

# JMIR Public Health and Surveillance

Impact Factor (2022): 8.5  
 Volume 7 (2021), Issue 7 ISSN 2369-2960 Editor in Chief: Travis Sanchez, PhD, MPH

## Contents

### Original Papers

<p>Telepsychiatry Consultation for Primary Care Treatment of Children and Adolescents Receiving Child Protective Services in Chile: Mixed Methods Feasibility Study (<a href="#">e25836</a>)                      Adrian Mundt, Matías Irrázaval, Pablo Martínez, Olga Fernández, Vania Martínez, Graciela Rojas. . . . .</p>	3
<p>Availability, Formulation, Labeling, and Price of Low-sodium Salt Worldwide: Environmental Scan (<a href="#">e27423</a>)                      Xuejun Yin, Hueiming Liu, Jacqui Webster, Kathy Trieu, Mark Huffman, J Miranda, Matti Marklund, Jason Wu, Laura Cobb, Ka Li, Sallie-Anne Pearson, Bruce Neal, Maoyi Tian. . . . .</p>	10
<p>The Reliability and Quality of YouTube Videos as a Source of Public Health Information Regarding COVID-19 Vaccination: Cross-sectional Study (<a href="#">e29942</a>)                      Calvin Chan, Viknesh Sounderajah, Elisabeth Daniels, Amish Acharya, Jonathan Clarke, Seema Yalamanchili, Pasha Normahani, Sheraz Markar, Hutan Ashrafian, Ara Darzi. . . . .</p>	21
<p>Measuring Public Concern About COVID-19 in Japanese Internet Users Through Search Queries: Infodemiological Study (<a href="#">e29865</a>)                      Zhiwei Gao, Sumio Fujita, Nobuyuki Shimizu, Kongmeng Liew, Taichi Murayama, Shuntaro Yada, Shoko Wakamiya, Eiji Aramaki. . . . .</p>	32
<p>The Influence of COVID-19 Information Sources on the Attitudes and Practices Toward COVID-19 Among the General Public of Saudi Arabia: Cross-sectional Online Survey Study (<a href="#">e28888</a>)                      Noor Alshareef, Ismaeel Yunusa, Mohammed Al-Hanawi. . . . .</p>	42
<p>The Saudi Ministry of Health’s Twitter Communication Strategies and Public Engagement During the COVID-19 Pandemic: Content Analysis Study (<a href="#">e27942</a>)                      Fatimah Alhassan, Sharifah AlDossary. . . . .</p>	57
<p>Pre-exposure Prophylaxis (PrEP) Information on Instagram: Content Analysis (<a href="#">e23876</a>)                      Eric Walsh-Buhi, Rebecca Houghton, Claire Lange, Ryli Hockensmith, John Ferrand, Lourdes Martinez. . . . .</p>	71
<p>The Effect of a Wordless, Animated, Social Media Video Intervention on COVID-19 Prevention: Online Randomized Controlled Trial (<a href="#">e29060</a>)                      Alain Vandormael, Maya Adam, Merlin Greuel, Jennifer Gates, Caterina Favaretti, Violetta Hachaturyan, Till Bärnighausen. . . . .</p>	83
<p>Building an Interactive Geospatial Visualization Application for National Health Care–Associated Infection Surveillance: Development Study (<a href="#">e23528</a>)                      Shuai Zheng, Jonathan Edwards, Margaret Dudeck, Prachi Patel, Lauren Wattenmaker, Muzna Mirza, Sheri Tejedor, Kent Lemoine, Andrea Benin, Daniel Pollock. . . . .</p>	95

---

The Psychosocial Impacts of COVID-19 on a Sample of Australian Adults: Cross-sectional Survey and Sentiment Analysis ( <a href="#">e29213</a> )	
Jillian Ryan, Hamza Sellak, Emily Brindal. ....	105
Risk Factors of Cholera Transmission in Al Hudaydah, Yemen: Case-Control Study ( <a href="#">e27627</a> )	
Abdulqawi Qaserah, Mohammed Al Amad, Abdulwahed Al Serouri, Yousef Khader. ....	118
Performance of the Severe Acute Respiratory Illness Sentinel Surveillance System in Yemen: Mixed Methods Evaluation Study ( <a href="#">e27621</a> )	
Ghamdan Alkholidy, Labiba Anam, Ali Almahaqri, Yousef Khader. ....	124

## Corrigenda and Addenda

Correction: Global Changes and Factors of Increase in Caloric/Salty Food Intake, Screen Use, and Substance Use During the Early COVID-19 Containment Phase in the General Population in France: Survey Study ( <a href="#">e31906</a> )	
Benjamin Rolland, Frédéric Haesebaert, Elodie Zante, Amine Benyamina, Julie Haesebaert, Nicolas Franck. ....	103

Original Paper

# Telepsychiatry Consultation for Primary Care Treatment of Children and Adolescents Receiving Child Protective Services in Chile: Mixed Methods Feasibility Study

Adrian P Mundt<sup>1,2</sup>, MD, PhD; Matías Irrázaval<sup>1,3,4</sup>, MD, MPH; Pablo Martínez<sup>1,4,5</sup>, PhD; Olga Fernández<sup>4,6</sup>, MSc, PhD; Vania Martínez<sup>4,5,7</sup>, MD, PhD; Graciela Rojas<sup>1,4,5</sup>, Dr med

<sup>1</sup>Departamento de Psiquiatría y Salud Mental, Hospital Clínico Universidad de Chile, Santiago, Chile

<sup>2</sup>Facultad de Medicina, Universidad Diego Portales, Santiago, Chile

<sup>3</sup>Departamento de Salud Mental, Ministerio de Salud, Santiago, Chile

<sup>4</sup>Millennium Institute for Depression and Personality Research (MIDAP), Santiago, Chile

<sup>5</sup>Millennium Nucleus to Improve the Mental Health of Adolescents and Youths, Imhay, Santiago, Chile

<sup>6</sup>Unidad de Psiquiatría Infantil y del Adolescente, Departamento Psiquiatría y Salud Mental, Universidad de Chile, Santiago, Chile

<sup>7</sup>Centro de Salud Reproductiva y Desarrollo Integral del Adolescente (CEMERA), Universidad de Chile, Santiago, Chile

**Corresponding Author:**

Adrian P Mundt, MD, PhD

Departamento de Psiquiatría y Salud Mental

Hospital Clínico Universidad de Chile

Av La Paz 1003

Santiago, 8420000

Chile

Phone: 56 950033439

Email: [adrian.mundt@gmail.com](mailto:adrian.mundt@gmail.com)

## Abstract

**Background:** Children and adolescents living under the supervision of child protective services have complex mental health care needs. The scarcity and uneven distribution of specialized mental health teams in Chile may limit the provision and quality of care for this vulnerable population. Telepsychiatry can address such health inequities.

**Objective:** The objective of this study was to evaluate the feasibility of a telepsychiatry consultation program for primary health care (PHC) treatment of children and adolescents living under the supervision of child protective services.

**Methods:** We developed a telepsychiatry consultation program for two rural PHC clinics located in central Chile (Valparaíso Region) and evaluated its implementation using a mixed methods study design. The program consisted of videoconferencing mental health consultation sessions scheduled twice per month (each 90 minutes long), over a 6-month period, delivered by child and adolescent psychiatrists based in Santiago, Chile. We described the number of mental health consultation sessions, participant characteristics, perceived usefulness and acceptability, and experiences with the telepsychiatry consultation program.

**Results:** During the 6-month study period, 15 videoconferencing mental health consultation sessions were held. The telepsychiatry consultation program assisted PHC clinicians in assigning the most adequate diagnoses and making treatment decisions on pharmacotherapy and/or psychotherapy of 11 minors with complex care needs. The intervention was perceived to be useful by PHC clinicians for improving the resolution capacity in the treatments of this patient population. Limitations such as connectivity issues were resolved in most sessions.

**Conclusions:** The telepsychiatry consultation program was feasible and potentially useful to support PHC clinicians in the management of institutionalized children and adolescents with complex psychosocial care needs living in a poorly resourced setting. A larger scale trial should assess clinical outcomes in the patient population. Regulations and resources for this service model are needed to facilitate sustainability and large-scale implementation.

(*JMIR Public Health Surveill* 2021;7(7):e25836) doi:[10.2196/25836](https://doi.org/10.2196/25836)

**KEYWORDS**

telemedicine; psychiatry; primary health care; child protective services

## Introduction

About half of the children and adolescents living under the supervision of child protective services (CPS) in high-income countries have mental disorders [1]. In Chile—a Latin American developing country—69% of the population living under CPS has mental disorders and complex psychosocial care needs, but the availability and uneven distribution of specialized mental health care (SMHC) in Chile limit the provision and quality of mental health care for this vulnerable population [2-4]. This prompted the implementation of an intersectoral mental health care plan aimed to improve the collaboration between CPS, primary health care (PHC), and SMHC [5].

Telepsychiatry, that is, videoconferencing to deliver mental health care, might address the shortage of SMHC by facilitating remote and timely access to quality care for underserved populations (eg, rural populations) [6-8]. Technology-based solutions increase the access to care, are acceptable by culturally diverse populations, and provide mental health outcomes comparable to in-person care [8,9]. These characteristics have also contributed to the role of telepsychiatry in emergency response to the current COVID-19 pandemic [10,11]. In Chile [12] and other countries [13-15], child and adolescent telepsychiatry services have shown promising results for the treatment of mental disorders in PHC. These initiatives have received positive feedback from health care providers serving CPS populations and reported increased use of outpatient resources for children with complex mental health care needs [16,17]. Since PHC professionals usually lack training in the treatment of complex mental health problems of CPS populations, and CPS populations lack funds for traveling to receive SMHC far from their residencies, telepsychiatry solutions are promising to improve access and quality of care.

To our knowledge, no previous study has examined the collaboration between SMHC and PHC to specifically treat CPS populations. The objective of this study was to evaluate the feasibility of a telepsychiatry consultation program (TCP) for the PHC treatment of children and adolescents living under the supervision of CPS.

## Methods

### Study Participants and Procedures

A feasibility study combining descriptive quantitative and qualitative methods was conducted in two rural PHC clinics located in the Valparaíso Region of Chile. PHC clinics were conveniently selected as local health services, and staff were open to engage with this pilot project. Four psychologists and one physician working in PHC clinics were recruited. These clinicians treated children and adolescents living under the supervision of CPS as part of their usual case load. We provided the technical infrastructure, security, and protocols needed to connect the rural PHC clinicians with the SMHC team comprising 3 child and adolescent psychiatrists who were working at the Department of Psychiatry and Mental Health of Hospital Clínico Universidad de Chile, Faculty of Medicine of Universidad de Chile, and Department of Psychiatry of Hospital El Pino, Santiago. The following measures were taken to

implement the intervention: (1) securing institutional permissions to move and install technological equipment, (2) installing computers and teleconferencing devices in the intervention sites, (3) ensuring that the PHC clinicians and the SMHC teams had compatible schedules, (4) training the PHC clinicians and the SMHC teams in the technical aspects of the interventions, and (5) running connectivity tests between the study sites to pilot the intervention. This study was approved by the institutional review board at Hospital Clínico Universidad de Chile. PHC clinicians provided written informed consent and participated voluntarily at any time. Patients (ie, children and adolescents living under the supervision of CPS) were not directly involved in the feasibility testing due to complicated consent procedures required for their participation.

### Intervention

The TCP used a computer-based teleconferencing service that met communication and security standards for voice over transmission control protocol/IP. Videoconferencing mental health consultation (MHC) sessions were held between PHC clinicians and the SMHC team through a secure web-based treatment platform, that is, closed network, personalized access passwords, and disabled recording function, by using the desktop video-conferencing program Vidyo (Vidyo, Inc).

Videoconferencing MHC sessions were scheduled twice per month over a 6-month period, reserving time slots of 90 minutes each. Electronic health records (EHRs) were implemented to share patient information (eg, medical records, tests, and professional reports) between PHC clinicians and the SMHC team. EHRs were updated after each videoconferencing MHC session. Neither in-person visits nor emergency treatments were included in the study, as PHC clinicians referred such cases to local SMHC consultants. The TCP was not provided to other populations.

The TCP followed the protocol of the Chilean Ministry of Health for MHC [18], providing diagnostic assistance, management recommendations for treatment-resistant cases, and SMHC referral assessment. A protocol of the Chilean Ministry of Health defines MHC as the joint and interactive activity between SMHC and PHC teams, occurring at least once per month, to facilitate the shared continuity of care [18]. The MHC should include supervision, support and training regarding clinical cases, and clinical and administrative coordination to guarantee care continuity [18]. Before the implementation of videoconferencing MHC sessions, in-person MHC was considered as the established clinical routine.

### Measurements and Analysis

The number of videoconferencing MHC sessions held during the 6-month study period was registered and compared to in-person standard MHC sessions that were provided by local SMHC consultants in the 6 months prior to TCP implementation. The following data were collected for each videoconferencing MHC session: (1) number and types of health care providers participating, duration (in minutes) of the session, and network connection technology used; (2) clinician-rated usefulness and acceptability (assessed using six closed Likert-type items, two questions with three answer choices, and an open-ended box

for comments); and (3) clinical patient information and actions taken or agreed for patient management (retrieved through inspection of the shared EHRs). The aforementioned information was analyzed using descriptive statistics.

After the 6-month study period, an open-ended, email questionnaire was used to explore the experiences of PHC clinicians and SMHC teams with the TCP (ie, satisfaction, facilitators and barriers, and recommendations). Following a grounded theory approach [19], two researchers independently coded the data using open coding. Consensus was reached through discussion and the involvement of a third researcher to ensure intersubjective consensus.

## Results

Half of the MHC sessions planned (4/8, 50%) were held in the 6 months prior to the implementation of TCP. In contrast, 15 of the 24 (63%) videoconferencing MHC sessions planned were held during the 6-month study period. Nine videoconferencing MHC sessions were cancelled due to a strike of the health care workers, incompatible schedules, and low clinical demand.

**Table 1.** Characteristics of 11 children and adolescents living under the supervision of child protective services in rural Chile who were treated in a telepsychiatry consultation program.

Variable	Participants
Sex (female), n (%)	9 (82)
Age in years, mean (SD)	14.1 (3.1)
<b>Living situation, n (%)</b>	
Residency without family member	5 (45)
Residency with family member	3 (27)
Family (nuclear or extended)	3 (27)
<b>Reason for living under child protective services<sup>a</sup>, n (%)</b>	
Abandonment	1 (6)
Sexual abuse	2 (12)
Maltreatment	2 (12)
Parental neglect	7 (41)
Sexual exploitation	2 (12)
Rape	1 (6)
Family violence	3 (18)

<sup>a</sup>May include more than one reason per case.

Data collected during each videoconferencing MHC session showed that these were clinically useful to PHC clinicians (Table 2); its duration was deemed adequate in 92% of occasions; and that, after each session, interest in participating in an additional session did not decrease. Qualitative data analysis (Table 3) revealed that PHC clinicians and the SMHC team perceived the TCP as useful and that it helped to improve the quality of mental health care for minors living under the supervision of CPS. PHC clinicians expressed the need to maintain the TCP over time and stated that the intervention met their expectations regarding the methodology used and guidance on pharmacological

On average, videoconferencing MHC sessions were conducted every 3.7 weeks, lasted 66 minutes (range 30-100 minutes), covered 1.1 cases, and involved the participation of 2.1 PHC clinicians. The SMHC team mostly used wireless connections in about 73% (11/15) of the sessions, whereas the PHC clinicians used wired connections in 73% (11/15) of the sessions. On two occasions, participants communicated over telephone due to important and unresolved internet connectivity problems. Connectivity issues caused shortening and/or interruption of 12 videoconferencing MHC sessions.

Most cases of minors included in the TCP were girls who lived under CPS (ie, not with their relatives) due to parental neglect (Table 1). The most frequent diagnostic hypothesis was mixed behavioral and emotional disorder, reported in 5 of the 11 (45%) participants. Furthermore, psychiatric comorbidity was reported in 7 (64%) patients (ie, case load of the participating PHC professionals). The SMHC team mostly made recommendations about pharmacological schemes (n=8, 73%), psychotherapies (n=4, 36%), and psychopathological assessments (n=4, 36%).

treatments. The SMHC team emphasized that the TCP allowed them to train PHC clinicians in general aspects of mental health care, and to increase the resolution capacity of PHC clinicians. Key facilitators of the implementation were positive attitudes of the authorities of PHC, clinicians, and the SMHC team toward TCP. Incomplete patient information in EHRs was one of the main barriers to implementation, followed by technological and logistic difficulties. Recommendations made for future TCP implementations included the availability of a support technician and more guidance on psychotherapy.

**Table 2.** Perceived usefulness of a telepsychiatry consultation program, as rated by primary health care clinicians (N=8).

Item	Score <sup>a</sup> , mean (SD)	Range
Improved clinical understanding of the case presented	2.6 (0.5)	2-3
Learning experience applicable to similar cases	2.5 (0.5)	2-3
Guidance about medications and dosage	2.6 (0.5)	2-3
Orientation about psychosocial interventions	2.5 (0.7)	1-3
Benefit for the children and adolescents	2.4 (0.6)	1-3
Benefit for the primary health clinician	2.8 (0.4)	2-3

<sup>a</sup>Scores range from 0 to 3, with higher scores reflecting higher perceived usefulness.

**Table 3.** Synthesis of the main categories found in the qualitative data.

Category and subcategory	Illustrative quote
Usefulness of the intervention and improved quality of care	“[Did the TCP help?] <i>Certainly, all the consultations and especially the last one... the person in charge of the team stated how useful the telepsychiatry consultations had been</i> ” [Psychiatrist #1]
Need for continued support	“ <i>It would be very useful to keep this program available, for more months</i> ” [Primary health care psychologist #4] “ <i>...It would be good for it to be a regular program, it would greatly contribute to mental health teams in primary care</i> ” [Primary health care psychologist #3]
<b>Meeting expectations</b>	
On the methodology used	“ <i>There was a clear methodology and a well-defined way of handling each case</i> ” [Primary health care physician #1] “ <i>The clinical advice enabled us to develop teamwork, which added dynamism to the care we delivered, ...communication became more effective, and that strengthened teamwork</i> ” [Primary health care psychologist #3]
On guiding pharmacological treatments	“ <i>The goal of the consultations was met, especially regarding the clarification of doubts about the pharmacotherapy</i> ” [Psychiatrist #1]
Training primary health care practitioners and increased resolution capacity	“ <i>It not only enabled us to give them advice about specific cases; we also trained them about more general aspects. I sent them complementary material for psychotherapy approaches</i> ” [Psychiatrist #3] “ <i>I specifically helped them by increasing their self-confidence and facilitating their decision-making processes</i> ” [Psychiatrist #1]
<b>Facilitators and barriers to implementation</b>	
Facilitators to implementation	“ <i>Openness of the directors, motivation of staff members</i> ” [Primary health care psychologist #1] “ <i>The willingness to help of the psychiatrist</i> ” [Primary health care psychologist #4] “ <i>Punctuality, willingness to participate and interest in the activity, they were open to my suggestions and advice as a consultant</i> ” [Psychiatrist #1] “ <i>I enjoy teamwork and have a personal interest in telepsychiatry</i> ” [Psychiatrist #2]
Barriers to implementation	“ <i>[On electronic health records] They don't include all the information that we'd need for an adequate teleconsultation session, for example, the medications that the patients were taking</i> ” [Psychiatrist #3] “ <i>The [residence] does not have full clinical records...about the patient</i> ” [Primary health care physician #1] “ <i>Interruptions due to technical problems such as internet access, availability of an office and/or computer</i> ” [Psychiatrist #1] “ <i>Sometimes we had no access to a computer and a suitable room..., which delayed the process</i> ” [Primary health care psychologist #1]

## Discussion

### Main Findings

The implementation of TCP was feasible and supported PHC clinicians in diagnosing and/or treating children and adolescents with complex psychosocial care needs living under the supervision of CPS. The intervention was perceived to be clinically useful for the patient population, and it increased the resolution capacity of PHC clinicians and their willingness to receive a future TCP.

### Strengths and Limitations

To our knowledge, this is the first study to report a TCP connecting SMHC teams with PHC clinicians to improve mental health care among minors living under the supervision of CPS. A protocol for future implementations was developed. To improve the generalizability of the study results, larger and geographically more varied samples are needed to reflect the heterogeneous scenario of PHC and CPS in resource-poor Latin American contexts. A limitation of the study was that it did not directly assess mental health outcomes in the minors living under the supervision CPS. The common connectivity issues experienced during the study period were a further limitation, although they were fixed in a timely manner and usually did not affect the workflow of the MHC. Nevertheless, they may be important barriers to the adoption of TCPs in resource-constrained rural areas. Other important limitations of the implementation process included the lack of patient information and history in the EHRs, reflecting problems of sharing essential information among CPS, PHC, and SMHC.

### Comparison With Literature

In Chile, the use of technology-based solutions for mental health care of remote and underserved populations has been important due to the geography, and it has become increasingly relevant during the COVID-19 pandemic—with a rapid shift to this mode of mental health service delivery to ensure access and continuity of care [11]. For instance, partnerships between centralized SMHC teams and PHC clinicians through the implementation of call centers and EHRs have been reported to be acceptable and satisfactory for PHC clinicians as well as adults and adolescents with depression [12,20]. This study is in line with these experiences and provides initial evidence for the feasibility of telepsychiatry—in the form of videoconferencing MHC and EHRs—to support PHC clinicians in the mental health care of children and adolescents with complex psychosocial care needs.

Evidence from the United States has provided strong foundations for the feasibility and potential efficacy of community-based

child and adolescent telepsychiatry [13-15]. Digital health interventions, including videoconferencing MHC, have also been found to be satisfactory and acceptable for health care providers, caregivers, and youths in juvenile community-based justice settings [21-23]. Many of these youths came from CPS and have important behavioral health needs, yet telepsychiatry interventions specifically aimed at addressing the needs of CPS population and their providers in community settings are lacking.

Experiences reported from the United States resemble those reported in the present study in that they have applied technology-based solutions to build partnerships between SMHC and community or PHC providers, having included data from CPS population in their analyses [16,17]. These statewide initiatives in Washington [16] and Wyoming [17] considered centralized child psychiatric telephone and videoconferencing MHCs for health care providers of the population covered by Medicaid. In these studies, MHC were initiated by community or PHC providers in need of guidance [16,17]. These studies demonstrated higher satisfaction among those providers treating institutionalized children and adolescents, significant declines in high-dose pediatric psychotropic prescribing, as well as an increase in the use of outpatient and community-based treatments among this population.

Although our findings are consistent with these experiences [16,17], this feasibility study was specifically aimed at supporting the PHC treatment of children and adolescents living under the supervision of CPS. The TCP provided a fixed schedule for the linkage between SMHC teams and PHC clinicians within the context of a specific intersectoral mental health care plan, including CPS, PHC, and SMHC. These intersectoral alliances create an opportunity for further implementation, evaluation, and refinement of telepsychiatry.

### Concluding Remarks

Implementing a TCP was feasible and potentially useful for addressing mental health challenges in PHC of children and adolescents living under the supervision of CPS. Future programs may consider incorporating a child and adolescent psychologist to the SMHC team. To demonstrate the effectiveness of these interventions, a future clinical trial should assess the clinical outcomes in the children and adolescents living under the supervision of CPS. Furthermore, future research should consider scaling the TCP to more PHC clinics, including a larger patient population. Health care policies should be developed to provide regulation and resources for this model of mental health care, to be sustainable over time.

### Acknowledgments

The authors wish to thank the Municipal Corporation of Quilpué and its health care centers for their participation in this study. They also acknowledge the Psychiatry Department of El Pino Hospital, San Bernardo, Metropolitan Region Santiago, Chile, for allowing a child-adolescent psychiatrist to participate as a consultant. This study was conducted as part of the e-mental health network of the Universidad de Chile (VID U-REDES-C\_2018\_07). This study was supported by Funds for developing a telemedicine project at the Hospital Clínico Universidad de Chile. MI, PM, OF, VM, and GR received funding from the Agencia Nacional de Investigación y Desarrollo (ANID) Millennium Science Initiative /Millennium Institute for Research on Depression and

Personality-MIDAP (ICS13\_005). PM, VM, and GR received funding from ANID – Millennium Science Initiative Program (NCS17\_035). APM received funding from FONDECYT Regular 1190613, ANID.

## Conflicts of Interest

None declared.

## References

1. Bronsard G, Alessandrini M, Fond G, Loundou A, Auquier P, Tordjman S, et al. The prevalence of mental disorders among children and adolescents in the child welfare system: a systematic review and meta-analysis. *Medicine (Baltimore)* 2016 Feb;95(7):e2622 [FREE Full text] [doi: [10.1097/MD.0000000000002622](https://doi.org/10.1097/MD.0000000000002622)] [Medline: [26886603](https://pubmed.ncbi.nlm.nih.gov/26886603/)]
2. Comité de los Derechos del Niño. Informe de la investigación relacionada en Chile en virtud del artículo 13 del Protocolo facultativo de la Convención sobre los Derechos del Niño relativo a un procedimiento de comunicaciones. Webpage in Spanish.: Oficina del Alto Comisionado de Derechos Humanos de la Organización de las Naciones Unidas; 2018 Jun 1. URL: [https://www.ohchr.org/Documents/HRBodies/CRC/CRC\\_C\\_CHL\\_INQ\\_1.pdf](https://www.ohchr.org/Documents/HRBodies/CRC/CRC_C_CHL_INQ_1.pdf) [accessed 2021-07-19]
3. Minoletti A, Alvarado R, Rayo X, Minoletti M. Evaluación del sistema de salud mental en Chile. Informe sobre la base del Instrumento de evaluación del sistema de salud mental de OMS (OMS IESM/WHO AIMS). Report in Spanish. Santiago, Chile: Ministerio de Salud; 2014. URL: [https://www.who.int/mental\\_health/who\\_aims\\_country\\_reports/who\\_aims\\_report\\_chile.pdf](https://www.who.int/mental_health/who_aims_country_reports/who_aims_report_chile.pdf) [accessed 2021-07-19]
4. Saldivia S, Vicente B, Kohn R, Rioseco P, Torres S. Use of mental health services in Chile. *Psychiatr Serv* 2004 Jan;55(1):71-76. [doi: [10.1176/appi.ps.55.1.71](https://doi.org/10.1176/appi.ps.55.1.71)] [Medline: [14699204](https://pubmed.ncbi.nlm.nih.gov/14699204/)]
5. Valenzuela Azócar C, Santander Cortéz X, Véliz Rojas V, Cordero Rojas A, De Ferrari Fontecilla I. Informe de Gestión SENAME de la Subsecretaría de Redes Asistenciales Años 2016-2018. Report in Spanish. Santiago, Chile: Ministerio de Salud; 2018 Feb. URL: <https://www.minsal.cl/wp-content/uploads/2018/03/Informe-de-Gesti%C3%B3n-SENAME-de-la-Subsecretar%C3%ADa-de-Redes-Asistenciales-A%C3%B1os-2016-2018.pdf> [accessed 2021-07-19]
6. Gardner J, Plaven BE, Yellowlees P, Shore J. Remote telepsychiatry workforce: a solution to psychiatry's workforce issues. *Curr Psychiatry Rep* 2020 Jan 27;22(2):8. [doi: [10.1007/s11920-020-1128-7](https://doi.org/10.1007/s11920-020-1128-7)] [Medline: [31989417](https://pubmed.ncbi.nlm.nih.gov/31989417/)]
7. Myers CR. Using telehealth to remediate rural mental health and healthcare disparities. *Issues Ment Health Nurs* 2019 Mar;40(3):233-239. [doi: [10.1080/01612840.2018.1499157](https://doi.org/10.1080/01612840.2018.1499157)] [Medline: [30508400](https://pubmed.ncbi.nlm.nih.gov/30508400/)]
8. Hilty DM, Sunderji N, Suo S, Chan S, McCarron RM. Telepsychiatry and other technologies for integrated care: evidence base, best practice models and competencies. *Int Rev Psychiatry* 2018 Dec;30(6):292-309. [doi: [10.1080/09540261.2019.1571483](https://doi.org/10.1080/09540261.2019.1571483)] [Medline: [30821540](https://pubmed.ncbi.nlm.nih.gov/30821540/)]
9. Hilty DM, Rabinowitz T, McCarron RM, Katzelnick DJ, Chang T, Bauer AM, et al. An update on telepsychiatry and how it can leverage collaborative, stepped, and integrated services to primary care. *Psychosomatics* 2018;59(3):227-250. [doi: [10.1016/j.psych.2017.12.005](https://doi.org/10.1016/j.psych.2017.12.005)] [Medline: [29544663](https://pubmed.ncbi.nlm.nih.gov/29544663/)]
10. Kinoshita S, Cortright K, Crawford A, Mizuno Y, Yoshida K, Hilty D, et al. Changes in telepsychiatry regulations during the COVID-19 pandemic: 17 countries and regions' approaches to an evolving healthcare landscape. *Psychol Med* 2020 Nov 27:1-8. [doi: [10.1017/s0033291720004584](https://doi.org/10.1017/s0033291720004584)]
11. Fischman P, Irrarrazaval M. Debate: Mental health, social crisis and the COVID-19 pandemic in Chile. *Child Adolesc Ment Health* 2020 Nov;25(4):256-257 [FREE Full text] [doi: [10.1111/camh.12424](https://doi.org/10.1111/camh.12424)] [Medline: [33063384](https://pubmed.ncbi.nlm.nih.gov/33063384/)]
12. Martínez V, Rojas G, Martínez P, Zitko P, Irrarrazaval M, Luttgés C, et al. Remote collaborative depression care program for adolescents in Araucanía Region, Chile: randomized controlled trial. *J Med Internet Res* 2018 Jan 31;20(1):e38 [FREE Full text] [doi: [10.2196/jmir.8021](https://doi.org/10.2196/jmir.8021)] [Medline: [29386172](https://pubmed.ncbi.nlm.nih.gov/29386172/)]
13. American Academy of Child Adolescent Psychiatry (AACAP) Committee on Telepsychiatry and AACAP Committee on Quality Issues. Clinical Update: Telepsychiatry With Children and Adolescents. *J Am Acad Child Adolesc Psychiatry* 2017 Oct;56(10):875-893. [doi: [10.1016/j.jaac.2017.07.008](https://doi.org/10.1016/j.jaac.2017.07.008)] [Medline: [28942810](https://pubmed.ncbi.nlm.nih.gov/28942810/)]
14. Myers K, Nelson E, Rabinowitz T, Hilty D, Baker D, Barnwell SS, et al. American telemedicine association practice guidelines for telemental health with children and adolescents. *Telemed J E Health* 2017 Oct;23(10):779-804. [doi: [10.1089/tmj.2017.0177](https://doi.org/10.1089/tmj.2017.0177)] [Medline: [28930496](https://pubmed.ncbi.nlm.nih.gov/28930496/)]
15. Marcus S, Malas N, Dopp R, Quigley J, Kramer AC, Tengelitsch E, et al. The Michigan Child Collaborative Care Program: building a telepsychiatry consultation service. *Psychiatr Serv* 2019 Sep 01;70(9):849-852. [doi: [10.1176/appi.ps.201800151](https://doi.org/10.1176/appi.ps.201800151)] [Medline: [31272335](https://pubmed.ncbi.nlm.nih.gov/31272335/)]
16. Hilt R, Romaine MA, McDonnell M, Sears J, Krupski A, Thompson J, et al. The Partnership Access Line: evaluating a child psychiatry consult program in Washington State. *JAMA Pediatr* 2013 Feb;167(2):162-168. [doi: [10.1001/2013.jamapediatrics.47](https://doi.org/10.1001/2013.jamapediatrics.47)] [Medline: [23247331](https://pubmed.ncbi.nlm.nih.gov/23247331/)]
17. Hilt R, Barclay R, Bush J, Stout B, Anderson N, Wignall J. A statewide child telepsychiatry consult system yields desired health system changes and savings. *Telemed J E Health* 2015 Jul;21(7):533-537. [doi: [10.1089/tmj.2014.0161](https://doi.org/10.1089/tmj.2014.0161)] [Medline: [25799043](https://pubmed.ncbi.nlm.nih.gov/25799043/)]



18. Orientaciones Técnicas: Consultorías en Salud Mental. Report in Spanish. Santiago, Chile: Ministerio de Salud; 2016. URL: <https://www.minsal.cl/wp-content/uploads/2015/09/OT-CONSULTORIAS-EN-SALUD-MENTAL.pdf> [accessed 2021-07-19]
19. Strauss A, Corbin J. Bases de la investigación cualitativa. Técnicas y procedimientos para desarrollar la teoría fundamentada. Book in Spanish. Medellín, Colombia: Universidad de Antioquia; 2016.
20. Rojas G, Guajardo V, Martínez P, Castro A, Fritsch R, Moessner M, et al. A remote collaborative care program for patients with depression living in rural areas: open-label trial. *J Med Internet Res* 2018 Apr 30;20(4):e158 [FREE Full text] [doi: [10.2196/jmir.8803](https://doi.org/10.2196/jmir.8803)] [Medline: [29712627](https://pubmed.ncbi.nlm.nih.gov/29712627/)]
21. Batastini AB. Improving rehabilitative efforts for juvenile offenders through the use of telemental healthcare. *J Child Adolesc Psychopharmacol* 2016 Apr;26(3):273-277. [doi: [10.1089/cap.2015.0011](https://doi.org/10.1089/cap.2015.0011)] [Medline: [26288198](https://pubmed.ncbi.nlm.nih.gov/26288198/)]
22. Tolou-Shams M, Yonek J, Galbraith K, Bath E. Text messaging to enhance behavioral health treatment engagement among justice-involved youth: qualitative and user testing study. *JMIR Mhealth Uhealth* 2019 Apr 05;7(4):e10904 [FREE Full text] [doi: [10.2196/10904](https://doi.org/10.2196/10904)] [Medline: [30950808](https://pubmed.ncbi.nlm.nih.gov/30950808/)]
23. Folk J, Harrison A, Rodriguez C, Wallace A, Tolou-Shams M. Feasibility of social media-based recruitment and perceived acceptability of digital health interventions for caregivers of justice-involved youth: mixed methods study. *J Med Internet Res* 2020 Apr 30;22(4):e16370 [FREE Full text] [doi: [10.2196/16370](https://doi.org/10.2196/16370)] [Medline: [32352388](https://pubmed.ncbi.nlm.nih.gov/32352388/)]

## Abbreviations

**CPS:** child protective services  
**EHR:** electronic health record  
**MHC:** mental health consultation  
**PHC:** primary health care  
**SMHC:** specialized mental health care  
**TCP:** telepsychiatry consultation program

*Edited by A Lansky; submitted 19.11.20; peer-reviewed by J Torales, L Seuren, C Rea; comments to author 03.01.21; revised version received 24.02.21; accepted 11.04.21; published 22.07.21.*

### *Please cite as:*

Mundt AP, Irarrázaval M, Martínez P, Fernández O, Martínez V, Rojas G  
*Telepsychiatry Consultation for Primary Care Treatment of Children and Adolescents Receiving Child Protective Services in Chile: Mixed Methods Feasibility Study*  
*JMIR Public Health Surveill* 2021;7(7):e25836  
URL: <https://publichealth.jmir.org/2021/7/e25836>  
doi: [10.2196/25836](https://doi.org/10.2196/25836)  
PMID: [34292164](https://pubmed.ncbi.nlm.nih.gov/34292164/)

©Adrian P Mundt, Matías Irarrázaval, Pablo Martínez, Olga Fernández, Vania Martínez, Graciela Rojas. Originally published in *JMIR Public Health and Surveillance* (<https://publichealth.jmir.org>), 22.07.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Public Health and Surveillance*, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

# Availability, Formulation, Labeling, and Price of Low-sodium Salt Worldwide: Environmental Scan

Xuejun Yin<sup>1</sup>, BA, MPH; Hueiming Liu<sup>1,2</sup>, BA, MPH, MBBS, PhD; Jacqui Webster<sup>1</sup>, PhD, RPHNutri; Kathy Trieu<sup>1</sup>, BAppSc, MPH, PhD; Mark D Huffman<sup>1,3</sup>, MPH, MD; J Jaime Miranda<sup>1,4,5</sup>, MSc, MD, PhD, FFPH; Matti Marklund<sup>1,6,7</sup>, MSE, PhD; Jason H Y Wu<sup>1</sup>, PhD, FAHA; Laura K Cobb<sup>8</sup>, MS, DrPH; Ka Chun Li<sup>1</sup>, MSc; Sallie-Anne Pearson<sup>9</sup>, BSc (Hons), PhD; Bruce Neal<sup>1,10</sup>, ChB, MB, PhD, FAHMS, FAHA, FRCP; Maoyi Tian<sup>1,11</sup>, MSc, PhD

<sup>1</sup>The George Institute for Global Health, University of New South Wales, Sydney, Australia

<sup>2</sup>Sydney Institute for Women, Children and their Families, Sydney Local Health District, Sydney, Australia

<sup>3</sup>Feinberg School of Medicine, Northwestern University, Evanston, IL, United States

<sup>4</sup>CRONICAS Centre of Excellence in Chronic Diseases, Universidad Peruana Cayetano Heredia, Lima, Peru

<sup>5</sup>Department of Medicine, School of Medicine, Universidad Peruana Cayetano Heredia, Lima, Peru

<sup>6</sup>Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

<sup>7</sup>Department of Public Health and Caring Sciences, Uppsala University, Uppsala, Sweden

<sup>8</sup>Resolve to Save Lives, Initiative of Vital Strategies, New York City, NY, United States

<sup>9</sup>Centre for Big Data Research in Health, University of New South Wales, Sydney, Australia

<sup>10</sup>School of Public Health, Imperial College London, London, United Kingdom

<sup>11</sup>The George Institute for Global Health at Peking University Health Science Center, Beijing, China

**Corresponding Author:**

Maoyi Tian, MSc, PhD

The George Institute for Global Health

University of New South Wales

Level 5, 1 King St, Newtown NSW

Sydney, 2042

Australia

Phone: 61 18610501623

Email: [mtian@georgeinstitute.org.cn](mailto:mtian@georgeinstitute.org.cn)

## Abstract

**Background:** Regular salt is about 100% sodium chloride. Low-sodium salts have reduced sodium chloride content, most commonly through substitution with potassium chloride. Low-sodium salts have a potential role in reducing the population's sodium intake levels and blood pressure, but their availability in the global market is unknown.

**Objective:** The aim of this study is to assess the availability, formulation, labeling, and price of low-sodium salts currently available to consumers worldwide.

**Methods:** Low-sodium salts were identified through a systematic literature review, Google search, online shopping site searches, and inquiry of key informants. The keywords “salt substitute,” “low-sodium salt,” “potassium salt,” “mineral salt,” and “sodium reduced salt” in six official languages of the United Nations were used for the search. Information about the brand, formula, labeling, and price was extracted and analyzed.

