloma-virus-and-cancer [accessed 2025-09-20]

XSL•FO
RenderX

2. Cancers linked with HPV each year. CDC. 2025 Dec 3. URL: https://www.cdc.gov/cancer/hpv/cases.html [accessed 2024-11-22]
3. Elissa M, Julianne G, Elizabeth U, Lauri M. Human papillomavirus. In: Pink Book: Epidemiology and Prevention of Vaccine-Preventable Diseases. Washington, DC: Centers for Disease Control and Prevention; 2021.
4. Luxembourg A, Moeller E. 9-Valent human papillomavirus vaccine: a review of the clinical development program. Expert Rev Vaccines 2017;16(11):1119-1139. [doi: 10.1080/14760584.2017.1383158] [Medline: 28956458]
5. Vaccination coverage among adults in the United States, National Health Interview Survey, 2022. CDC. 2024 Oct 4. URL: https://www.cdc.gov/adultvaxview/publications-resources/adult-vaccination-coverage-2022.html [accessed 2024-11-22]
6. Fokom Domgue J, Cunningham SA, Yu RK, Shete S. Reasons for not receiving the HPV vaccine among eligible adults: lack of knowledge and of provider recommendations contribute more than safety and insurance concerns. Cancer Med 2020;9(14):5281-5290 [FREE Full text] [doi: 10.1002/cam4.3192] [Medline: 32483891]
7. Wang X, Cohen RA. Health information technology use among adults: United States, July-December 2022. In: NCHS Data Briefs, no. 482. Hyattsville, MD: National Center for Health Statistics; Oct 2023.
8. Bhagavathula AS, Massey PM. Google Trends on human papillomavirus vaccine searches in the United States from 2010 to 2021: infodemiology study. JMIR Public Health Surveill 2022;8(8):e37656 [FREE Full text] [doi: 10.2196/37656] [Medline: 36036972]
9. Kearney MD, Selvan P, Hauer MK, Leader AE, Massey PM. Characterizing HPV vaccine sentiments and content on Instagram. Health Educ Behav 2019;46(2_suppl):37-48. [doi: 10.1177/1090198119859412] [Medline: 31742459]
10. Massey PM, Kearney MD, Hauer MK, Selvan P, Koku E, Leader AE. Dimensions of misinformation about the HPV vaccine on Instagram: content and network analysis of social media characteristics. J Med Internet Res 2020;22(12):e21451 [FREE Full text] [doi: 10.2196/21451] [Medline: 33270038]
11. Covolo L, Ceretti E, Passeri C, Boletti M, Gelatti U. What arguments on vaccinations run through YouTube videos in Italy? A content analysis. Hum Vaccin Immunother 2017;13(7):1693-1699 [FREE Full text] [doi: 10.1080/21645515.2017.1306159] [Medline: 28362544]
12. Kornides ML, Badlis S, Head KJ, Putt M, Cappella J, Gonzalez-Hernadez G. Exploring content of misinformation about HPV vaccine on Twitter. J Behav Med 2023;46(1-2):239-252 [FREE Full text] [doi: 10.1007/s10865-022-00342-1] [Medline: 35896853]
13. Grimes DR, Brennan LJ, O'Connor R. Establishing a taxonomy of potential hazards associated with communicating medical science in the age of disinformation. BMJ Open 2020;10(7):e035626 [FREE Full text] [doi: 10.1136/bmjopen-2019-035626] [Medline: 32624466]
14. Dib F, Mayaud P, Longfier L, Chauvin P, Launay O. Effect of internet use for searching information on vaccination on the uptake of human papillomavirus vaccine in France: a path-analysis approach. Prev Med 2021;149:106615. [doi: 10.1016/j.ypmed.2021.106615] [Medline: 33989671]
15. Urgent health challenges for the next decade. World Health Organization. 2020 Jan. URL: https://www.who.int/news-room/photo-story/photo-story-detail/urgent-health-challenges-for-the-next-decade [accessed 2023-11-15]
16. Pho A, Bakken S, Lunn M, Lubensky ME, Flentje A, Dastur Z, et al. Online health information seeking, health literacy, and human papillomavirus vaccination among transgender and gender-diverse people. J Am Med Inform Assoc 2022;29(2):285-295 [FREE Full text] [doi: 10.1093/jamia/ocab150] [Medline: 34383916]
17. Waser M, Heiss R, Borena W. Factors affecting children's HPV vaccination in Austria: evidence from a parent survey. Hum Vaccin Immunother 2022;18(6):2126251 [FREE Full text] [doi: 10.1080/21645515.2022.2126251] [Medline: 36251011]
18. Walker KK, Owens H, Zimet G. The role of the media on maternal confidence in provider HPV recommendation. BMC Public Health 2020;20(1):1765 [FREE Full text] [doi: 10.1186/s12889-020-09877-x] [Medline: 33228674]
19. Thompson EL, Wheldon CW, Rosen BL, Maness SB, Kasting ML, Massey PM. Awareness and knowledge of HPV and HPV vaccination among adults ages 27-45 years. Vaccine 2020;38(15):3143-3148. [doi: 10.1016/j.vaccine.2020.01.053] [Medline: 32029321]
20. Yacouti A, Elkhoudri N, El Got A, Benider A, Hadrya F, Baddou R, et al. Awareness, attitudes and acceptability of the HPV vaccine among female university students in Morocco. PLoS One 2022;17(4):e0266081 [FREE Full text] [doi: 10.1371/journal.pone.0266081] [Medline: 35395019]
21. Lakneh EA, Mersha EA, Asresie MB, Belay HG. Knowledge, attitude, and uptake of human papilloma virus vaccine and associated factors among female preparatory school students in Bahir Dar City, Amhara Region, Ethiopia. PLoS One 2022;17(11):e0276465 [FREE Full text] [doi: 10.1371/journal.pone.0276465] [Medline: 36409675]
22. Sidiropoulou M, Gerogianni G, Kourti FE, Pappa D, Zartaloudi A, Koutelekos I, et al. Perceptions, knowledge and attitudes among young adults about prevention of HPV infection and immunization. Healthcare (Basel) 2022;10(9):1721 [FREE Full text] [doi: 10.3390/healthcare10091721] [Medline: 36141333]
23. Koman J, Fauvelle K, Schuck S, Texier N, Mebarki A. Physicians' perceptions of the use of a chatbot for information seeking: qualitative study. J Med Internet Res 2020;22(11):e15185 [FREE Full text] [doi: 10.2196/15185] [Medline: 33170134]