**Results:** A total of 87 low-sodium salts were available in 47 out of 195 (24%) countries worldwide, including 28 high-income countries, 13 upper-middle-income countries, and 6 lower-middle-income countries. The proportion of sodium chloride varied from 0% (sodium-free) to 88% (as percent of weight; regular salt is 100% sodium chloride). Potassium chloride was the most frequent component with levels ranging from 0% to 100% (potassium chloride salt). A total of 43 (49%) low-sodium salts had labels with the potential health risks, and 33 (38%) had labels with the potential health benefits. The median price of low-sodium salts in high-income, upper-middle-income, and lower-middle-income countries was US \$15.00/kg (IQR 6.4-22.5), US \$2.70/kg (IQR 1.7-5.5), and US \$2.90/kg (IQR 0.50-22.2), respectively. The price of low-sodium salts was between 1.1 and 14.6 times that of regular salts.

**Conclusions:** Low-sodium salts are not widely available and are commonly more expensive than regular salts. Policies that promote the availability, affordability, and labeling of low-sodium salts should increase uptake, helping populations reduce blood pressure and prevent cardiovascular diseases.

**International Registered Report Identifier (IRRID):** RR2-10.1111/jch.14054

(*JMIR Public Health Surveill* 2021;7(7):e27423) doi:[10.2196/27423](https://doi.org/10.2196/27423)

## KEYWORDS

low-sodium salt; salt substitute; availability; formulation; labeling; price; sodium; salt; blood pressure; cardiology

## Introduction

Cardiovascular diseases (CVDs) are the leading causes of death worldwide, and high blood pressure is a leading risk factor for CVDs [1,2]. Dietary sodium intake is a strong causal determinant of blood pressure levels [3,4]. The World Health Organization (WHO) recommends reducing sodium intake as one of the *best-buy* strategies for lowering the population-level risk of CVDs and set a goal of 30% reduction in the population [5]. Different strategies are proposed to reduce dietary sodium intake. Although the effect of sodium reduction interventions on blood pressure has been demonstrated in controlled research settings, there are few examples of sodium reduction at a population scale [6,7]. Innovative strategies are needed to reduce sodium intake from processed foods, food eaten outside the home, and home cooked foods.

Regular salt is about 100% sodium chloride. Low-sodium salt has reduced sodium chloride content, most commonly through substitution with potassium chloride or magnesium sulphate. There is a growing body of evidence supporting the use of low-sodium salts as an effective intervention to reduce dietary sodium intake, lower blood pressure, and thereby prevent the adverse consequences of high blood pressure [8,9]. Although one concern with low-sodium salts enriched with potassium was the potential risk for people who have advanced chronic kidney disease (CKD) due to hyperkalemia, a comparative risk assessment model estimated that nationwide replacement of the salt supply with potassium-enriched, low-sodium salt in China could result in the prevention of 1 in every 9 deaths from CVD and a net benefit with the use of low-sodium salts even in individuals with CKD [10]. The approach of integrating low-sodium salts as a public health intervention can potentially reduce dietary sodium intake at the population level through reformulating manufactured foods with low-sodium salts, replacing cooking salt used at home or in restaurants with low-sodium salts, or a combination thereof. Thus, population uptake of low-sodium salts has been recommended as one of seven priority strategies to reduce population sodium consumption [11].

Low-sodium salts may have a potential role in enhancing equitable access to effective CVD prevention worldwide as long as it is affordable and has sufficient reach to countries regardless of the income levels. However, little is known about the availability and accessibility of this emerging product in the global market or about factors that may affect equitable uptake including formulation, price, or labeling. We, therefore, performed a systematic search of low-sodium salts to understand

better their availability, formulation, labeling, and the price in different countries.

## Methods

This study was a systematic search of low-sodium salts conducted from October 2019 to September 2020. The study protocol has been published previously [12].

### Definition of Low-sodium Salt

Low-sodium salts were defined as table salt or cooking salt that replaced sodium chloride content with other minerals such as potassium chloride or magnesium sulphate. The sodium content and terminology for low-sodium salts can vary. For example, in some cases, the term *salt substitute* is used as a synonym for low-sodium salt. In this study, we use the term *low-sodium salts* as a category, including both sodium-reduced and sodium-free salts. All low-sodium salts available for retail purchase were eligible for inclusion, but those that had ceased production were excluded.

### Search Strategy

Low-sodium salts were identified from four sources. First, a systematic literature search was conducted in MEDLINE, Embase, and Cochrane Library from inception through March 2020 without language restrictions. The search strategies are listed in [Multimedia Appendix 1](#) and include the keywords of “salt substitute,” “low sodium salt,” “potassium salt,” “mineral salt,” and “sodium reduced salt.” Second, we searched major global online shopping sites, including Amazon, eBay, Walmart, JD, and RedMart, to identify low-sodium salts. Third, we executed a search using Google advanced engine with a Google Chrome browser from Australia. Initial keywords included different terms describing low-sodium salts. The search strategies are shown in [Textbox 1](#). Initial keywords and country names were combined for a Google search to identify the availability of low-sodium salts in different countries. The first 25 results of each search were examined for eligibility. This was primarily because the search results after the first 25 were generally unrelated to low-sodium salts. Six official languages of the United Nations (Arabic, Chinese, English, French, Russian, and Spanish) were used in the Google searches and online shopping website searches considering that the language would influence the ranking of resulting pages. Google searches were conducted using a Google Chrome browser. The translation from other languages to English was performed by the built-in translation service within the Google Chrome browser [13]. Fourth, we conducted semistructured interviews with key informants who have been involved in the research,

manufacturing, implementation, and promotion of low-sodium salts. We used purposive and snowball sampling to recruit key informants. Academic representatives were identified through the systematic review, which was part of the broader environmental scan. Corresponding authors of eligible studies that were identified from the systematic review were contacted via corresponding authors' email addresses. Academic

representatives who participated in the interview were asked to refer the study to potential key informants from relevant government agencies or the salt manufacturing industry through a snowball sampling strategy. A total of 18 key informants from 9 countries representing all WHO regions provided information about the low-sodium salts they knew. Results from four data sources were combined, and duplicates were then removed.

**Textbox 1.** Keywords and terms in Google search strategy.

#### Initial keywords

- “low-sodium salt,” “salt substitute,” “potassium salt,” “mineral salt,” “sodium reduced salt”

#### Terms of different aspects relevant to the salt substitutes

- Formula: “formula,” “composition,” “ingredient”
- Price: “price,” “cost”
- Country: for example, “Finland,” “China,” “Italy”
- Safety issue: “safety,” “adverse events,” “warning,” “danger,” “threatening”

To identify the low-sodium salts used in research studies, we had broad inclusion criteria on the study population and study types. All titles and abstracts of records identified in the databases were screened for eligibility. Duplicated studies were excluded. Full texts of papers after eligibility screening were obtained for independent review by two authors (XY and KCL). Differences between reviewers were resolved through discussion or consultation with a third senior investigator (MT) when necessary. Information about the low-sodium salts was extracted. Low-sodium salts were excluded if they were manufactured exclusively for a study but not available for a general population or ceased production and were no longer available in the market.

#### Data Extraction and Analysis

Data were extracted by one researcher (XY) to a predesigned Microsoft Excel (Microsoft Corporation) sheet, including data source, search date, product website, brand, formulation, nutrition facts, safety warnings, and price. The price of regular salt in the same brand was also extracted from online shopping websites for price comparison. Information on the low-sodium salt was summarized by country. Price was converted into US dollars per kilogram (US \$/kg) using the exchange rate at the time of the search. The international dollar of low-sodium salts

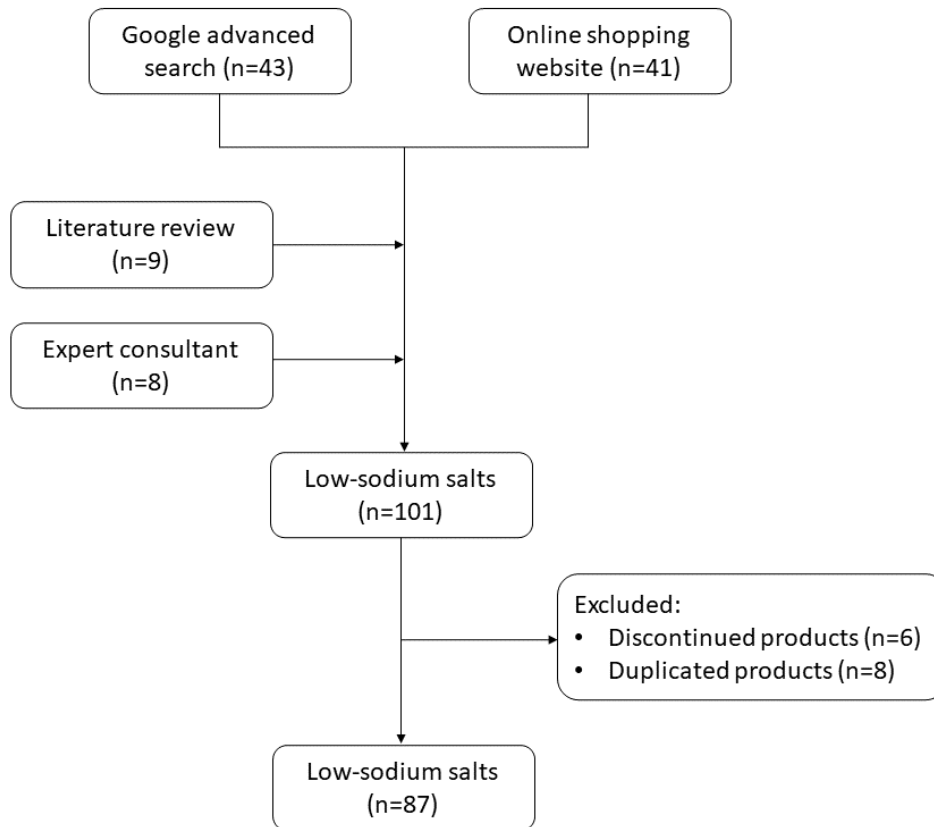
by the exchange rate in the year of 2020 was also reported for considering purchasing power parities and the commodity prices [14]. The average price of low-sodium salts was compared by country income levels. The income levels were divided according to the World Bank classification [15]. Statistical analyses were performed using Microsoft Excel and SPSS (IBM Corp, Version 26).

## Results

### Summary of Identified Low-sodium Salts

Our final analyses comprised 87 low-sodium salts that were identified across 47 countries. Google searches identified 43 products. Another 41 products were identified from online shopping sites. From the published research studies, 4440 abstracts were screened, 67 of them were related to low-sodium salts, and 9 brands of low-sodium salts were identified from full-text papers. A total of 8 low-sodium salts were complemented from consulting experts. There were 8 duplicated products and 6 discontinued products, which were removed from the final analysis (Figure 1). Key characteristics of all included low-sodium salts are described in [Multimedia Appendix 2](#).

**Figure 1.** Flowchart of low-sodium salt identification.

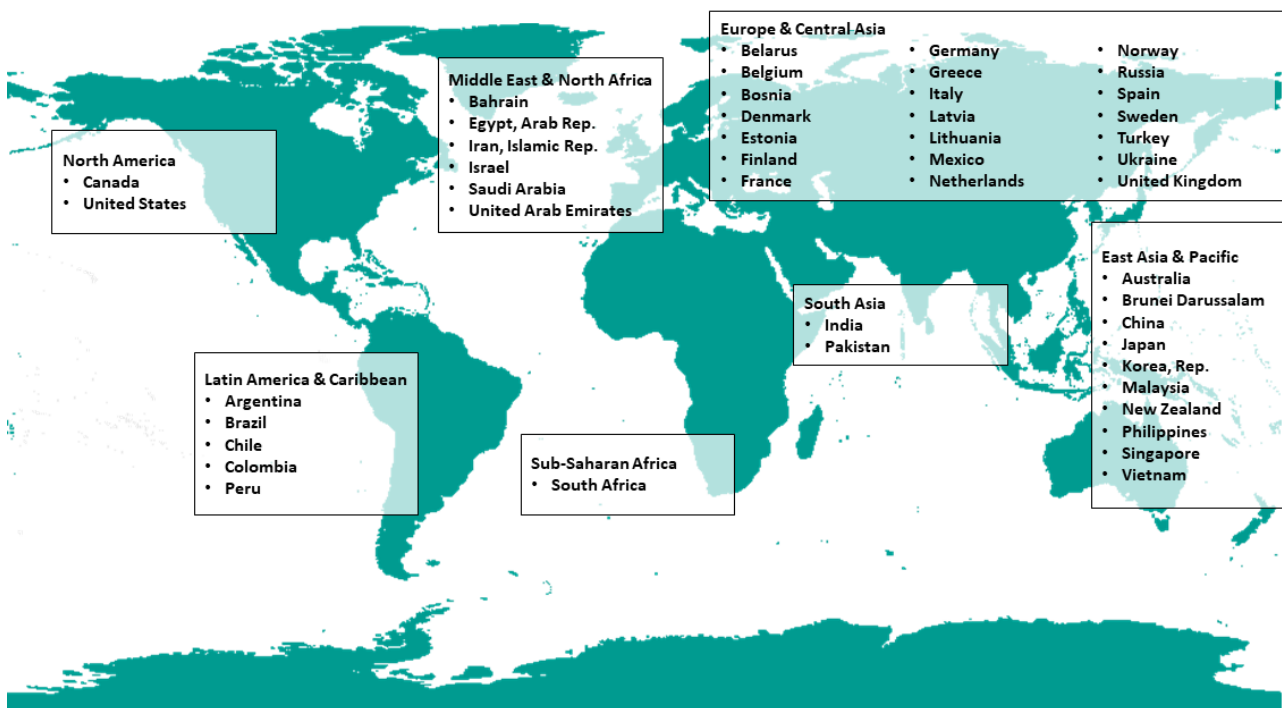


**Availability of Low-sodium Salts**

Low-sodium salts were available in 47 countries, including 28 high-income countries, 13 upper-middle-income countries, and 6 lower-middle-income countries. The United States, Canada, the United Kingdom, Italy, China, India, France, Sweden,

Argentina, and Russia produced more than 1 brand of low-sodium salt. Low-sodium salts were found in all seven world regions defined by the World Bank. Most countries with low-sodium salts were located in the Europe and Central Asia countries (n=21) and East Asia and Pacific countries (n=10; [Figure 2](#)).

**Figure 2.** Availability of low-sodium salt in different countries.



### Formulation and Labeling

Among low-sodium salts, the sodium chloride varied from 0% (sodium-free) to 88% (as percent of weight, regular salt is 100% sodium chloride). Potassium chloride varied from 0% to 100%. A total of 51 low-sodium salts reported both sodium and potassium levels whose composition are shown in Figure 3. The

composition of low-sodium salts varied by different countries. For example, less than 20% of the sodium chloride was replaced in low-sodium salts produced in India, while at least 50% of the sodium chloride was replaced in countries in North America, the Middle East, and Latin America. Sodium-free salt substitutes were mostly available in the United State and Canada (n=7).

**Figure 3.** The proportions of sodium chloride (NaCl), potassium chloride (KCl), and other minerals for low-sodium salts by regions (N=51).



\* Pansalt was available in multiple countries

Among the 87 products, 41 (46%) were iodized, 46 (53%) had nutrition facts information on their labels, and 55 (63%) had a front of package labeling indicating less sodium in low-sodium salt than regular salt. A total of 43 (49%) products had labels advising potential health risks, 33 (38%) had labels advising potential health benefits, 16 (18%) had labels with both, and 27 (31%) had labels with neither. The labels advising potential health risk was signed to warn high-risk populations who might

potentially be harmed by excessive potassium intake, but the contents of these advisory labels varied. Examples of advisory warning labels on low-sodium salt packages are shown in Textbox 2. The labels advising potential health benefits included lowering sodium intake without compromising taste, being additional sources of potassium, and lowering blood pressure. Examples of health benefit labels on low-sodium salt packages are shown in Textbox 3.

**Textbox 2.** Examples of advisory warning labels found on low-sodium salt packages.

- “A good source of potassium. Should not be used by persons on a sodium or potassium restricted diet unless approved by a physician.” (United States, Morton and Canada, Windsor)
- “For normal healthy people. Persons having diabetes, heart or kidney disease, or receiving medical treatment should consult a physician before using a salt alternative or substitute.” (United States, Nu-salt)
- “It should be used with caution among people who are not suitable for high potassium intakes, such as high-temperature workers, heavy-labour workers, renal dysfunction and hypertension patients taking antihypertensive drugs.” (China, China Salt)
- “It is recommended to use the product under medical supervision.” (Italy, Novo sal)
- “Medical advice should be sought for diet requiring low or restricted sodium or potassium intake. Not suitable for use with a certain diuretic.” (Finland, Pansalt)
- “In case of kidney or heart failure or high blood pressure, you should consult your doctor before use.” (Russia, Mediterra salt)

**Textbox 3.** Examples of health benefit labels found on low-sodium salt packages.

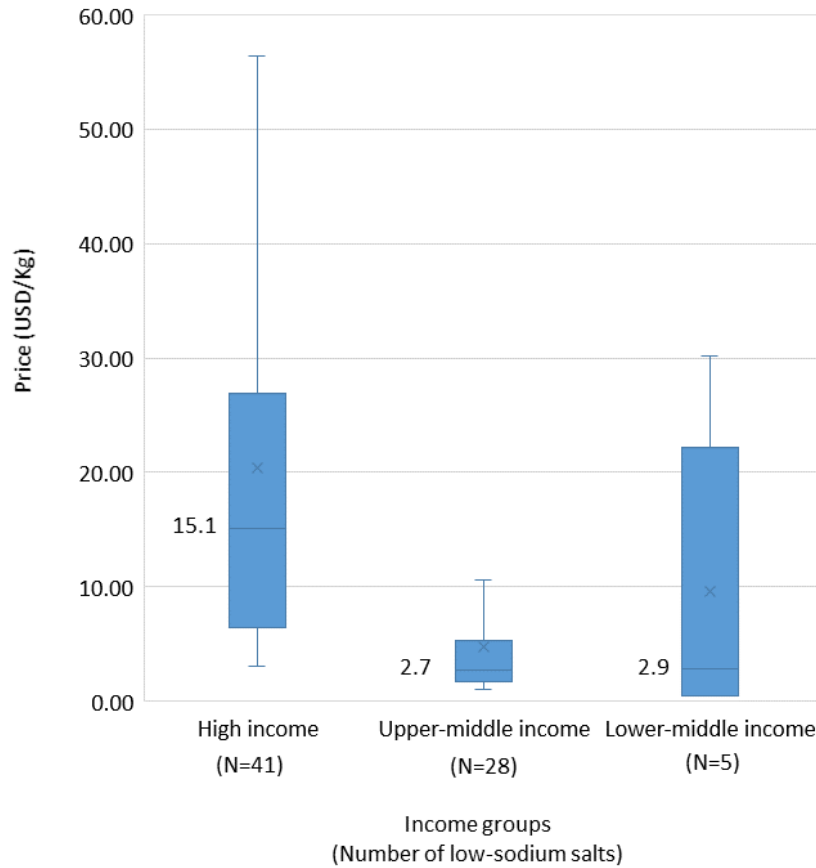
- “Pansalt is a clinically-proven low-sodium salt with added essential minerals for a healthy heart. The low sodium mineral salt is the result of many years of intensive scientific research and is ideal for consumers who want to follow a healthier diet without compromising on good taste. Apart from the overall health benefits. Pansalt delivers the true taste of salt but without the harmful effects resulting in a high consumer approval rate.” (Finland, Pansalt)
- “TATA salt lite has been specially formulated to provide 15% lower sodium than ordinary salt. It is generally accepted that lower sodium in diets may assist in management of high blood pressure.” (India, TATA)
- “It is being a good source of potassium; helps you maintain normal blood pressure.” (United Kingdom, Lo salt)
- “Good for your health – potassium and a reduced consumption of sodium contribute to the maintenance of normal blood pressure. Potassium further contributes to the normal functioning of the nervous system and muscles.” (Netherlands, Nezo Light Salt)
- “Replace part of sodium chloride with potassium chloride. The taste is like regular salt. It can help reduce the intake of sodium and increase potassium intake.” (China, Yi Yan Tang)
- “It can significantly reduce excessive sodium intake and increase potassium and magnesium intake. Potassium is necessary to maintain the work of the heart, nervous system and muscle system, clear excess fluid from the body, and maintain normal blood pressure” (Russia, Valetex Prodimpex)

**Price of Low-sodium Salts**

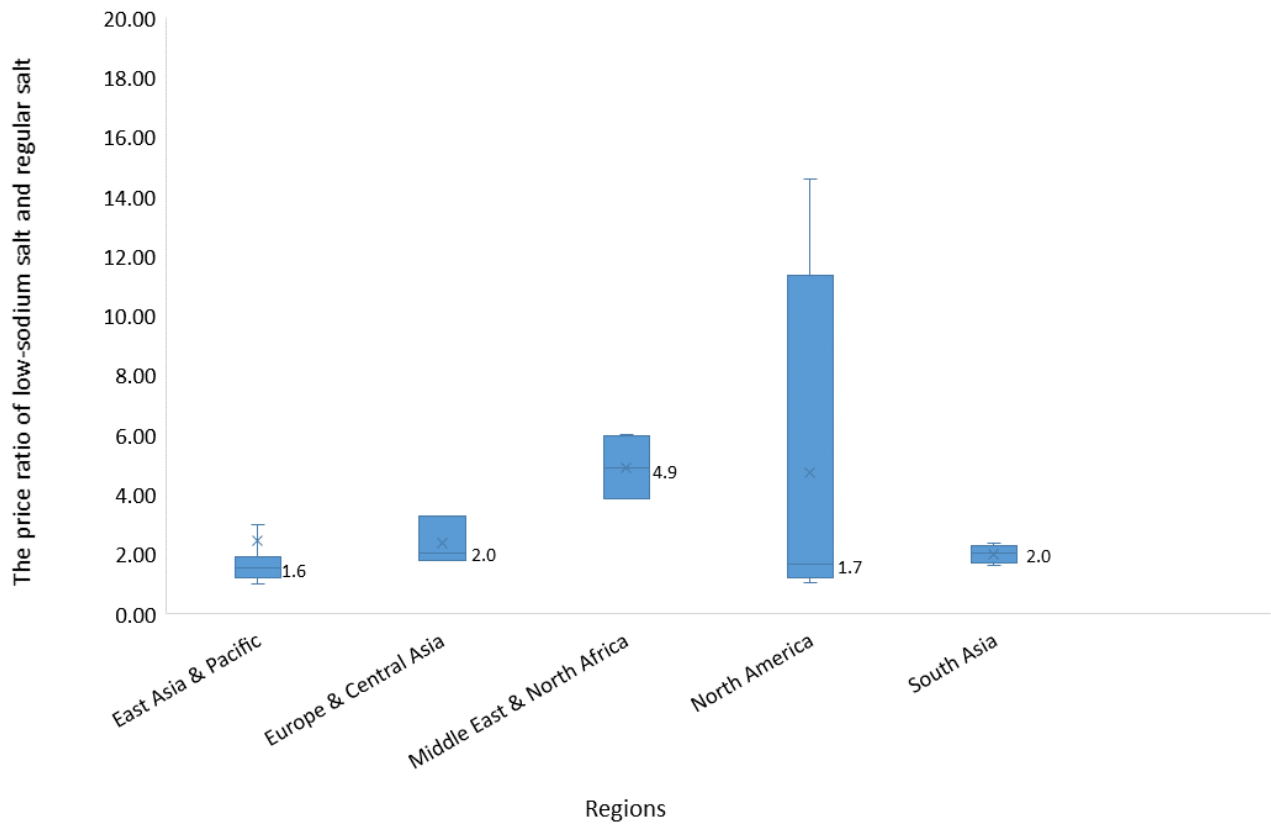
The price of low-sodium salts varied from US \$0.46/kg to US \$87.00/kg. The median price of low-sodium salts in high-income, upper-middle-income, and lower-middle-income countries was US \$15.10/kg (IQR 6.4-26.9), US \$2.70/kg (IQR 1.7-5.3), and US \$2.90/kg (IQR 0.5-22.2), respectively (Figure 4). After converting to the international dollar, the median international dollar of low-sodium salts in high-income and

upper-middle-income was US \$14.80/kg (IQR 8.5-32.9) and US \$11.60/kg (IQR 7.4-18.4), respectively. The median international dollar of low-sodium salts in lower-middle-income countries surged to US \$111.10/kg (IQR 11.0-469.9). Among salt manufacturers producing both low-sodium salts and regular salts (N=38), the price of low-sodium salts was 1.7 times the price of the regular salts. The price difference between low-sodium salt and regular salt by regions is presented in Figure 5.

**Figure 4.** Price for low-sodium salts across country income groups.



**Figure 5.** Price differential among salt manufacturers producing both low-sodium salts and regular salts by different regions.





## Discussion

The findings of this study indicated that low-sodium salts were only available in 24% of the countries worldwide. Noticeably, 60% of these countries were high-income countries. Prices varied between US \$0.46/kg and US \$87.00/kg but were consistently higher than regular salts. Sodium chloride content ranged from 0% to 88%. Advisory labels on benefits and harms of using low-sodium salts were not standardized.

The availability of low-sodium salts was limited in low- and middle-income countries. Many low- and middle-income countries are facing disproportionately large burdens of CVD due to unhealthy diet [2,16]. However, there are few effective and affordable strategies to combat the increasing burden in resource-constrained settings. Salt reduction was listed as a *best-buy* intervention by the WHO, as it is an effective intervention in low- and middle-income countries where the primary source of sodium is from salt added in cooking and food preparation, and the proportion of discretionary salt to total sodium intake is higher [17,18]. A low-sodium salt population intervention could potentially enhance equity of access to effective CVD prevention. Important next steps include the increase in production and market demand of low-sodium salts. One example of this was in 2010, the Beijing city government strengthened the supply chain to ensure that low-sodium salts were available in 28 different supermarket chains as part of a citywide salt reduction initiative [19]. Low-sodium salt is a simple and low-cost intervention, which makes it a compelling proposal for low- and middle-income countries. In high-income countries, where CVD is also a major burden and low-sodium salts are generally easy to access, low-sodium salts can serve as a complementary strategy for product reformulation and an ideal replacement added during home cooking or at the table. The approach to integrating low-sodium salts in public health interventions may vary by country depending on the major dietary source of sodium. Feasibility studies in countries with different diet patterns are needed to understand better how to introduce low-sodium salts onto the market and how to promote low-sodium salt use in different contexts.

Low-sodium salts have higher prices than regular salts in all countries where low-sodium salts are commercially available. The price of food is a vital factor that influences consumers' behavior [20]. One cluster randomized control trial revealed that even with the same health education program, the adoption of low-sodium salts was higher in villages where the price of low-sodium salts was subsidized compared with villages without a price subsidy [21]. Results of this study showed that the prices of low-sodium salts were 1.0 to 14.6 times higher than the price of regular salts. Price differences were highest in North America, likely because low-sodium salts produced in this area had a larger proportion of potassium chloride (potassium costs about four times more than sodium) [22]. Some countries, like India, do not have readily available sources of potassium, which may further increase costs of potassium supplementation [23]. Narrowing the price difference between low-sodium salts and regular salts would promote low-sodium salt use, which can motivate manufacturers to improve low-sodium salt availability. Government subsidies and regulation can increase the

affordability and reduce the price differences between low-sodium salts and regular salts, especially among individuals with limited resources who may face a disproportionate risk and burden of CVD. The low-sodium salt intervention ought to be implemented by reformulation or policies without passing the costs onto individuals who will be less likely to afford these higher prices. One real-world example of subsidies was in Beijing, where in addition to improving the supply chain, an additional 75g of low-sodium salt was provided in the 400g package [19]. Further quantitative market surveys and qualitative studies can be designed to understand the context and mechanism of the price of low-sodium salts in the future.

Low-sodium salts varied in composition. Overall, the sodium chloride was mostly replaced by potassium chloride. The acceptability of low-sodium salt taste depends on individual preferences (palate) and concentration of potassium [24]. There have been reports that unpleasant off-tastes associated with potassium chloride may prevent customers from choosing low-sodium salts [25]. However, several studies found low-sodium salts with about 30% of sodium chloride replaced by potassium chloride had a similar flavor to regular salts when used in cooking or processed food manufacturing [24,26,27]. Reducing sodium and increasing potassium in the salt used for cooking and processed food reformulation could make the salt less *salty* in taste, leading consumers to use more low-sodium salts. Effective taste-improving agents can be introduced to overcome potential sensory drawbacks, such as taste masking or umami ingredients [28]. The acceptable range of preferences for compositions of low-sodium salts used in cooking by the general population in different culture context would need to be confirmed in future research. In addition to taste acceptability, further studies regarding the benefits and risks of low-sodium salts with different sodium and potassium proportion are still necessary.

Food labeling was designed to provide information and help consumers make healthier food choices. The effect of low-sodium salts on reducing sodium intake and blood pressure has been proven in clinical trials [8,9]. The modeling study showed that the net benefits of a national implementation of low-sodium salts were substantial in preventing CVDs, and the net benefit still presented for individuals with CKD [10]. Only 33 brands of low-sodium salts were identified as having health benefit labels. Evidence-based benefits of low-sodium salt should be reported on the package to encourage people to make healthier choices in salt. Despite limited data on the dose-response relationship between the use of salt substitutes and serum potassium levels, one concern with scaling up low-sodium salts is the potential to increase the risk of hyperkalemia and sudden cardiac arrest for a small number of people who are advised to limit dietary potassium due to impaired renal function and potassium excretion [29]. Although there is little research investigating the relative benefits and risks of salt substitute use in such populations, people who have advanced CKD or who take medications that interfere with the renin-angiotensin-aldosterone axis, including potassium-sparing diuretics, may be suboptimal candidates for low-sodium salts that are enriched with potassium. Precautionary labeling on the potassium level and health warning information are important

in response to these concerns, but the optimal threshold and format of such labeling are uncertain. For labels on low-sodium salts to be useful, they must be both credible and accurate. There were 43 low-sodium salts with labels advising potential risks, but 27 of them did not mention the health benefits. An overcautious label not based on scientific evidence or labels only focused on health risks might discourage people who would receive health benefits by using the low-sodium salts. The knowledge and awareness of low-sodium salts is generally low in recent baseline surveys done for the trials with low-sodium salts in China and India [30,31]. A survey conducted in India to assess the knowledge and awareness of low-sodium salts among doctors indicated 71.5% of participants did not know about their contraindications in patients with severe renal disease, cardiac problems, and patients on potassium-sparing diuretics [32]. Adequate education campaigns together with regulations on the labeling are imperative to ensure that regulators, the salt industry, food scientists, clinicians, and consumers are aware of the level of potassium additives and how to read the labeling information appropriately.

One strength of this study lies in the careful searches across diverse data sources to describe the landscape of low-sodium salts. We complemented conventional systematic review methods with online searches and key informant consultation to identify as many products as possible. However, results from Google searches are dynamic and can be difficult to replicate, and thus, such a method can only reflect the landscape for a

short period. The changes in availability, formulation, labeling, and prices were outside the scope of this study. Additionally, only low-sodium salts having an online presence were accessed in the study. The availability of low-sodium salts may be underestimated in low- and middle-income countries due to the challenges of access to available online information in these settings. Keywords in all six major languages of the United Nations were used to maximize the possibility of identifying low-sodium salts in non-English speaking countries. However, this may still miss many countries. We tried to minimize the risk of omitting low-sodium salts by using complementary information sources, including systematic reviews and consulting key informants from a diverse set of countries based on income categories.

This study's results provide evidence on the availability, formulation, labeling, and price of low-sodium salts worldwide. These results can be used to inform efforts to scale up low-sodium salt use as an effective public health intervention in different countries. We therefore recommend making low-sodium salts more widely available, providing accurate and standardized labeling, and subsidizing the price of low-sodium salt. Future studies on how to promote the production, sale, adoption, spread, scale-up, and sustainability of low-sodium salts in diverse settings will also help governments take the necessary steps to reduce population-level dietary sodium consumption through the widespread use of low-sodium salt.

---

## Acknowledgments

This research received no specific grant from any funding agency in public, commercial, or not-for-profit sectors. XY and KCL are supported by the University of New South Wales Scientia PhD scholarship. JW is supported by a National Heart Foundation Future Leader II Fellowship (APP102039) and receives funding for work on salt reduction from the National Health and Medical Research Council, the WHO, and the Victorian Health Promotion Foundation. KT was supported by an Early Career Fellowship (APP1161597) from the National Health and Medical Research Council of Australia and a Postdoctoral Fellowship (Award ID: 102140) from the National Heart Foundation of Australia. JHYW is supported by a University of New South Wales Scientia Fellowship. HL is supported by a National Health and Medical Research Council of Australia program grant fellowship.

---

## Conflicts of Interest

In the past 3 years, MDH received funding from the World Heart Federation to serve as its senior program advisor for the Emerging Leaders program, which is supported by Boehringer Ingelheim and Novartis with previous support from BUPA and AstraZeneca. MDH also receives support from the American Heart Association, Verily, AstraZeneca, and the American Medical Association for food system surveillance-related research. The George Institute for Global Health's wholly owned enterprise, George Health Enterprises, has received investment funds to develop fixed-dose combination products containing aspirin, statin, and blood pressure-lowering drugs. MDH plans to submit patents for heart failure polypills, including for heart failure with reduced ejection fraction. BN has received in kind support (salt substitute supplies) from Nutek and the Beijing Salt Manufacturing Corporation for studies of the effects of salt substitutes. JW is the Director of the World Health Organization Collaborating Centre on Population Salt Reduction at the George Institute for Global Health.

---

### Multimedia Appendix 1

Database search strategy.

[[DOCX File, 16 KB - publichealth\\_v7i7e27423\\_app1.docx](#) ]

---

### Multimedia Appendix 2

Key characteristics of all included low-sodium salts.

[[XLSX File \(Microsoft Excel File\), 24 KB - publichealth\\_v7i7e27423\\_app2.xlsx](#) ]

## References

1. GBD 2017 Causes of Death Collaborators. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2018 Nov 10;392(10159):1736-1788 [FREE Full text] [doi: [10.1016/S0140-6736\(18\)32203-7](https://doi.org/10.1016/S0140-6736(18)32203-7)] [Medline: [30496103](https://pubmed.ncbi.nlm.nih.gov/30496103/)]
2. GBD 2017 Diet Collaborators. Health effects of dietary risks in 195 countries, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2019 May 11;393(10184):1958-1972 [FREE Full text] [doi: [10.1016/S0140-6736\(19\)30041-8](https://doi.org/10.1016/S0140-6736(19)30041-8)] [Medline: [30954305](https://pubmed.ncbi.nlm.nih.gov/30954305/)]
3. Huang L, Trieu K, Yoshimura S, Neal B, Woodward M, Campbell NRC, et al. Effect of dose and duration of reduction in dietary sodium on blood pressure levels: systematic review and meta-analysis of randomised trials. *BMJ* 2020 Feb 24;368:m315 [FREE Full text] [doi: [10.1136/bmj.m315](https://doi.org/10.1136/bmj.m315)] [Medline: [32094151](https://pubmed.ncbi.nlm.nih.gov/32094151/)]
4. Forouhi NG, Unwin N. Global diet and health: old questions, fresh evidence, and new horizons. *Lancet* 2019 May 11;393(10184):1916-1918 [FREE Full text] [doi: [10.1016/S0140-6736\(19\)30500-8](https://doi.org/10.1016/S0140-6736(19)30500-8)] [Medline: [30954306](https://pubmed.ncbi.nlm.nih.gov/30954306/)]
5. A global brief on hypertension: silent killer, global public health crisis: World Health Day 2013. World Health Organization. 2013. URL: <https://apps.who.int/iris/handle/10665/79059> [accessed 2021-07-07]
6. Barberio AM, Sumar N, Trieu K, Lorenzetti DL, Tarasuk V, Webster J, et al. Population-level interventions in government jurisdictions for dietary sodium reduction: a Cochrane Review. *Int J Epidemiol* 2017 Oct 01;46(5):1551-1405 [FREE Full text] [doi: [10.1093/ije/dyw361](https://doi.org/10.1093/ije/dyw361)] [Medline: [28204481](https://pubmed.ncbi.nlm.nih.gov/28204481/)]
7. Ponce-Lucero V, Saavedra-García L, Cateriano-Arévalo E, Perez-Leon S, Villarreal-Zegarra D, Horna-Alva D, et al. Parents' perceptions about salt consumption in urban areas of Peru: formative research for a social marketing strategy. *Nutrients* 2020 Jan 08;12(1):176 [FREE Full text] [doi: [10.3390/nu12010176](https://doi.org/10.3390/nu12010176)] [Medline: [31936312](https://pubmed.ncbi.nlm.nih.gov/31936312/)]
8. Peng Y, Li W, Wen X, Li Y, Hu J, Zhao L. Effects of salt substitutes on blood pressure: a meta-analysis of randomized controlled trials. *Am J Clin Nutr* 2014 Dec;100(6):1448-1454. [doi: [10.3945/ajcn.114.089235](https://doi.org/10.3945/ajcn.114.089235)] [Medline: [25411279](https://pubmed.ncbi.nlm.nih.gov/25411279/)]
9. Hernandez AV, Emonds EE, Chen BA, Zavala-Loayza AJ, Thota P, Pasupuleti V, et al. Effect of low-sodium salt substitutes on blood pressure, detected hypertension, stroke and mortality. *Heart* 2019 Jun;105(12):953-960. [doi: [10.1136/heartjnl-2018-314036](https://doi.org/10.1136/heartjnl-2018-314036)] [Medline: [30661034](https://pubmed.ncbi.nlm.nih.gov/30661034/)]
10. Marklund M, Singh G, Greer R, Cudhea F, Matsushita K, Micha R, et al. Estimated population wide benefits and risks in China of lowering sodium through potassium enriched salt substitution: modelling study. *BMJ* 2020 Apr 22;369:m824 [FREE Full text] [doi: [10.1136/bmj.m824](https://doi.org/10.1136/bmj.m824)] [Medline: [32321724](https://pubmed.ncbi.nlm.nih.gov/32321724/)]
11. Ide N, Ajenikoko A, Steele L, Cohn J, J Curtis C, Frieden TR, et al. Priority actions to advance population sodium reduction. *Nutrients* 2020 Aug 22;12(9):2543 [FREE Full text] [doi: [10.3390/nu12092543](https://doi.org/10.3390/nu12092543)] [Medline: [32842580](https://pubmed.ncbi.nlm.nih.gov/32842580/)]
12. Yin X, Liu H, Trieu K, Webster J, Farrand C, Li K, et al. The effectiveness, feasibility, and acceptability of low-sodium salts worldwide: an environmental scan protocol. *J Clin Hypertens (Greenwich)* 2020 Dec;22(12):2258-2265. [doi: [10.1111/jch.14054](https://doi.org/10.1111/jch.14054)] [Medline: [32966696](https://pubmed.ncbi.nlm.nih.gov/32966696/)]
13. Google Help. URL: <https://www.google.com/support/chrome/bin/answer.py?hlrm=ru&answer=173424> [accessed 2021-04-01]
14. Implied PPP conversion rate. International Monetary Fund. 2020. URL: <https://www.imf.org/external/datamapper/PPPEX@WEO/OEMDC/ADVEC/WEOWORLD/DA/IND?year=2020> [accessed 2021-04-01]
15. World Bank country and lending groups. World Bank Data Help Desk. URL: <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups> [accessed 2021-03-29]
16. Gheorghe A, Griffiths U, Murphy A, Legido-Quigley H, Lamptey P, Perel P. The economic burden of cardiovascular disease and hypertension in low- and middle-income countries: a systematic review. *BMC Public Health* 2018 Aug 06;18(1):975 [FREE Full text] [doi: [10.1186/s12889-018-5806-x](https://doi.org/10.1186/s12889-018-5806-x)] [Medline: [30081871](https://pubmed.ncbi.nlm.nih.gov/30081871/)]
17. Thout SR, Yu J, Tian M, Huffman MD, Arnott C, Li Q, et al. Rationale, design, and baseline characteristics of the Salt Substitute in India Study (SSiS): the protocol for a double-blinded, randomized-controlled trial. *J Clin Hypertens (Greenwich)* 2020 Aug;22(8):1504-1512. [doi: [10.1111/jch.13947](https://doi.org/10.1111/jch.13947)] [Medline: [32710677](https://pubmed.ncbi.nlm.nih.gov/32710677/)]
18. Bhat S, Marklund M, Henry ME, Appel LJ, Croft KD, Neal B, et al. A systematic review of the sources of dietary salt around the world. *Adv Nutr* 2020 May 01;11(3):677-686 [FREE Full text] [doi: [10.1093/advances/nmz134](https://doi.org/10.1093/advances/nmz134)] [Medline: [31904809](https://pubmed.ncbi.nlm.nih.gov/31904809/)]
19. Shao S, Hua Y, Yang Y, Liu X, Fan J, Zhang A, et al. Salt reduction in China: a state-of-the-art review. *Risk Manag Healthc Policy* 2017;10:17-28. [doi: [10.2147/RMHP.S75918](https://doi.org/10.2147/RMHP.S75918)] [Medline: [28260957](https://pubmed.ncbi.nlm.nih.gov/28260957/)]
20. Wolfson JA, Ramsing R, Richardson CR, Palmer A. Barriers to healthy food access: associations with household income and cooking behavior. *Prev Med Rep* 2019 Mar;13:298-305 [FREE Full text] [doi: [10.1016/j.pmedr.2019.01.023](https://doi.org/10.1016/j.pmedr.2019.01.023)] [Medline: [30792944](https://pubmed.ncbi.nlm.nih.gov/30792944/)]
21. Li N, Yan LL, Niu W, Yao C, Feng X, Zhang J, et al. The effects of a community-based sodium reduction program in rural China - a cluster-randomized trial. *PLoS One* 2016;11(12):e0166620 [FREE Full text] [doi: [10.1371/journal.pone.0166620](https://doi.org/10.1371/journal.pone.0166620)] [Medline: [27935977](https://pubmed.ncbi.nlm.nih.gov/27935977/)]
22. Drewnowski A, Rehm C, Maillot M, Monsivais P. The relation of potassium and sodium intakes to diet cost among U.S. adults. *J Hum Hypertens* 2015 Jan;29(1):14-21 [FREE Full text] [doi: [10.1038/jhh.2014.38](https://doi.org/10.1038/jhh.2014.38)] [Medline: [24871907](https://pubmed.ncbi.nlm.nih.gov/24871907/)]

23. Saha M, Maurya BR, Bahadur I, Kumar A, Meena VS. Can potassium-solubilising bacteria mitigate the potassium problems in India? In: Meena VS, Maurya BR, Verma JP, Meena RS, editors. Potassium Solubilizing Microorganisms for Sustainable Agriculture. New Delhi: Springer; 2016:127-136.
24. Saavedra-Garcia L, Bernabe-Ortiz A, Gilman R, Diez-Canseco F, Cárdenas MK, Sacksteder K, et al. Applying the triangle taste test to assess differences between low sodium salts and common salt: evidence from Peru. PLoS One 2015;10(7):e0134700 [FREE Full text] [doi: [10.1371/journal.pone.0134700](https://doi.org/10.1371/journal.pone.0134700)] [Medline: [26225848](https://pubmed.ncbi.nlm.nih.gov/26225848/)]
25. Sinopoli DA, Lawless HT. Taste properties of potassium chloride alone and in mixtures with sodium chloride using a check-all-that-apply method. J Food Sci 2012 Sep;77(9):S319-S322. [doi: [10.1111/j.1750-3841.2012.02862.x](https://doi.org/10.1111/j.1750-3841.2012.02862.x)] [Medline: [22901084](https://pubmed.ncbi.nlm.nih.gov/22901084/)]
26. Li N, Prescott J, Wu Y, Barzi F, Yu X, Zhao L, China Salt Substitute Study Collaborative Group. The effects of a reduced-sodium, high-potassium salt substitute on food taste and acceptability in rural northern China. Br J Nutr 2009 Apr;101(7):1088-1093. [doi: [10.1017/S0007114508042360](https://doi.org/10.1017/S0007114508042360)] [Medline: [18710605](https://pubmed.ncbi.nlm.nih.gov/18710605/)]
27. Braschi A, Gill L, Naismith DJ. Partial substitution of sodium with potassium in white bread: feasibility and bioavailability. Int J Food Sci Nutr 2009 Sep;60(6):507-521. [doi: [10.1080/09637480701782118](https://doi.org/10.1080/09637480701782118)] [Medline: [18608540](https://pubmed.ncbi.nlm.nih.gov/18608540/)]
28. Capanec K, Vugrinec S, Cvetković T, Ranilović J. Potassium Chloride-Based Salt Substitutes: A Critical Review with a Focus on the Patent Literature. Compr Rev Food Sci Food Saf 2017 Sep;16(5):881-894. [doi: [10.1111/1541-4337.12291](https://doi.org/10.1111/1541-4337.12291)] [Medline: [33371617](https://pubmed.ncbi.nlm.nih.gov/33371617/)]
29. Kovesdy CP, Matsushita K, Sang Y, Brunskill NJ, Carrero JJ, Chodick G, et al. Serum potassium and adverse outcomes across the range of kidney function: a CKD Prognosis Consortium meta-analysis. Eur Heart J 2018 May 01;39(17):1535-1542 [FREE Full text] [doi: [10.1093/eurheartj/ehy100](https://doi.org/10.1093/eurheartj/ehy100)] [Medline: [29554312](https://pubmed.ncbi.nlm.nih.gov/29554312/)]
30. Neal B, Tian M, Li N, Elliott P, Yan L, Labarthe D, et al. Rationale, design, and baseline characteristics of the Salt Substitute and Stroke Study (SSaSS)-A large-scale cluster randomized controlled trial. Am Heart J 2017 Jun;188:109-117. [doi: [10.1016/j.ahj.2017.02.033](https://doi.org/10.1016/j.ahj.2017.02.033)] [Medline: [28577665](https://pubmed.ncbi.nlm.nih.gov/28577665/)]
31. Yu J, Thout S, Li Q, Tian M, Marklund M, Arnott C, et al. Effects of a reduced-sodium added-potassium salt substitute on blood pressure in rural Indian hypertensive patients: a randomized, double-blind, controlled trial. Am J Clin Nutr 2021 Jul 01;114(1):185-193. [doi: [10.1093/ajcn/nqab054](https://doi.org/10.1093/ajcn/nqab054)] [Medline: [33782684](https://pubmed.ncbi.nlm.nih.gov/33782684/)]
32. Fathima K, Bhargava M. Salt reduction and low-sodium salt substitutes: awareness among health-care providers in Mangalore, Karnataka. Indian J Community Med 2018;43(4):266-269 [FREE Full text] [doi: [10.4103/ijcm.IJCM\\_22\\_18](https://doi.org/10.4103/ijcm.IJCM_22_18)] [Medline: [30662177](https://pubmed.ncbi.nlm.nih.gov/30662177/)]

## Abbreviations

**CKD:** chronic kidney disease

**CVD:** cardiovascular disease

**WHO:** World Health Organization

*Edited by Y Khader; submitted 25.01.21; peer-reviewed by P Banik, M Tan, P Sarajlic, Y Xu; comments to author 21.03.21; revised version received 19.04.21; accepted 13.05.21; published 14.07.21.*

*Please cite as:*

*Yin X, Liu H, Webster J, Trieu K, Huffman MD, Miranda JJ, Marklund M, Wu JHY, Cobb LK, Li KC, Pearson SA, Neal B, Tian M Availability, Formulation, Labeling, and Price of Low-sodium Salt Worldwide: Environmental Scan*

*JMIR Public Health Surveill 2021;7(7):e27423*

*URL: <https://publichealth.jmir.org/2021/7/e27423>*

*doi: [10.2196/27423](https://doi.org/10.2196/27423)*

*PMID: [33985938](https://pubmed.ncbi.nlm.nih.gov/33985938/)*

©Xuejun Yin, Hueiming Liu, Jacqui Webster, Kathy Trieu, Mark D Huffman, J Jaime Miranda, Matti Marklund, Jason H Y Wu, Laura K Cobb, Ka Chun Li, Sallie-Anne Pearson, Bruce Neal, Maoyi Tian. Originally published in JMIR Public Health and Surveillance (<https://publichealth.jmir.org>), 14.07.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

# The Reliability and Quality of YouTube Videos as a Source of Public Health Information Regarding COVID-19 Vaccination: Cross-sectional Study

Calvin Chan<sup>1</sup>, BSc; Viknesh Sounderajah<sup>1,2</sup>, MBBS, MSc, MRCS; Elisabeth Daniels<sup>1</sup>, BSc; Amish Acharya<sup>1,2</sup>, BSc, MBBS, MRCS; Jonathan Clarke<sup>1,2,3</sup>, MBBS, MPH, PhD; Seema Yalamanchili<sup>1,2</sup>, MBBS, MSc, MRCS; Pasha Normahani<sup>1</sup>, MBBS, MSc, MRCS; Sheraz Markar<sup>1</sup>, MBBChir, MA, PhD; Hutan Ashrafian<sup>1,2</sup>, MBBS, MBA, PhD, MRCS; Ara Darzi<sup>1,2</sup>, PC, KBE, FRS, FMEDSCI, HONFRENG

<sup>1</sup>Department of Surgery & Cancer, Imperial College London, London, United Kingdom

<sup>2</sup>Institute of Global Health Innovation, Imperial College London, London, United Kingdom

<sup>3</sup>Department of Mathematics, Imperial College London, London, United Kingdom

**Corresponding Author:**

Viknesh Sounderajah, MBBS, MSc, MRCS

Department of Surgery & Cancer

Imperial College London

10th floor, Queen Elizabeth Queen Mother Building, St. Mary's Hospital, South Wharf Road

London, W2 1NY

United Kingdom

Phone: 44 02033126666

Email: [vs1108@imperial.ac.uk](mailto:vs1108@imperial.ac.uk)

## Abstract

**Background:** Recent emergency authorization and rollout of COVID-19 vaccines by regulatory bodies has generated global attention. As the most popular video-sharing platform globally, YouTube is a potent medium for the dissemination of key public health information. Understanding the nature of available content regarding COVID-19 vaccination on this widely used platform is of substantial public health interest.

**Objective:** This study aimed to evaluate the reliability and quality of information on COVID-19 vaccination in YouTube videos.

**Methods:** In this cross-sectional study, the phrases “coronavirus vaccine” and “COVID-19 vaccine” were searched on the UK version of YouTube on December 10, 2020. The 200 most viewed videos of each search were extracted and screened for relevance and English language. Video content and characteristics were extracted and independently rated against Health on the Net Foundation Code of Conduct and DISCERN quality criteria for consumer health information by 2 authors.

**Results:** Forty-eight videos, with a combined total view count of 30,100,561, were included in the analysis. Topics addressed comprised the following: vaccine science (n=18, 58%), vaccine trials (n=28, 58%), side effects (n=23, 48%), efficacy (n=17, 35%), and manufacturing (n=8, 17%). Ten (21%) videos encouraged continued public health measures. Only 2 (4.2%) videos made nonfactual claims. The content of 47 (98%) videos was scored to have low (n=27, 56%) or moderate (n=20, 42%) adherence to Health on the Net Foundation Code of Conduct principles. Median overall DISCERN score per channel type ranged from 40.3 (IQR 34.8-47.0) to 64.3 (IQR 58.5-66.3). Educational channels produced by both medical and nonmedical professionals achieved significantly higher DISCERN scores than those of other categories. The highest median DISCERN scores were achieved by educational videos produced by medical professionals (64.3, IQR 58.5-66.3) and the lowest median scores by independent users (18, IQR 18-20).

**Conclusions:** The overall quality and reliability of information on COVID-19 vaccines on YouTube remains poor. Videos produced by educational channels, especially by medical professionals, were higher in quality and reliability than those produced by other sources, including health-related organizations. Collaboration between health-related organizations and established medical and educational YouTube content producers provides an opportunity for the dissemination of high-quality information on COVID-19 vaccination. Such collaboration holds potential as a rapidly implementable public health intervention aiming to engage a wide audience and increase public vaccination awareness and knowledge.

**KEYWORDS**

COVID-19; infodemiology; public health; quality; reliability; social media; vaccination; vaccine; video; web-based health information; YouTube

## **Introduction**

The recent emergency authorization and rollout of COVID-19 vaccines by regulatory bodies has generated global media attention. Unsurprisingly, internet searches related to COVID-19 vaccination increased drastically during November-December 2020, as the public attempted to source information amid a surge in media coverage [1]. Many internet users turned to YouTube, the second-most visited website globally after Google, for further information [2].

YouTube is the most popular video-sharing platform worldwide. Over 1 billion hours' worth of video is streamed each day on the website, and it is visited by over 2 billion unique users monthly [3]. It has strong penetrance globally and across all major sociodemographic groups. YouTube provides a potent means of disseminating real-time information across a population; users are able to curate video content from sources varying from individual users, through celebrities, to media outlets. Aware of their central role in the dissemination of key public health information, YouTube has implemented a COVID-19 medical misinformation policy, which forbids COVID-19-related content that contradicts local health authorities and risks public safety [4].

There have been, however, high-profile instances of internet-propagated misinformation regarding COVID-19, including the ingestion of cleaning products as potential treatment, which have had severe consequences [5]. Despite the aforementioned measures implemented by YouTube, there remains a concern for COVID-19 vaccination programs to remain an easy target for misinformation content. Previous studies have highlighted that vaccination programs, such as the human papillomavirus vaccination program, have been a common target of high-profile YouTube videos propagated by a community of vocal users who are critical of vaccination programs [6]. Furthermore, there remains a concern that anti-COVID-19 vaccination videos could (1) pose a significant threat to compliance with the vaccination program, especially among those who are disproportionately affected by the illness; (2) create spill-over dissonance toward other critical COVID-19 public health measures; and (3) displace notifications regarding other emerging time-sensitive information from official public health sources. In fact, from the beginning of the pandemic until now, COVID-19 vaccine hesitancy has been steadily increasing [7]. Understanding the nature of available content regarding the COVID-19 vaccination program on this widely used platform is, therefore, of substantial public health interest and forms a foundation on which strategies for misinformation counteraction can be based.

To date, no studies have evaluated the quality and reliability of COVID-19 vaccination-related information available on YouTube. The objective of this study, therefore, was to evaluate

the reliability and quality of information of YouTube's most prominent videos on COVID-19 vaccination, using 2 validated criteria (DISCERN and HONcode).

## **Methods**

### **Methods Overview**

Ethical approval for this study was waived because all gathered data are freely available in the public domain. The phrases "coronavirus vaccine" and "COVID-19 vaccine" were searched on the UK version of YouTube on December 10, 2020. The search was conducted in an incognito browser (Google Chrome) to avoid biased suggestions based on cookies. Search results were sorted by view count to identify videos that had achieved the greatest impact and were most likely to trend, thereby reaching further viewers. The 200 most viewed videos (10 pages) of each search were subsequently extracted.

Video titles and channels were first screened for relevance and English language before full-video screening. Videos were included if they described 1 or more of the following: mechanisms of action of vaccines, clinical trial procedures, manufacturing processes, side effects or safety, and vaccine efficacy. Descriptions of the criteria are provided in [Multimedia Appendix 1](#), Table S1. If there was uncertainty regarding whether a video should be included, a consensus was sought between authors, with a predisposition to include the video for full assessment. Additionally, videos were assessed for their promotion of public health measures such as hand-washing, wearing masks, or social distancing. Finally, instances of nonfactual content in these videos were noted. Nonfactual information (ie, misinformation) was defined as nonscientific corroborated content that contradicted medical information provided by the current local health authority or the World Health Organization. Examples of misinformation are available on YouTube's medical misinformation policy [4]. Duplicate videos and non-English-language videos were excluded. Video content screening was completed independently by 2 authors (CC and ED). Any discrepancies were resolved through discussion with a third author (VS).

Characteristics (video URL, channel, country of origin, view count, duration, video age, and the number of likes, dislikes, and comments) of the included videos were extracted. Videos were placed in 6 main categories by YouTube channel type: educational channels produced by medical professionals, educational channels produced by nonmedical individuals (eg, science education or explanatory media), independent nonmedical users (eg, vloggers with no obvious affiliations), internet media (eg, newsmagazine shows or talk shows), news agencies (ie, clips uploaded from network news), and nonprofit or medical organizations (eg, hospitals, government organizations, or universities). Descriptions and examples of

channel types are provided in [Multimedia Appendix 1](#), Table S2.

Reliability of the video content (ie, the extent to which the source of information, and therefore the information itself, could be relied upon, evident from clearly referenced and scientifically corroborated content) was assessed against a modified Health on the Net Foundation Code of Conduct (HONcode) checklist [8] and modified DISCERN quality criteria for consumer health information [8,9], which have previously been used to assess the quality of health information on YouTube. The quality of video content (ie, completeness, understandability, relevance, depth, and accuracy of information provided) was also assessed using the DISCERN quality criteria. Video rating was completed independently by 2 authors (CC and ED).

**Textbox 1.** Modified DISCERN quality criteria for assessing the reliability and quality of YouTube content on COVID-19 vaccination. Each question was rated from 1 (“worst”) to 5 (“best”). Sections 2 and 3 only applied to videos on vaccine science (ie, how the treatment [vaccination] works).

**Section 1: is the video reliable?**

1. Are the aims clear?
2. Does it achieve its aims?
3. Is it relevant?
4. Is it clear what sources of information were used to compile the video?
5. Is it clear when the information used or reported in the video was produced?
6. Is it balanced and unbiased?
7. Does it provide details of additional sources of support and information?
8. Does it refer to areas of uncertainty?

**Section 2: how good is the quality of information on treatment choices?**

9. Does it describe how each treatment works?
10. Does it describe the benefits of each treatment?
11. Does it describe the risks of each treatment?
12. Does it describe what would happen if no treatment is used?
13. Does it describe how the treatment choices affect overall quality of life?
14. Is it clear that there may be more than one possible treatment choice?
15. Does it provide support for shared decision-making?

**Section 3: overall rating of the video**

16. Based on the answers to all of the above questions, rate the overall quality of the video as a source of information about treatment choices

HONcode consists of 8 principles that evaluate the reliability and credibility of health information [10]. Videos were rated with a score of 1 (adherent) or 0 (nonadherent) for each of the 8 principles. The DISCERN instrument consists of 16 questions rated from 1 to 5, which assess health content across 3 domains including video reliability (8 questions, 40 points), treatment information quality (7 questions, 35 points), and an overall reviewer rating (5 points), thus yielding a maximum cumulative score of 80 ([Textbox 1](#)) [11]. The first 8 questions were applied to all videos, and 8 additional questions were applied to videos specifically on vaccine science (ie, how the treatment works). The questionnaire was adapted from Goobie et al [8] and Loeb et al [9].

## Statistical Analysis

Statistical analysis was performed using Stata (version 13, StataCorp). Intercategory differences were assessed using Kruskal–Wallis tests and the post hoc Dunn test. Interrater reliability was assessed using the Cohen  $\kappa$  statistic. A DISCERN score of +1 or –1 point was considered agreement. Associations between engagement metrics and DISCERN scores were evaluated using linear regression. Significance was set at  $P < .05$ . Data are presented as median (IQR) values.

## Availability of Data and Material

The data sets used or analyzed in this study are available from the corresponding author on reasonable request.

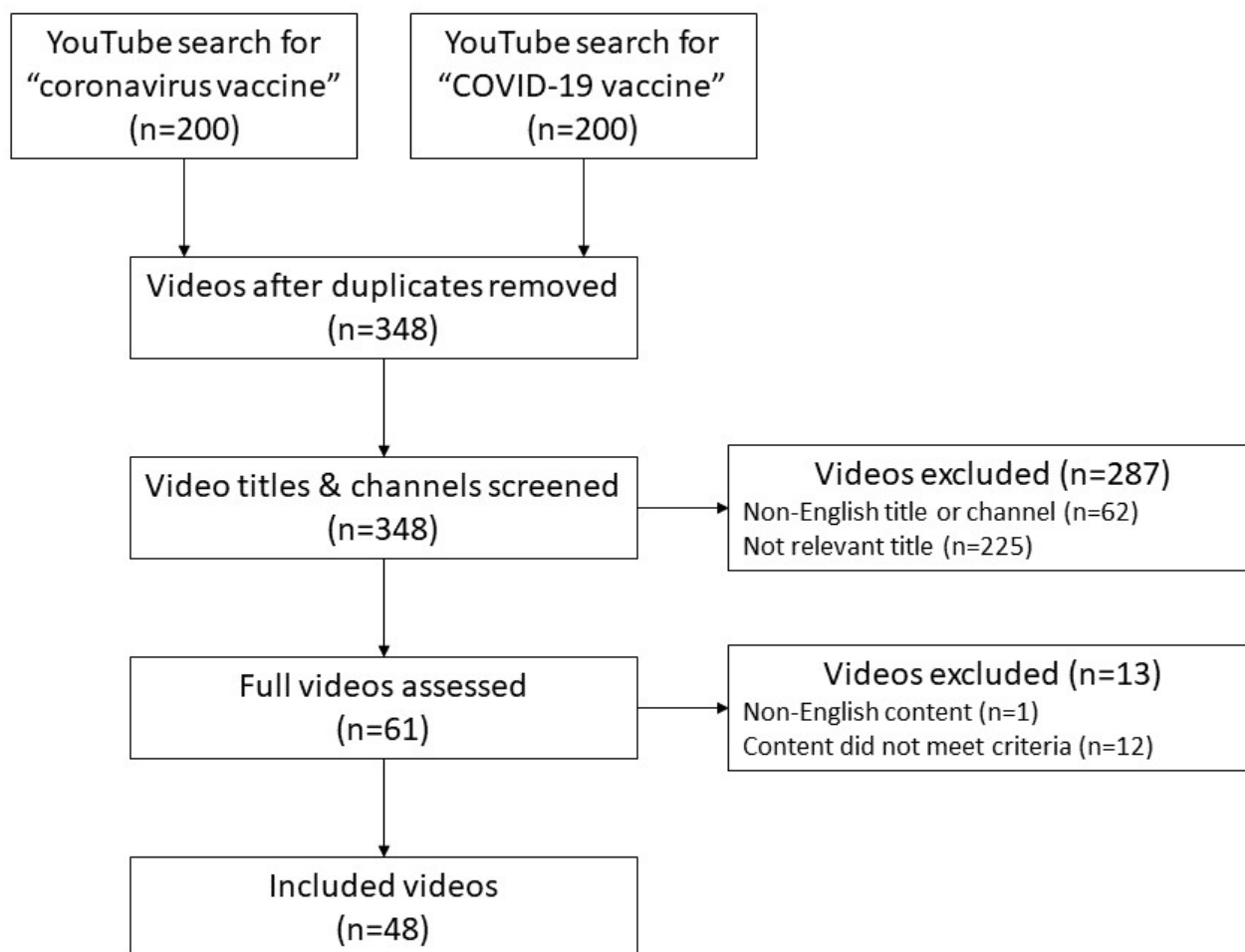
## Results

### Video Characteristics

The video review process is illustrated in [Figure 1](#). From the 200 results of each search, 52 duplicate videos were removed, yielding 348 unique videos. After the video title and channel were screened and the full video was assessed, 48 videos were included for data extraction, with a combined total view count of 30,100,561. Videos that were not in English ( $n=62$ ) or did not meet the study inclusion criteria ( $n=225$ )—describing topics such as vaccination priority, national distribution plans, politics, or pandemic mortality figures—were excluded. The characteristics of the included videos are summarized in [Table 1](#). The majority (75%) of videos were produced by US channels.

The median number of views per video was 236,064 (IQR 152,082-596,234).

**Figure 1.** Flow diagram for the results for searches on COVID-19 vaccine-related videos on YouTube and the video selection process for inclusion in the study. The 2 searches were performed on December 10, 2020.





**Table 1.** Characteristics of COVID-19 vaccine-related YouTube videos included in the study.

Characteristics	Values
<b>Country of origin, n (%)</b>	
United States	34 (75)
United Kingdom	6 (13)
Canada	3 (6)
Germany	2 (4)
Australia	1 (2)
Switzerland	1 (2)
<b>Channel type, n (%)</b>	
Educational (nonmedical)	10 (21)
Educational (medical)	6 (13)
Independent users	5 (10)
Internet media	10 (21)
News agencies	13 (27)
Nonprofit or medical organizations	4 (8)
Video duration (minutes), median (IQR)	9:18 (4:42-11:51)
Video age (days since upload), median (IQR)	54 (19-191)
<b>Engagement, median (IQR)</b>	
Views	236,064 (152,082-596,234)
Views per day since upload	6364 (3031-11,717)
Likes	5600 (2300-9200)
Dislikes	545 (236-1200)
Likes:dislikes ratio	14.3 (3-24)
Comments	1891 (1059-3424)

**Source**

The 48 videos were segregated into 6 categories by their YouTube channel type. The contribution of each channel type

toward the total view count is detailed in [Table 2](#). The most viewed video (6,668,737 views, 22% of total views) described the mechanisms of action of COVID-19 vaccines and was produced by a medical organization (JAMA Network).

**Table 2.** Engagement metrics, content, adherence to the Health on the Net Foundation Code of Conduct (HONCode), and DISCERN score of COVID-19 vaccine-related videos on YouTube, stratified by channel type.

Parameters	Educational (nonmedical)	Educational (medical)	Independent nonmedical users	Internet media	News agencies	Nonprofit or medical organizations	Overall	<i>P</i> value <sup>a</sup>
Videos, n (%)	10 (21)	6 (13)	5 (10)	10 (21)	13 (27)	4 (8)	48 (100)	N/A <sup>b</sup>
Total views, n (%)	5,473,987 (18)	1,467,003 (5)	1,766,051 (6)	4,205,972 (14)	9,798,419 (33)	7,389,129 (25)	30,100,561 (100)	N/A
Views per video, median (IQR)	367,560 (162,635-623,924)	213,707 (143,570-302,387)	221,299 (209,150-334,553)	197,365 (135,953-557,721)	275,615 (137,721-644,970)	247,977 (241,877-1,853,382)	236,064 (152,082-596,234)	N/A
Views per day since upload, median (IQR)	7604 (3681-10,802)	8309 (5179-9699)	8852 (7406-27,879)	1984 (880-5470)	5560 (4279-10,656)	11,734 (2532-37,832)	6364 (3031-11,717)	N/A
Likes, median (IQR)	15,050 (6050-25,250)	5600 (5000-6725)	10,000 (5800-10,000)	4200 (2000-6275)	2600 (1300-5100)	3400 (1998-15,200)	5600 (2300-9200)	N/A
Dislikes, median (IQR)	656 (261-1325)	268 (213-405)	557 (342-701)	651 (223-1200)	716 (404-1200)	1700 (970-5700)	545 (236-1200)	N/A
Likes:dislikes ratio, median (IQR)	23.8 (18-26)	20.9 (18-25)	18 (6-30)	12.3 (3-18)	4.3 (2-11)	2.8 (2-8)	14.3 (3-24)	N/A
Comments, median (IQR)	1633 (1047-3444)	1932 (1691-2253)	2133 (1680-6402)	1430 (622-3480)	2204 (1194-3611)	1258 (735-1858)	1891 (1059-3424)	N/A
<b>Content<sup>c</sup>, n (%)</b>								N/A
Vaccine science	9 (90)	6 (100)	1 (20)	5 (50)	4 (31)	3 (75)	28 (58)	
Trial process	5 (50)	5 (83)	1 (20)	7 (70)	8 (62)	2 (50)	28 (58)	
Manufacturing process	2 (20)	0 (0)	0 (0)	4 (40)	2 (15)	0 (0)	8 (17)	
Side effects and safety	3 (30)	5 (83)	5 (100)	1 (10)	7 (54)	2 (50)	23 (48)	
Vaccine efficacy	2 (20)	5 (83)	3 (60)	2 (20)	4 (31)	1 (25)	17 (35)	
Public health information	3 (30)	1 (17)	1 (20)	3 (30)	1 (8)	1 (25)	10 (21)	
<b>HONcode adherence (/8), n (%)</b>								N/A
Low (0-2)	2 (20)	1 (17)	5 (100)	5 (50)	10 (77)	3 (75)	27 (56)	
Moderate (3-5)	7 (70)	5 (83)	0 (0)	4 (40)	3 (23)	1 (25)	20 (42)	
High (6-8)	1 (10)	0 (0)	0 (0)	1 (10)	0 (0)	0 (0)	1 (2)	
<b>DISCERN score, median (IQR)</b>								
Reliability (/40, n=48)	28.5 (25.9-32.5)	35.5 (32.0-39.4)	18 (18-20)	24.8 (23.3-26.8)	23.8 (20.5-24.5)	23.5 (22.9-24.5)	25 (22.9-29.1)	<i>&lt;.001</i>
Treatment quality (/35, n=30)	16.5 (14-18)	21.3 (20.1-25.8)	19 (19-19)	17.5 (17.0-18.5)	14.8 (12.6-17.4)	20.5 (16.8-21.3)	18.3 (15.1-20.5)	.052
Overall quality judgement (/5, n=30)	3 (3-4)	4.8 (4.1-5.0)	3.5 (3.5-3.5)	3 (2.5-4.0)	2.5 (2.0-3.1)	3 (3.0-3.5)	3.3 (3-4)	.07
Total score (/80, n=30)	48.5 (46.5-53.0)	64.3 (58.5-66.3)	43.5 (43.5-43.5)	49 (45.5-52.0)	40.3 (34.8-47.0)	52 (45.5-54.3)	50 (45.1-53.1)	.03

<sup>a</sup>*P* values were produced using Kruskal-Wallis tests. Significant figures are in italics.

<sup>b</sup>N/A: not applicable.

<sup>c</sup>Percentage values have been calculated relative to the total number of videos of each channel type: educational (nonmedical) (n=10), educational (medical) (n=6), independent nonmedical users (n=5), internet media (n=10), news agencies (n=13), and nonprofit or medical organizations (n=4).

## Content

Twenty-eight of 48 (58%) videos addressed vaccine science and mechanisms of action. Twenty-eight videos also discussed vaccine trials, and 23 videos discussed vaccine safety or side effects. Ten videos advocated the importance of continued traditional public health measures to reduce COVID-19 transmission (eg, hand-washing, mask-wearing, and social distancing).

Regarding nonfactual content, 2 videos (1 by internet media and 1 independently produced) contained unsubstantiated vaccine safety concerns, despite YouTube's aforementioned COVID-19 misinformation policy. Both videos were interviews with a single prominent antivaccination advocate. These 2 nonfactual videos accounted for 390,927 views (1.3% of total viewership).

**Table 3.** Description of principles of the Health on the Net Foundation Code of Conduct [8] and the number of COVID-19 vaccine-related YouTube videos that met each criterion.

Principle	Description	Videos, n (%)
Authoritative	Any medical or health advice provided in this video will only be given by medically trained and qualified professionals unless a clear statement is made that the advice offered is from a nonmedically qualified individual or organization.	27 (56)
Complementary	The information provided is designed to support, not replace, the relationship that exists between a patient and his/her existing physician.	38 (79)
Privacy	The information in the video maintains the right to confidentiality and respect of the individual patient featured.	0 (0)
Attribution	Where appropriate, information contained in the video will be supported by clear references to source data and, where possible, have specific links to those data.	20 (42)
Justifiability	Any claims relating to the benefits or performance of a specific treatment, commercial product, or service will be supported by appropriate, balanced evidence in the manner outlined in attribution principle.	20 (42)
Transparency	The designers of the video will seek to provide information in the clearest possible manner and provide contact addresses for viewers who seek further information or support.	17 (35)
Financial disclosure	Support for this video will be clearly identified, including the identities of commercial and noncommercial organizations that have contributed funding, services, or material for the video.	7 (15)
Advertising policy	If advertising is a source of funding, it will be clearly stated. Advertising and other promotional material will be presented to viewers in a manner and context that facilitate differentiation between it and the original content.	2 (4)

There were significant differences in DISCERN reliability and overall scores among different channel types ( $P<.001$  and  $P=.03$ , respectively). Educational channels produced by medical professionals attained the greatest median (IQR) DISCERN scores for reliability (median 35.5, IQR 32.0-39.4), quality (median 21.3, IQR 20.1-25.8), user judgement (median 4.8, IQR 4.1-5), and overall score (median 64.3, IQR 58.5-66.3). Videos produced by independent nonmedical individuals achieved the lowest reliability score (median 18, IQR 18-20). The post hoc Dunn test revealed that educational channels produced by medical professionals attained significantly higher overall DISCERN scores than nonmedical educational ( $P=.01$ ), independent ( $P=.02$ ), internet media ( $P=.01$ ), and news channels ( $P<.001$ ). Additionally, both educational channels produced by medical and nonmedical professionals achieved significantly higher DISCERN reliability scores than those produced by independent nonmedical users ( $P<.001$  and  $P<.001$ ,

## Adherence With HONcode and DISCERN Principles

There was strong interrater agreement for both HONcode principles (median 94%, IQR 93%-97%; median  $\kappa=0.81$ , IQR 0.73-0.87) and DISCERN (median 88%, IQR 82-94%; median  $\kappa=0.83$ , IQR 0.76-0.91).

Forty-seven of 48 (98%) videos had either low (56%) or moderate (42%) adherence with HONcode principles. In general, videos scored poorly regarding the disclosure of financial sources and advertising (Table 3). Regarding the "authoritative" domain, only approximately half of the videos had an input from a medical professional or relevant scientist. Additionally, only a minority of videos fulfilled criteria relating to the "attribution," "justifiability," and "transparency" of the data presented.

respectively), internet media ( $P=.007$  and  $P=.04$ , respectively), and news channels ( $P<.001$  and  $P=.003$ , respectively).

Regression analysis revealed no significant association between engagement metrics and DISCERN or HONcode scores (Multimedia Appendix 1, Table S3). However, there was a significant positive association between DISCERN and HONcode ratings for all videos included in the study ( $P<.001$ ;  $r^2=.583$ ).

## Discussion

### Principal Findings

This study highlights the importance of YouTube as a medium for sharing of curated COVID-19-related information. We demonstrate growing public interest in extracting vaccine-related content from this resource, with the videos shortlisted in this study having been viewed over 30 million times globally thus

far, and an average of 1890 comments in the discussion thread of each video. The available content appears favorably received, with a mean of 14.3 likes per dislike per video. This study, however, demonstrates the varying quality of information provided on YouTube, with 98% of reviewed content with low to moderate adherence to HONcode principles and DISCERN reliability scores ranging from 18 (nonmedical individuals) to 35.5 (educational channels) out of 40.

Despite variable video quality, our search identified only 2 (4.2%) videos that would constitute mis- or disinformation, which accounted for only 1.3% of the total viewership. In comparison, studies evaluating misinformation on YouTube, published in March and June 2020, highlighted a significantly higher proportion of videos containing misleading or nonfactual information [12,13]. This is probably owing to YouTube's COVID-19 medical misinformation policy, which came into effect on October 14, 2020. YouTube now operates a “three strikes” system to prevent users from uploading unsubstantiated videos, which parallels the COVID-19 misinformation policies of Facebook and Twitter [14,15]. While the policy is explicitly directed at videos that contain unreal claims, such as the COVID-19 vaccines “kill people who receive them” or “contain a microchip,” tackling more insidious forms of misinformation in a timely manner has proven difficult. Recent criticism of these policies has highlighted their reliance on scientific consensus from health authorities to determine what exactly constitutes misinformation [16]. In such a rapidly developing field, with limited longitudinal evidence, this consensus cannot be readily achieved, which allows time for inaccurate social media content to be shared. This is of particular concern since previous studies on vaccine hesitancy have reported that videos of a negative tone are more likely to be shared and liked, thus perpetuating misinformation and confirmation biases [17]. In an effort to rapidly combat this, there has been increased government engagement and centralization of initiatives to reduce and prevent the spread of misinformation. The World Health Organization, in partnership with government agencies, has introduced several initiatives to improve public awareness of and to tackle vaccine-related misinformation (the so-called “infodemic”) on the internet [18,19]. Additionally, social media companies and the government of the United Kingdom have agreed on a battery of measures to reduce vaccine disinformation through rapid removal of flagged content and increased cooperation with public health bodies to ensure that authoritative messages regarding vaccine safety are disseminated to as many individuals as possible [20].

Along with user-directed policies, several other strategies have been suggested to limit the dissemination of false health information on social media platforms. These include mobilizing medical professionals as advocates to counter the propagation of misinformation [21]. Among the videos reviewed in our study, less than one-third were posted by nonprofit or medical organizations or medical professionals, which accounts for a lower proportion than that posted by news agencies. Furthermore, videos from established health-related organizations such as JAMA Network or World Health Organization only accounted for 25% of the total viewership. While there has been an exponential growth in medical

YouTubers, and despite the fact that videos produced by these individuals achieved the highest reliability and quality scores, concurrent with other studies, we found that their current role remains limited [22]. Of note, the most viewed shortlisted video was developed by JAMA Network, which may suggest the importance of brand recognition or marketing in attracting audiences.

In addition to the varied provenance of available vaccine-related content, we identified a paucity of reliable information on YouTube. Information in videos produced by reputable health-related organizations was significantly more reliable than that obtained from videos produced by only nonmedical individual users ( $P=.007$ ), and its reliability was comparable to videos from all other categories. The majority also only achieved “low adherence” to HONcode principles. Even though most of the videos produced by these nonprofit or medical organizations explained vaccine concepts in a clear and approachable manner, often utilizing the “infographic” format, they did not cite sources or provide links to further information, a common phenomenon in videos produced by established educational channels on YouTube. Thus, these videos were unable to fulfil both DISCERN reliability indicators (eg, “referencing of information” and “directing viewers to additional sources of knowledge”) and HONcode principles (eg, attribution and transparency) and were unable to attain high scores.