XSL•FO
RenderX

24. Hopkins AM, Logan JM, Kichenadasse G, Sorich MJ. Artificial intelligence chatbots will revolutionize how cancer patients access information: ChatGPT represents a paradigm-shift. JNCI Cancer Spectr 2023;7(2):pkad010 [FREE Full text] [doi: 10.1093/jncics/pkad010] [Medline: 36808255]
25. Roo, a sexual health chatbot. Planned Parenthood. URL: https://www.plannedparenthood.org/learn/roo-sexual-health-chatbot [accessed 2025-05-27]
26. Fetrati H, Chan G, Orji R. Chatbots for sexual health improvement: a systematic review. Int J Human Computer Interact 2024;41(4):1997-2019. [doi: 10.1080/10447318.2024.2376808]
27. Nadarzynski T, Miles O, Cowie A, Ridge D. Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: a mixed-methods study. Digit Health 2019;5:2055207619871808 [FREE Full text] [doi: 10.1177/2055207619871808] [Medline: 31467682]
28. Alberts IL, Mercolli L, Pyka T, Prenosil G, Shi K, Rominger A, et al. Large language models (LLM) and ChatGPT: what will the impact on nuclear medicine be? Eur J Nucl Med Mol Imaging 2023;50(6):1549-1552 [FREE Full text] [doi: 10.1007/s00259-023-06172-w] [Medline: 36892666]
29. Shankar R, Bundele A, Mukhopadhyay A. Barriers and enablers for the deployment of large language model-based conversational robots for older adults: a protocol for a systematic review of qualitative studies. PLoS One 2025;20(4):e0321093 [FREE Full text] [doi: 10.1371/journal.pone.0321093] [Medline: 40168400]
30. Ray PP. ChatGPT: a comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. Internet Things Cyber-Phys Syst 2023;3:121-154. [doi: 10.1016/j.iotcps.2023.04.003]
31. Chakraborty C, Pal S, Bhattacharya M, Dash S, Lee S. Overview of chatbots with special emphasis on artificial intelligence-enabled ChatGPT in medical science. Front Artif Intell 2023;6:1237704 [FREE Full text] [doi: 10.3389/frai.2023.1237704] [Medline: 38028668]
32. Chopde A, Agrawal M. Chatbot using deep learning. SSRN J 2022. [doi: 10.2139/ssrn.4140506]
33. What is DeepSeek AI? DeepSeek AI. URL: https://deepseek.ai/ [accessed 2025-04-19]
34. Docus, AI-powered health platform. Docus. URL: https://docus.ai [accessed 2025-04-21]
35. Shahsavar Y, Choudhury A. User intentions to use ChatGPT for self-diagnosis and health-related purposes: cross-sectional survey study. JMIR Hum Factors 2023;10:e47564 [FREE Full text] [doi: 10.2196/47564] [Medline: 37195756]
36. Al Shboul MKI, Alwreikat A, Alotaibi FA. Investigating the use of ChatGPT as a novel method for seeking health information: a qualitative approach. Sci Technol Libr 2023;43(3):225-234. [doi: 10.1080/0194262x.2023.2250835]
37. Nadarzynski T, Bayley J, Llewellyn C, Kidsley S, Graham CA. Acceptability of artificial intelligence (AI)-enabled chatbots, video consultations and live webchats as online platforms for sexual health advice. BMJ Sex Reprod Health 2020;46(3):210-217 [FREE Full text] [doi: 10.1136/bmjsrh-2018-200271] [Medline: 31964779]