While nonfactual claims were limited to a small minority of videos, the absence of key information—particularly regarding basic vaccinology and the importance of concurrent public health measures—currently limits the utility of YouTube videos as robust sources of public health information. These findings echo those of previous studies reporting that reliability and quality of non-vaccine-related information on COVID-19 on YouTube is unsatisfactory [13,23]. As such, viewers are provided with incomplete evidence as to how the COVID-19 vaccine fits into the larger public health effort and are not provided with curated resources that could potentially provide these pertinent details.

### Limitations and Future Directions

There are some limitations to this study. First, the subset of videos examined were limited to the English language only, given that 88% of the videos were from the YouTube channels from the United States or the United Kingdom. While this represents a language bias and limits the generalizability of our findings to different languages and non-English-speaking countries, we note that similar findings have been reported with respect to COVID-19-related information in other languages [13].

Second, the search strategy was limited to 2 search phrases (“coronavirus vaccine” and “COVID-19 vaccine”). These phrases would not encompass the various searches the general public may make on this topic (eg, “covid vaccine,” “coronavirus vaccination,” or “covid vaccination”), which could yield different video results. Additionally, the search terms used were “neutral” and may not reflect searches made by individuals who (1) have already been previously subjected to misinformation, (2) have a network that shares similar

misinformation content, or (3) are part of groups that are more likely to seek misinformation.

Third, videos were first screened by title relevance and for the purpose of pragmatism, only those videos that had titles relevant to 5 domains of COVID-19 vaccine information (mechanisms of action of vaccines, clinical trial procedures, manufacturing processes, side effects and safety, and vaccine efficacy) were considered for full video analysis. However, videos with nonrelevant titles may still contain relevant information on COVID-19 vaccination and given that they are accessible by the public through neutral search terms, they could contribute toward the dissemination of incorrect or low-quality information. Additionally, a large majority of search results were excluded at the screening stage, which resulted in a relatively small sample size. This process may have introduced a selection bias, limiting the generalizability of our findings. For completeness of reporting in the future, all videos in the search results should be analyzed for low-quality or incorrect information.

Fourth, the search was conducted at a single timepoint (December 10, 2020), which was relatively early in the timeline of global COVID-19 vaccine distribution. Given the dynamic nature of the COVID-19 pandemic and vaccine development, knowledge and attitudes regarding vaccination may evolve with increased scientific understanding and public health interventions. Moreover, the search results used in this study led to the inclusion of videos produced in late March and April when vaccine development was still in its early stages. Therefore, topics such as vaccine manufacturing, efficacy, or safety were not discussed in this early sample. These videos instead concentrated on the explanation of vaccine science and the methodology of clinical trials. A cross-sectional analysis of YouTube videos through searches at multiple timepoints or of those stratified by the age of the video could be conducted in the future to assess the progression of video content and quality.

Fifth, although used in previous studies that assessed YouTube as a source of medical information, DISCERN and HONcode principles were developed and validated for the assessment of written medical information. However, a strong interrater agreement between the scoring systems suggests they are reasonable tools to use in the absence of a validated alternative.

Finally, although efforts were made to reduce selection bias by performing the search in an incognito window, the physical search location could still be revealed to YouTube through the IP address. As such, further studies should consider assessing the nature of content that users are exposed to at different locations, perhaps carrying out stratifying analysis using socioeconomic markers such as index of multiple deprivation. Additionally, routes of misinformation may vary depending on culture, education level, and even at a national level. Social and ethnic determinants have been demonstrated to impact vaccine hesitancy [24,25]. It is important to understand the drivers of vaccine hesitancy and develop high-quality, widely available educational resources to target these demographic groups and improve vaccine uptake. Creating videos in collaboration with medical professionals and taking advantage of YouTube's widespread reach represents one potential solution.

## Conclusions

Our findings demonstrate that YouTube videos produced by educational channels, especially those produced by medical professionals, achieve the highest quality and reliability metrics. Consistent with previous similar studies, this suggests that there is currently a missed opportunity in collaboration between respected health-related organizations and established educational YouTube content producers to disseminate high-quality information on COVID-19 vaccination [26]. This could potentially be a rapidly implementable public health intervention to engage a wider audience and increase public awareness and knowledge.

---

## Acknowledgments

JC acknowledges support from The Wellcome Trust (grant# 215938/Z/19/Z), which supports the Sir Henry Wellcome Postdoctoral fellowship. Infrastructure support for this study was provided by the Imperial Biomedical Research Centre, National Institute for Health Research.

---

## Conflicts of Interest

None declared.

---

Multimedia Appendix 1

Supplementary tables.

[DOCX File, 24 KB - [publichealth\\_v7i7e29942\\_app1.docx](#) ]

---

## References

1. COVID-19. Google Trends. URL: <https://trends.google.com/trends/explore?q=covid-19> [accessed 2020-12-17]
2. The top 500 sites on the web. Alexa. 2021. URL: <https://www.alexa.com/topsites> [accessed 2021-01-06]
3. YouTube for Press. YouTube About. URL: <https://www.youtube.com/intl/en-GB/about/press/> [accessed 2020-12-17]
4. Criddle C. Coronavirus: YouTube bans misleading Covid-19 vaccine videos. BBC News. 2020 Oct 14. URL: <https://www.bbc.co.uk/news/technology-54541517> [accessed 2020-12-17]
5. Coleman A. 'Hundreds dead' because of Covid-19 misinformation. BBC News. 2020 Aug 12. URL: <https://www.bbc.co.uk/news/world-53755067> [accessed 2021-01-31]

6. Keelan J, Pavri-Garcia V, Tomlinson G, Wilson K. YouTube as a source of information on immunization: a content analysis. *JAMA* 2007 Dec 05;298(21):2482-2484. [doi: [10.1001/jama.298.21.2482](https://doi.org/10.1001/jama.298.21.2482)] [Medline: [18056901](https://pubmed.ncbi.nlm.nih.gov/18056901/)]
7. Szilagyi PG, Thomas K, Shah MD, Vizueta N, Cui Y, Vangala S, et al. National Trends in the US Public's Likelihood of Getting a COVID-19 Vaccine-April 1 to December 8, 2020. *JAMA* 2020 Dec 29;325(4):396-398. [doi: [10.1001/jama.2020.26419](https://doi.org/10.1001/jama.2020.26419)] [Medline: [33372943](https://pubmed.ncbi.nlm.nih.gov/33372943/)]
8. Goobie GC, Guler SA, Johannson KA, Fisher JH, Ryerson CJ. YouTube Videos as a Source of Misinformation on Idiopathic Pulmonary Fibrosis. *Ann Am Thorac Soc* 2019 May;16(5):572-579. [doi: [10.1513/AnnalsATS.201809-644OC](https://doi.org/10.1513/AnnalsATS.201809-644OC)] [Medline: [30608877](https://pubmed.ncbi.nlm.nih.gov/30608877/)]
9. Loeb S, Sengupta S, Butaney M, Macaluso JN, Czarniecki SW, Robbins R, et al. Dissemination of Misinformative and Biased Information about Prostate Cancer on YouTube. *Eur Urol* 2019 Apr;75(4):564-567. [doi: [10.1016/j.eururo.2018.10.056](https://doi.org/10.1016/j.eururo.2018.10.056)] [Medline: [30502104](https://pubmed.ncbi.nlm.nih.gov/30502104/)]
10. Discover the 8 principles of the HONcode in 35 languages. Health On the Net Foundation. URL: <https://www.hon.ch/cgi-bin/HONcode/principles.pl?English> [accessed 2020-12-18]
11. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. *J Epidemiol Community Health* 1999 Mar;53(2):105-111 [FREE Full text] [doi: [10.1136/jech.53.2.105](https://doi.org/10.1136/jech.53.2.105)] [Medline: [10396471](https://pubmed.ncbi.nlm.nih.gov/10396471/)]
12. Moon H, Lee GH. Evaluation of Korean-Language COVID-19-Related Medical Information on YouTube: Cross-Sectional Infodemiology Study. *J Med Internet Res* 2020 Aug 12;22(8):e20775 [FREE Full text] [doi: [10.2196/20775](https://doi.org/10.2196/20775)] [Medline: [32730221](https://pubmed.ncbi.nlm.nih.gov/32730221/)]
13. Khatri P, Singh SR, Belani NK, Yeong YL, Lohan R, Lim YW, et al. YouTube as source of information on 2019 novel coronavirus outbreak: a cross sectional study of English and Mandarin content. *Travel Med Infect Dis* 2020;35:101636 [FREE Full text] [doi: [10.1016/j.tmaid.2020.101636](https://doi.org/10.1016/j.tmaid.2020.101636)] [Medline: [32205267](https://pubmed.ncbi.nlm.nih.gov/32205267/)]
14. Jin K. Keeping People Safe and Informed About the Coronavirus - About Facebook. Facebook, Inc. 2020 Dec 18. URL: <https://about.fb.com/news/2020/12/coronavirus/#misinformation-update> [accessed 2021-01-31]
15. COVID-19 medical misinformation policy. YouTube Help. URL: [https://support.google.com/youtube/answer/9891785?hl=en&hl=en&ref\\_topic=9282436](https://support.google.com/youtube/answer/9891785?hl=en&hl=en&ref_topic=9282436) [accessed 2021-01-31]
16. Wardle C, Singerman E. Too little, too late: social media companies' failure to tackle vaccine misinformation poses a real threat. *BMJ* 2021 Jan 21;372:n26. [doi: [10.1136/bmj.n26](https://doi.org/10.1136/bmj.n26)] [Medline: [33478950](https://pubmed.ncbi.nlm.nih.gov/33478950/)]
17. Donzelli G, Palomba G, Federigi I, Aquino F, Cioni L, Verani M, et al. Misinformation on vaccination: A quantitative analysis of YouTube videos. *Hum Vaccin Immunother* 2018 Jul 03;14(7):1654-1659 [FREE Full text] [doi: [10.1080/21645515.2018.1454572](https://doi.org/10.1080/21645515.2018.1454572)] [Medline: [29553872](https://pubmed.ncbi.nlm.nih.gov/29553872/)]
18. How to report misinformation online. World Health Organization. URL: <https://www.who.int/campaigns/connecting-the-world-to-combat-coronavirus/how-to-report-misinformation-online> [accessed 2021-02-01]
19. Be careful what you share. Things aren't always what they seem online. SHARE Checklist: Government of the United Kingdom. URL: <https://sharechecklist.gov.uk/> [accessed 2021-02-01]
20. Social media giants agree package of measures with UK Government to tackle vaccine disinformation. Government of the United Kingdom. 2020 Nov 08. URL: <https://www.gov.uk/government/news/social-media-giants-agree-package-of-measures-with-uk-government-to-tackle-vaccine-disinformation> [accessed 2021-02-01]
21. Trethewey SP. Strategies to combat medical misinformation on social media. *Postgrad Med J* 2020 Jan;96(1131):4-6 [FREE Full text] [doi: [10.1136/postgradmedj-2019-137201](https://doi.org/10.1136/postgradmedj-2019-137201)] [Medline: [31732511](https://pubmed.ncbi.nlm.nih.gov/31732511/)]
22. Stokel-Walker C. Covid-19: The doctors turned YouTubers. *BMJ* 2020 May 28;369:m1563. [doi: [10.1136/bmj.m1563](https://doi.org/10.1136/bmj.m1563)] [Medline: [32467087](https://pubmed.ncbi.nlm.nih.gov/32467087/)]
23. Dutta A, Beriwal N, Van Breugel LM, Sachdeva S, Barman B, Saikia H, et al. YouTube as a Source of Medical and Epidemiological Information During COVID-19 Pandemic: A Cross-Sectional Study of Content Across Six Languages Around the Globe. *Cureus* 2020 Jun 15;12(6):e8622 [FREE Full text] [doi: [10.7759/cureus.8622](https://doi.org/10.7759/cureus.8622)] [Medline: [32685293](https://pubmed.ncbi.nlm.nih.gov/32685293/)]
24. Bertonecello C, Ferro A, Fonzo M, Zanovello S, Napoletano G, Russo F, et al. Socioeconomic Determinants in Vaccine Hesitancy and Vaccine Refusal in Italy. *Vaccines (Basel)* 2020 Jun 05;8(2):276 [FREE Full text] [doi: [10.3390/vaccines8020276](https://doi.org/10.3390/vaccines8020276)] [Medline: [32516936](https://pubmed.ncbi.nlm.nih.gov/32516936/)]
25. Robertson E, Reeve KS, Niedzwiedz CL, Moore J, Blake M, Green M, et al. Predictors of COVID-19 vaccine hesitancy in the UK household longitudinal study. *Brain Behav Immun* 2021 May;94:41-50 [FREE Full text] [doi: [10.1016/j.bbi.2021.03.008](https://doi.org/10.1016/j.bbi.2021.03.008)] [Medline: [33713824](https://pubmed.ncbi.nlm.nih.gov/33713824/)]
26. Li HO, Bailey A, Huynh D, Chan J. YouTube as a source of information on COVID-19: a pandemic of misinformation? *BMJ Glob Health* 2020 May;5(5):e002604 [FREE Full text] [doi: [10.1136/bmjgh-2020-002604](https://doi.org/10.1136/bmjgh-2020-002604)] [Medline: [32409327](https://pubmed.ncbi.nlm.nih.gov/32409327/)]

## Abbreviations

**HONcode:** Health on the Net Foundation Code of Conduct

*Edited by T Sanchez; submitted 26.04.21; peer-reviewed by S Trethewey, A Shamsi, H Moon; comments to author 12.05.21; revised version received 17.05.21; accepted 21.05.21; published 08.07.21.*

*Please cite as:*

*Chan C, Sounderajah V, Daniels E, Acharya A, Clarke J, Yalamanchili S, Normahani P, Markar S, Ashrafian H, Darzi A  
The Reliability and Quality of YouTube Videos as a Source of Public Health Information Regarding COVID-19 Vaccination:  
Cross-sectional Study*

*JMIR Public Health Surveill 2021;7(7):e29942*

*URL: <https://publichealth.jmir.org/2021/7/e29942>*

*doi: [10.2196/29942](https://doi.org/10.2196/29942)*

*PMID: [34081599](https://pubmed.ncbi.nlm.nih.gov/34081599/)*

©Calvin Chan, Viknesh Sounderajah, Elisabeth Daniels, Amish Acharya, Jonathan Clarke, Seema Yalamanchili, Pasha Normahani, Sheraz Markar, Hutan Ashrafian, Ara Darzi. Originally published in JMIR Public Health and Surveillance (<https://publichealth.jmir.org>), 08.07.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

# Measuring Public Concern About COVID-19 in Japanese Internet Users Through Search Queries: Infodemiological Study

Zhiwei Gao<sup>1</sup>, BEng; Sumio Fujita<sup>2</sup>, PhD; Nobuyuki Shimizu<sup>2</sup>, PhD; Kongmeng Liew<sup>1</sup>, PhD; Taichi Murayama<sup>1</sup>, MSc; Shuntaro Yada<sup>1</sup>, PhD; Shoko Wakamiya<sup>1</sup>, PhD; Eiji Aramaki<sup>1</sup>, PhD

<sup>1</sup>Social Computing Laboratory, Graduate School of Science and Technology, Nara Institute of Science and Technology, Ikoma, Japan

<sup>2</sup>Yahoo Japan Corporation, Tokyo, Japan

**Corresponding Author:**

Eiji Aramaki, PhD

Social Computing Laboratory

Graduate School of Science and Technology

Nara Institute of Science and Technology

Takayama-cho 8916-5

Ikoma, Nara 630-0192

Japan

Phone: 81 743 72 6053

Email: [aramaki@is.naist.jp](mailto:aramaki@is.naist.jp)

## Abstract

**Background:** COVID-19 has disrupted lives and livelihoods and caused widespread panic worldwide. Emerging reports suggest that people living in rural areas in some countries are more susceptible to COVID-19. However, there is a lack of quantitative evidence that can shed light on whether residents of rural areas are more concerned about COVID-19 than residents of urban areas.

**Objective:** This infodemiology study investigated attitudes toward COVID-19 in different Japanese prefectures by aggregating and analyzing Yahoo! JAPAN search queries.

**Methods:** We measured COVID-19 concerns in each Japanese prefecture by aggregating search counts of COVID-19–related queries of Yahoo! JAPAN users and data related to COVID-19 cases. We then defined two indices—the localized concern index (LCI) and localized concern index by patient percentage (LCIPP)—to quantitatively represent the degree of concern. To investigate the impact of emergency declarations on people's concerns, we divided our study period into three phases according to the timing of the state of emergency in Japan: before, during, and after. In addition, we evaluated the relationship between the LCI and LCIPP in different prefectures by correlating them with prefecture-level indicators of urbanization.

**Results:** Our results demonstrated that the concerns about COVID-19 in the prefectures changed in accordance with the declaration of the state of emergency. The correlation analyses also indicated that the differentiated types of public concern measured by the LCI and LCIPP reflect the prefectures' level of urbanization to a certain extent (ie, the LCI appears to be more suitable for quantifying COVID-19 concern in urban areas, while the LCIPP seems to be more appropriate for rural areas).

**Conclusions:** We quantitatively defined Japanese Yahoo users' concerns about COVID-19 by using the search counts of COVID-19–related search queries. Our results also showed that the LCI and LCIPP have external validity.

(*JMIR Public Health Surveill* 2021;7(7):e29865) doi:[10.2196/29865](https://doi.org/10.2196/29865)

**KEYWORDS**

COVID-19; search query; infodemiology; quantitative analysis; concern; rural; urban; Internet; information-seeking behavior; attitude; Japan

## Introduction

The COVID-19 pandemic has been threatening global health since the end of December 2019. The outbreak has created critical challenges for public health, research, and medical

communities [1]. As of July 12, 2021, COVID-19 has affected 220 countries and territories, with over 187 million confirmed cases, and has claimed over 4 million lives [2]. COVID-19 has also disrupted many lives and caused psychological trauma on a large scale [3,4].



As with any outbreak of an infectious disease, the population's psychological reactions play a critical role in shaping the spread of the disease and the occurrence of emotional distress and social disorder during and after the outbreak [5]. Recently, Ahorsu et al [6] developed the Fear of COVID-19 Scale (FCV-19S) by conducting qualitative interviews to assess individuals' fear of COVID-19. Gao et al [7] found that greater concern about COVID-19 (frequent exposure to COVID-19-related social media) was positively associated with adverse mental health outcomes. Furthermore, Su et al's [8] Twitter-based analysis revealed that spatial-temporal and socioeconomic disparities shaped US residents' response to COVID-19.

Infodemiology is the science of distribution and determinants of information in an electronic medium, specifically the internet, or in a population, with the ultimate aim to inform public health and public policy [9]. The underlying objective of this field is to measure the pulse of public opinion, attention, behavior, knowledge, and attitudes by tracking what people do and write on the internet [10], such as by analyzing queries from internet search engines to predict disease outbreaks. Bernardo et al [11] published a scoping review in which they assessed the current state of knowledge regarding the use of search queries and social media for disease surveillance, showing their usability. Daughton et al [12] conducted an infodemiology study using social media data from Twitter to identify human behaviors associated with COVID-19 transmission and the perceived impact of COVID-19 on individuals. Mavragani et al [13] provided a methodological framework for using Google Trends in infodemiology by analyzing the value and validity of using Google Trends data.

Worryingly, use of the internet and social media has increased dramatically due to the enforcement of social distancing and stay-at-home orders in many areas, which has led to more search queries for updates on the local COVID-19 situation [14]. However, there is a lack of quantitative evidence about the relationship between these searches for updates and the psychological reactions of the population toward COVID-19. Additionally, one may think that residents of urban areas would show more concern about COVID-19 due to larger crowds and easier access to public transportation. However, concern about COVID-19 is also prevalent in rural areas, given that people who live in rural areas may be more vulnerable to COVID-19 than residents of urban areas [15,16].

Therefore, we aimed to quantitatively reflect the different types of concern in rural and urban areas by analyzing the Japanese public's psychological reactions toward COVID-19 using an infodemiology approach (ie, concern about COVID-19 in search queries). We used Yahoo! JAPAN to get the search queries in this study because it is the largest portal site in Japan, with a significant user base; there are about 70 million and 20 million monthly active users for smartphone and PC, respectively. This covers about 85% of smartphone users and 61% of PC users in Japan. The proportion of female and male smartphone users is

52% and 48%, respectively. About 32% of users are aged 20-39 years, 40% of users are aged 40-59 years, and 28% of users are aged  $\geq 60$  years [17]. We first developed two concern indices based on search queries to measure generalized COVID-19 concerns, which we called the localized concern index (LCI) and the localized concern index by patient percentage (LCIPP). We then used these indices to investigate COVID-19 concerns in relation to prefecture urbanization. To evaluate the feasibility of these concern indices, we examined the prefecture-level correlations with several indicators of ruralization and public health outcomes. The Methods section details the process of defining the LCI and LCIPP equations.

## Methods

### Target Queries

First, we explored people's COVID-19 concerns by analyzing search queries over different time periods. We initially established a baseline of common search queries prior to and during the pandemic by selecting the search queries of Yahoo! JAPAN's users from April to May 2019 and from April to May 2020 (we chose to examine queries during April and May due to the Japanese government's declaration of a state of emergency in April 2020 [18]). Additionally, since older adults (those aged  $>65$  years) are at a significantly greater risk of adverse COVID-19 outcomes [19], we speculated that there might be more COVID-19-related search queries from this group. Therefore, we started the analysis by targeting the search queries of people over 65 years of age. We extracted 100,000 search queries of this older population for the two aforementioned time periods and ranked them in reverse order according to the search counts. A change in time period results in a change in search query ranking; thus, we quantified the difference in ranking by defining the following rank change index:

$$\frac{R_{2020} - R_{2019}}{R_{2019} + 2}$$

The larger the index, the fewer the search counts for this query in 2019 or the greater the number of search counts in 2020, and vice versa. A constant of two was added to both the numerator and denominator, to avoid instances of zeros (uncountable) in the search counts. Table 1 shows the top five rapidly increasing search queries in April 2020 compared to April 2019. As expected, these terms appeared to be COVID-19-related.

Consistently, of the top 100 queries in the ascending query list, 76 queries contained COVID-19-related keywords (eg, コロナ [corona], マスク [mask]); of these, 33 queries contained prefecture names combined with コロナ感染者 (corona cases). A query pattern, such as the prefecture name plus "coronavirus cases," clearly displays the prefecture's information and it also reflects the user's concern about COVID-19 to some extent. Therefore, we chose this query pattern as our target query and sorted the queries by region according to the prefecture names mentioned in the query. Table 2 lists some query samples used in this study.

**Table 1.** Top five rapidly ascending search queries and their rank change index in April 2020 compared to April 2019. The first one, “novel\_coronavirus,” is the original data we retrieved from the database.

Search query	Rank change index
novel_coronavirus	18.87
シャープ マスク (Sharp's face mask)	18.74
新型コロナウイルス (novel coronavirus)	17.99
コロナ 感染者数 (corona cases)	17.96
東京都 コロナウイルス感染者 (Tokyo coronavirus cases)	17.63

**Table 2.** Samples of target queries.

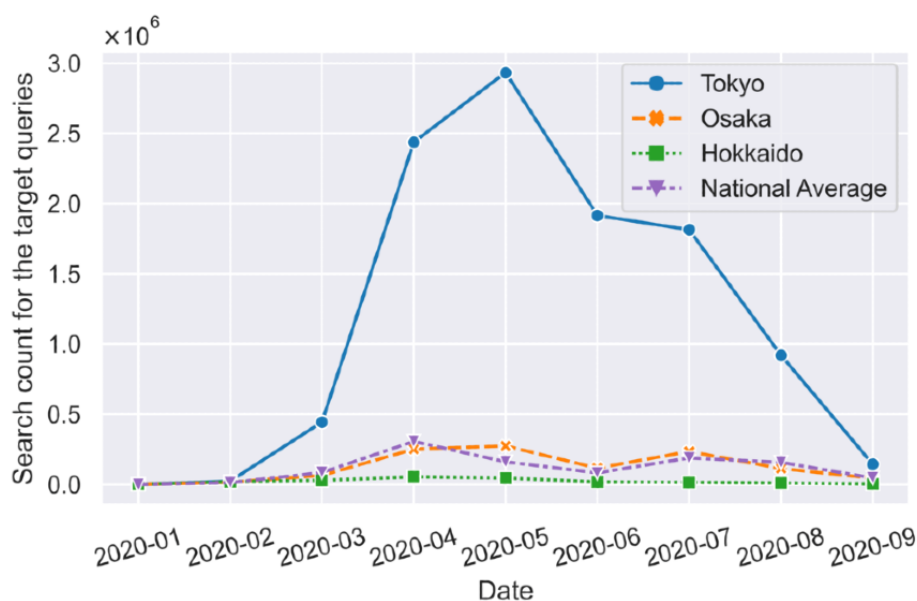
Search query	Rank change index
東京都 コロナウイルス感染者 (Tokyo coronavirus cases)	17.63
神奈川県 コロナ感染者 (Kanagawa corona cases)	16.26
埼玉県 コロナウイルス感染者 (Saitama coronavirus cases)	16.18
福岡県 コロナウイルス感染者 (Fukuoka coronavirus cases)	16.11
茨城県 コロナウイルス感染者 (Ibaraki coronavirus cases)	15.84

## Baseline Queries

However, when we calculated the search counts of target queries that comprised the prefecture name and “coronavirus cases” from January to September 2020, we found that the search counts in Tokyo were much higher than in the other prefectures, as shown in Figure 1. It is possible that Tokyo’s larger population meant that the frequencies of the general search counts were higher than in less populated prefectures. We speculated that this excessive disparity would inevitably have an impact on our subsequent calculations. To mitigate this effect, we introduced baseline queries for each prefecture that were frequently searched for in that prefecture and that had a relatively stable search count for a certain period. The baseline queries contained the prefecture name and “X,” where “X”

referred to any keywords as long as the sample variance of their monthly search counts was as small as possible. For example, for Tokyo, 東京 23区 (Tokyo 23 wards) and 東京 天気 過去 (past weather in Tokyo) were some of the baseline queries, and their sample variances of search counts in these nine months were 499.6 and 683, respectively. In summary, we compared the search counts of possible queries with the above pattern from January to September 2020, and then identified the top three queries with the smallest sample variance during those nine months to form the baseline queries. This was done to balance the impact of excessive disparities by using the quotient of the baseline query and our target query, as the search counts of the target queries were high for urbanized prefectures (such as Tokyo), where the search count frequencies of the baseline queries were also high.

**Figure 1.** Search counts of targeted queries in Tokyo, Osaka, and Hokkaido, as well as the national average from January to September 2020.



## COVID-19 Concern Indices

Next, our query-based equation to quantify the level of concern about COVID-19—the LCI equation—was defined for each prefecture *pref* as follows:

$$LCI_{pref} = \frac{Count(tq_{pref})}{Count(bq_{pref})}$$

where  $tq_{pref}$  and  $bq_{pref}$  are the target query and the baseline query, respectively, and  $Count(.)$  is a function that counts the occurrences of a query. We used the logarithmic result for the LCI calculation, which suggests that a higher LCI means a higher frequency of baseline-controlled, prefecture-specified COVID-19-related queries, which in turn reflects a greater level of concern about COVID-19 in that prefecture.

However, the LCI appeared to be greatly influenced by the severity of the COVID-19 situation in each prefecture. For example, urbanized areas, such as Tokyo and Osaka, which are the first and second most populous cities in Japan and the hardest hit, naturally had the highest search counts for target queries and a higher LCI. The COVID-19 pandemic has raised people's anxiety levels [20], and the LCI appears to reflect this attitude in hard-hit, largely urban areas. However, heightened concern about COVID-19 in rural areas [21] does not seem to be reflected in the LCI. Rather, given the lower number of cases in rural areas, it is likely that concern about COVID-19 in rural areas may be related to socioecological variables unrelated to the risk of infection.

For this reason, we attempted to improve our LCI equation to be able to examine COVID-19 concerns beyond the direct influence of actual COVID-19 cases. We argue that this would quantify the prevalence of concern about COVID-19 beyond the risk of actual infection. To do so, we modified the LCI to account for the number of new cases per month in each prefecture by the population of that prefecture to calculate the percentage of infected patients; that is, *Number of monthly new cases/population*.

In summary, the LCIPP for each prefecture *pref* was calculated as follows:

$$LCIPP_{pref} = \frac{LCI_{pref}}{Number\ of\ monthly\ new\ cases/population}$$

Similarly, we used the logarithmic result for the LCIPP calculation.

## Correlation Indicators

To examine whether the differentiated patterns of associations for the LCI and LCIPP reflected the COVID-19 concerns in rural and urban areas, we examined prefecture-level correlations with both indices. Accordingly, we included the corresponding measures and examined the following aspects, with the corresponding indicators in parentheses: (1) the prevalence of farming (number of farming households, percentage of farmland, rice production), (2) population change (rate of population change), (3) ease of accessibility (reachable areas within one

hour from major stations of each prefecture, travel time from Tokyo to major stations in each prefecture), and (4) public health outcomes (reported symptoms, daily outpatients, COVID-19 cases per million residents).

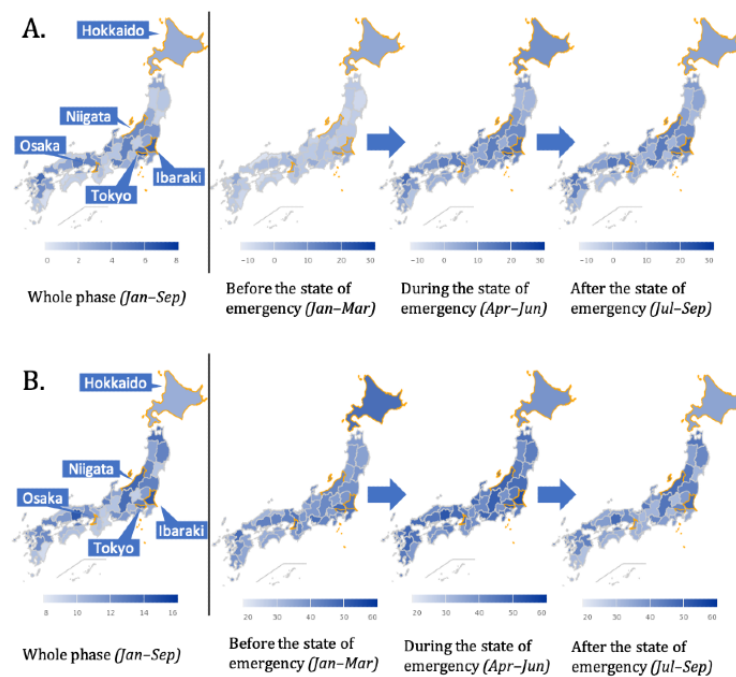
We argue that rural areas can be identified by the higher prevalence of farming households, more farming area, and higher rice production. Conversely, urban areas can be identified by their higher population density, ease of accessibility (the proportion of reachable areas within one hour from major stations of each prefecture), and rate of population change (urban areas should exhibit population growth, while rural areas should show a population decline). Finally, prefectural public health was measured as the number of cumulative COVID-19 cases per 1 million residents (as of September 30, 2020), the proportion of the population that reported general symptoms (non-COVID-19-related), the number of ambulance dispatches, and the number of daily outpatients in 2019. Public health outcomes are able to reflect the current state of local health care to some extent. Since COVID-19 outbreaks would take over existing health care resources, we aimed to find the relationship between public concern about COVID-19 and the current state of local health care through the public health outcomes indicators.

## Results

### LCI and LCIPP Results by Prefecture

Figure 2A and Figure 2B show the geographical results of the LCI and LCIPP from January to September 2020, respectively. Table 3 shows the top prefectures with the highest and lowest LCI and LCIPP. These are also shown as the darkest and lightest regions in Figure 2, respectively. In addition, we examined the LCI and LCIPP across three phases following the timing of the state of emergency: (1) before the state of emergency (January–March), (2) during the state of emergency (April–June), and (3) after the state of emergency (July–September). For a breakdown of the LCI and LCIPP by prefecture, please refer to our online supplementary material [22]. The LCI results show that before the declaration of the state of emergency, there was a low level of concern about COVID-19 nationwide. However, with the state of emergency declaration, some prefectures (including Tokyo and Ibaraki) began to show a high level of concern. At the end of the state of emergency, the overall level of concern in the country did not change much, though prefectures such as Hokkaido showed a slight decrease. The LCIPP results show a high level of concern for COVID-19 in Hokkaido even before the national state of emergency was declared. This is because Hokkaido activated COVID-19 response measures at an earlier stage (March 2020) [23], which resulted in a high level of concern for COVID-19 among local residents. With the declaration of a nationwide state of emergency in April 2020, most prefectures also showed a high level of concern, and with the end of the state of emergency, the level of concern declined accordingly.

**Figure 2.** Geographical results. The prefectures with orange boundary lines are the ones mentioned in this study: Hokkaido, Niigata, Ibaraki, Tokyo, and Osaka. (A) Localized concern index (LCI) results. (B) Localized concern index by patient percentage (LCCIP) results.



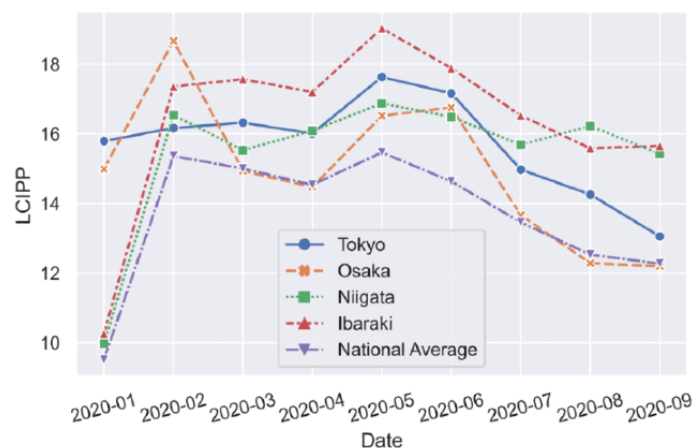
**Table 3.** Prefectures with the highest and lowest localized concern index and localized concern index by patient percentage.

Search query	Rank change index
東京都 コロナウイルス感染者 (Tokyo coronavirus cases)	17.63
神奈川県 コロナ感染者 (Kanagawa corona cases)	16.26
埼玉県 コロナウイルス感染者 (Saitama coronavirus cases)	16.18
福岡県 コロナウイルス感染者 (Fukuoka coronavirus cases)	16.11
茨城県 コロナウイルス感染者 (Ibaraki coronavirus cases)	15.84

Of all the prefectures, we selected four to be displayed in Figure 3: Tokyo, Osaka, Niigata, and Ibaraki. We used the number of farming households as a criterion for distinguishing between urban and rural areas. According to a survey by the Statistics Bureau of Japan [24], Tokyo and Osaka have the lowest numbers of farming households in Japan, while Ibaraki and Niigata have

the highest numbers. This suggests that Niigata and Ibaraki have relatively lower urbanization rates than Tokyo and Osaka. Nevertheless, Figure 3 indicates that some urban areas (ie, Tokyo and Osaka) have a relatively lower LCIPP than some rural areas (ie, Ibaraki) in terms of the general trends.

**Figure 3.** LCIPP results of Tokyo, Osaka, Niigata, and Ibaraki, as well as the national average from January to September 2020. LCIPP: localized concern index by patient percentage.



### Correlations with the LCI and LCIPP

By calculating the correlation coefficients of LCI and LCIPP on the aforementioned indicators, our results showed that the LCI was significantly correlated with the COVID-19 infection risk (cumulative cases) and urbanized prefectures. The latter was identified through population density and ease of

accessibility. The LCI was higher in prefectures that have a high proportion of reachable areas within one hour from their major stations and in prefectures that had growing populations. In contrast, the LCIPP was higher in prefectures that were more rural, as determined by their higher number of farming households and rice production. Table 4 displays the correlation results.

**Table 4.** Correlation coefficients of the whole phase localized concern index and localized concern index by patient percentage and some indicators.

Index and prefecture	Value
<b>Highest localized concern index</b>	
Tokyo	7.04
Ibaraki	5.90
Fukuoka	5.54
Saitama	5.31
Okayama	4.92
<b>Lowest localized concern index</b>	
Miyazaki	0.63
Ehime	0.87
Mie	0.96
Miyagi	1.02
Wakayama	1.07
<b>Highest localized concern index by patient percentage</b>	
Okayama	14.36
Ibaraki	14.31
Niigata	13.78
Nagano	13.48
Aomori	13.46
<b>Lowest localized concern index by patient percentage</b>	
Miyazaki	8.68
Ehime	9.18
Gunma	9.23
Okinawa	9.30
Shiga	9.43

### Age-Based LCIPP Results

We also investigated the LCIPP results for three different age groups: (1) age 25-44 years, (2) age 45-64 years, and (3) age >65 years. At this point, *population* in the LCIPP equation was

replaced by the number of people in each age group. Figure 4 shows that older people over the age of 65 years had fewer concerns than those aged 25-44 years and 45-64 years, while those aged 25-44 years and 45-64 years had almost the same level of concern about COVID-19.

**Figure 4.** LCIPP results for those aged 25-44 years, 45-64 years, and over 65 years. LCIPP: localized concern index by patient percentage.



## Discussion

### Principal Findings

#### *Nature of the LCI and LCIPP*

As expected, the LCI appeared to represent the overall concern about COVID-19. To a large extent, this was influenced by the actual prevalence of COVID-19 within the prefecture. Similarly, concern about COVID-19 was heightened in prefectures that are dense, highly accessible, and growing in population. Most likely, these are reflections of urbanized prefectures that have a highly developed infrastructure and that attract the migration of younger workers due to a higher number of job opportunities. As such, these fast-paced, young, and dense prefectures naturally present a greater risk of COVID-19 infection, and the increased concern shown in the search queries that was observed in the LCI is not surprising.

However, once we removed the variance explained by the inclusion of daily increases in COVID-19 cases from the equation, we found the opposite effect. The LCIPP ceased to reflect the COVID-19 risk in urban areas, and there was no significant relationship with the cumulative COVID-19 cases. Furthermore, the pattern of significant correlations revealed an association with the rural prefectures. The larger the proportion of farming households and the greater the rice production, the higher the public concern about COVID-19, which is beyond that explained by the risk of infection. Why does the LCIPP reflect increased ruralization? We posit an explanation in the form of collectivism afforded by farming, specifically in rice-farming societies [25,26]. These societies tend to be more collectivistic and residents have a greater psychological desire to protect the community from internal and external threats. COVID-19 is one such threat; thus, rural communities may have greater vigilance and concern about preventing COVID-19 from becoming prevalent in their community. This is also consistent with prior studies that have established links between COVID-19

concern and collectivism [27] and collectivism with COVID-19 prevention behavior [28].