38. Jeyaraman M, Jeyaraman N, Nallakumarasamy A, Yadav S, Bondili SK. ChatGPT in medical education and research: a boon or a bane? Cureus 2023;15(8):e44316 [FREE Full text] [doi: 10.7759/cureus.44316] [Medline: 37779749]
39. Asch DA. An interview with ChatGPT about health care. NEJM Catalyst 2023;4(2). [doi: 10.1056/CAT.23.0043]
40. Kooli C. Chatbots in education and research: a critical examination of ethical implications and solutions. Sustainability 2023;15(7):5614. [doi: 10.3390/su15075614]
41. Ayers JW, Poliak A, Dredze M, Leas EC, Zhu Z, Kelley JB, et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. JAMA Intern Med 2023;183(6):589-596 [FREE Full text] [doi: 10.1001/jamainternmed.2023.1838] [Medline: 37115527]
42. Li Y, Li J, Li M, Yu E, Rhee D, Amith M, et al. VaxBot-HPV: a GPT-based chatbot for answering HPV vaccine-related questions. JAMIA Open 2025;8(1):ooaf005. [doi: 10.1093/jamiaopen/ooaf005] [Medline: 39975811]
43. Goodman RS, Patrinely JR, Stone CA, Zimmerman E, Donald RR, Chang SS, et al. Accuracy and reliability of chatbot responses to physician questions. JAMA Netw Open 2023;6(10):e2336483 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.36483] [Medline: 37782499]
44. Andrikyan W, Sametinger SM, Kosfeld F, Jung-Poppe L, Fromm MF, Maas R, et al. Artificial intelligence-powered chatbots in search engines: a cross-sectional study on the quality and risks of drug information for patients. BMJ Qual Saf 2025;34(2):100-109 [FREE Full text] [doi: 10.1136/bmjqs-2024-017476] [Medline: 39353736]
45. Shapiro GK, Holding A, Perez S, Amsel R, Rosberger Z. Validation of the vaccine conspiracy beliefs scale. Papillomavirus Res 2016;2:167-172 [FREE Full text] [doi: 10.1016/j.pvr.2016.09.001] [Medline: 29074176]
46. HPV vaccine. Google Trends. URL: https://trends.google.com/trends/explore?geo=US&q=%2Fm%2F094j5f&hl=en-US [accessed 2024-02-07]
47. Kan M. DeepSeek sees huge surge to become second most popular AI chatbot. PCMag. 2025 Feb 3. URL: https://www.pcmag.com/news/deepseek-sees-huge-surge-to-become-second-most-popular-ai-chatbot [accessed 2025-04-21]
48. Flick U. The SAGE Handbook of Qualitative Data Analysis. London: SAGE Publications; 2014.
49. SSA - POMS: NL 10605.105 - What is the Flesch-Kincaid readability test? Social Security Administration (USA). 2015. URL: https://secure.ssa.gov/poms.nsf/lnx/0910605105 [accessed 2025-04-22]
50. Agar A, Sahin A. Kyphosis-related information on the internet is the quality, content and readability sufficient for the patients? Global Spine J 2022;12(3):476-482 [FREE Full text] [doi: 10.1177/21925682211015955] [Medline: 33977763]