This explanation has potential implications for policy. Specifically, the discrepancy between the LCI and LCIPP shows that more care must be taken when comparing public attitudes toward COVID-19 between rural and urban communities. If the LCIPP is indeed a result of collectivism and greater vigilance against COVID-19, this may also indicate a broader adoption of preventive measures, such as handwashing and mask wearing. In contrast, the LCI does not seem to be similarly indicative of such concerns. For greater effectiveness, public campaigns that promote such behaviors should therefore use different strategies when targeting rural and urban communities.

Interestingly, we note that the public health measures in Table 3 did not correlate significantly with the LCIPP. This suggests that the preexisting or general health of a prefecture's population does not appear to affect public concern about COVID-19.

#### *Preliminary Analysis of the Age-Based Results*

Japan, which is one of the fastest aging countries in the world, has the highest proportion of older people worldwide [29]. Emerging studies suggested that older people are more susceptible to COVID-19 and likely to have poor outcomes [19,30,31]. However, we found that individuals aged over 65 years had reduced LCIPP scores, which suggests that their concern might be lower than that of those aged 25-44 years and 45-64 years. Alternatively, this could be a consequence of internet literacy, as users above 65 years of age may have less proficiency in using the internet, or are simply less accustomed to using search terms and queries for topics of concern. However, more research is needed to contextualize this result.

#### *Effectiveness of Search Queries as Public Concern Indicators*

Finally, we evaluated the usefulness of our method by extracting search queries and combining them with actual COVID-19

infection rates to quantify public concern. Our results showed the correlations for the LCI and LCIPP demonstrate external validity, as they were both associated with constructs that could be explained by previous research. This study joins a growing body of literature that uses web-based search queries to track public health (eg, Murayama et al [32]), and the LCIPP adds the dimension of using prefecture-level infection rates to control for expected outcomes. Therefore, this study was able to effectively quantify public concern at a deeper level, which we propose is explained by the collectivistic psychological tendencies of a society.

### Limitations

We note the limitations of our approach. Our method of extracting prefecture information from search queries relied on searches for the prefecture name + “coronavirus cases,” and not location-based information, such as IP addresses. This may not necessarily represent queries that are only from residents of these prefectures; it may also reflect queries from nonresidents who are interested in the COVID-19 situation in these prefectures (eg, a user who may have family in these prefectures). In addition, although Yahoo! JAPAN has a large user base in Japan, in the broader search engine market, Google

still has a significant share. Search queries and results might vary between search engines due to user preferences and service provider settings. Therefore, from the perspective of data diversity, it would be a better choice to consider a combination of search queries from multiple search engines.

### Conclusions

In summary, this study used search queries from Yahoo! JAPAN users to quantify the degree of concern about COVID-19 in rural and urban areas. We first established that Yahoo search queries could be used to quantify COVID-19–related concerns. We then defined the LCI and LCIPP as quantitative indicators of prefecture-level COVID-19 concern. The LCI was indicative of COVID-19 concern in urban areas, whereas the LCIPP appeared to be indicative of COVID-19 concern in rural prefectures. By investigating the relationships between these concern indices and prefecture-level information, we showed that the LCI and LCIPP have external validity. These results suggest that one potential application could be conducting differentiated public campaigns about COVID-19 prevention and misinformation. In light of the different sources of concern in rural and urban areas, such campaigns could adopt different risk communication strategies in these areas.

### Acknowledgments

This study was supported in part by JSPS KAKENHI Grant Number JP19K20279 and Health and Labor Sciences Research Grant Number H30-shinkougousei-shitei-004.

### Conflicts of Interest

SF is a current employee of Yahoo Japan Corporation, which provides the search log data analyzed in the paper. The other authors declare no conflicts of interest.

### References

1. Fauci AS, Lane HC, Redfield RR. Covid-19 - Navigating the Uncharted. *N Engl J Med* 2020 Mar 26;382(13):1268-1269 [FREE Full text] [doi: [10.1056/NEJMe2002387](https://doi.org/10.1056/NEJMe2002387)] [Medline: [32109011](https://pubmed.ncbi.nlm.nih.gov/32109011/)]
2. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases* 2020 May;20(5):533-534. [doi: [10.1016/s1473-3099\(20\)30120-1](https://doi.org/10.1016/s1473-3099(20)30120-1)]
3. Wang C, Tee M, Roy AE, Fardin MA, Srichokchatchawan W, Habib HA, et al. The impact of COVID-19 pandemic on physical and mental health of Asians: A study of seven middle-income countries in Asia. *PLoS One* 2021 Feb 11;16(2):e0246824 [FREE Full text] [doi: [10.1371/journal.pone.0246824](https://doi.org/10.1371/journal.pone.0246824)] [Medline: [33571297](https://pubmed.ncbi.nlm.nih.gov/33571297/)]
4. Zacher H, Rudolph CW. Individual differences and changes in subjective wellbeing during the early stages of the COVID-19 pandemic. *Am Psychol* 2021 Jan 23;76(1):50-62. [doi: [10.1037/amp0000702](https://doi.org/10.1037/amp0000702)] [Medline: [32700938](https://pubmed.ncbi.nlm.nih.gov/32700938/)]
5. Cullen W, Gulati G, Kelly BD. Mental health in the COVID-19 pandemic. *QJM* 2020 May 01;113(5):311-312 [FREE Full text] [doi: [10.1093/qjmed/hcaa110](https://doi.org/10.1093/qjmed/hcaa110)] [Medline: [32227218](https://pubmed.ncbi.nlm.nih.gov/32227218/)]
6. Ahorsu DK, Lin C, Imani V, Saffari M, Griffiths MD, Pakpour AH. The Fear of COVID-19 Scale: Development and Initial Validation. *Int J Ment Health Addict* 2020 Mar 27:1-9 [FREE Full text] [doi: [10.1007/s11469-020-00270-8](https://doi.org/10.1007/s11469-020-00270-8)] [Medline: [32226353](https://pubmed.ncbi.nlm.nih.gov/32226353/)]
7. Gao J, Zheng P, Jia Y, Chen H, Mao Y, Chen S, et al. Mental health problems and social media exposure during COVID-19 outbreak. *PLoS One* 2020;15(4):e0231924 [FREE Full text] [doi: [10.1371/journal.pone.0231924](https://doi.org/10.1371/journal.pone.0231924)] [Medline: [32298385](https://pubmed.ncbi.nlm.nih.gov/32298385/)]
8. Su Y, Venkat A, Yadav Y, Puglisi LB, Fodeh SJ. Twitter-based analysis reveals differential COVID-19 concerns across areas with socioeconomic disparities. *Comput Biol Med* 2021 May;132:104336. [doi: [10.1016/j.compbiomed.2021.104336](https://doi.org/10.1016/j.compbiomed.2021.104336)] [Medline: [33761419](https://pubmed.ncbi.nlm.nih.gov/33761419/)]
9. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. *J Med Internet Res* 2009 Mar 27;11(1):e11 [FREE Full text] [doi: [10.2196/jmir.1157](https://doi.org/10.2196/jmir.1157)] [Medline: [19329408](https://pubmed.ncbi.nlm.nih.gov/19329408/)]
10. Eysenbach G. Infodemiology and infoveillance tracking online health information and cyberbehavior for public health. *Am J Prev Med* 2011 May;40(5 Suppl 2):S154-S158. [doi: [10.1016/j.amepre.2011.02.006](https://doi.org/10.1016/j.amepre.2011.02.006)] [Medline: [21521589](https://pubmed.ncbi.nlm.nih.gov/21521589/)]

11. Bernardo TM, Rajic A, Young I, Robiadek K, Pham MT, Funk JA. Scoping review on search queries and social media for disease surveillance: a chronology of innovation. *J Med Internet Res* 2013 Jul 18;15(7):e147 [FREE Full text] [doi: [10.2196/jmir.2740](https://doi.org/10.2196/jmir.2740)] [Medline: [23896182](https://pubmed.ncbi.nlm.nih.gov/23896182/)]
12. Daughton AR, Shelley CD, Barnard M, Gerts D, Watson Ross C, Crooker I, et al. Mining and Validating Social Media Data for COVID-19-Related Human Behaviors Between January and July 2020: Infodemiology Study. *J Med Internet Res* 2021 May 25;23(5):e27059 [FREE Full text] [doi: [10.2196/27059](https://doi.org/10.2196/27059)] [Medline: [33882015](https://pubmed.ncbi.nlm.nih.gov/33882015/)]
13. Mavragani A, Ochoa G. Google Trends in Infodemiology and Infoveillance: Methodology Framework. *JMIR Public Health Surveill* 2019 May 29;5(2):e13439 [FREE Full text] [doi: [10.2196/13439](https://doi.org/10.2196/13439)] [Medline: [31144671](https://pubmed.ncbi.nlm.nih.gov/31144671/)]
14. Koeze E, Popper N. The Virus Changed the Way We Internet. *The New York Times*. 2020 Apr 07. URL: <https://www.nytimes.com/interactive/2020/04/07/technology/coronavirus-internet-use.html> [accessed 2021-01-25]
15. Leatherby L. The Worst Virus Outbreaks in the U.S. Are Now in Rural Areas. *The New York Times*. 2020 Oct 22. URL: <https://www.nytimes.com/interactive/2020/10/22/us/covid-rural-us.html> [accessed 2021-01-25]
16. Singh KD, Gettleman J. 'Rural Surge' Propels India Toward More Covid-19 Infections Than U.S. *Japan Times*. 2020 Oct 9. URL: <https://www.japantimes.co.jp/news/2020/10/09/asia-pacific/rural-surge-india-coronavirus-infections/> [accessed 2021-01-25]
17. Yahoo! Japan Overview. Yahoo Japan Corporation. 2021. URL: [https://s.yimg.jp/images/listing/pdfs/yj\\_mediaguide\\_eng.pdf](https://s.yimg.jp/images/listing/pdfs/yj_mediaguide_eng.pdf) [accessed 2021-05-26]
18. COVID-19 Guidelines for Lifting the State of Emergency. Prime Minister of Japan and His Cabinet. 2020 May 14. URL: [https://japan.kantei.go.jp/ongoingtopics/\\_00025.html](https://japan.kantei.go.jp/ongoingtopics/_00025.html) [accessed 2021-01-18]
19. Mueller AL, McNamara MS, Sinclair DA. Why does COVID-19 disproportionately affect older people? *Aging (Albany NY)* 2020 May 29;12(10):9959-9981 [FREE Full text] [doi: [10.18632/aging.103344](https://doi.org/10.18632/aging.103344)] [Medline: [32470948](https://pubmed.ncbi.nlm.nih.gov/32470948/)]
20. Gamonal LR, Mateos E, Fernández J, Roncero C. Anxiety, worry and perceived stress in the world due to the COVID-19 pandemic, March 20. Preliminary results. medRxiv. Preprint published online on April 6, 2020. [doi: [10.1101/2020.04.03.20043992](https://doi.org/10.1101/2020.04.03.20043992)]
21. (Japanese) Returning home after the Corona disaster. Seidansha. 2020 Sep 01. URL: <https://bunshun.jp/articles/-/39985> [accessed 2020-01-25]
22. Gao Z. LCI and LCIPP Scores. URL: <https://aoi.naist.jp/concern> [accessed 2021-01-30]
23. Hokkaido set to lift coronavirus state of emergency. *Kyodo*. 2020 Mar 19. URL: <https://www.japantimes.co.jp/news/2020/03/19/national/hokkaido-set-to-lift-coronavirus-emergency/> [accessed 2021-01-23]
24. Japan Statistical Yearbook 2020. Statistics Bureau of Japan. 2020. URL: <https://www.stat.go.jp/english/data/nenkan/69nenkan/1431-08.html> [accessed 2021-01-20]
25. Liu SS, Morris MW, Talhelm T, Yang Q. Ingroup vigilance in collectivistic cultures. *Proc Natl Acad Sci USA* 2019 Jun 27;116(29):14538-14546. [doi: [10.1073/pnas.1817588116](https://doi.org/10.1073/pnas.1817588116)]
26. Uchida Y, Takemura K, Fukushima S, Saizen I, Kawamura Y, Hitokoto H, et al. Farming cultivates a community-level shared culture through collective activities: Examining contextual effects with multilevel analyses. *J Pers Soc Psychol* 2019 Jan;116(1):1-14. [doi: [10.1037/pspa0000138](https://doi.org/10.1037/pspa0000138)] [Medline: [30382740](https://pubmed.ncbi.nlm.nih.gov/30382740/)]
27. Germani A, Buratta L, Delvecchio E, Mazzeschi C. Emerging Adults and COVID-19: The Role of Individualism-Collectivism on Perceived Risks and Psychological Maladjustment. *Int J Environ Res Public Health* 2020 May 17;17(10):3497 [FREE Full text] [doi: [10.3390/ijerph17103497](https://doi.org/10.3390/ijerph17103497)] [Medline: [32429536](https://pubmed.ncbi.nlm.nih.gov/32429536/)]
28. Huang F, Ding H, Liu Z, Wu P, Zhu M, Li A, et al. How fear and collectivism influence public's preventive intention towards COVID-19 infection: a study based on big data from the social media. *BMC Public Health* 2020 Nov 16;20(1):1707 [FREE Full text] [doi: [10.1186/s12889-020-09674-6](https://doi.org/10.1186/s12889-020-09674-6)] [Medline: [33198699](https://pubmed.ncbi.nlm.nih.gov/33198699/)]
29. Muramatsu N, Akiyama H. Japan: super-aging society preparing for the future. *Gerontologist* 2011 Aug;51(4):425-432. [doi: [10.1093/geront/gnr067](https://doi.org/10.1093/geront/gnr067)] [Medline: [21804114](https://pubmed.ncbi.nlm.nih.gov/21804114/)]
30. Jordan RE, Adab P, Cheng KK. Covid-19: risk factors for severe disease and death. *BMJ* 2020 Mar 26;368:m1198. [doi: [10.1136/bmj.m1198](https://doi.org/10.1136/bmj.m1198)] [Medline: [32217618](https://pubmed.ncbi.nlm.nih.gov/32217618/)]
31. Aw D, Woodrow L, Ogliari G, Harwood R. Association of frailty with mortality in older inpatients with Covid-19: a cohort study. *Age Ageing* 2020 Oct 23;49(6):915-922 [FREE Full text] [doi: [10.1093/ageing/afaa184](https://doi.org/10.1093/ageing/afaa184)] [Medline: [32778870](https://pubmed.ncbi.nlm.nih.gov/32778870/)]
32. Murayama T, Shimizu N, Fujita S, Wakamiya S, Aramaki E. Robust two-stage influenza prediction model considering regular and irregular trends. *PLoS One* 2020;15(5):e0233126 [FREE Full text] [doi: [10.1371/journal.pone.0233126](https://doi.org/10.1371/journal.pone.0233126)] [Medline: [32437380](https://pubmed.ncbi.nlm.nih.gov/32437380/)]

## Abbreviations

**LCI:** localized concern index

**LCIPP:** localized concern index by patient percentage



*Edited by G Eysenbach; submitted 23.04.21; peer-reviewed by S Matsuda, A Mavragani; comments to author 11.05.21; revised version received 01.06.21; accepted 13.06.21; published 20.07.21.*

*Please cite as:*

*Gao Z, Fujita S, Shimizu N, Liew K, Murayama T, Yada S, Wakamiya S, Aramaki E*

*Measuring Public Concern About COVID-19 in Japanese Internet Users Through Search Queries: Infodemiological Study*

*JMIR Public Health Surveill 2021;7(7):e29865*

*URL: <https://publichealth.jmir.org/2021/7/e29865>*

*doi: [10.2196/29865](https://doi.org/10.2196/29865)*

*PMID: [34174781](https://pubmed.ncbi.nlm.nih.gov/34174781/)*

©Zhiwei Gao, Sumio Fujita, Nobuyuki Shimizu, Kongmeng Liew, Taichi Murayama, Shuntaro Yada, Shoko Wakamiya, Eiji Aramaki. Originally published in JMIR Public Health and Surveillance (<https://publichealth.jmir.org>), 20.07.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

# The Influence of COVID-19 Information Sources on the Attitudes and Practices Toward COVID-19 Among the General Public of Saudi Arabia: Cross-sectional Online Survey Study

Noor Alshareef<sup>1,2\*</sup>, PhD; Ismaeel Yunusa<sup>3\*</sup>, PhD; Mohammed Khaled Al-Hanawi<sup>1,2\*</sup>, PhD

<sup>1</sup>Department of Health Services and Hospital Administration, Faculty of Economics and Administration, King Abdulaziz University, Jeddah, Saudi Arabia

<sup>2</sup>Health Economics Research Group, King Abdulaziz University, Jeddah, Saudi Arabia

<sup>3</sup>College of Pharmacy, University of South Carolina, Columbia, SC, United States

\* all authors contributed equally

## Corresponding Author:

Mohammed Khaled Al-Hanawi, PhD

Department of Health Services and Hospital Administration

Faculty of Economics and Administration

King Abdulaziz University

University Building 125 Faculties Street

Jeddah, 80200

Saudi Arabia

Phone: 966 556522222

Email: [mkalhanawi@kau.edu.sa](mailto:mkalhanawi@kau.edu.sa)

## Abstract

**Background:** The COVID-19 pandemic has resulted in panic among the general public, leading many people to seek out information related to COVID-19 through various sources, including social media and traditional media. Identifying public preferences for obtaining such information may help health authorities to effectively plan successful health preventive and educational intervention strategies.

**Objective:** The aim of this study was to examine the impact of the types of sources used for obtaining COVID-19 information on the attitudes and practices of the general public in Saudi Arabia during the pandemic, and to identify the socioeconomic and demographic factors associated with the use of different sources of information.

**Methods:** This study used data from a cross-sectional online survey conducted on residents of Saudi Arabia from March 20 to 24, 2020. Data were analyzed using descriptive, bivariate, and multivariable logistic regression analyses. Bivariate analysis of categorical variables was performed to determine the associations between information sources and socioeconomic and demographic factors. Multivariable logistic regression analyses were employed to examine whether socioeconomic and demographic variables were associated with the source of information used to obtain information about COVID-19. Moreover, univariable and multivariable logistic regression analyses were conducted to examine how sources of information influence attitudes and practices of adhering to preventive measures.

**Results:** In this analysis of cross-sectional survey data, 3358 participants were included. Most participants reported using social media, followed by the Ministry of Health (MOH) of the Kingdom of Saudi Arabia, as their primary source of information. Seeking information via social media was significantly associated with lower odds of having an optimistic attitude (adjusted odds ratio [aOR] 0.845, 95% CI 0.733-0.974;  $P=.02$ ) and adhering to preventive measures (aOR 0.725, 95% CI 0.630-0.835;  $P<.001$ ) compared to other sources of information. Participants who obtained their COVID-19 information via the MOH had greater odds of having an optimistic attitude (aOR 1.437, 95% CI 1.234-1.673;  $P<.001$ ) and adhering to preventive measures (aOR 1.393, 95% CI 1.201-1.615;  $P<.001$ ) than those who obtained information via other sources.

**Conclusions:** This study provides evidence that different sources of information influence attitudes and preventive actions differently within a pandemic crisis context. Health authorities in Saudi Arabia should pay attention to the use of appropriate social media channels and sources to allow for more effective dissemination of critical information to the public.

(*JMIR Public Health Surveill* 2021;7(7):e28888) doi:[10.2196/28888](https://doi.org/10.2196/28888)

**KEYWORDS**

attitudes; communications media; COVID-19; information-seeking behavior; pandemics; practices; Saudi Arabia; social media; sources

## *Introduction*

COVID-19, which was caused by the novel SARS-CoV-2 virus, was first reported in Wuhan, China, and has since spread extensively across the globe [1]. By the end of January 2020, the World Health Organization (WHO) announced a public health emergency of international concern and called for the collaborative effort of all countries to prevent the rapid spread of the virus. Subsequently, on March 11, 2020, the WHO declared COVID-19 a global pandemic [2]. During emerging infectious diseases, the adoption of preventive measures is one of the most crucial interventions to halt the spread of infection [3,4]. Prior research has demonstrated the role of media in providing health information and promoting preventive behaviors [5,6]. Engagement in preventive behaviors is most likely to occur when there is clear, reliable, and frequent communication. By contrast, weak or inconsistent communication has the potential to reduce public trust and, consequently, increase the likelihood of adverse social and economic impacts. It is evident how rapid communication and publication contributed to the recognition of the severity and magnitude of COVID-19 by the general public [7].

In health crisis situations, information demand is usually high; there are many unknowns and people tend to resort to sources they trust [8]. Several information sources are currently available for obtaining health-related information. As a consequence, the importance of these sources during a global health crisis intensifies. For instance, traditional media, such as television and newspapers, play a role in communicating evidence-based information to the public [9]. People may also rely on their family members, friends, and coworkers for information on COVID-19 [10]. Social media also represents a vital platform for people to express their opinions, perceptions, and attitudes toward COVID-19 public health policies [11].

Given the wide range of available communication channels, evidence shows that people rely on a variety of different sources of information [12] and that utilization of these sources tends to change over time [13]. For instance, in 2017, Spence et al [14] showed that traditional news media were the preferred channels of information during a health crisis, whereas a more recent study showed that social media platforms or online news sites were the most predominant sources for obtaining information during a crisis [15]. The WHO declared that the COVID-19 pandemic has been accompanied by a so-called “infodemic” of misinformation that renders finding clear information, reliable guidance, and trustworthy sources difficult [16]. Misinformation about the pandemic can pose a significant risk to public health and public actions, thereby undermining public health efforts in mitigating the profound impact of the virus. In this sense, a major challenge is to ensure that people have access to accurate information that allows them to act appropriately [9].

In response to the high demand for timely and trustworthy information in chaotic situations such as the COVID-19 pandemic, the Ministry of Health (MOH) of the Kingdom of Saudi Arabia (KSA) acted as a key and official source responsible for communicating information, and for sustaining a reliable and consistent flow of information pertinent to COVID-19 to the public. To communicate and engage the community, traditional channels, such as television, radio, and text messages, as well as technology and digital health programs were all utilized under the MOH leadership. High-quality media materials developed by the MOH, Ministry of Media, and the Center for Government Communication were produced. Further, joint press conferences providing updates have been held on a daily basis during this pandemic. Video messages of ministers and other prominent public figures recommending that the public follow precautionary measures were widely shared. Additionally, unconventional awareness-raising champions to advise people to stay indoors were created [17].

In addition, the MOH relied on the use of different platforms to disseminate information about the disease, including the route of transmission and effective prevention practice measures [18]. To ensure widespread dissemination, this information was translated into other languages, such as Filipino, Urdu, Portuguese, Russian, and French [19]. In addition, the MOH, together with other ministries, has utilized SMS text messaging in spreading awareness and emphasizing the importance of adopting precautionary measures [18]. The WHO Information Network for Epidemics was also established to provide simplified, timely, and accurate information from trusted sources [8]. Taken together, these local and international communication efforts were intended to raise awareness and encourage preventive actions.

Given that only half of the Saudi population have adequate health literacy [20], which may pose some challenges in acquiring and obtaining health information that is particularly important in decision making during outbreaks, it is essential to understand the sources from which people obtain public health crisis information. Thereby, this study aims to (1) understand sources of COVID-19–related information among the Saudi public, (2) examine factors associated with selection of information sources, and (3) investigate the impact of the utilized information sources on public attitudes and prevention practices toward COVID-19. Gaining a deeper understanding of preferred information sources can inform public health officials on to where to extend efforts to reach a broader audience. Moreover, it can also assist in enriching and supporting the ongoing response to the pandemic and in preparing for future pandemics.

## *Methods*

### **Study Design and Sample**

This study used data from a cross-sectional survey conducted on residents of the KSA from March 20 to 24, 2020. Given the

social and physical distancing measures implemented in the country during that time because of the COVID-19 pandemic, data were collected via an online self-reported questionnaire using SurveyMonkey. Invitations to participate in the study were distributed via social media platforms (ie, Twitter and WhatsApp). Participants were recruited using a simplified snowball-sampling technique, where the invited participants were requested to pass on the invitations to their WhatsApp contacts. A link to the questionnaire was also posted on the King Abdulaziz University website. This online approach was being used to avoid any physical contact.

The original cross-sectional survey included individuals aged 18 years or older living in the KSA at the time of data collection. Online informed consent was obtained from all participants before proceeding with the survey. A total of 3427 participants completed the questionnaire. After excluding respondents who reported living outside the KSA and all responses with missing data on outcome variables, the final sample for analysis comprised responses from 3358 participants. The survey instrument used, data collection procedures, and sample size determination method are described in detail elsewhere [21].

## Measures

The online self-reported questionnaire was developed according to community guidelines for preventing the spread of COVID-19 from the US Centers for Disease Control and Prevention (CDC) [22]. The questionnaire was initially drafted in English, translated to Arabic to ensure the meaning of the content, and then back-translated to English. The Arabic version of the survey was administered for this study. The questionnaire consisted of four primary sections. The first section gathered information on respondents' socioeconomic and demographic characteristics. The second section assessed respondents' knowledge of COVID-19. The third section assessed respondents' attitudes toward COVID-19 using a 5-point Likert scale. The final section of the questionnaire assessed the respondents' infection prevention practices. A full description of the survey instrument can be found elsewhere [21].

The primary outcome of this study was sources of COVID-19 information. This was based on a survey question that asked for the main source of the respondents' knowledge related to COVID-19. Responses to this question were coded into six categories: social media, MOH, television and newspapers, friends and/or family, self-learning, and official sources, including the WHO, medical journals and articles, or the CDC. To examine the relationship between potential predictors and COVID-19 information sources, we dichotomized each of the six response variables into social media versus other information sources (others), MOH versus others, television and newspapers versus others, friends and/or family versus others, self-learning versus others, and official sources versus others.

We also examined covariates that might be associated with the type of information sources related to sociodemographic characteristics, including age, gender, marital status, education level, employment status, monthly income, and nationality. The age variable was divided into five categories: 18-29 (reference category), 30-39, 40-49, 50-59, and  $\geq 60$  years. Gender was coded as a binary variable, with 1 for male and 0 for female.

Marital status was also captured as a binary variable, with 1 for unmarried, including single, widowed, and divorced, and 0 for married. Education level was divided into three categories: high school or below (reference category), college or university degree, and postgraduate degree. Employment status was categorized into five groups: government employee (reference group), nongovernment employee, self-employed, retired, and unemployed. Monthly income (in Saudi Riyal [SR]) was grouped into eight categories (a currency exchange rate of SR 1=US \$0.27 is applicable):  $<3000$  (reference category); 3000 to less than 5000; 5000 to less than 7000; 7000 to less than 10,000; 10,000 to less than 15,000; 15,000 to less than 20,000; 20,000 to less than 30,000; and  $\geq 30,000$ . Nationality was coded as a binary variable, with 1 for Saudi and 0 for non-Saudi.

Information was also collected on attitudes toward COVID-19 using a 5-point Likert scale. Respondents were asked to state their level of agreement for six statements, including "It is important to put distance between myself and other people to avoid transmission of COVID-19," "Handwashing is important to protect myself from COVID-19," "I stay home if I am sick, except to get medical care to protect myself from getting COVID-19," "COVID-19 will eventually be successfully controlled; Saudi Arabia's strict measures can help win the battle against COVID-19," and "Following all precautions from the Ministry of Health will prevent the spread of COVID-19." Scores for attitudes were calculated based on the respondents' answers to each attitudinal statement as follows: 1=strongly disagree, 2=disagree, 3=undecided, 4=agree, and 5=strongly agree. The total attitude score was computed by adding up the points for each of the respondent's answers to the six statements, yielding a total attitude score between 6 and 30, with high scores indicating more optimistic attitudes or beliefs that individuals hold internally. Using the median total attitude score of 29, we dichotomized the attitude into optimistic attitudes for scores  $\geq 29$  and nonoptimistic attitudes for scores  $< 29$ .

Information was also collected on practices related to COVID-19. There were five questions related to practices and behavior, including (1) going to social events with large numbers of people, (2) going to crowded places, (3) avoiding cultural behaviors such as shaking hands, (4) practicing social distancing, and (5) washing hands after sneezing, coughing, nose-blowing, and having recently been in a public place. Respondents were asked to respond "yes" or "no" to each item. A score of 1 was given for answers that reflected a positive practice and a score of 0 was given for answers that reflected a negative practice. The total practice score ranged from 0 to 5, with high scores indicating better practices. We also used the median total practice score to dichotomize practices into positive practices that indicated adherence to COVID-19 precautionary measures (scores  $> 4$ ) and negative practices that indicated nonadherence to COVID-19 precautionary measures (scores  $\leq 4$ ).

## Data Analyses

Descriptive, bivariate, and multivariable logistic regression analyses were employed in this study. The proportion of participants who used each source to obtain information about COVID-19 is presented in terms of frequency and percentage. Mean and SD values are used to describe the continuous

variables, whereas frequencies and percentages are used to describe the categorical variables. Bivariate analysis of categorical variables was performed using chi-square tests to determine the associations between sources of information and independent variables. Additionally, an independent *t* test was performed to assess differences in mean values for attitudes and practice scores.

Furthermore, logistic regression analysis for each COVID-19 information source was also conducted. Multivariable logistic regression analyses were employed to examine whether sociodemographic variables were associated with the source of COVID-19 information. Moreover, univariable and multivariable logistic regressions were conducted to examine the relationship between information sources and attitudes and practices toward COVID-19 prevention, controlling for sociodemographic characteristics. In the regression models, we estimated crude odds ratios (ORs) for univariable analyses and adjusted ORs (aORs) for multivariable analyses with their respective 95% CIs. Statistical significance was determined if a *P* value was less than .05. All analyses were performed using Stata 15.1 software (StataCorp LLC).

### Ethical Clearance

All procedures performed in this study involving human participants complied with institutional or national research committee ethical standards, as well as the 1964 Helsinki Declaration and subsequent amendments or equivalent ethical

standards. This study has been reviewed and was given a favorable opinion by the King Abdulaziz University Research Ethics Committee and was designed and performed in accordance with the ethical principles established by the university. Ethical approval was obtained from the Biomedical Ethics Research Committee, Faculty of Medicine, King Abdulaziz University (Ref-180-20).

### Data Availability Statement

The data sets generated and/or analyzed during this study are not publicly available due to privacy and confidentiality agreements as well as other restrictions but are available from the corresponding author (MKA) on reasonable request.

## Results

### Characteristics of Study Participants

**Table 1** presents the descriptive statistics of respondents' socioeconomic and demographic characteristics and the sources they used for obtaining COVID-19 information. Most of the participants were married, had a college or university degree, and were Saudi citizens. Approximately two-thirds of the respondents had optimistic attitudes and more than half of them adhered to preventive measures. The majority of respondents reported social media as their main source of COVID-19 information, followed by the MOH, and then television and newspapers.

**Table 1.** Descriptive statistics of respondents' socioeconomic and demographic characteristics and the sources they used for obtaining COVID-19 information.

Variable	Value (N=3358)
Knowledge score <sup>a</sup> , mean (SD)	17.98 (2.22)
Attitude score <sup>b</sup> , mean (SD)	28.24 (2.74)
Nonoptimistic respondents, n (%)	1280 (38.12)
Optimistic respondents, n (%)	2078 (61.88)
Practice score <sup>c</sup> , mean (SD)	4.34 (0.87)
Respondents with negative practices, indicating nonadherence to COVID-19 precautionary measures, n (%)	1554 (46.28)
Respondents with positive practices, indicating adherence to COVID-19 precautionary measures, n (%)	1804 (53.72)
<b>Age (years), n (%)</b>	
18 to 29	1005 (29.93)
30 to 39	934 (27.81)
40 to 49	686 (20.43)
50 to 59	468 (13.94)
≥60	265 (7.89)
<b>Gender, n (%)</b>	
Female	1945 (57.92)
Male	1413 (42.08)
<b>Marital status, n (%)</b>	
Married	2129 (63.40)
Unmarried	1229 (36.60)
<b>Nationality, n (%)</b>	
Non-Saudi	264 (7.86)
Saudi	3094 (92.14)
<b>Monthly income (Saudi Riyal<sup>d</sup>), n (%)</b>	
<3000	835 (24.87)
3000 to <5000	290 (8.64)
5000 to <7000	254 (7.56)
7000 to <10,000	352 (10.48)
10,000 to <15,000	580 (17.27)
15,000 to <20,000	471 (14.03)
20,000 to <30,000	332 (9.89)
≥30,000	244 (7.27)
<b>Education level, n (%)</b>	
High school or below	530 (15.78)
College or university degree	1886 (56.16)
Postgraduate degree	942 (28.05)
<b>Employment status, n (%)</b>	
Government employee	1308 (38.95)
Private sector employee	544 (16.20)
Retired	312 (9.29)
Self-employed	135 (4.02)

Variable	Value (N=3358)
Unemployed	1059 (31.54)
<b>Sources of COVID-19 information, n (%)</b>	
Social media	1567 (46.66)
Ministry of Health	1138 (33.89)
Television and newspapers	432 (12.86)
Friends and/or family	55 (1.64)
Self-learning	98 (2.92)
Official sources	68 (2.03)

<sup>a</sup>Knowledge scores: Each statement had the responses "true," "false," and "don't know." Incorrect or uncertain (don't know) responses were given a score of 0, and correct answers were assigned a score of 1. The total score for knowledge ranged from 0 to 22, with high scores indicating better knowledge of COVID-19.

<sup>b</sup>Attitude scores: Each item ranged from 1 (strongly disagree) to 5 (strongly agree); the total attitude score ranged from 6 and 30, with high scores indicating more optimistic attitudes or beliefs.

<sup>c</sup>Practice scores were 1 (reflected a positive practice) or 0 (reflected a negative practice) for each item; the total practice score ranged from 0 to 5, with high scores indicating better practices.

<sup>d</sup>A currency exchange rate of 1 Saudi Riyal=US \$0.27 is applicable.

### Relationship Between Sources of COVID-19 Information, Socioeconomic and Demographic Factors, Attitudes, and Practices Toward COVID-19

Table 2 shows the results of bivariate analysis for assessing the association between different variables and sources of COVID-19-related information. The age category was significantly associated with all information sources except for official sources. Most of the respondents who indicated social media as their main information source were aged 18 to 29 years (539/1567, 34.40%), whereas respondents who indicated television and newspapers as their main source were aged 50 to 59 years (124/432, 28.7%). As age increased, the proportion of respondents who used social media as their main source of information decreased. Moreover, there was a significant association between employment status and all sources of information except for official sources.

Level of education was significantly associated with seeking information from family and/or friends (Table 2). More than half of those who indicated obtaining information from family and/or friends had a college education or higher (32/55, 58%). Nationality was associated with seeking information from the MOH, with the great majority being Saudi citizens (1068/1138,

93.85%). Income level was also significantly associated with seeking information from social media, television and newspapers, family and/or friends, self-learning, and official sources. Most of the participants who indicated obtaining information from the MOH reported earning less than SR 3000 (278/1138, 24.43%). Moreover, more than half of the participants who sourced information from family and/or friends earned less than SR 3000 (30/55, 55%). Among those who used social media, the percentage of respondents with optimistic attitudes was higher than that of respondents with nonoptimistic attitudes (935/1567, 59.67% vs 632/1567, 40.33%). Moreover, among respondents who sourced information from friends and/or family, a higher number had nonoptimistic attitudes than optimistic attitudes (29/55, 53% vs 26/55, 47%). On the contrary, among those who used the MOH as their source of information, more respondents had optimistic attitudes than nonoptimistic attitudes (770/1138, 67.66% vs 368/1138, 32.34%). However, among the respondents who used social media, those who did not adhere to COVID-19 preventive measures were higher in number than their counterparts (789/1567, 50.35% vs 778/1567, 49.65%). On the other hand, among respondents who used the MOH as a main source of information, the majority adhered to preventive measures than did not adhere (669/1138, 58.79% vs 469/1138, 41.21%).

**Table 2.** Bivariate analysis of socioeconomic and demographic factors, attitudes, and practices toward COVID-19 with sources of information.