XSL•FO
RenderX

51. Xie Y, Seth I, Hunter-Smith DJ, Rozen WM, Seifman MA. Investigating the impact of innovative AI chatbot on post-pandemic medical education and clinical assistance: a comprehensive analysis. ANZ J Surg 2024;94(1-2):68-77. [doi: 10.1111/ans.18666] [Medline: 37602755]
52. Raja H, Lodhi S. Assessing the readability and quality of online information on anosmia. Ann R Coll Surg Engl 2024;106(2):178-184 [FREE Full text] [doi: 10.1308/rcsann.2022.0147] [Medline: 37051757]
53. Dykstra C, Laily A, Marsh EE, Kasting ML, DeMaria AL. "I think people should be more aware": uterine fibroid experiences among women living in Indiana, USA. Patient Educ Couns 2023;107:107584 [FREE Full text] [doi: 10.1016/j.pec.2022.107584] [Medline: 36473404]
54. Laily A, Jo S, Duncan RJ, Hubach RD, LaRoche KJ, Rodriguez NM, et al. Human papillomavirus awareness in a digital world: the intersection of faith and online media. Prev Med Rep 2025 Oct;58:103241 [FREE Full text] [doi: 10.1016/j.pmedr.2025.103241] [Medline: 41020203]
55. LaRoche KJ, Gross E, Sheehy G, Foster AM. Put a ring in it: exploring women's experiences with the contraceptive vaginal ring in Ontario. Womens Health Issues 2018;28(5):415-420. [doi: 10.1016/j.whi.2018.04.009] [Medline: 30061032]
56. LaRoche KJ, Jozkowski KN, Crawford BL, Jackson F. Can someone be both pro-life and pro-choice? Results from a national survey of US adults. Perspect Sex Reprod Health 2024;56(2):106-115. [doi: 10.1111/psrh.12262] [Medline: 38661077]
57. Hubach RD, Owens C. Findings on the monkeypox exposure mitigation strategies employed by men who have sex with men and transgender women in the United States. Arch Sex Behav 2022;51(8):3653-3658 [FREE Full text] [doi: 10.1007/s10508-022-02423-3] [Medline: 36103027]
58. Hubach RD, Zipfel R, Muñoz FA, Brongiel I, Narvarte A, Servin AE. Barriers to sexual and reproductive care among cisgender, heterosexual and LGBTQIA+ adolescents in the border region: provider and adolescent perspectives. Reprod Health 2022;19(1):93 [FREE Full text] [doi: 10.1186/s12978-022-01394-x] [Medline: 35414000]
59. Kasting ML, Christy SM, Stout ME, Zimet GD, Mosher CE. Attitudinal correlates of HPV vaccination in college women. Clin Nurs Res 2022;31(5):826-835 [FREE Full text] [doi: 10.1177/10547738211045227] [Medline: 34538128]
60. Laily A, Duncan R, Gabhart KM, Nephew LD, Christy SM, Vadaparampil ST, et al. Differences in provider hepatitis C virus screening recommendations by patient risk status. Prev Med Rep 2024;38:102602 [FREE Full text] [doi: 10.1016/j.pmedr.2024.102602] [Medline: 38375175]
61. Laily A, Nair I, Shank SE, Wettschurack C, Khamis G, Dykstra C, et al. Enhancing uterine fibroid care: clinician perspectives on diagnosis, disparities, and strategies for improving health care. Womens Health Rep (New Rochelle) 2024;5(1):293-304 [FREE Full text] [doi: 10.1089/whr.2023.0113] [Medline: 38558944]