Variable	Social media (no: n=1791; yes: n=1567)		Ministry of Health (no: n=2220; yes: n=1138)		Television and newspapers (no: n=2926; yes: n=432)		Friends and/or family (no: n=3303; yes: n=55)		Self-learning (no: n=3260; yes: n=98)		Official sources (no: n=3290; yes: n=68)	
	n (%)		n (%)		n (%)		n (%)		n (%)		n (%)	
	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
<b>Age (years)</b>	<.001		<.001		<.001		<.001		<.001		.46	
18 to 29	466 (26)	539 (34)	659 (30)	346 (30)	959 (33)	46 (11)	970 (29)	35 (64)	987 (30)	18 (18)	984 (30)	21 (31)
30 to 39	487 (27)	447 (29)	570 (26)	364 (32)	864 (30)	70 (16)	927 (28)	7 (13)	910 (28)	24 (24)	912 (28)	22 (32)
40 to 49	363 (20)	323 (21)	470 (21)	216 (19)	576 (20)	110 (25)	680 (21)	6 (11)	665 (20)	21 (21)	676 (21)	10 (15)
50 to 59	301 (17)	167 (11)	319 (14)	149 (13)	344 (12)	124 (29)	464 (14)	4 (7)	451 (14)	17 (17)	461 (14)	7 (10)
≥60	174 (10)	91 (6)	202 (9)	63 (6)	183 (6)	82 (19)	262 (8)	3 (5)	247 (8)	18 (18)	257 (8)	8 (12)
<b>Gender</b>	.047		.65		.02		.81		.38		.40	
Female	109 (56)	936 (60)	122 (58)	653 (57)	1717 (59)	228 (53)	1914 (58)	31 (56)	1884 (58)	61 (62)	1909 (58)	36 (53)
Male	782 (44)	631 (40)	928 (42)	485 (43)	1209 (41)	204 (47)	1389 (42)	24 (44)	1376 (42)	37 (38)	1381 (42)	32 (47)
<b>Marital status</b>	.08		.09		<.001		<.001		.69		.07	
Married	1160 (65)	969 (62)	1430 (64)	699 (61)	1791 (61)	338 (78)	2106 (64)	23 (42)	2165 (63)	64 (65)	2083 (64)	36 (53)
Unmarried	631 (35)	598 (38)	790 (36)	439 (39)	1135 (39)	94 (22)	1197 (36)	32 (58)	1195 (37)	34 (35)	1197 (36)	32 (47)
<b>Nationality</b>	.16		.01		.25		.50		.21		.77	
Non-Saudi	130 (7)	134 (9)	194 (9)	70 (6)	224 (8)	40 (9)	261 (8)	3 (5)	253 (8)	11 (11)	258 (8)	6 (9)
Saudi	1661 (93)	1433 (91)	2106 (91)	1088 (94)	2702 (92)	392 (91)	3042 (92)	52 (95)	3107 (92)	87 (89)	3182 (92)	62 (91)
<b>Monthly income (Saudi Riyal<sup>a</sup>)</b>	.002		.12		<.001		<.001		<.001		<.001	
<3000	407 (23)	428 (27)	557 (25)	278 (24)	772 (26)	63 (15)	805 (24)	30 (55)	813 (25)	22 (22)	821 (25)	14 (21)
3000 to <5000	141 (8)	149 (10)	198 (9)	92 (8)	255 (9)	35 (8)	286 (9)	4 (7)	284 (9)	6 (6)	286 (9)	4 (6)
5000 to <7000	127 (7)	127 (8)	162 (7)	92 (8)	225 (8)	29 (7)	250 (8)	4 (7)	254 (8)	0 (0)	252 (8)	2 (3)
7000 to <10,000	202 (11)	150 (10)	229 (10)	123 (11)	288 (10)	64 (15)	349 (11)	3 (5)	345 (11)	7 (7)	347 (11)	5 (7)
10,000 to <15,000	309 (17)	271 (17)	365 (16)	215 (19)	516 (18)	64 (15)	576 (17)	4 (7)	565 (17)	15 (15)	569 (17)	11 (16)
15,000 to <20,000	262 (15)	209 (13)	307 (14)	164 (14)	402 (14)	69 (16)	467 (14)	4 (7)	451 (14)	20 (20)	466 (14)	5 (7)
20,000 to <30,000	199 (11)	133 (8)	221 (10)	111 (10)	269 (9)	63 (15)	329 (10)	3 (5)	321 (10)	11 (11)	321 (10)	11 (16)
≥30,000	144 (8)	100 (6)	181 (8)	63 (6)	199 (7)	45 (10)	241 (7)	3 (5)	227 (7)	17 (17)	228 (7)	16 (24)
<b>Education level</b>	.22		.45		.70		<.001		.03		<.001	



Variable	Social media (no: n=1791; yes: n=1567)		Ministry of Health (no: n=2220; yes: n=1138)		Television and newspapers (no: n=2926; yes: n=432)		Friends and/or family (no: n=3303; yes: n=55)		Self-learning (no: n=3260; yes: n=98)		Official sources (no: n=3290; yes: n=68)	
	n (%)		n (%)		n (%)		n (%)		n (%)		n (%)	
	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
High school or below	278 (16)	252 (16)	363 (16)	167 (15)	456 (16)	74 (17)	513 (16)	17 (31)	515 (16)	15 (15)	525 (16)	5 (7)
College or university degree	988 (55)	898 (57)	1240 (56)	646 (57)	1649 (56)	237 (55)	1854 (56)	32 (58)	1842 (57)	44 (45)	1857 (56)	29 (43)
Postgraduate degree	525 (29)	417 (27)	617 (28)	325 (29)	821 (28)	121 (28)	936 (28)	6 (11)	903 (28)	39 (40)	908 (28)	34 (50)
<b>Employment status</b>	<.001		<.001		<.001		<.001		.02		.35	
Government employee	727 (41)	581 (37)	819 (37)	489 (43)	1154 (39)	154 (36)	1302 (39)	6 (11)	1262 (39)	46 (47)	1276 (39)	32 (47)
Private sector	271 (15)	273 (17)	360 (16)	184 (16)	485 (17)	59 (14)	532 (16)	12 (22)	536 (16)	8 (8)	536 (16)	8 (12)
Retired	194 (11)	118 (8)	236 (11)	76 (7)	214 (7)	98 (23)	310 (9)	2 (4)	297 (9)	15 (15)	309 (9)	3 (4)
Self-employed	79 (4)	56 (4)	92 (4)	43 (4)	104 (4)	31 (7)	135 (4)	0 (0)	134 (4)	1 (1)	131 (4)	4 (6)
Unemployed	520 (29)	539 (34)	713 (32)	346 (30)	969 (33)	90 (21)	1024 (31)	35 (64)	1081 (32)	28 (29)	1088 (32)	21 (31)
<b>Attitudes</b>	.01		<.001		.07		.02		.78		.07	
Nonoptimistic	648 (36)	632 (40)	912 (41)	368 (32)	1098 (37)	182 (42)	1251 (38)	29 (53)	1244 (38)	36 (37)	1247 (38)	33 (49)
Optimistic	1143 (64)	935 (60)	1308 (59)	770 (68)	1828 (63)	250 (58)	2162 (62)	26 (47)	2016 (62)	62 (63)	2043 (62)	35 (51)
<b>Practices</b>	<.001		<.001		.61		.90		.94		.91	
Negative	765 (43)	789 (51)	1085 (49)	469 (41)	1359 (46)	195 (45)	1529 (46)	25 (45)	1509 (46)	45 (46)	1523 (46)	31 (46)
Positive	1026 (57)	778 (49)	1135 (51)	669 (59)	1567 (54)	237 (55)	1774 (54)	30 (55)	1751 (54)	53 (54)	1767 (54)	37 (54)

<sup>a</sup>A currency exchange rate of 1 Saudi Riyal=US \$0.27 is applicable.

### Association of Sources of COVID-19 Information With Participants' Socioeconomic and Demographic Characteristics

Table 3 shows the results of the regression analysis of the association between COVID-19 sources of information and socioeconomic and demographic characteristics. Compared with those 18 to 29 years old, participants aged 30 to 39 years (aOR 0.723, 95% CI 0.574-0.910;  $P=.006$ ), 40 to 49 years (aOR 0.696, 95% CI 0.538-0.899;  $P=.006$ ), 50 to 59 years (aOR 0.410, 95% CI 0.303-0.553;  $P<.001$ ), and  $\geq 60$  years (aOR 0.364, 95% CI 0.243-0.545;  $P<.001$ ) had significantly lower odds of using social media to obtain information related to COVID-19. These odds decreased as age increased, suggesting that young adults

are more likely to use social media to obtain information about COVID-19. By contrast, the odds of using television and newspapers significantly increased among respondents aged 30 to 39 years (aOR 1.857, 95% CI 1.186-2.909;  $P=.007$ ), 40 to 49 years (aOR 4.421, 95% CI 2.823-6.924;  $P<.001$ ), 50 to 59 years (aOR 7.787, 95% CI 4.875-12.430;  $P<.001$ ), and  $\geq 60$  years (aOR 8.519, 95% CI 4.940-14.690;  $P<.001$ ) compared to those aged 18 to 29 years, further supporting that older individuals are more likely to use television and newspapers as information sources. Moreover, respondents aged 60 years or above had significantly greater odds of learning by themselves compared to those aged 18 to 20 years (aOR 4.270, 95% CI 1.536-11.860;  $P=.005$ ).

**Table 3.** Associations between sources of COVID-19 information and socioeconomic and demographic characteristics.

Variable	Social media		Ministry of Health		Television and newspapers		Friends and/or family		Self-learning		Official sources	
	aOR <sup>a</sup> (95% CI)	<i>P</i> value	aOR (95% CI)	<i>P</i> value	aOR (95% CI)	<i>P</i> value	aOR (95% CI)	<i>P</i> value	aOR (95% CI)	<i>P</i> value	aOR (95% CI)	<i>P</i> value
<b>Age (years)</b>												
18 to 29	Ref <sup>b</sup>											
30 to 39	0.723 (0.574- 0.910)	.006	1.193 (0.938- 1.516)	.15	1.857 (1.186- 2.909)	.007	0.377 (0.136- 1.043)	.06	1.792 (0.820- 3.916)	.14	0.924 (0.404- 2.111)	.85
40 to 49	0.696 (0.538- 0.899)	.006	0.877 (0.668- 1.151)	.34	4.421 (2.823- 6.924)	<.001	0.456 (0.147- 1.415)	.18	1.907 (0.827- 4.394)	.13	0.507 (0.186- 1.376)	.18
50 to 59	0.410 (0.303- 0.553)	<.001	1.010 (0.741- 1.378)	.95	7.787 (4.875- 12.43)	<.001	0.473 (0.127- 1.760)	.27	2.125 (0.854- 5.287)	.11	0.522 (0.168- 1.623)	.26
≥60	0.364 (0.243- 0.545)	<.001	0.796 (0.516- 1.227)	.30	8.519 (4.940- 14.69)	<.001	0.816 (0.157- 4.221)	.81	4.270 (1.536- 11.86)	.005	1.174 (0.346- 3.976)	.80
<b>Gender</b>												
Female	Ref											
Male	0.939 (0.804- 1.097)	.43	1.096 (0.932- 1.289)	.27	0.961 (0.753- 1.228)	.76	1.770 (0.965- 3.246)	.07	0.619 (0.383- 1.001)	.05	1.254 (0.723- 2.176)	.42
<b>Marital status</b>												
Married	Ref											
Unmarried	0.809 (0.674- 0.971)	.02	1.166 (0.965- 1.409)	.11	0.873 (0.648- 1.175)	.37	0.810 (0.369- 1.780)	.60	1.475 (0.874- 2.487)	.15	2.352 (1.270- 4.352)	.006
<b>Nationality</b>												
Non-Saudi	Ref											
Saudi	0.881 (0.677- 1.146)	.35	1.485 (1.107- 1.990)	.008	0.736 (0.498- 1.084)	.12	2.088 (0.631- 6.907)	.28	0.473 (0.236- 0.945)	.03	0.565 (0.229- 1.388)	.21
<b>Monthly income (Saudi Riyal<sup>c</sup>)</b>												
<3000	Ref											
3000 to <5000	1.127 (0.827- 1.535)	.45	0.897 (0.645- 1.248)	.52	1.006 (0.609- 1.663)	.98	0.542 (0.166- 1.766)	.31	0.812 (0.294- 2.239)	.69	1.565 (0.471- 5.201)	.46
5000 to <7000	1.060 (0.750- 1.496)	.74	1.039 (0.723- 1.494)	.84	0.987 (0.565- 1.726)	.96	0.784 (0.221- 2.777)	.71	1 (N/A <sup>d</sup> )	N/A	1.172 (0.233- 5.895)	.85
7000 to <10,000	0.857 (0.615- 1.192)	.36	0.986 (0.696- 1.397)	.94	1.371 (0.829- 2.268)	.22	0.624 (0.155- 2.515)	.51	0.660 (0.216- 2.014)	.47	2.130 (0.611- 7.428)	.24
10,000 to <15,000	1.094 (0.793- 1.510)	.58	1.046 (0.745- 1.469)	.80	0.655 (0.389- 1.100)	.11	0.758 (0.195- 2.933)	.69	0.784 (0.282- 2.173)	.64	2.897 (0.984- 9.076)	.07

Variable	Social media		Ministry of Health		Television and newspapers		Friends and/or family		Self-learning		Official sources	
	aOR <sup>a</sup> (95% CI)	<i>P</i> value	aOR (95% CI)	<i>P</i> value	aOR (95% CI)	<i>P</i> value	aOR (95% CI)	<i>P</i> value	aOR (95% CI)	<i>P</i> value	aOR (95% CI)	<i>P</i> value
15,000 to <20,000	1.089 (0.776-1.526)	.62	1.013 (0.708-1.448)	.94	0.705 (0.416-1.193)	.19	0.951 (0.235-3.842)	.95	1.261 (0.464-3.420)	.65	1.717 (0.458-6.434)	.42
20,000 to <30,000	0.931 (0.643-1.349)	.71	0.938 (0.635-1.386)	.75	0.998 (0.575-1.731)	>.99	1.022 (0.218-4.783)	.98	0.981 (0.326-2.951)	.97	4.947 (1.490-16.42)	.009
≥30,000	1.046 (0.703-1.556)	.83	0.671 (0.435-1.034)	.05	0.788 (0.437-1.420)	.43	1.076 (0.226-5.124)	.93	2.772 (0.957-8.032)	.06	10.68 (3.299-34.55)	<.001
<b>Education level</b>												
High school or below	Ref											
College or university degree	0.979 (0.799-1.199)	.83	1.083 (0.873-1.345)	.47	1.019 (0.748-1.388)	.90	0.673 (0.357-1.267)	.22	0.756 (0.404-1.412)	.38	1.720 (0.646-4.581)	.28
Postgraduate degree	0.988 (0.769-1.268)	.92	1.082 (0.831-1.410)	.56	0.855 (0.586-1.246)	.42	0.464 (0.157-1.371)	.17	0.832 (0.397-1.741)	.63	3.324 (1.141-9.684)	.03
<b>Employment status</b>												
Government employee	Ref											
Private sector	1.124 (0.890-1.418)	.33	0.901 (0.706-1.150)	.40	1.066 (0.738-1.538)	.73	4.364 (1.464-13.01)	.008	0.456 (0.195-1.064)	.07	0.606 (0.253-1.446)	.26
Retired	1.256 (0.899-1.755)	.18	0.639 (0.445-0.919)	.02	1.515 (1.038-2.209)	.03	0.813 (0.119-5.518)	.83	0.810 (0.361-1.815)	.61	0.377 (0.094-1.507)	.17
Self-employed	0.969 (0.663-1.418)	.87	0.917 (0.616-1.365)	.67	1.753 (1.090-2.822)	.02	1 (N/A)	N/A	0.120 (0.015-0.924)	.04	0.975 (0.310-3.067)	.97
Unemployed	1.071 (0.809-1.415)	.63	0.824 (0.613-1.108)	.20	1.078 (0.702-1.655)	.73	4.528 (1.264-16.21)	.20	0.755 (0.310-1.842)	.54	1.843 (0.689-4.934)	.23

<sup>a</sup>aOR: adjusted odds ratio.

<sup>b</sup>Ref: reference.

<sup>c</sup>A currency exchange rate of 1 Saudi Riyal=US \$0.27 is applicable.

<sup>d</sup>N/A: not applicable because aOR=1.

In terms of gender, men had significantly lower odds of learning by themselves (aOR 0.619, 95% CI 0.383-1.001;  $P=.04$ ) compared to women. Unmarried individuals had significantly lower odds of using social media than married individuals (aOR 0.809, 95% CI 0.674-0.971;  $P=.03$ ), whereas unmarried individuals had significantly greater odds of using official sources (aOR 2.352, 95% CI 1.270-4.352;  $P=.006$ ). Interestingly, Saudi nationals had significantly greater odds of using the MOH as a main source of information compared to non-Saudi nationals (aOR 1.485, 95% CI 1.107-1.990;  $P=.008$ ), whereas Saudi nationals had significantly lower odds of learning by themselves compared to non-Saudi nationals (aOR 0.473,

95% CI 0.236-0.945;  $P=.03$ ). Those who earned SR 30,000 or more, compared with those who earned less than SR 3000, had significantly lower odds of sourcing information from the MOH (aOR 0.671, 95% CI 0.435-1.034;  $P=.05$ ). By contrast, those who earned SR 30,000 or more (aOR 10.680, 95% CI 3.299-34.550;  $P<.001$ ) and those who earned SR 20,000 to less than SR 30,000 (aOR 4.947, 95% CI 1.490-16.420;  $P=.009$ ) had significantly greater odds of obtaining COVID-19-related information via official sources compared to those who earned less than SR 3000. Respondents with a postgraduate education had significantly greater odds of obtaining information related to COVID-19 via official sources compared to those with a high

school education or below (aOR 3.324, 95% CI 1.141-9.684;  $P=.03$ ).

Retired respondents had significantly lower odds of using the MOH as a main source of information than those who were government employees (aOR 0.639, 95% CI 0.445-0.919;  $P=.02$ ). By contrast, retired respondents had significantly greater odds of using television and newspapers as an information source than government employees (aOR 1.515, 95% CI 1.038-2.209;  $P=.03$ ). Private sector employees had significantly greater odds of seeking information from family and/or friends than government employees (aOR 4.364, 95% CI 1.464-13.010;  $P=.008$ ). Self-employed respondents had significantly greater odds of using television and newspapers as an information source than government employees (aOR 1.753, 95% CI 1.090-2.822;  $P=.02$ ). Unemployed people had greater odds of seeking information from family and/or friends compared to government employees (aOR 4.528, 95% CI 1.264-16.210;  $P=.02$ ).

### Association of Attitudes and Practices Toward COVID-19 With Sources of Information

Table 4 presents the relationship between the source of information used by the respondents and their attitudes and

practices toward COVID-19. We found that respondents who used social media had significantly lower odds of having optimistic attitudes than respondents who used other sources (aOR 0.845, 95% CI 0.733-0.974;  $P=.02$ ). Moreover, those who used social media as a main source of information had significantly lower odds of adhering to preventive measures (aOR 0.725, 95% CI 0.630-0.835;  $P<.001$ ) compared to those who obtained their information from other sources. Interestingly, participants who obtained information from the MOH had significantly greater odds of having optimistic attitudes than those who obtained information from other sources (aOR 1.437, 95% CI 1.234-1.673;  $P<.001$ ). Also, those who obtained information from the MOH had significantly greater odds of adhering to preventive measures than those who obtained information from other sources (aOR 1.393, 95% CI 1.201-1.615;  $P<.001$ ). This study also found that participants who obtained their information from television and newspapers had significantly lower odds of having optimistic attitudes compared to those who obtained information from other sources (aOR 0.801, 95% CI 0.645-0.994;  $P=.045$ ).

**Table 4.** Association between attitudes and practices toward COVID-19 and sources of information.

Source of information	Attitudes				Practices			
	Unadjusted OR <sup>a</sup> (95% CI)	<i>P</i> value	aOR <sup>b</sup> (95% CI)	<i>P</i> value	Unadjusted OR (95% CI)	<i>P</i> value	aOR (95% CI)	<i>P</i> value
Social media	0.838 (0.729-0.964)	.01	0.845 (0.733-0.974)	.02	0.735 (0.641-0.842)	<.001	0.725 (0.630-0.835)	<.001
Ministry of Health	1.458 (1.255-1.695)	<.001	1.437 (1.234-1.673)	<.001	1.363 (1.180-1.575)	<.001	1.393 (1.201-1.615)	<.001
Television and newspapers	0.825 (0.672-1.013)	.07	0.801 (0.645-0.994)	.045	1.054 (0.860-1.291)	.61	0.999 (0.805-1.240)	>.99
Friends and/or family	0.546 (0.320-0.932)	.03	0.624 (0.363-1.072)	.09	1.034 (0.605-1.766)	.90	1.279 (0.737-2.222)	.38
Self-learning	1.062 (0.700-1.612)	.78	1.070 (0.701-1.633)	.75	1.015 (0.678-1.519)	.94	1.011 (0.669-1.519)	.96
Official sources	0.647 (0.400-1.046)	.08	0.648 (0.397-1.055)	.08	1.028 (0.635-1.666)	.91	1.061 (0.646-1.741)	.82

<sup>a</sup>OR: odds ratio.

<sup>b</sup>aOR: adjusted odds ratio; adjusted for age, gender, marital status, nationality, monthly income, education level, and employment status.

## Discussion

### Principal Findings

In analyzing data collected during the early stage of the COVID-19 global health crisis in the KSA, the goal of this study was to advance knowledge regarding where the Saudi public obtain their information about COVID-19, and whether acquisition of these different sources of information influence the attitudes and adaptation of preventive actions. Our results indicate that during the pandemic, the respondents mainly relied on social media to obtain information related to COVID-19,

followed by the MOH and television and newspapers. Social media as a main source of information was shown to negatively influence attitudes and compliance to preventive measures; on the contrary, a positive influence on both attitudes and practices was found among respondents using the MOH as a main source of information. Television and newspapers as a source of information was also shown to have a negative influence on participants' attitudes.

## Meaning of Findings and Comparison With Existing Literature

Given the increasing significance of social media and its substantial relevance in a time of global health crises [23], this study shows that the sample population mainly relies on social media platforms for obtaining information about COVID-19. Evidence shows that in times of crisis, many people turn to social media to obtain crisis-related information [24]. The speed at which information disseminates through social media may drive the public to utilize the most accessible source [25]. In contrast to previous findings [26], this study found that social media was associated with lower odds of adoption of preventive behaviors and optimistic attitudes. This is of interest for several reasons. First, this finding might be a reflection of health information-related consumer perceptions of social media. Individuals tend to expose themselves only to communication channels they perceive as trustworthy [27], which, in turn, may impact their behaviors and decisions based on the received information [28]. Second, the tendency of social media to spread misinformation, rumors, and inaccurate data can exacerbate the consequences of behavioral responses. It has been shown that exposure to COVID-19 misinformation on social media adversely impacts the adoption of preventive behaviors [29]. Third, this finding could be, in part, explained by the high acceptability of risky behaviors among young adults [30,31]. Our finding that social media was the primary source of information among young adults supports this assumption.

Of particular interest, however, is the finding that the MOH was the second most preferred source of COVID-19-related information. In this study, respondents who obtained their information from the MOH had more positive attitudes toward COVID-19 and were more likely to adopt preventive behaviors compared to those who obtained information from other sources. An earlier study conducted in Saudi Arabia showed that the MOH was perceived as a reliable source among the Saudi public [32]. This could partially explain why respondents who obtained their information from the MOH had positive attitudes and adopted preventive behaviors. This could also explain why international official sources, such as the WHO, CDC, and medical journals, were the least dominant sources of information among respondents in this study. Taken together, these findings suggest that the wide-reaching efforts of the MOH had significant potential in communicating risk and crisis information and, in turn, promoting preventive actions. Although the MOH has established its presence, sought to control interactions, and responded to misinformation on social media, efforts aimed at leveraging social media to reach a broader and diverse audience, particularly the young, are required. In other words, to effectively raise awareness and encourage adaptation of preventive behaviors, the MOH should utilize channels that reach a larger audience and that are trusted by the public.

In this study, age, nationality, and employment status were found to be significantly associated with using the MOH as an information source. Saudi nationals tend to use the MOH as their primary source of COVID-19-related information. It is important to note that daily press conferences providing COVID-19 updates by the MOH spokesperson may have resulted in gaining the public's trust and offering reassurance

that the country is working hard to mitigate effects from the pandemic. At times of chaos and confusion, people want both localized information relevant to them and international information from trusted international health organization sources such as the WHO [33]. However, non-Saudi respondents were less likely to rely on the MOH in obtaining COVID-19-related information; while the reason behind this is unclear, one possible reason could be that they heard stories about the strain of the pandemic on family and friends in their country and they were more accepting and receptive to the information they shared with them.

The role of social media as an information source becomes salient when limited information is provided by traditional media [23]. In this study, traditional media such as newspapers and television was the third most common information source among the Saudi public. During the H1N1 flu crisis, mass media was found to play a positive role in reducing panic through increasing public knowledge and the likelihood of adopting preventive measures [5]. Lin and Lagoe [34] also demonstrated that the use of television and newspapers increased the intention to vaccinate against H1N1 among media users. However, our study showed that respondents who used television and newspapers as their primary source for obtaining information related to COVID-19 had lower odds of optimistic attitudes toward COVID-19. Given the importance of television as a source of information among the Saudi public, it could be used to support the public health agenda through the placement of doctors on television news. It has been reported that television news viewing is largely determined by perceived informative value; thus, television can be a particularly significant source of information during outbreak crises [35]. In line with other studies, age was positively correlated with receiving information from traditional media [10].

Interpersonal sources, such as friends and/or family and self-learning, were the least preferred information sources among the Saudi public. In comparative nationally representative samples from the United Kingdom, the United States, Italy, Australia, New Zealand, and South Korea, interpersonal sources were among the most trusted sources of COVID-19-related information [36]. Furthermore, in this study, some individuals did not depend on any particular source of information about COVID-19 but instead looked for the information themselves. Perhaps those concerned about their health and disease prevention prefer to obtain COVID-19-related information themselves through the internet. In contrast to the findings of previous studies demonstrating a positive association between active information acquisition and engaging in healthy behaviors [37], we found that self-learning does not influence either attitudes or preventive behaviors.

## Limitations

The findings of this study should be interpreted with some limitations. Data from this study were collected online via the snowball-sampling method. Thus, the participant responses may reflect a possible like-minded social network of those who received the survey link. In addition, given the survey's online nature, our findings may not be representative of the opinions of those living in rural areas where internet access is relatively

limited or among older adults who are less likely to use the internet. However, a community-based national sampling approach would not have been feasible given the imposed lockdown during the study period. Moreover, the study's cross-sectional nature with data collected in a snapshot cannot be used to characterize the change in information-seeking behavior over a longer period of time. Nevertheless, these findings provide useful guidance for health authorities and policy makers to identify the influence of sources of information on attitudes and practices toward COVID-19, and to deploy appropriate actions where necessary. Finally, it is worth noting that the survey included only Arabic speakers. Given that the KSA has a significant non-Arabic-speaking foreign population, it is important to acknowledge that the information sources and relationships reported in this study may not apply to a large group within Saudi society.

## Conclusions

This study provides evidence that different communication channels exert different influences on attitudes and preventive actions taken in a pandemic crisis context. Among the

communication channels examined in this study, the MOH was a widely consumed information source that has potential to shape public attitudes toward COVID-19 and enhance engagement in preventive actions. Both social and traditional media were information sources that were associated with lower odds of having optimistic attitudes. The findings of this study might help inform current local COVID-19 responses by offering insight into how the Saudi public uses sources of information, and further identifies which communication channels have the greatest impact on public attitudes and consequent prevention behaviors toward COVID-19. Moreover, by identifying the most predominantly used communication channels, public health authorities might be able to acknowledge differences in preferences of information sources and respond more effectively to future pandemics. Future research should examine changes in information-seeking behaviors over time and consider probability sampling of both social media and non-social media users who predominantly live in rural areas to provide further perspective on disparities in seeking information.

## Acknowledgments

We are grateful to all respondents who participated in this study.

## Conflicts of Interest

None declared.

## References

1. Wu F, Zhao S, Yu B, Chen Y, Wang W, Song Z, et al. A new coronavirus associated with human respiratory disease in China. *Nature* 2020 Mar;579(7798):265-269 [FREE Full text] [doi: [10.1038/s41586-020-2008-3](https://doi.org/10.1038/s41586-020-2008-3)] [Medline: [32015508](https://pubmed.ncbi.nlm.nih.gov/32015508/)]
2. Tang D, Tou J, Wang J, Chen Q, Wang W, Huang J, et al. Prevention and control strategies for emergency, limited-term, and elective operations in pediatric surgery during the epidemic period of COVID-19. *World J Pediatr Surg* 2020 Mar 26;3(1):e000122. [doi: [10.1136/wjps-2020-000122](https://doi.org/10.1136/wjps-2020-000122)]
3. Morrison LG, Yardley L. What infection control measures will people carry out to reduce transmission of pandemic influenza? A focus group study. *BMC Public Health* 2009 Jul 23;9:258 [FREE Full text] [doi: [10.1186/1471-2458-9-258](https://doi.org/10.1186/1471-2458-9-258)] [Medline: [19627568](https://pubmed.ncbi.nlm.nih.gov/19627568/)]
4. Aiello AE, Coulborn RM, Aragon TJ, Baker MG, Burrus BB, Cowling BJ, et al. Research findings from nonpharmaceutical intervention studies for pandemic influenza and current gaps in the research. *Am J Infect Control* 2010 May;38(4):251-258. [doi: [10.1016/j.ajic.2009.12.007](https://doi.org/10.1016/j.ajic.2009.12.007)] [Medline: [20226569](https://pubmed.ncbi.nlm.nih.gov/20226569/)]
5. Zhang L, Kong Y, Chang H. Media use and health behavior in H1N1 flu crisis: The mediating role of perceived knowledge and fear. *Atl J Commun* 2015 Apr 30;23(2):67-80. [doi: [10.1080/15456870.2015.1013101](https://doi.org/10.1080/15456870.2015.1013101)]
6. Xavier AJ, d'Orsi E, Wardle J, Demakakos P, Smith SG, von Wagner C. Internet use and cancer-preventive behaviors in older adults: Findings from a longitudinal cohort study. *Cancer Epidemiol Biomarkers Prev* 2013 Nov;22(11):2066-2074 [FREE Full text] [doi: [10.1158/1055-9965.EPI-13-0542](https://doi.org/10.1158/1055-9965.EPI-13-0542)] [Medline: [24148971](https://pubmed.ncbi.nlm.nih.gov/24148971/)]
7. Vallejo BM, Ong RAC. Policy responses and government science advice for the COVID 19 pandemic in the Philippines: January to April 2020. *Prog Disaster Sci* 2020 Oct;7:100115 [FREE Full text] [doi: [10.1016/j.pdisas.2020.100115](https://doi.org/10.1016/j.pdisas.2020.100115)] [Medline: [34173440](https://pubmed.ncbi.nlm.nih.gov/34173440/)]
8. About EPI-WIN. World Health Organization. 2021. URL: <https://www.who.int/teams/risk-communication/about-epi-win> [accessed 2021-02-25]
9. Zarocostas J. How to fight an infodemic. *Lancet* 2020 Feb 29;395(10225):676 [FREE Full text] [doi: [10.1016/S0140-6736\(20\)30461-X](https://doi.org/10.1016/S0140-6736(20)30461-X)] [Medline: [32113495](https://pubmed.ncbi.nlm.nih.gov/32113495/)]
10. Wang P, Lu W, Ko N, Chen Y, Li D, Chang Y, et al. COVID-19-related information sources and the relationship with confidence in people coping with COVID-19: Facebook survey study in Taiwan. *J Med Internet Res* 2020 Jun 05;22(6):e20021 [FREE Full text] [doi: [10.2196/20021](https://doi.org/10.2196/20021)] [Medline: [32490839](https://pubmed.ncbi.nlm.nih.gov/32490839/)]

11. Pérez-Escoda A, Jiménez-Narros C, Perlado-Lamo-de-Espinosa M, Pedrero-Esteban LM. Social networks' engagement during the COVID-19 pandemic in Spain: Health media vs healthcare professionals. *Int J Environ Res Public Health* 2020 Jul 21;17(14):5261 [FREE Full text] [doi: [10.3390/ijerph17145261](https://doi.org/10.3390/ijerph17145261)] [Medline: [32708231](https://pubmed.ncbi.nlm.nih.gov/32708231/)]
12. Freberg K, Palenchar MJ, Veil SR. Managing and sharing H1N1 crisis information using social media bookmarking services. *Public Relat Rev* 2013 Sep;39(3):178-184. [doi: [10.1016/j.pubrev.2013.02.007](https://doi.org/10.1016/j.pubrev.2013.02.007)]
13. Jang K, Baek YM. When information from public health officials is untrustworthy: The use of online news, interpersonal networks, and social media during the MERS outbreak in South Korea. *Health Commun* 2019 Aug;34(9):991-998. [doi: [10.1080/10410236.2018.1449552](https://doi.org/10.1080/10410236.2018.1449552)] [Medline: [29558170](https://pubmed.ncbi.nlm.nih.gov/29558170/)]
14. Spence PR, Lachlan K, Burke JM, Seeger MW. Media use and information needs of the disabled during a natural disaster. *J Health Care Poor Underserved* 2007 May;18(2):394-404. [doi: [10.1353/hpu.2007.0047](https://doi.org/10.1353/hpu.2007.0047)] [Medline: [17483567](https://pubmed.ncbi.nlm.nih.gov/17483567/)]
15. Liu BF, Fraustino JD, Jin Y. Social media use during disasters. *Communic Res* 2015 Jan 13;43(5):626-646. [doi: [10.1177/0093650214565917](https://doi.org/10.1177/0093650214565917)]
16. Infodemic management. World Health Organization. 2021. URL: <https://www.who.int/teams/risk-communication/infodemic-management> [accessed 2021-02-26]
17. The Kingdom of Saudi Arabia's Experience in Health Preparedness and Response to COVID-19 Pandemic. Riyadh, Saudi Arabia: Ministry of Health; 2020 Aug. URL: <https://www.moh.gov.sa/en/Ministry/MediaCenter/Publications/Documents/COVID-19-NATIONAL.pdf> [accessed 2021-03-01]
18. Hassounah M, Raheel H, Alhefzi M. Digital response during the COVID-19 pandemic in Saudi Arabia. *J Med Internet Res* 2020 Sep 01;22(9):e19338 [FREE Full text] [doi: [10.2196/19338](https://doi.org/10.2196/19338)] [Medline: [32790642](https://pubmed.ncbi.nlm.nih.gov/32790642/)]
19. MOH issues COVID-19 awareness guidelines in many languages. Ministry of Health. 2020 Mar 17. URL: <https://www.moh.gov.sa/en/Ministry/MediaCenter/News/Pages/News-2020-03-17-001.aspx> [accessed 2021-02-24]
20. Almubark R, Basyouni M, Alghanem A, Althumairi N, Alkhamis D, Alharbi LS, et al. Health literacy in Saudi Arabia: Implications for public health and healthcare access. *Pharmacol Res Perspect* 2019 Aug;7(4):e00514 [FREE Full text] [doi: [10.1002/prp2.514](https://doi.org/10.1002/prp2.514)] [Medline: [31397117](https://pubmed.ncbi.nlm.nih.gov/31397117/)]
21. Al-Hanawi MK, Angawi K, Alshareef N, Qattan AMN, Helmy HZ, Abudawood Y, et al. Knowledge, attitude and practice toward COVID-19 among the public in the Kingdom of Saudi Arabia: A cross-sectional study. *Front Public Health* 2020;8:217 [FREE Full text] [doi: [10.3389/fpubh.2020.00217](https://doi.org/10.3389/fpubh.2020.00217)] [Medline: [32574300](https://pubmed.ncbi.nlm.nih.gov/32574300/)]
22. COVID-19. Centers for Disease Control and Prevention. 2021. URL: <https://www.cdc.gov/coronavirus/2019-nCoV/index.html> [accessed 2021-03-01]
23. Seo M. Amplifying panic and facilitating prevention: Multifaceted effects of traditional and social media use during the 2015 MERS crisis in South Korea. *Journal Mass Commun Q* 2019 Jul 26;98(1):221-240. [doi: [10.1177/1077699019857693](https://doi.org/10.1177/1077699019857693)]
24. Veil SR, Buehner T, Palenchar MJ. A work-in-process literature review: Incorporating social media in risk and crisis communication. *J Conting Crisis Manage* 2011;19(2):110-122. [doi: [10.1111/j.1468-5973.2011.00639.x](https://doi.org/10.1111/j.1468-5973.2011.00639.x)]
25. Zeng L, Starbird K, Spiro ES. #Unconfirmed: Classifying rumor stance in crisis-related social media messages. In: Proceedings of the 10th International AAAI Conference on Web and Social Media. 2016 Presented at: 10th International AAAI Conference on Web and Social Media; May 17-20, 2016; Cologne, Germany p. 747-750 URL: <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM16/paper/view/13075/12845>
26. Liu PL. COVID-19 information seeking on digital media and preventive behaviors: The mediation role of worry. *Cyberpsychol Behav Soc Netw* 2020 Oct;23(10):677-682. [doi: [10.1089/cyber.2020.0250](https://doi.org/10.1089/cyber.2020.0250)] [Medline: [32498549](https://pubmed.ncbi.nlm.nih.gov/32498549/)]
27. Tsfati Y, Cappella JN. Do people watch what they do not trust? *Communic Res* 2016 Jun 30;30(5):504-529. [doi: [10.1177/0093650203253371](https://doi.org/10.1177/0093650203253371)]
28. Xiao N, Sharman R, Rao H, Upadhyaya S. Factors influencing online health information search: An empirical analysis of a national cancer-related survey. *Decis Support Syst* 2014 Jan;57:417-427. [doi: [10.1016/j.dss.2012.10.047](https://doi.org/10.1016/j.dss.2012.10.047)]
29. Lee JJ, Kang K, Wang MP, Zhao SZ, Wong JYH, O'Connor S, et al. Associations between COVID-19 misinformation exposure and belief with COVID-19 knowledge and preventive behaviors: Cross-sectional online study. *J Med Internet Res* 2020 Nov 13;22(11):e22205 [FREE Full text] [doi: [10.2196/22205](https://doi.org/10.2196/22205)] [Medline: [33048825](https://pubmed.ncbi.nlm.nih.gov/33048825/)]
30. van Nieuwenhuijzen M, Junger M, Velderman MK, Wiefferink KH, Paulussen TW, Hox J, et al. Clustering of health-compromising behavior and delinquency in adolescents and adults in the Dutch population. *Prev Med* 2009 Jun;48(6):572-578. [doi: [10.1016/j.ypmed.2009.04.008](https://doi.org/10.1016/j.ypmed.2009.04.008)] [Medline: [19389423](https://pubmed.ncbi.nlm.nih.gov/19389423/)]
31. Jahangard L, Behmanesh H, Ahmadpanah M, Poormoosavi SM, Solitani A, Highighi M. Risky behaviors and health-promoting behaviors in young adults: An epidemiological study. *Iran J Psychiatry* 2019 Oct;14(4):302-308 [FREE Full text] [Medline: [32071604](https://pubmed.ncbi.nlm.nih.gov/32071604/)]
32. Alanezi F, Aljadhali A, Alyousef S, Alrashed H, Alshaikh W, Mushcab H, et al. Implications of public understanding of COVID-19 in Saudi Arabia for fostering effective communication through awareness framework. *Front Public Health* 2020;8:494 [FREE Full text] [doi: [10.3389/fpubh.2020.00494](https://doi.org/10.3389/fpubh.2020.00494)] [Medline: [33042941](https://pubmed.ncbi.nlm.nih.gov/33042941/)]
33. Lupton D, Lewis S. Learning about COVID-19: A qualitative interview study of Australians' use of information sources. *BMC Public Health* 2021 Apr 07;21(1):662 [FREE Full text] [doi: [10.1186/s12889-021-10743-7](https://doi.org/10.1186/s12889-021-10743-7)] [Medline: [33823843](https://pubmed.ncbi.nlm.nih.gov/33823843/)]
34. Lin CA, Lagoe C. Effects of news media and interpersonal interactions on H1N1 risk perception and vaccination intent. *Commun Res Rep* 2013 Apr;30(2):127-136. [doi: [10.1080/08824096.2012.762907](https://doi.org/10.1080/08824096.2012.762907)]

35. Southwell BG, Boudewyns V, Hwang Y, Yzer M. Entertainment tonight? The value of informative TV news among US viewers. *Electron News* 2008 Jul;2(3):123-137. [doi: [10.1080/19312430802213043](https://doi.org/10.1080/19312430802213043)]
36. Browne B. *Global Attitudes to COVID-19 Pandemic and Response*. Canberra, Australia: The Australia Institute; 2020 Apr. URL: <http://australiainstitute.org.au/wp-content/uploads/2020/12/April-2020-Global-attitudes-to-COVID-19-pandemic-and-response-WEB.pdf> [accessed 2021-03-15]
37. Kelly B, Hornik R, Romantan A, Schwartz JS, Armstrong K, DeMichele A, et al. Cancer information scanning and seeking in the general population. *J Health Commun* 2010 Oct;15(7):734-753 [FREE Full text] [doi: [10.1080/10810730.2010.514029](https://doi.org/10.1080/10810730.2010.514029)] [Medline: [21104503](https://pubmed.ncbi.nlm.nih.gov/21104503/)]

## Abbreviations

**aOR:** adjusted odds ratio  
**CDC:** Centers for Disease Control and Prevention  
**KSA:** Kingdom of Saudi Arabia  
**MOH:** Ministry of Health  
**OR:** odds ratio  
**SR:** Saudi Riyal  
**WHO:** World Health Organization

*Edited by G Eysenbach; submitted 17.03.21; peer-reviewed by N Alharbi, K Reuter, S Ghaddar; comments to author 08.04.21; revised version received 21.04.21; accepted 19.05.21; published 30.07.21.*

*Please cite as:*

*Alshareef N, Yunusa I, Al-Hanawi MK*

*The Influence of COVID-19 Information Sources on the Attitudes and Practices Toward COVID-19 Among the General Public of Saudi Arabia: Cross-sectional Online Survey Study*

*JMIR Public Health Surveill* 2021;7(7):e28888

URL: <https://publichealth.jmir.org/2021/7/e28888>

doi: [10.2196/28888](https://doi.org/10.2196/28888)

PMID: [34081610](https://pubmed.ncbi.nlm.nih.gov/34081610/)

©Noor Alshareef, Ismaeel Yunusa, Mohammed Khaled Al-Hanawi. Originally published in *JMIR Public Health and Surveillance* (<https://publichealth.jmir.org>), 30.07.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Public Health and Surveillance*, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.