62. Schwab-Reese L, Short C, Jacobs L, Fingerman M. Rapport building in written crisis services: qualitative content analysis. J Med Internet Res 2024;26:e42049 [FREE Full text] [doi: 10.2196/42049] [Medline: 38748472]
63. Schwab-Reese L, Renner LM, Quang D, Carlyle K, Nair I, Chase K, et al. "If I stay silent, it doesn't help anyone": understanding how people use TikTok to talk about sexual violence. Violence Against Women 2025;31(12-13):3311-3329. [doi: 10.1177/10778012241279106] [Medline: 39262394]
64. Owens C, Gray SJ, Carter K, Hoffman M, Mullen C, Hubach RD. Implementation facilitators and barriers for primary care providers prescribing daily oral PrEP to adolescents in the United States. AIDS Patient Care STDS 2023;37(8):379-393. [doi: 10.1089/apc.2023.0090] [Medline: 37566536]
65. Kasting ML, Laily A, Nephew LD, Shields CG, Shedd-Steele R, Rawl SM. Development and feasibility testing of a multilevel intervention to increase hepatitis C virus screening among baby boomers in primary care. J Cancer Educ 2023;38(2):718-729 [FREE Full text] [doi: 10.1007/s13187-023-02268-x] [Medline: 36800082]
66. Zhou J, Chen AZ, Shah D, Schwab-Reese LM, De Choudhury M. A risk taxonomy and reflection tool for large language model adoption in public health. Proc ACM Hum Comput Interact 2025;9(7):CSCW363. [doi: 10.1145/3757544]
67. Zhou J, Alfonso M, Schwab-Reese L. Ethical use of AI in public health. 2025 Presented at: APHA Annual Meeting 2025; November 2-5, 2025; Washington, DC URL: https://apha.confex.com/apha/2025/meetingapp.cgi/Session/75396
68. Brennan LP, Rodriguez NM, Head KJ, Zimet GD, Kasting ML. Obstetrician/gynecologists' HPV vaccination recommendations among women and girls 26 and younger. Prev Med Rep 2022;27:101772 [FREE Full text] [doi: 10.1016/j.pmedr.2022.101772] [Medline: 35321213]
69. Owens HN, Brownstein NC, Whiting J, Arevalo M, Kasting ML, Vadaparampil ST, et al. Factors associated with human papillomavirus (HPV) patient-clinician communication and HPV vaccine uptake among adults in the United States. Prev Med 2025;195:108280. [doi: 10.1016/j.ypmed.2025.108280] [Medline: 40188881]
70. Feminizing hormone therapy. Mayo Clinic. URL: https://www.mayoclinic.org/tests-procedures/feminizing-hormone-therapy/about/pac-20385096 [accessed 2026-01-21]
71. Hoekstra EJ, van den Ent MMVX, Dao H, Khalaf H, Salovaara A. Measles supplementary immunization activities and GAVI funds as catalysts for improving injection safety in Africa. J Infect Dis 2011;204(Suppl 1):S190-S197. [doi: 10.1093/infdis/jir073] [Medline: 21666161]
72. Rao A, Pang M, Kim J, Kamineni M, Lie W, Prasad AK, et al. Assessing the utility of ChatGPT throughout the entire clinical workflow: development and usability study. J Med Internet Res 2023;25:e48659 [FREE Full text] [doi: 10.2196/48659] [Medline: 37606976]