Original Paper

# The Saudi Ministry of Health's Twitter Communication Strategies and Public Engagement During the COVID-19 Pandemic: Content Analysis Study

Fatimah Mohammed Alhassan<sup>1,2</sup>, BSc, MSc; Sharifah Abdullah AlDossary<sup>1,2</sup>, BSc, MSc, PhD

<sup>1</sup>Department of Health Informatics, College of Public Health and Health Informatics, King Saud bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia

<sup>2</sup>King Abdullah International Medical Research Center, Riyadh, Saudi Arabia

**Corresponding Author:**

Fatimah Mohammed Alhassan, BSc, MSc

Department of Health Informatics

College of Public Health and Health Informatics

King Saud bin Abdulaziz University for Health Sciences

PO Box 3660

Riyadh, 11481

Saudi Arabia

Phone: 966 11 4299999

Email: [fatimah.m.alhasan@gmail.com](mailto:fatimah.m.alhasan@gmail.com)

## Abstract

**Background:** During a public health crisis such as the current COVID-19 pandemic, governments and health authorities need quick and accurate methods of communicating with the public. While social media can serve as a useful tool for effective communication during disease outbreaks, few studies have elucidated how these platforms are used by the Ministry of Health (MOH) during disease outbreaks in Saudi Arabia.

**Objective:** Guided by the Crisis and Emergency Risk Communication model, this study aimed to explore the MOH's use of Twitter and the public's engagement during different stages of the COVID-19 pandemic in Saudi Arabia.

**Methods:** Tweets and corresponding likes and retweets were extracted from the official Twitter account of the MOH in Saudi Arabia for the period of January 1 through August 31, 2020. Tweets related to COVID-19 were identified; subsequently, content analysis was performed, in which tweets were coded for the following message types: risk messages, warnings, preparations, uncertainty reduction, efficacy, reassurance, and digital health responses. Public engagement was measured by examining the numbers of likes and retweets. The association between outbreak stages and types of messages was assessed, as well as the effect of these messages on public engagement.

**Results:** The MOH posted a total of 1393 original tweets during the study period. Of the total tweets, 1293 (92.82%) were related to COVID-19, and 1217 were ultimately included in the analysis. The MOH posted the majority of its tweets (65.89%) during the initial stage of the outbreak. Accordingly, the public showed the highest level of engagement (as indicated by numbers of likes and retweets) during the initial stage. The types of messages sent by the MOH significantly differed across outbreak stages, with messages related to uncertainty reduction, reassurance, and efficacy being prevalent among all stages. Tweet content, media type, and crisis stage influenced the level of public engagement. Engagement was negatively associated with the inclusion of hyperlinks and multimedia files, while higher level of public engagement was associated with the use of hashtags. Tweets related to warnings, uncertainty reduction, and reassurance received high levels of public engagement.

**Conclusions:** This study provides insights into the Saudi MOH's communication strategy during the COVID-19 pandemic. Our results have implications for researchers, governments, health organizations, and practitioners with regard to their communication practices during outbreaks. To increase public engagement, governments and health authorities should consider the public's need for information. This, in turn, could raise public awareness regarding disease outbreaks.

(*JMIR Public Health Surveill* 2021;7(7):e27942) doi:[10.2196/27942](https://doi.org/10.2196/27942)

**KEYWORDS**

COVID-19; Crisis and Emergency Risk Communication; effective communication; health authorities; outbreak; pandemic; public engagement; public health; social media; Twitter

## *Introduction*

### **Background**

Coronaviruses are a large family of viruses that cause diseases ranging from those with common cold symptoms to more severe pneumonia-like illnesses [1]. On December 31, 2019, the World Health Organization (WHO) Country Office in China declared that a new coronavirus, SARS-CoV-2, had been detected in Wuhan. Within a few weeks, the virus had spread from Wuhan to many provinces within China. It subsequently spread outside China, reaching over 200 countries. The rapid and continuous spread of the virus led the WHO to declare COVID-19, caused by SARS-CoV-2, a public health emergency of international concern on January 30, 2020, and a pandemic on March 11, 2020 [2-5].

Saudi Arabia is the second largest Arab country with a population of over 34 million people [6]. The Ministry of Health (MOH) in Saudi Arabia is the largest provider of health care services, providing approximately 60% of the health care services nationwide, while the remainder is covered by other governmental and private facilities [7]. Since the confirmation of the first case of COVID-19 in Saudi Arabia on March 2, 2020, the government has taken prompt and decisive measures to combat the outbreak. These measures included, but were not limited to, closures of borders, schools, mosques, and Umrah (the minor pilgrimage to Mecca, which can be undertaken any time of the year); cessation of international flights; mandatory quarantine periods for returning travelers; workplace closures, with individuals working from home (apart from essential workers); and partial to complete lockdowns [8]. Digital health measures were also implemented and effectively utilized during the pandemic [9,10]. As of March 25, 2020, the MOH designated 25 hospitals with 80,000 hospital beds and 8000 intensive care unit beds for the treatment of COVID-19 cases. An additional 2200 beds were allocated for the isolation of suspected and quarantined cases [11]. Saudi Arabia has robust preparedness and response capabilities that have been strengthened through prior experience with the Middle East respiratory syndrome coronavirus and decades of planning and managing religious mass gatherings of Hajj (the annual pilgrimage to Mecca) and Umrah, which can serve as a model for other countries in the region.

In public health emergencies such as the COVID-19 pandemic, effective communication is crucial for informing the public about disease situation updates, motivating them to adopt preventive measures, and reassuring them that the government is in control of the outbreak [12-14]. Such communication requires timely dissemination of accurate and reliable information. Traditionally, governments and public health authorities have relied on websites, print media, and television as the main platforms for disseminating outbreak-related information to the public. However, the evolution of digital communications technologies such as social media has

facilitated increased sharing of information for both public health authorities and the general public.

In recent years, social media has developed rapidly. Both individuals and health care organizations are using these platforms increasingly to communicate and share information [15]. Social media facilitates 2-way communication and direct engagement with audiences. Safko and Brake [16] define social media as “activities, practices, and behaviors among communities of people who gather online to share information, knowledge, and opinions using conversational media.” Today, there are over 3.8 billion active social media users worldwide across many different platforms [17]. In the field of health education and promotion, the use of social media has established its effectiveness by providing access to information, delivering health campaigns, and offering social support [18]. Many government agencies and public health organizations (eg, the WHO, the Centers for Disease Control and Prevention [CDC], and other local health departments) have adopted social media to enhance their communication with the public [19].

Social media can serve as a useful tool to relay outbreak-related updates and critical information effectively to the public. Existing research suggests that people often turn to social media for information during infectious disease outbreaks, which can influence their decision-making and subsequent behaviors [19]. The WHO calls for more proactive use of social media to disseminate health messages to journalists, physicians, and the general public, particularly to counteract misinformation regarding infectious diseases [20].

Several studies have investigated the use of social media platforms such as Twitter and Facebook during infectious disease outbreaks. For example, Chen et al [21] studied the temporal variability in the CDC’s response during different stages of the Zika epidemic and public engagement on Twitter. They reported that the CDC was more active in the early warning stages of the Zika epidemic and successfully gained public attention, particularly in the first quarter of 2016. However, when the number of Zika cases increased sharply in the second and third quarters of 2016, the CDC’s efforts on Twitter decreased substantially.

Lwin et al [22] examined the strategic use of Facebook in communicating the Zika epidemic by three main Singapore health authorities: the National Environment Agency, the Health Promotion Board, and the MOH. The researchers found that Facebook was used strategically for Zika-related communication. They also found that preparedness messages (eg, posts mentioning responders and providing recommendations to reduce harm) may have been the most effective, as evidenced by greater levels of public engagement.

Guidry et al [23] examined Ebola-related posts on Instagram and Twitter from three key health organizations: the CDC, the WHO, and Me´decins Sans Frontie`res (ie, Doctors Without Borders). They found Instagram to be a particularly useful

platform for communicating with the public during crises. It was further suggested that social media messaging is more effective when it is utilized by health organizations with which the public is already familiar, and when it is based on strategic use of risk communication principles.

A recent study by Raamkumar [24] examined the use of Facebook for COVID-19-related outreach by public health authorities in Singapore, the United States, and England, and the corresponding public response to these efforts. They reported that the Singapore MOH was the most active in terms of posting frequency, while the CDC elicited the most responses. Furthermore, they reported that posts on preventive and safety measures and situation updates were the most frequently employed by public health authorities in these 3 countries.

### Theoretical Framework

Crisis events such as the COVID-19 pandemic requires unique health communication and education strategies in which public health authorities must meet the public needs for information [14]. Theories suggest that the public has various information needs at different stages of a crisis [25]. The Crisis and Emergency Risk Communication (CERC) model serves as a useful tool to guide authorities' communication strategies during different stages of a crisis. It specifies a broad set of communication activities that vary throughout the life cycle of the crisis. The CERC model was originally developed by the CDC after the 2001 anthrax attacks and the events of September 11, 2001, in the United States [26,27]. It is an integrated model that draws elements from risk communication theories (persuading individuals to take action to limit risks), crisis communication theories (responding to the public's immediate need for information), and theories of health communication [27].

The CERC model describes five general stages of a crisis: precrisis, initial event, maintenance, resolution, and evaluation. For each stage, a set of recommended communication activities is also described. According to the model, specific and distinct communication activities should be carried out in each stage.

The first stage of the CERC model is the precrisis period. In this stage, the crisis has yet to occur. Communication messages at this stage should focus on risk information, warnings, and preparation.

The second stage of the model is the initial event, when the crisis actually occurs. This stage is initiated by a clear trigger event that signals the beginning of a crisis. Communication messages in this stage should focus on reducing public uncertainty by providing timely updates regarding the crisis, messages of self-efficacy, and reassurance from authority-initiated measures.

The third stage, or maintenance stage, begins when "most or all of the direct harm is contained, and the intensity of the crisis begins to subside" [28]. It further echoes many of the communication activities from earlier stages, including uncertainty reduction, reassurance, and self-efficacy messages.

The fourth stage is the resolution stage, in which the crisis continues to wind down and new understandings of risk emerge.

Communication messages at this stage involve updates about ongoing resolutions, discussions about causes, and new understandings of risk.

The final stage is the evaluation stage. This stage occurs when the crisis itself is over. Communication during this stage should focus on assessing the adequacy and efficacy of the response and reaching a consensus on the lessons learned from the crisis [26,28].

### Objective

With over 15 million Twitter users in Saudi Arabia [29], the microblogging and social media platform Twitter presents an opportunity to examine the Saudi MOH's use of social media in crisis communication during pandemics. According to a recent national survey, approximately 78% of respondents reported the MOH as their main source of information about COVID-19 [30]. The objective of this study is to investigate the use of Twitter by the MOH and the associated public engagement during different stages of the COVID-19 pandemic in Saudi Arabia.

### Methods

#### Data Collection

All tweets from the Saudi MOH (@SaudiMOH) posted between January 1 and August 31, 2020, were collected and included in the study. The tweets were collected on September 30, 2020, via the GET statuses/user\_timeline endpoint of Twitter's application programming interface [31] by using the python library Tweepy [32]. For each tweet, the following data were collected: tweet ID, tweet text (body), number of likes, number of retweets, and date posted. Only original tweets, rather than retweets, were considered. Tweets written in languages other than Arabic or English were removed to avoid misinterpretation. Tweets unrelated to COVID-19 were manually excluded by scanning the content of the tweets. Daily confirmed case counts of COVID-19 in Saudi Arabia were obtained from Our World in Data [33].

#### Crisis Stages

The various stages of the CERC model have not been clearly defined or operationalized within the context of infectious disease outbreaks. The CERC model assumes that "crises will develop in largely predictable and systematic ways" [26]. However, infectious disease outbreaks such as the current COVID-19 pandemic may last for months or even years, without clear boundaries (compared to other crises, such as extreme weather events) [34,35]. In this study, outbreak stages were determined on the basis of the CERC model and specific events of the COVID-19 outbreak in Saudi Arabia.

The precrisis stage was determined to span from January 1 to March 1, 2020, when the outbreak began in China and some European countries, but when no cases had yet been identified in Saudi Arabia.

The initial event stage was determined to last from March 2 to June 20, 2020. The first confirmed case of COVID-19 in Saudi Arabia was reported on March 2. During this period, there were 157,600 confirmed cases of COVID-19 in the country.

The maintenance stage was determined to span from June 21 to August 31, 2020. On June 21, 2020, Saudi Arabian authorities lifted the nationwide curfew and allowed the resumption of all activities, except in Mecca.

### Content Analysis and Coding Categories

This study used a content analysis approach to investigate the Saudi MOH's communication on Twitter regarding the COVID-19 pandemic. Berg [36] explained that content analysis is a "careful, detailed, systematic examination and interpretation of a particular body of material in an effort to identify patterns, themes, biases, and meanings." COVID-19-related tweets were coded using a codebook adapted from one that was developed on the basis of the CERC framework [22]. The codebook consisted of the following categories: (1) risk messages, including tweets containing information about the disease, its transmission mechanisms, and its symptoms; (2) warnings, including tweets highlighting risk factors and dangers associated with COVID-19; (3) preparations, including tweets mentioning responders and providing response recommendations and advice; (4) uncertainty reduction, including tweets containing information about case reports and providing the public with reliable sources of information; (5) efficacy, including tweets containing information about personal preventive measures and highlighting the common responsibility for disease prevention; and (6) reassurance, including tweets that calmed the public by providing information about government interventions and expressing gratitude and regards to the health staff and the public. Through initial scanning of the tweets, digital responses were found to be frequently mentioned in the MOH tweets. Given the importance of digital health during the COVID-19 pandemic, an additional category called "digital health responses" was introduced. Messages of this category pertained to tweets promoting digital health services, ranging from digital screening to surveillance, contact tracing, and follow-up apps. The codebook is presented in [Multimedia Appendix 1](#).

Each tweet was primarily categorized on the basis of the content within its 280 characters. If the content was not clear, linked visuals such as photographs, videos, and other media were analyzed. It was beyond the scope of this study to analyze the content of hyperlinks. The content analysis of the MOH's tweets was conducted using Excel (version 16.43, Microsoft Inc), which was later imported into SPSS (version 27, IBM Corp) for statistical analysis.

### Intercoder Reliability

Intercoder reliability was established by 2 independent coders (FH and SD). Each author independently coded a randomly selected subsample of 122 (10%) tweets. This meets the Neuendorf recommendation of coding 10%-20% of the total sample for reliability [37]. Reliability was assessed using the ReCal statistical program with the Cohen  $\kappa$  statistic [38]. The  $\kappa$  values for all categories were greater than 0.8, which indicated "almost perfect" agreement, except for the coding category of "Responders," which had a  $\kappa$  value of 0.545, indicating "moderate agreement" [39]. Coding discrepancies for the "Responders" category were resolved through discussion. Once intercoder reliability was established, the first coder coded the remaining tweets.

### Statistical Analysis

Statistical analysis was performed using SPSS (version 27) [40]. Counts and percentages were used to summarize categorical variables. The median was used to summarize the distribution of continuous variables owing to the skewed nature of those included. The chi-square test of independence was used to assess the associations between outbreak stages and message types. Significant chi-square outcomes were further subjected to multiple post hoc Z-tests to compare each pair of outbreak stages. *P* values were adjusted for the false discovery rate by using the Benjamini-Hochberg adjustment.

The nonparametric Kruskal-Wallis test was used to compare the distribution of likes and retweets (engagement indicators) between outbreak stages. The Mann-Whitney *U* test was used to test differences in engagement between tweets including and those not including different message types.

A negative binomial regression analysis was used to examine associations among tweet content, media type, crisis stage, and public engagement. Negative binomial regression was used as engagement variables demonstrated positive skew and overdispersion. The incidence rate ratio (IRR) was calculated as the exponent of the regression coefficients. All statistical tests were performed at a significance level of .05.

### Ethics Approval

Ethics approval was not required for this study as the study did not involve any human subjects. All data analyzed in this study were publicly available and collected from a governmental public Twitter account.

### Data Availability

The data that support our findings are available on request from the corresponding author.

## Results

### Results Overview

The MOH posted a total of 1393 original tweets (an average of 5.85 tweets per day) during the study period. Overall, 1293 (92.8%) tweets were related to COVID-19, of which 1217 were included in the analysis. The other tweets ( $n=76$ ) were removed because they were in languages other than Arabic or English. The results are presented in three sections: the MOH response to COVID-19 on Twitter across stages, message types across stages, and public engagement with tweets from the MOH.

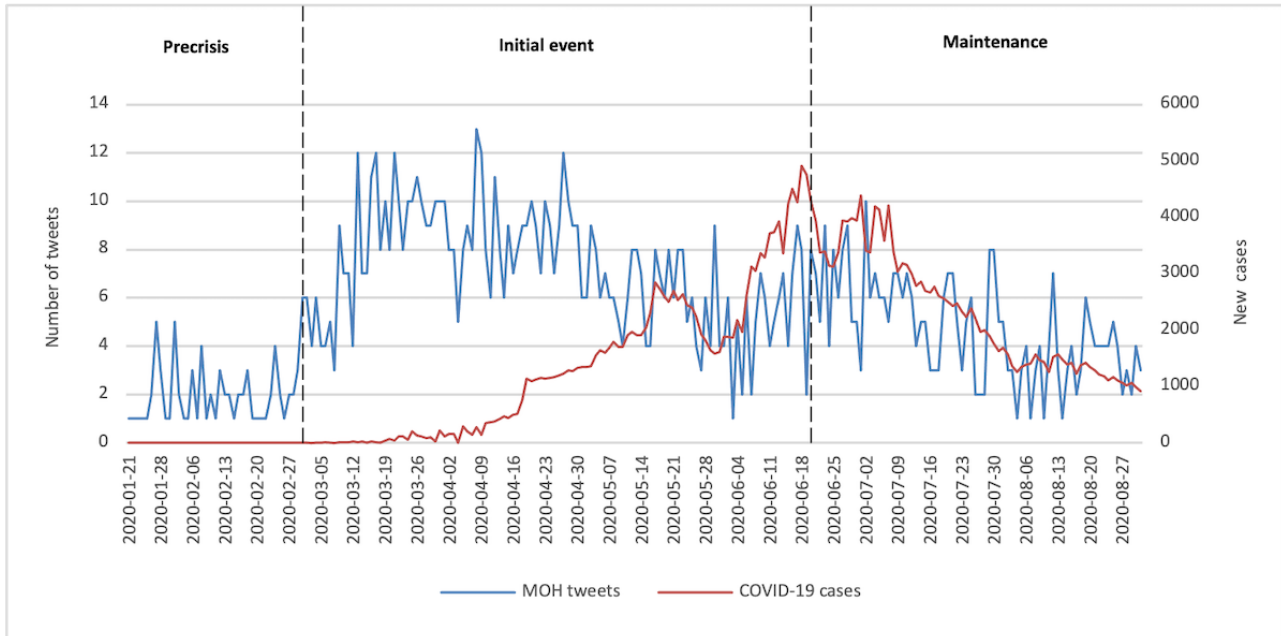
### MOH Response to COVID-19 on Twitter Across Stages of the COVID-19 Outbreak

Confirmed and reported COVID-19 cases were plotted in relation to COVID-19-related tweets posted by the MOH ([Figure 1](#)). The first COVID-19-related tweet was posted on January 21, 2020. Overall, 79 (6.5%) tweets were posted during the precrisis stage, when the outbreak began in China and some European countries, but when no cases had yet been identified in Saudi Arabia (2.03 tweets per day on average). On March 2, 2020, the MOH confirmed the first COVID-19 case in Saudi Arabia, which signaled the start of the crisis. COVID-19-related tweets were consistently posted as the number of cases

increased, with an average of 7.23 daily tweets. As of June 20, 2020, the MOH had posted 802 tweets (66% of the total), which is a >3-fold increase in its average daily tweets compared to the precrisis stage. On June 21, 2020, the country lifted curfew

restrictions and resumed all economic and commercial activities [41]. From that date until the end of August 2020, the MOH continued to provide ongoing information regarding COVID-19, with an average of 4.67 daily tweets (n=336, 27.6%).

**Figure 1.** The Saudi Ministry of Health's Twitter communication in relation to confirmed COVID-19 cases (January 1 to August 31, 2020). MOH: Ministry of Health.



### Message Types Across Stages of the COVID-19 Outbreak

Message types across stages of the outbreak are summarized in Table 1. Of the 1217 tweets, nearly half (49.47%) contained uncertainty reduction information, 28.35% contained efficacy information, and one-fifth (21.53%) of all tweets contained reassuring information.

Tweets about warning messages accounted for a low proportion of all tweets in the precrisis stage (2.5%), which significantly increased during the initial stage ( $\chi^2_2=6.2$ ;  $P=.046$ ) (Table 1 and Multimedia Appendix 2). Tweets about preparation

messages accounted for a low proportion of tweets in the precrisis stage (11.4%) and the initial stage (11.2%), and the proportion increased significantly in the maintenance stage (22.3%;  $\chi^2_2=24.4$ ;  $P<.001$ ). Conversely, the percentage of reassurance tweets peaked in the precrisis stage (67.1%) and significantly decreased in later stages ( $\chi^2_2=103.9$ ;  $P<.001$ ). The frequency of efficacy tweets was higher in the initial stage (34.4%) than in the precrisis (15.2%) and maintenance (16.9%) stages ( $\chi^2_2=42.7$ ;  $P<.001$ ). Lastly, tweets promoting digital health services increased significantly in frequency, from 8.6% in the initial stage to 16.7% in the maintenance stage ( $\chi^2_2=26.4$ ;  $P<.001$ ).

**Table 1.** Categories of message types across outbreak stages in Saudi Arabia (January 1 to August 31, 2020).

Message type <sup>a</sup>	Precrisis stage (n=79), n (%)	Initial event stage (n=802), n (%)	Maintenance stage (n=336), n (%)	Overall (n=1217), n (%)	P value	$\chi^2$ (df=2)
<b>Risk messages</b>	5 (6.3)	62 (7.7)	21 (6.3)	88 (7.2)	.65	0.876
Disease information	4 (5.1)	24 (3.0)	8 (2.4)	36 (3.0)	.45	1.613
Symptoms	3 (4.8)	44 (5.5)	15 (4.5)	62 (5.1)	.67	0.805
<b>Warnings</b>	2 (2.5)	91 (11.3)	33 (9.8)	126 (10.4)	.046	6.162
Risk factor	2 (2.5)	40 (5.0)	29 (8.6)	71 (5.8)	.03	7.399
Danger	0 (0.0)	53 (6.6)	4 (1.2)	57 (4.7)	<.001	19.722
<b>Preparations</b>	9 (11.4)	90 (11.2)	75 (22.3)	174 (14.3)	<.001	24.390
Responders	8 (10.1)	27 (3.4)	11 (3.3)	46 (3.8)	.009	9.363
Recommendations	1 (1.3)	64 (8.0)	64 (19.0)	129 (10.6)	<.001	38.376
<b>Uncertainty reduction</b>	41 (51.9)	382 (47.6)	179 (53.3)	602 (49.5)	.20	3.216
Case report	30 (38.0)	157 (19.6)	72 (21.4)	259 (21.3)	.001	14.538
Information resources	12 (15.2)	225 (28.1)	107 (31.8)	344 (28.3)	.01	8.802
<b>Efficacy</b>	12 (15.2)	276 (34.4)	57 (17.0)	345 (28.3)	<.001	42.699
Personal prevention	12 (15.2)	226 (28.2)	51 (15.2)	289 (23.7)	<.001	25.520
Common responsibility	0 (0.0)	86 (10.7)	9 (2.7)	95 (7.8)	<.001	28.447
<b>Reassurance</b>	53 (67.1)	145 (18.1)	64 (19.0)	262 (21.5)	<.001	103.938
Calming	51 (64.6)	91 (11.3)	41 (12.2)	183 (15)	<.001	162.297
Thanks and regards	0 (0.0)	41 (5.1)	23 (6.8)	64 (5.3)	.047	6.117
Government interventions	4 (5.1)	57 (7.1)	26 (7.7)	87 (7.1)	.71	0.696
Digital health responses	0 (0.0)	69 (8.6)	56 (16.7)	125 (10.3)	<.001	26.375

<sup>a</sup>A tweet can have more than 1 category.

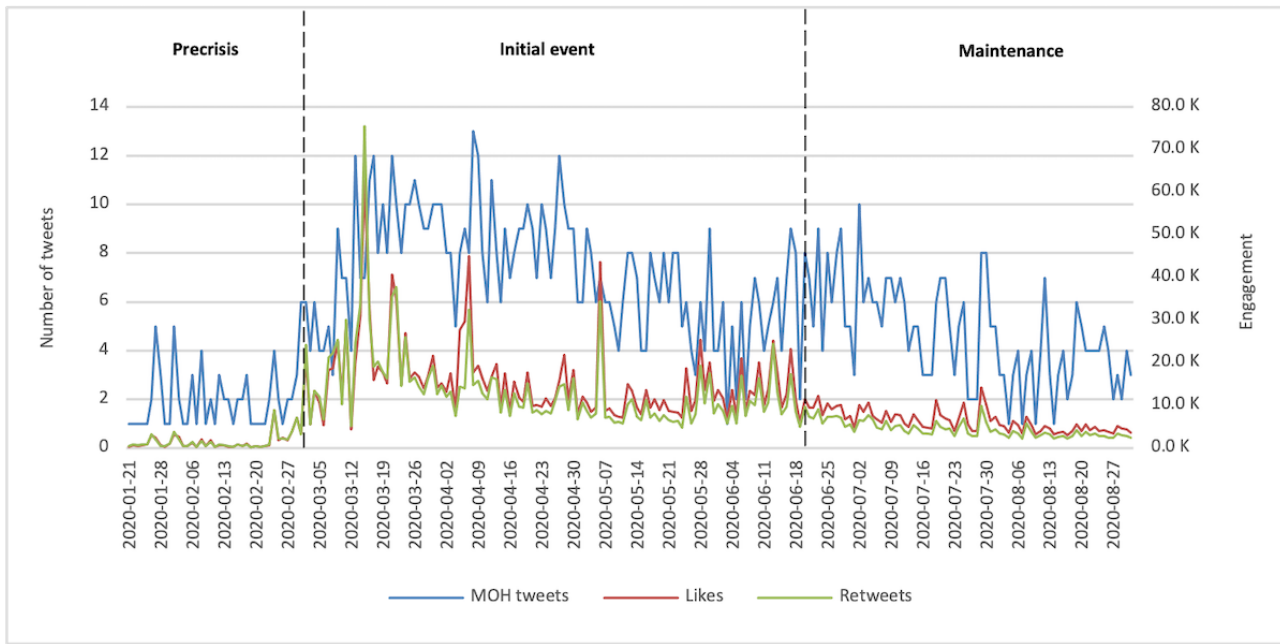
### Public Engagement With MOH Tweets

Figure 2 demonstrates public engagement (represented by frequencies of likes and retweets) in relation to MOH tweets. Public engagement was not accurately aligned with the development of the COVID-19 outbreak in the country; rather, the public was most engaged when the initial cases appeared in March 2020.

Since engagement variables were not normally distributed, median values (rather than mean values) were used for statistical

comparisons. MOH tweets were associated with median values of 819 likes and 603 retweets. A Kruskal–Wallis  $H$  test revealed significant differences in the frequencies of likes ( $\chi^2_2=70.344$ ;  $P<.001$ ) and retweets ( $\chi^2_2=59.764$ ;  $P<.001$ ) among different outbreak stages. The Mann–Whitney  $U$  test was used for post hoc comparisons of average ranks. Post hoc pairwise comparisons (Multimedia Appendix 3) revealed that the distribution of retweets was not significantly different between the precrisis and maintenance stages ( $Z=-2.14$ ;  $P=.96$ ).

**Figure 2.** Public engagement in relation to the Saudi Ministry of Health's tweets (January 1 to August 31, 2020). The left y-axis shows the daily number of tweets. The right y-axis shows the number of likes and retweets (in thousands). MOH: Ministry of Health.



Median public engagement levels across different stages of the outbreak and message types are summarized in Table 2. Overall, the public engaged most with tweets from the MOH during the initial stage, with median values of 974 likes and 753 retweets. Further analysis using Mann–Whitney *U* tests (Table 3) revealed that both like and retweet frequencies during the initial stage were significantly higher for tweets that provided uncertainty reduction messages ( $Z=-3.133$ ;  $P=.002$ ) and reassurance messages ( $Z=-5.843$ ;  $P<.001$ ) than those that did not. Tweets containing risk messages at this stage received fewer likes ( $Z=-4.219$ ;  $P<.001$ ) and retweets ( $Z=-4.252$ ;  $P<.001$ ) than those

that did not. In contrast, tweets containing risk messages during the precrisis stage received more likes ( $Z=-2.034$ ;  $P=.04$ ) and retweets ( $Z=-2.15$ ;  $P=.03$ ) than those that did not. Among tweets in the maintenance stage, the frequencies of both likes ( $Z=-3.708$ ;  $P<.001$ ) and retweets ( $Z=-3.605$ ;  $P<.001$ ) were significantly higher for tweets that contained uncertainty reduction information than for those that did not. Conversely, frequencies of both likes ( $Z=-4.534$ ;  $P<.001$ ) and retweets ( $Z=-4.547$ ;  $P<.001$ ) were significantly lower for tweets that promoted digital health services than for those that did not.

**Table 2.** Median public engagement across outbreak stages and message types (January 1 to August 31, 2020).

Message type <sup>a</sup>	Likes				Retweets			
	All stages	Precrisis stage	Initial event stage	Maintenance stage	All stages	Precrisis stage	Initial event stage	Maintenance stage
All	819.0	437.0	974.0	582.0	603.0	473.0	753.0	351.0
Risk messages	598.5	686.0	475.5	630.0	409.5	766.0	385.5	381.0
Warnings	785.0	1678.0	913.0	513.0	614.5	1801.0	710.0	335.0
Preparations	594.5	343.0	974.0	464.0	397.5	438.0	708.0	273.0
Uncertainty reduction	841.5	389.0	1025.0	678.0	658.5	438.0	770.0	409.0
Efficacy	852.0	560.0	950.5	590.0	683.0	623.5	755.5	422.0
Reassurance	982.0	418.0	1629.0	568.5	740.0	401.0	1198.0	316.0
Digital health responses	650.0	0.0	899.0	452.0	444.0	0.0	620.0	273.5

<sup>a</sup>Values presented are medians.

**Table 3.** Median public engagement for tweets with or without message type at different stages of the COVID-19 outbreak in Saudi Arabia (January 1 to August 31, 2020).

Variable <sup>a</sup>	Median present	Median absent	U value	Z-value	P value
<b>Precrisis</b>					
<b>Risk messages</b>					
Retweets	766.00	432.00	78.000	-2.154	.03
Likes	686.00	403.50	84.000	-2.034	.04
<b>Initial event</b>					
<b>Risk messages</b>					
Retweets	385.50	781.50	15490.000	-4.252	<.001
Likes	974.50	1008.00	15547.500	-4.219	<.001
<b>Uncertainty reduction</b>					
Retweets	770.00	723.50	70667.500	-2.915	.004
Likes	1025.00	964.50	69956.000	-3.133	.002
<b>Reassurance</b>					
Retweets	1198.00	659.00	34581.000	-5.169	<.001
Likes	1629.00	856.00	32880.500	-5.843	<.001
<b>Maintenance</b>					
<b>Uncertainty reduction</b>					
Retweets	409.00	317.50	10781.500	-3.681	<.001
Likes	678.50	524.50	10680.000	-3.795	<.001
<b>Digital health responses</b>					
Retweets	273.50	404.50	4823.000	-4.547	<.001
Likes	452.00	643.50	4831.000	-4.534	<.001

<sup>a</sup>Only significant variables are reported.

Negative binomial regression outcomes are summarized in [Table 4](#) (complete models are provided in [Multimedia Appendix 4](#)). The model was significantly better than the null model, which indicated that the variables (as a set) predicted the number of likes ( $n=1217$ ;  $\chi^2_{12}=458.627$ ;  $P<.001$ ). The negative binomial regression model for retweets ([Table 4](#)) was also significantly better than the null model ( $n=1217$ ;  $\chi^2_{12}=575.495$ ;  $P<.001$ ). The Wald test revealed that all of predictor variables were significant ( $P<.05$ ), except for risk messages and preparations.

Our results indicate that the use of hashtags was significantly associated with higher levels of engagement (likes:  $IRR=2.470$ ;  $P<.001$ ; retweets:  $IRR=2.813$ ;  $P<.001$ ), whereas the use of hyperlinks was significantly associated with lower levels of engagement (likes:  $IRR=0.839$ ;  $P=.045$ ; retweets:  $IRR=0.727$ ;  $P<.001$ ). Compared to text-only content, the use of photographs and videos was associated with significantly lower numbers of

likes and retweets (photographs: likes,  $IRR=0.530$ ;  $P<.001$ ; retweets,  $IRR=0.476$ ;  $P<.001$ ; videos: likes,  $IRR=0.698$ ;  $P=.006$ ; retweets,  $IRR=0.576$ ;  $P<.006$ ).

With respect to the impact of content type on public engagement, tweets with content related to warnings (likes:  $IRR=1.334$ ;  $P=.005$ ; retweets:  $IRR=1.544$ ;  $P<.001$ ), uncertainty reduction (likes:  $IRR=2.210$ ;  $P<.001$ ; retweets:  $IRR=2.197$ ;  $P<.001$ ), and reassurance (likes:  $IRR=1.551$ ;  $P<.001$ ; retweets:  $IRR=1.517$ ;  $P<.001$ ) were significantly associated with higher levels of engagement.

Regarding crisis stages, tweets posted during the initial and maintenance stages were significantly associated with higher levels of engagement than those posted during the precrisis stage (initial stage: likes,  $IRR=2.931$ ;  $P<.001$ ; retweets,  $IRR=2.471$ ;  $P<.001$ ; maintenance stage: likes,  $IRR=2.355$ ;  $P<.001$ ; retweets,  $IRR=1.623$ ;  $P<.001$ ).



**Table 4.** Associations of tweet content, media type, and crisis stage with public engagement (January 1 to August 31, 2020).

Variables	Likes			Retweets		
	Incidence rate ratio	95% CI	P value	Incidence rate ratio	95% CI	P value
Intercept	320.737	227.465-452.255	<.001	325.265	231.247-457.509	<.001
Hashtags	2.470	2.172-2.810	<.001	2.813	2.469-3.206	<.001
Hyperlinks	0.839	0.707-0.996	.045	0.727	0.611-0.865	<.001
<b>Media type</b>						
Text only	Reference	N/A <sup>a</sup>	N/A	Reference	N/A	N/A
Photographs	0.530	0.428-0.656	<.001	0.476	0.384-0.591	<.001
Videos	0.698	0.540-0.901	.006	0.576	0.445-0.745	<.001
<b>Message type</b>						
Risk messages	1.177	0.934-1.482	.167	1.088	0.863-1.371	.48
Warnings	1.334	1.089-1.634	.005	1.544	1.261-1.891	<.001
Preparations	1.007	0.836-1.212	.944	1.008	0.838-1.213	.93
Uncertainty reduction	2.210	1.882-2.595	<.001	2.197	1.870-2.583	<.001
Efficacy	1.096	0.926-1.297	.288	1.200	1.014-1.420	.03
Reassurance	1.551	1.320-1.821	<.001	1.517	1.296-1.776	<.001
<b>Crisis stage</b>						
Precrisis	Reference	N/A	N/A	Reference	N/A	N/A
Initial event	2.931	2.309-3.721	<.001	2.471	1.945-3.141	<.001
Maintenance	2.355	1.825-3.039	<.001	1.623	1.256-2.096	.001

<sup>a</sup>N/A: not applicable.

## Discussion

### Principal Findings

This study shows that Twitter was used as a real-time communication channel to share, communicate, and disseminate information during the COVID-19 pandemic. Overall, the MOH's response to COVID-19 on Twitter was aligned with the developments of the outbreak in Saudi Arabia. During the precrisis and maintenance stages, the MOH's use of Twitter was partially consistent with the CERC model. While in the initial stage, the MOH's communication was in line with the CERC model. The results of public engagement showed that the levels of engagement were different as the pandemic evolved. The tweets in the initial stage elicited the most engagement by far.

The number of tweets pertaining to the COVID-19 pandemic was relatively low (n=79) in the precrisis stage. However, when COVID-19 emerged in Saudi Arabia on March 2, 2020, and began to spread throughout the country, the MOH increased its Twitter activity. The MOH posted most of its COVID-19-related tweets (n=802, 65.89%) during this initial stage. This likely corresponded with the public's need for information, given that individuals increasingly use social media to seek information during crises [34,42]. The MOH regularly posted tweets regarding COVID-19 (n=336) throughout the maintenance stage and until the end of the study period.

The results concerning public engagement showed that the tweets in the initial stage received the most engagement from the public. This result was expected, considering the lockdowns and curfew restrictions imposed during this stage. Recent studies have even reported an increase in the usage of the internet and social media during COVID-19 lockdowns [43].

The MOH's use of Twitter was partially consistent with the CERC model. During precrisis, the CERC model suggests that communication should focus on risk information, warnings, and preparations. This is because during the precrisis stage, the public tends to seek information regarding the nature of the risk itself. However, our results show that a large proportion of the MOH's tweets in this stage included reassurance (67.1%) or uncertainty reduction (51.9%), while those including risk messages and warnings represented only 8.8% of all tweets.

Agwa [44], in her study of the Egyptian MOH's use of Facebook during the COVID-19 pandemic, also observed a lack of risk information during early stages of the pandemic. One potential explanation for this finding is the novelty and scientific uncertainty associated with COVID-19 [45,46].

During the initial stage, the CERC model suggests that communication messages should focus on reducing public uncertainty and providing messages regarding efficacy and reassurance. Consistent with the model, the MOH reduced uncertainty by updating case reports, holding press conferences, and providing information sources that accounted for 47.6% of tweets. The MOH tweets also emphasized efficacy (34.4%) by

highlighting common responsibility and personal prevention measures to limit the spread of SARS-CoV-2 while providing reassurance (18.1%). The MOH's communication during this stage was in line with the CERC model.

As a crisis continues into the maintenance stage, the CERC model requires ongoing communication of uncertainty reduction and reassurance. Additional efficacy messages inform members of the public about the expected course of action at this stage.

In accordance with the model, the MOH's tweets during this stage provided information sources to reduce uncertainty in addition to efficacy and reassurance information. However, the results showed that a considerable proportion (22.3%) of the MOH's tweets in this stage included preparation messages (mostly recommendations). This could potentially be part of the MOH's efforts to prepare the public for the "new normal," especially as the country lifted its nationwide curfew and began its gradual reopening during this stage. While the CERC model suggests that health communicators should offer reassurance in the maintenance stage, crisis communication experts do not recommend that messages be overly reassuring, as such information may reduce an authority's credibility [47]. This is particularly the case in unexpected and unpredictable events such as the COVID-19 pandemic. In addition, it should be noted that some infectious disease outbreaks and epidemics commonly become chronic crises that develop into crisis stages for longer periods, making it difficult to make precise predictions [26].

Most importantly, our findings indicate that different types of messages received different levels of engagement as the outbreak evolved. In the precrisis stage, the public showed a high level of interest in warnings and risk messages, as indicated by a high level of engagement. Based on this finding, it may be inferred that the public wanted to understand the full scope of the risk. In the initial stage, members of the public engaged more with certain messages (such as uncertainty reduction and reassurance) than others; indicating their simultaneous need for increased understanding and reduced anxiety regarding the outbreak. During the maintenance stage, the public also showed a high level of interest in information related to uncertainty reduction, which indicates that they may have still felt uncertain during this stage.

Our findings also identified a number of factors associated with greater public engagement during the pandemic. First, our results show that the use of hyperlinks was negatively associated with public engagement. This further supports the inferences by Chung [48] that hyperlinks increase the complexity of a message by requiring an extra action by the audience, thereby reducing engagement. Another potential explanation is that hyperlinks direct people to another webpage, at which point they may forget about the original message. Second, in contrast with previous reports supporting the positive effect of media on public engagement [49-52], we found that the inclusion of multimedia content (eg, photographs and videos) was negatively associated with public engagement. This finding is similar to that of Chen et al [53], who also found that media richness was negatively associated with public engagement with government social media during the COVID-19 pandemic. This discrepancy can be attributed to the differences in the events examined, as most

studies supporting the positive effect were conducted in noncrisis situations [53].

The content of the tweets was also significantly associated with public engagement, where tweets related to uncertainty reduction, public reassurance, and warnings received higher levels of engagement. This is consistent with the results of Tang et al [54], who examined the public health agencies' tweets in Texas, where tweets that provided information about COVID-19 or described the government's actions in containing the spread of COVID-19 were found to be more likely to be retweeted. This suggests that as the CERC model indicates, people need this type of information during a public health crisis.

## Limitations

There are several limitations and future considerations of this study. First, while this study focused on Twitter, the MOH used other social media platforms (such as Facebook and Instagram) during the COVID-19 pandemic. Twitter is a microblogging platform that allows users to post short messages that contain up to 280 characters. Facebook allows for much longer posts than Twitter, while Instagram is centered on images rather than text. Since social media platforms vary greatly in their characteristics, different communication strategies may be adopted and employed. Future studies should consider focusing on the same topic on Facebook, Instagram, and other popular social media platforms.

Second, given that the pandemic is still ongoing, this study focused on the first 3 stages of the crisis (precrisis, initial, and maintenance stages) and did not examine the resolution or evaluation stage. Future studies should expand the scope of the analysis to provide a more comprehensive description of the MOH's crisis communication on Twitter.

Only tweets written in Arabic and English were included in the analysis. In addition, this study did not attempt to analyze public replies to the MOH's tweets. Further studies are needed to examine the content of the public's replies to understand their responses and opinions regarding tweets from the MOH. A final limitation is the time interval between the posting date of a tweet and the date of data collection. Since older tweets may take a longer time to accumulate engagement, future studies should consider such temporal effects. Future studies should also incorporate more appropriate measures of engagement, beyond simply the numbers of likes and retweets.

## Conclusions

The COVID-19 pandemic is an extreme crisis and has generated significant challenges for governments. Effective communication with the public is of crucial importance. This study provided some insight into the Saudi MOH's outbreak communication strategy. Our findings identified differences in MOH communication practices during different stages of the COVID-19 outbreak in Saudi Arabia, in terms of both types of message content and levels of public engagement. Uncertainty reduction, efficacy, and reassurance were the most common types of messages in MOH tweets. Our results provide several implications for crisis communication by researchers, governments, health organizations, and practitioners to engage their external public. Governments and health authorities should

consider the public's information needs to promote their engagement; this, in turn, could raise the public's awareness of a health crisis. Effective communication during disease outbreaks and other public health emergencies has the potential to change outcomes and save lives.

---

## Acknowledgments

The authors gratefully acknowledge King Abdullah International Medical Research Center and King Saud bin Abdulaziz University for Health Sciences for supporting and funding the publication of this paper.

---

## Conflicts of Interest

None declared.

---

### Multimedia Appendix 1

Codebook.

[\[DOCX File, 19 KB - publichealth\\_v7i7e27942\\_app1.docx\]](#)

---

### Multimedia Appendix 2

Categories of message types by outbreak stages, post-hoc pairwise comparisons.

[\[DOCX File, 17 KB - publichealth\\_v7i7e27942\\_app2.docx\]](#)

---

### Multimedia Appendix 3

Public engagement across outbreak stages, post-hoc pairwise comparisons.

[\[DOCX File, 14 KB - publichealth\\_v7i7e27942\\_app3.docx\]](#)

---

### Multimedia Appendix 4

Negative binomial regression results for associations of tweet content, media type, and the crisis stage with public engagement.

[\[DOCX File, 21 KB - publichealth\\_v7i7e27942\\_app4.docx\]](#)

---

## References

1. Fehr AR, Perlman S. Coronaviruses: an overview of their replication and pathogenesis. *Methods Mol Biol* 2015;1282:1-23 [[FREE Full text](#)] [doi: [10.1007/978-1-4939-2438-7\\_1](https://doi.org/10.1007/978-1-4939-2438-7_1)] [Medline: [25720466](https://pubmed.ncbi.nlm.nih.gov/25720466/)]
2. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis* 2020 May;20(5):533-534 [[FREE Full text](#)] [doi: [10.1016/S1473-3099\(20\)30120-1](https://doi.org/10.1016/S1473-3099(20)30120-1)] [Medline: [32087114](https://pubmed.ncbi.nlm.nih.gov/32087114/)]
3. Giovanetti M, Benvenuto D, Angeletti S, Ciccozzi M. The first two cases of 2019-nCoV in Italy: Where they come from? *J Med Virol* 2020 May;92(5):518-521 [[FREE Full text](#)] [doi: [10.1002/jmv.25699](https://doi.org/10.1002/jmv.25699)] [Medline: [32022275](https://pubmed.ncbi.nlm.nih.gov/32022275/)]
4. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, China Novel Coronavirus Investigating and Research Team. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* 2020 Feb 20;382(8):727-733 [[FREE Full text](#)] [doi: [10.1056/NEJMoa2001017](https://doi.org/10.1056/NEJMoa2001017)] [Medline: [31978945](https://pubmed.ncbi.nlm.nih.gov/31978945/)]
5. Wu F, Zhao S, Yu B, Chen Y, Wang W, Song Z, et al. A new coronavirus associated with human respiratory disease in China. *Nature* 2020 Mar;579(7798):265-269 [[FREE Full text](#)] [doi: [10.1038/s41586-020-2008-3](https://doi.org/10.1038/s41586-020-2008-3)] [Medline: [32015508](https://pubmed.ncbi.nlm.nih.gov/32015508/)]
6. Population Estimates. General Authority for Statistics, Kingdom of Saudi Arabia. URL: <https://www.stats.gov.sa/en/43> [accessed 2021-04-12]
7. Statistical Yearbook. Ministry of Health, Kingdom of Saudi Arabia. URL: <https://www.moh.gov.sa/en/Ministry/Statistics/book/Pages/default.aspx> [accessed 2021-04-12]
8. Algaissi AA, Alharbi NK, Hassanain M, Hashem AM. Preparedness and response to COVID-19 in Saudi Arabia: Building on MERS experience. *J Infect Public Health* 2020 Jun;13(6):834-838 [[FREE Full text](#)] [doi: [10.1016/j.jiph.2020.04.016](https://doi.org/10.1016/j.jiph.2020.04.016)] [Medline: [32451260](https://pubmed.ncbi.nlm.nih.gov/32451260/)]
9. Hassounah M, Raheel H, Alhefzi M. Digital Response During the COVID-19 Pandemic in Saudi Arabia. *J Med Internet Res* 2020 Sep 01;22(9):e19338 [[FREE Full text](#)] [doi: [10.2196/19338](https://doi.org/10.2196/19338)] [Medline: [32790642](https://pubmed.ncbi.nlm.nih.gov/32790642/)]
10. National Digital Transformation Semiannual Report. National Digital Transformation Unit. 2020. URL: <https://ndu.gov.sa/report/2020-bareport.pdf> [accessed 2021-04-06]
11. WHO, Saudi Arabia join forces to fight COVID-19 nationally, regionally and globally. World Health Organization. 2020 Mar 25. URL: <https://www.emro.who.int/media/news/who-saudi-arabia-join-forces-to-fight-covid-19-nationally-regionally-and-globally.html> [accessed 2021-04-15]
12. Jones SC, Waters L, Holland O, Bevins J, Iverson D. Developing pandemic communication strategies: Preparation without panic. *J Bus Res* 2010 Feb;63(2):126-132. [doi: [10.1016/j.jbusres.2009.02.009](https://doi.org/10.1016/j.jbusres.2009.02.009)]

13. Zhao Y, Cheng S, Yu X, Xu H. Chinese Public's Attention to the COVID-19 Epidemic on Social Media: Observational Descriptive Study. *J Med Internet Res* 2020 May 04;22(5):e18825 [FREE Full text] [doi: [10.2196/18825](https://doi.org/10.2196/18825)] [Medline: [32314976](https://pubmed.ncbi.nlm.nih.gov/32314976/)]
14. Strekalova YA. Health Risk Information Engagement and Amplification on Social Media. *Health Educ Behav* 2017 Apr;44(2):332-339. [doi: [10.1177/1090198116660310](https://doi.org/10.1177/1090198116660310)] [Medline: [27413028](https://pubmed.ncbi.nlm.nih.gov/27413028/)]
15. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. *J Med Internet Res* 2013 Apr 23;15(4):e85 [FREE Full text] [doi: [10.2196/jmir.1933](https://doi.org/10.2196/jmir.1933)] [Medline: [23615206](https://pubmed.ncbi.nlm.nih.gov/23615206/)]
16. Safko L, Brake DK. *The Social Media Bible: Tactics, Tools, and Strategies for Business Success*. Hoboken, NJ: Wiley; 2009.
17. Kemp S. Digital 2020: Global Digital Overview. DataReportal. 2020 Jan 30. URL: <https://datareportal.com/reports/digital-2020-global-digital-overview> [accessed 2020-12-28]
18. Korda H, Itani Z. Harnessing social media for health promotion and behavior change. *Health Promot Pract* 2013 Jan;14(1):15-23. [doi: [10.1177/1524839911405850](https://doi.org/10.1177/1524839911405850)] [Medline: [21558472](https://pubmed.ncbi.nlm.nih.gov/21558472/)]
19. Tang L, Bie B, Park S, Zhi D. Social media and outbreaks of emerging infectious diseases: A systematic review of literature. *Am J Infect Control* 2018 Sep;46(9):962-972 [FREE Full text] [doi: [10.1016/j.ajic.2018.02.010](https://doi.org/10.1016/j.ajic.2018.02.010)] [Medline: [29628293](https://pubmed.ncbi.nlm.nih.gov/29628293/)]
20. Developing tools for strategic communication to the media on emerging infectious diseases (EID) : report of an informal consultation, WHO-SEARO, New Delhi, 26 May 2011. World Health Organization. URL: <https://apps.who.int/iris/handle/10665/204966> [accessed 2021-05-12]
21. Chen S, Xu Q, Buchenberger J, Bagavathi A, Fair G, Shaikh S, et al. Dynamics of Health Agency Response and Public Engagement in Public Health Emergency: A Case Study of CDC Tweeting Patterns During the 2016 Zika Epidemic. *JMIR Public Health Surveill* 2018 Nov 22;4(4):e10827 [FREE Full text] [doi: [10.2196/10827](https://doi.org/10.2196/10827)] [Medline: [30467106](https://pubmed.ncbi.nlm.nih.gov/30467106/)]
22. Lwin MO, Lu J, Sheldenkar A, Schulz PJ. Strategic Uses of Facebook in Zika Outbreak Communication: Implications for the Crisis and Emergency Risk Communication Model. *Int J Environ Res Public Health* 2018 Sep 10;15(9):1974 [FREE Full text] [doi: [10.3390/ijerph15091974](https://doi.org/10.3390/ijerph15091974)] [Medline: [30201929](https://pubmed.ncbi.nlm.nih.gov/30201929/)]
23. Guidry JP, Jin Y, Orr CA, Messner M, Meganck S. Ebola on Instagram and Twitter: How health organizations address the health crisis in their social media engagement. *Public Relat Rev* 2017 Sep;43(3):477-486. [doi: [10.1016/j.pubrev.2017.04.009](https://doi.org/10.1016/j.pubrev.2017.04.009)]
24. Sesagiri Raamkumar A, Tan SG, Wee HL. Measuring the Outreach Efforts of Public Health Authorities and the Public Response on Facebook During the COVID-19 Pandemic in Early 2020: Cross-Country Comparison. *J Med Internet Res* 2020 May 19;22(5):e19334 [FREE Full text] [doi: [10.2196/19334](https://doi.org/10.2196/19334)] [Medline: [32401219](https://pubmed.ncbi.nlm.nih.gov/32401219/)]
25. Reynolds B, Quinn Crouse S. Effective communication during an influenza pandemic: the value of using a crisis and emergency risk communication framework. *Health Promot Pract* 2008 Oct;9(4 Suppl):13S-17S. [doi: [10.1177/1524839908325267](https://doi.org/10.1177/1524839908325267)] [Medline: [18936255](https://pubmed.ncbi.nlm.nih.gov/18936255/)]
26. Reynolds B, Seeger MW. Crisis and emergency risk communication as an integrative model. *J Health Commun* 2005;10(1):43-55. [doi: [10.1080/10810730590904571](https://doi.org/10.1080/10810730590904571)] [Medline: [15764443](https://pubmed.ncbi.nlm.nih.gov/15764443/)]
27. Veil S, Reynolds B, Sellnow TL, Seeger MW. CERC as a theoretical framework for research and practice. *Health Promot Pract* 2008 Oct;9(4 Suppl):26S-34S. [doi: [10.1177/1524839908322113](https://doi.org/10.1177/1524839908322113)] [Medline: [18936257](https://pubmed.ncbi.nlm.nih.gov/18936257/)]
28. Centers for Disease Control and Prevention. URL: <https://emergency.cdc.gov/cerc/> [accessed 2021-06-12]
29. Leading countries based on number of Twitter users as of April 2020 (in Millions). Statista. URL: <https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/> [accessed 2020-05-28]
30. Aldarhami A, Bazaid AS, Althomali OW, Binsaleh NK. Public Perceptions and Commitment to Social Distancing "Staying-at-Home" During COVID-19 Pandemic: A National Survey in Saudi Arabia. *Int J Gen Med* 2020;13:677-686 [FREE Full text] [doi: [10.2147/IJGM.S269716](https://doi.org/10.2147/IJGM.S269716)] [Medline: [33061535](https://pubmed.ncbi.nlm.nih.gov/33061535/)]
31. Get Tweet timelines. Twitter. URL: <https://developer.twitter.com/en/docs/twitter-api/v1/tweets/timelines/api-reference/get-statuses-user-timeline> [accessed 2021-03-10]
32. Tweepy: An easy-to-use Python library for accessing the Twitter API. Tweepy. URL: <https://www.tweepy.org/> [accessed 2021-03-10]
33. Ritchie H, Ortiz-Ospina E, Beltekian D, Mathieu E, Hasell J, Macdonald B, et al. Statistics and Research: Coronavirus (COVID-19) Cases. Our World in Data. URL: <https://ourworldindata.org/covid-cases> [accessed 2021-07-05]
34. Lachlan KA, Spence PR, Lin X, Najarian K, Del Greco M. Social media and crisis management: CERC, search strategies, and Twitter content. *Comput Hum Behav* 2016 Jan;54:647-652. [doi: [10.1016/j.chb.2015.05.027](https://doi.org/10.1016/j.chb.2015.05.027)]
35. Panagiotopoulos P, Barnett J, Bigdeli AZ, Sams S. Social media in emergency management: Twitter as a tool for communicating risks to the public. *Technol Forecast Soc Change* 2016 Oct;111:86-96. [doi: [10.1016/j.techfore.2016.06.010](https://doi.org/10.1016/j.techfore.2016.06.010)]
36. Berg B. *Qualitative Research Methods for the Social Sciences*. Boston, MA: Allyn and Bacon; 2001.
37. Neuendorf K. *The content analysis guidebook*. Thousand Oaks, CA: Sage Publications; 2002.
38. Freelon D. ReCal: Intercoder Reliability Calculation as a Web Service. *Int J Internet Sci* 2010;5(1):20-33 [FREE Full text] [doi: [10.4135/9781412963947.n228](https://doi.org/10.4135/9781412963947.n228)]
39. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977 Mar;33(1):159-174. [Medline: [843571](https://pubmed.ncbi.nlm.nih.gov/843571/)]

40. IBM SPSS Statistics for Macintosh, version 27.0. Armonk, NY: IBM Corp. URL: <https://www.ibm.com/analytics/spss-statistics-software> [accessed 2021-07-05]
41. An Approval Issued to fully Lift the Curfew from 06 am on Sunday, Ban on Umrah, Visit, Int'l Flights to Continue. Saudi Press Agency. URL: <https://www.spa.gov.sa/viewfullstory.php?lang=en&newsid=2100088> [accessed 2021-05-12]
42. Lachlan KA, Spence PR, Lin X. Expressions of risk awareness and concern through Twitter: On the utility of using the medium as an indication of audience needs. *Comput Hum Behav* 2014 Jun;35:554-559. [doi: [10.1016/j.chb.2014.02.029](https://doi.org/10.1016/j.chb.2014.02.029)]
43. Donthu N, Gustafsson A. Effects of COVID-19 on business and research. *J Bus Res* 2020 Sep;117:284-289 [FREE Full text] [doi: [10.1016/j.jbusres.2020.06.008](https://doi.org/10.1016/j.jbusres.2020.06.008)] [Medline: [32536736](https://pubmed.ncbi.nlm.nih.gov/32536736/)]
44. Ajwa NA. Strategies for communicating health risks through social media against the Corona pandemic (analytical study on the official pages of the Egyptian Ministry of Health). *J Mass Commun Res* 2020 Aug 01;54(4):2433-2494. [doi: [10.21608/jsb.2020.108731](https://doi.org/10.21608/jsb.2020.108731)]
45. Tang JW. COVID-19: interpreting scientific evidence - uncertainty, confusion and delays. *BMC Infect Dis* 2020 Sep 08;20(1):653 [FREE Full text] [doi: [10.1186/s12879-020-05387-8](https://doi.org/10.1186/s12879-020-05387-8)] [Medline: [32895050](https://pubmed.ncbi.nlm.nih.gov/32895050/)]
46. Han PKJ, Scharnetzki E, Scherer AM, Thorpe A, Lary C, Waterston LB, et al. Communicating Scientific Uncertainty About the COVID-19 Pandemic: Online Experimental Study of an Uncertainty-Normalizing Strategy. *J Med Internet Res* 2021 Apr 22;23(4):e27832 [FREE Full text] [doi: [10.2196/27832](https://doi.org/10.2196/27832)] [Medline: [33769947](https://pubmed.ncbi.nlm.nih.gov/33769947/)]
47. Seeger MW. Best Practices in Crisis Communication: An Expert Panel Process. *J Appl Commun Res* 2006 Aug;34(3):232-244. [doi: [10.1080/00909880600769944](https://doi.org/10.1080/00909880600769944)]
48. Chung JE. Retweeting in health promotion: Analysis of tweets about Breast Cancer Awareness Month. *Comput Hum Behav* 2017 Sep;74:112-119. [doi: [10.1016/j.chb.2017.04.025](https://doi.org/10.1016/j.chb.2017.04.025)]
49. Guidry JPD, Meganck SL, Lovari A, Messner M, Medina-Messner V, Sherman S, et al. Tweeting about #Diseases and #PublicHealth: Communicating Global Health Issues across Nations. *Health Commun* 2020 Aug;35(9):1137-1145. [doi: [10.1080/10410236.2019.1620089](https://doi.org/10.1080/10410236.2019.1620089)] [Medline: [31131628](https://pubmed.ncbi.nlm.nih.gov/31131628/)]
50. Bhattacharya S, Srinivasan P, Polgreen P. Social media engagement analysis of U.S. Federal health agencies on Facebook. *BMC Med Inform Decis Mak* 2017 Apr 21;17(1):49 [FREE Full text] [doi: [10.1186/s12911-017-0447-z](https://doi.org/10.1186/s12911-017-0447-z)] [Medline: [28431582](https://pubmed.ncbi.nlm.nih.gov/28431582/)]
51. Siyam N, Alqaryouti O, Abdallah S. Mining government tweets to identify and predict citizens engagement. *Technol Soc* 2020 Feb;60:101211. [doi: [10.1016/j.techsoc.2019.101211](https://doi.org/10.1016/j.techsoc.2019.101211)]
52. Bonsón E, Perea D, Bednárová M. Twitter as a tool for citizen engagement: An empirical study of the Andalusian municipalities. *Gov Inf Q* 2019 Jul;36(3):480-489. [doi: [10.1016/j.giq.2019.03.001](https://doi.org/10.1016/j.giq.2019.03.001)]
53. Chen Q, Min C, Zhang W, Wang G, Ma X, Evans R. Unpacking the black box: How to promote citizen engagement through government social media during the COVID-19 crisis. *Comput Human Behav* 2020 Sep;110:106380 [FREE Full text] [doi: [10.1016/j.chb.2020.106380](https://doi.org/10.1016/j.chb.2020.106380)] [Medline: [32292239](https://pubmed.ncbi.nlm.nih.gov/32292239/)]
54. Tang L, Liu W, Thomas B, Tran HTN, Zou W, Zhang X, et al. Texas Public Agencies' Tweets and Public Engagement During the COVID-19 Pandemic: Natural Language Processing Approach. *JMIR Public Health Surveill* 2021 Apr 26;7(4):e26720 [FREE Full text] [doi: [10.2196/26720](https://doi.org/10.2196/26720)] [Medline: [33847587](https://pubmed.ncbi.nlm.nih.gov/33847587/)]

## Abbreviations

**CDC:** Centers for Disease Control and Prevention  
**CERC:** Crisis and Emergency Risk Communication  
**IRR:** incidence rate ratio  
**MOH:** Ministry of Health  
**WHO:** World Health Organization

*Edited by T Sanchez; submitted 14.02.21; peer-reviewed by K Lachlan, M Pobiruchin, A Sudaryanto; comments to author 07.03.21; revised version received 29.04.21; accepted 11.06.21; published 12.07.21.*

### *Please cite as:*

Alhassan FM, AlDossary SA

*The Saudi Ministry of Health's Twitter Communication Strategies and Public Engagement During the COVID-19 Pandemic: Content Analysis Study*

*JMIR Public Health Surveill* 2021;7(7):e27942

URL: <https://publichealth.jmir.org/2021/7/e27942>

doi: [10.2196/27942](https://doi.org/10.2196/27942)

PMID: [34117860](https://pubmed.ncbi.nlm.nih.gov/34117860/)

©Fatimah Mohammed Alhassan, Sharifah Abdullah AIDossary. Originally published in JMIR Public Health and Surveillance (<https://publichealth.jmir.org>), 12.07.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

# Pre-exposure Prophylaxis (PrEP) Information on Instagram: Content Analysis

Eric Walsh-Buhi<sup>1</sup>, MPH, PhD; Rebecca Fagen Houghton<sup>1\*</sup>, MPA; Claire Lange<sup>1\*</sup>; Ryli Hockensmith<sup>1\*</sup>; John Ferrand<sup>1\*</sup>, MPH; Lourdes Martinez<sup>2\*</sup>, MA, PhD

<sup>1</sup>Department of Applied Health Science, Indiana University School of Public Health-Bloomington, Bloomington, IN, United States

<sup>2</sup>School of Communication, San Diego State University, San Diego, CA, United States

\*these authors contributed equally

**Corresponding Author:**

Eric Walsh-Buhi, MPH, PhD

Department of Applied Health Science

Indiana University School of Public Health-Bloomington

1025 E. 7th Street, Room 116

Bloomington, IN, 47405

United States

Phone: 1 812 855 4867

Email: [erwals@iu.edu](mailto:erwals@iu.edu)

## Abstract

**Background:** There is still an HIV epidemic in the United States, which is a substantial issue for populations bearing a disproportionate burden of HIV infections. Daily oral pre-exposure prophylaxis (PrEP) has proven to be safe and effective in reducing HIV acquisition risk. However, studies document that PrEP awareness/usage is low. There is also limited understanding of social media platforms, such as Instagram, as PrEP information sources.

**Objective:** Given the paucity of research on PrEP-related Instagram posts and popularity of this social media platform, the purpose of this research is to describe the source characteristics, image types, and textual contents of PrEP-related posts on Instagram.

**Methods:** Using Crowdtangle Search, a public insights tool owned/operated by Facebook, we retrieved publicly accessible and English-language-only Instagram posts for the 12-month period preceding April 22, 2020, using the following terms: Truvada or “pre-exposure prophylaxis” or #truvada or #truvadaprep or #truvadawhore or #truvadaforprep. We employed a qualitative coding methodology to manually extract information from posts. Using a pretested codebook, we performed content analysis on 250 posts, examining message and source characteristics (ie, organization type [eg, government, news] and individual type [eg, physician]), including information about PrEP (eg, how it works, cost), and indicated users. Frequencies and percentages were calculated for all categorical variables. A Chi-square test was conducted to determine differences between source types on a variety of message characteristics.

**Results:** Three-quarters of the posts (193/250, 77.2%) were posted by organizations. Of the 250 posts reviewed, approximately two-thirds (174/250, 69.6%) included a photograph, more than half (142/250, 56.8%) included an infographic, and approximately one-tenth (30/250, 12%) included a video. More than half defined PrEP (137/250, 54.8%), but fewer posts promoted PrEP use, explained how PrEP works, and included information on the effectiveness of PrEP or who can use it. The most commonly hashtagged populations among posts were men who have sex with men (MSM), but not necessarily bisexual men. Few posts contained race-/ethnicity-related hashtags (11/250, 4.4%). Fewer posts contained transgender-associated tags (eg, #transgirl; 5/250, 2%). No posts contained tags related to heterosexuals or injection drug users. We found statistical differences between source types (ie, individual versus organization). Specifically, posts from organizations more frequently contained information about who can use PrEP, whereas posts from individuals more frequently contained information describing adverse effects.

**Conclusions:** This study is among the first to review Instagram for PrEP-related content, and it answers the National AIDS Strategy’s call for a clearer articulation of the science surrounding HIV risk/prevention through better understanding of the current public information environment. This study offers a snapshot of how PrEP is being discussed (and by whom) on one of the most popular social media platforms and provides a foundation for developing and implementing PrEP promotion interventions on Instagram.

**KEYWORDS**

digital health; social media; HIV; pre-exposure prophylaxis; Instagram; content analysis; communication

## Introduction

### Background

There is still an HIV epidemic in the United States, which is a substantial issue for priority populations bearing a disproportionate burden of HIV infections. In 2018, rates of HIV diagnoses remained higher among gay and bisexual men (GBM) and other men who have sex with men (MSM), with a particularly high burden among Black GBM; Black and Latinx women and men; people who inject drugs; people aged 20-34 years; people in the southern parts of the United States; and transgender women, with a high burden among Black transwomen [1,2].

### PrEP for HIV Prevention

Daily oral pre-exposure prophylaxis (PrEP), with a fixed-dose combination of tenofovir disoproxil fumarate (TDF) 300 mg and emtricitabine (FTC) 200 mg, has proven safe and effective in reducing the risk of sexual HIV acquisition in adults [3]. Several PrEP clinical trials have demonstrated its safety and a substantial reduction in the rate of HIV acquisition in MSM, men and women among heterosexual HIV-discordant couples, heterosexual men and women recruited as individuals, and persons who inject drugs [4-7]. Accordingly, PrEP is recommended for three main adult groups in the United States—MSM, heterosexually active adults, and persons who inject drugs—estimated in 2015 at 1.1 million [8].

### PrEP Awareness and Usage

Large national studies of GBM, transmasculine individuals, and Black people in the United States document that PrEP awareness and usage are low [9-11]. For example, in a probability-based cohort study of gay and bisexual men, Dodge et al reported that among HIV-negative/unknown/untested GBM, <7% reported using PrEP in the past 6 months [12]. Over half of the GBM reported not using a condom during anal sex with their most recent partner, with 94% of these men not on PrEP. Among high-risk Black individuals participating in another study, one-fifth knew about PrEP and the most reported reason for the lack of willingness to use PrEP was low self-perceived risk (65%). In urban areas, although PrEP awareness and use among MSM has increased, PrEP use remains lower among Black and Latinx MSM [13]. Lastly, PrEP awareness and use is critically low among persons who inject drugs [14-17].

In 2015, the National HIV/AIDS Strategy for the United States updated to 2020 was released by the White House and included four key areas [18]. One critical area is “full access to comprehensive PrEP services for those whom it is appropriate and desired, with support for medication adherence for those using PrEP.” The strategy outlined the need for providing “clear, specific, consistent, and scientifically up-to-date messages about HIV risks and prevention strategies” and utilization of “evidence-based social marketing and education campaigns,

and... (leveraging of) digital tools and new technologies,” such as social media platforms [18].

### PrEP and Social Media

Although PrEP is a novel and promising approach to HIV prevention, there is limited understanding of social media platforms as sources of information on PrEP. Social media platforms are interactive websites and applications that enable users to create, share, comment on, and modify content [19]. As of 2019, 9 in 10 US adults reported that they use the Internet, 8 in 10 reported that they own a smartphone, and 7 in 10 reported that they use social media [20]. Black and Latinx individuals; lesbians, gay men, and bisexuals; and young adults who are at the greatest risk for HIV are heavy users of social media, and they represent important priority populations for HIV risk/prevention messages [21,22].

Twitter and, to a lesser extent, YouTube appear to be the preferred social media platforms for most research on PrEP. For example, Keckojevic et al examined firsthand narratives of individuals detailing their PrEP experiences via YouTube [23,24]. They reported that the narratives covered a wide range of topics (eg, reasons to start PrEP, interacting with providers, PrEP adverse effects, insurance coverage, stigma); although the study focused on gay men, the authors emphasized a general lack of PrEP awareness in the lesbian, gay, bisexual, and transgender community. Using natural language processing techniques, Breen et al found increasing PrEP discussion on Twitter [25]. They further noted that people who mention PrEP on Twitter also talk about other issues, including general health, sexually transmitted diseases (STDs), stigma, and politics. McLaughlin et al performed a content analysis of PrEP for HIV prevention on Twitter and found that the most common categories of tweets pertained to recipients of PrEP, with the second and third most common categories capturing efficacy and moral judgment, respectively [26]. The authors also found that individuals (as opposed to organizational accounts) comprised the majority of tweet creators. In their content analysis of PrEP messaging on Twitter, Schwarz and Grimm found that 54% of the tweets included PrEP awareness/information, 15% discussed the barriers to use PrEP, 14% contained consequences/limitations, 9% included antistigma, and 6% mentioned the stigma of using PrEP [27]. They further reported that tweets mentioning race, gender, and sexual orientation were rare. Individuals were more likely to tweet about antistigma than organizations and media outlets.

Although research has examined PrEP messaging on Twitter, to date, we have found limited PrEP-related research focusing on popular image-sharing social media sites such as Instagram. With more than 1 billion monthly active users, Instagram is an increasingly popular photo- and video-based social media platform with a wide variety of users [28]. Each day, Instagram users upload more than 500 million photos, videos, and stories (a combination of videos, text, and photos) [29]. Instagram is



the third most popular social media platform in the United States, behind YouTube and Facebook, with 38% and 67% of US adults and 18-29-year-olds, respectively, using Instagram [21]. It is also an excellent platform for reaching out to a diverse population, as the platform has a balance of genders (52% female), and 40% of Black and 51% of Latinx Internet-using adults use Instagram [21,29]. This positions Instagram as a unique platform for examining health communication and, specifically, PrEP-related communication intended for and among racial and ethnic minority priority populations.

Instagram previously has been examined as a health promotion modality with some researchers highlighting the general utility of Instagram as a source of education and motivation [30] as well as users' experiences on receiving social support via Instagram [31]. Additionally, Instagram relies heavily on images in its posts and has been referred to as a forum for parasocial interaction or the one-sided feeling of connectedness between a fan and a celebrity [32]. Parasocial interaction can be examined in the context of Instagram posts owing to the large fan bases (ie, followers) attached to certain Instagram celebrities as well as those celebrities' depictions of their personal lives through images. Given that other social media sites are not exclusively image-driven, this positions Instagram as an optimal data source for this type of study.

In one study, Nobles et al examined Instagram posts from January 2017 to July 2018; although they did not focus specifically on PrEP, they did find that PrEP was mentioned in a very small percentage (6.2%) of the posts tagged "HIV" [33]. These authors further noted that the visual content of specific clinical interventions, such as PrEP promotion, are not well represented on Instagram relative to public health priorities. To date, only one study [34] has been undertaken to specifically examine Instagram posts in the context of PrEP promotion efforts, and its results describe surface-level engagements in Instagram posts as generally positive among Black MSM. Given the paucity of research on PrEP-related Instagram posts and the popularity of this social media platform, the purpose of this research is to build on previous research, such as the study by Nobles et al (which is broad in focus and characterizes the social media landscape regarding HIV risk and prevention messaging), with particular focus on PrEP-related communications on Instagram. Although past research suggests that PrEP awareness/information comprises the largest content focus on Twitter [27] and that most PrEP-related content on YouTube focuses on GBM or MSM [23,24,35], it remains unclear whether these result patterns exist on Instagram as well. Based on this, we pursued three research questions (RQs):

*RQ1: What is the textual content of PrEP-related Instagram posts?*

*RQ2: Which priority populations are the focus of PrEP-related Instagram posts?*

*RQ3: How does the textual content of PrEP-related posts vary by source (ie, individual accounts versus organizational accounts)?*

Attempting to answer these research questions will allow us to describe the landscape surrounding PrEP-related posts on a

social media platform that has been understudied in the context of sexual health research and how posts are designed to target specific users. These answers will be used to develop novel communication strategies for promoting PrEP uptake and adherence in populations at risk for HIV transmission.

## Methods

### Data Source

Using Crowdtangle Search [36], a Facebook-owned tool that tracks interactions on public content from Facebook pages and groups, verified profiles, Instagram accounts, and subreddits, we retrieved publicly accessible and English-language-only Instagram posts for the 12-month period preceding April 22, 2020, using the following search terms: Truvada or "pre-exposure prophylaxis" or #truvada or #truvadaprep or #truvadawhore or #truvadaforprep [36]. We selected these hashtags based on previous Twitter data research and a review of popular PrEP-related hashtags provided in the Instagram search textbox as well as those that were suggested by the search textbox when our initial terms were entered. The initial search returned 275 posts. A total of 250 posts were analyzed, as 20 posts were excluded for not being in English and 5 for nonfunctioning links at the time of coding. Data on the final 250 public posts included the post URL, account and username associated with the post, date and time of posting, type of post (eg, photo, video), number of followers at the time of posting, and text associated with the post. Engagement metrics were also available and analyzed, including the number of likes as well as the number of comments and views.

### Data Coding

Despite the growing use of machine learning methods, there is evidence that these methods do not always align well with social science objectives [37]. For this study, we opted to use a qualitative coding methodology through the development of a coding document adapted from sentiment and other content analysis research (eg, PrEP on YouTube, human papillomavirus [HPV] vaccine information on YouTube) to manually extract information from this sample of Instagram posts [23,24,38-40]. Using a pretested codebook, we performed content analysis on all the 250 posts. The codebook included variables related to the sources, image types, and caption characteristics, as well as hashtags used for each post. The coded source characteristics included whether the Instagram account/profile represented an individual or organization. Individual profiles were coded according to self-identification as a parent, child, or spiritual/religious person; political affiliations; and their reported professions (eg, journalist, physician). For organizational accounts, the organization type was coded using information displayed in the user profile or links embedded in the user profile or bio to the organization's website (eg, business, media outlet, nonprofit, government). As we extracted data from the biographies/profiles of the posters, these data are not considered anonymous. For example, some "personal" information was recorded, such as the name (eg, Prevention305) and affiliation of the poster.

Guided by previous research, we coded the Instagram posts for specific information about PrEP and the indicated users [23,24]

as follows: whether the posts defined PrEP, explained how PrEP works, who can use it, or how to obtain it; discussed the effectiveness of using PrEP to prevent HIV; mentioned the adverse effects of using PrEP; promoted PrEP use; and indicated

any stigmatization or antistigma sentiments regarding PrEP users (eg, “My pharmacist just called me a slut for taking Truvada.”) [23,27]. The definitions for the variables mentioned above are listed in Table 1.

**Table 1.** Definitions, examples, and frequencies of textual content characteristics and variables regarding pre-exposure prophylaxis-related messages (N=250).

Message characteristic	Definition	Example Instagram post	n	%
Defines PrEP <sup>a</sup>	Taking a prescription drug as a means of preventing HIV <sup>b</sup> infection in an HIV-negative person	“Truvada and Descovy are FDA <sup>c</sup> -approved medications to prevent HIV when taken once daily.”	137	54.8
Promotes PrEP use	Encouraging PrEP use among those at risk, can include action items, such as “ask your doctor about PrEP.”	“Talk to your doctor or local