XSL•FO
RenderX

73. Passanante A, Pertwee E, Lin L, Lee KY, Wu JT, Larson HJ. Conversational AI and vaccine communication: systematic review of the evidence. J Med Internet Res 2023;25:e42758 [FREE Full text] [doi: 10.2196/42758] [Medline: 37788057]
74. Singer E. Thousands of U.S. government web pages have been taken down since Friday. The New York Times. 2025 Feb. URL: https://www.nytimes.com/2025/02/02/upshot/trump-government-websites-missing-pages.html [accessed 2025-05-09]
75. Stone W. Judge orders restoration of federal health websites. NPR. 2025 Feb 11. URL: https://www.npr.org/sections/shots-health-news/2025/02/11/nx-s1-5293387/judge-orders-cdc-fda-hhs-websites-restored [accessed 2025-05-09]
76. Aljamaan F, Temsah M, Altamimi I, Al-Eyadhy A, Jamal A, Alhasan K, et al. Reference hallucination score for medical artificial intelligence chatbots: development and usability study. JMIR Med Inform 2024;12:e54345 [FREE Full text] [doi: 10.2196/54345] [Medline: 39083799]
77. Jamaluddin J, Gaffar NA, Din NSS. Hallucination: a key challenge to artificial intelligence-generated writing. Malays Fam Physician 2023;18:68 [FREE Full text] [doi: 10.51866/lte.527] [Medline: 38111832]
78. Rooney MK, Santiago G, Perni S, Horowitz DP, McCall AR, Einstein AJ, et al. Readability of patient education materials from high-impact medical journals: a 20-year analysis. J Patient Exp 2021;8:2374373521998847 [FREE Full text] [doi: 10.1177/2374373521998847] [Medline: 34179407]
79. Nematzadeh A, Ciampaglia GL, Ahn Y, Flammini A. Information overload in group communication: from conversation to cacophony in the Twitch chat. R Soc Open Sci 2019;6(10):191412 [FREE Full text] [doi: 10.1098/rsos.191412] [Medline: 31824736]
80. Mao B, Jia X, Huang Q. How do information overload and message fatigue reduce information processing in the era of COVID-19? An ability–motivation approach. J Inf Sci 2022;50(5):1242-1254. [doi: 10.1177/01655515221118047]
81. Liu Y, Yan W, Hu B, Li Z, Lai YL. Effects of personalization and source expertise on users' health beliefs and usage intention toward health chatbots: evidence from an online experiment. Digit Health 2022;8:20552076221129718 [FREE Full text] [doi: 10.1177/20552076221129718] [Medline: 36211799]
82. Park E. I trust you, but let me talk to AI: the role of the chat agents, empathy, and health issues in misinformation guidance. Int J Strateg Commun 2025;19(2):231-260. [doi: 10.1080/1553118x.2025.2462087]
83. Talyshinskii A, Naik N, Hameed BMZ, Juliebø-Jones P, Somani BK. Potential of AI-driven chatbots in urology: revolutionizing patient care through artificial intelligence. Curr Urol Rep 2024;25(1):9-18 [FREE Full text] [doi: 10.1007/s11934-023-01184-3] [Medline: 37723300]
84. Ishaaq N, Sohail SS. Re: Investigating the impact of innovative AI chatbot on post-pandemic medical education and clinical assistance: a comprehensive analysis. ANZ J Surg 2024;94(3):494. [doi: 10.1111/ans.18721] [Medline: 37784244]
85. Morgan JC, Badlis S, Head KJ, Zimet G, Cappella JN, Kornides ML. Types of HPV vaccine misinformation circulating on Twitter (X) that parents find most concerning: insights from a cross-sectional survey and content analysis. J Med Internet Res 2025;27:e54657 [FREE Full text] [doi: 10.2196/54657] [Medline: 40354114]
86. Huang L, Huschka TR, Barwise AK, Allen JP, Wolfersteig W, Hamm K, et al. Psychological reactance, misinformation, and distrust: a mixed methods analysis of COVID-19 vaccine uptake. J Clin Transl Sci 2024;8(1):e48 [FREE Full text] [doi: 10.1017/cts.2024.15] [Medline: 38510694]
87. Whitehead HS, French CE, Caldwell DM, Letley L, Mounier-Jack S. A systematic review of communication interventions for countering vaccine misinformation. Vaccine 2023;41(5):1018-1034 [FREE Full text] [doi: 10.1016/j.vaccine.2022.12.059] [Medline: 36628653]
88. Garcia Valencia OA, Suppadungsuk S, Thongprayoon C, Miao J, Tangpanithandee S, Craici IM, et al. Ethical implications of chatbot utilization in nephrology. J Pers Med 2023;13(9):1363 [FREE Full text] [doi: 10.3390/jpm13091363] [Medline: 37763131]
89. Cosma C, Radi A, Cattano R, Zanobini P, Bonaccorsi G, Lorini C, et al. Exploring chatbot contributions to enhancing vaccine literacy and uptake: a scoping review of the literature. Vaccine 2025;44:126559 [FREE Full text] [doi: 10.1016/j.vaccine.2024.126559] [Medline: 39615346]

Abbreviations

AI: artificial intelligence
CDC: Centers for Disease Control and Prevention
COREQ: Consolidated Criteria for Reporting Qualitative Research
HPV: human papillomavirus
LLM: large language model
VCBS: Vaccine Conspiracy Beliefs Scale
WHO: World Health Organization

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 26.Jun.2025; peer-reviewed by H Fetrati, G Cangelosi, MS Mekbib, S Sahai, X Liang; comments to author 01.Sep.2025; revised version received 04.Dec.2025; accepted 22.Dec.2025; published 18.Feb.2026.

Please cite as:
Laily A, Schwab-Reese LM, Davish M, Cahue E, LaRoche KJ, Rodriguez NM, Duncan RJ, Hubach RD, Kasting ML
Examining Artificial Intelligence Chatbots' Responses in Providing Human Papillomavirus Vaccine Information for Young Adults: Qualitative Content Analysis
JMIR Public Health Surveill 2026;12:e79720
URL: https://publichealth.jmir.org/2026/1/e79720
doi:10.2196/79720
PMID:

©Alfu Laily, Laura M Schwab-Reese, Megan Davish, Emily Cahue, Kathryn J LaRoche, Natalia M Rodriguez, Robert J Duncan, Randolph D Hubach, Monica L Kasting. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 18.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Retraction

Retraction: Secular Trends in Gastric and Esophageal Cancer Attributable to Dietary Carcinogens From 1990 to 2019 and Projections Until 2044 in China: Population-Based Study

JMIR Publications Editorial Office[1]

JMIR Publications, Toronto, ON

Corresponding Author:
JMIR Publications
130 Queens Quay East, Suite 1100-1102
Toronto, ON
Phone: 1 4165832040
Email: ed-support@jmir.org

Related Article:

Retraction of: https://publichealth.jmir.org/2023/1/e48449

(JMIR Public Health Surveill 2026;12:e101187) doi:10.2196/101187

The JMIR Publications Editorial Office is retracting the article "Secular Trends in Gastric and Esophageal Cancer Attributable to Dietary Carcinogens From 1990 to 2019 and Projections Until 2044 in China: Population-Based Study" [1].

This action is being taken due to identified concerns regarding the integrity of the peer-review process for this article. Following investigation, an author disclosed to the Editorial Office that they had impersonated an appointed reviewer for this article and fabricated a review.

Following this disclosure, the corresponding author Bing Hua Yin requested retraction of the article. This matter has since been referred to the institute for further investigation.

All of the authors agree with the decision to retract the article.

Reference

1. Guo HX, Wang Q, Wang C, Yin QC, Huo HZ, Yin BH. Retracted: Secular trends in gastric and esophageal cancer attributable to dietary carcinogens from 1990 to 2019 and projections until 2023, Vol. 9:e48449.

Submitted 12.May.2026; this is a non–peer-reviewed article;accepted 14.May.2026; published 16.Jun.2026.

Please cite as:
JMIR Publications Editorial Office
Retraction: Secular Trends in Gastric and Esophageal Cancer Attributable to Dietary Carcinogens From 1990 to 2019 and Projections Until 2044 in China: Population-Based Study
JMIR Public Health Surveill 2026;12:e101187
URL: https://publichealth.jmir.org/2026/1/e101187
doi:10.2196/101187
PMID:

©JMIR Publications Editorial Office. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 16.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Retraction

Retraction: Arterial Stiffness and Obesity as Predictors of Diabetes: Longitudinal Cohort Study

JMIR Publications Editorial Office[1]

JMIR Publications, Toronto, ON

Corresponding Author:
JMIR Publications
130 Queens Quay East, Suite 1100-1102
Toronto, ON
Phone: 1 4165832040
Email: ed-support@jmir.org

Related Article:

Retraction of: https://publichealth.jmir.org/2024/1/e46088

(JMIR Public Health Surveill 2026;12:e108369) doi:10.2196/108369

The JMIR Publications Editorial Office is retracting the article "Arterial Stiffness and Obesity as Predictors of Diabetes: Longitudinal Cohort Study" [1] due to concerns regarding the integrity of the peer review process.

Author Lin Liu responded to the concerns, stating that there were undisclosed conflicts of interest with the reviewers but asserted these did not impact the peer review process for the article.

However, the Editorial Office and editor in chief have determined that there is sufficient evidence indicating that the peer review process for this article was compromised.

The authors did not respond to the decision to retract the article.

Reference

1. Cui C, Wu Z, Zheng J, Chu J, Zhu J, Qi Y, et al. Arterial stiffness and obesity as predictors of diabetes: longitudinal cohort study. JMIR Public Health Surveill 2024 Feb 08;10:e46088 [FREE Full text] [doi: 10.2196/46088] [Medline: 38329798]

Submitted 30.Jul.2026; this is a non–peer-reviewed article;accepted 30.Jul.2026; published 19.Aug.2026.

Please cite as:
JMIR Publications Editorial Office
Retraction: Arterial Stiffness and Obesity as Predictors of Diabetes: Longitudinal Cohort Study
JMIR Public Health Surveill 2026;12:e108369
URL: https://publichealth.jmir.org/2026/1/e108369
doi:10.2196/108369
PMID:

©JMIR Publications Editorial Office. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 19.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX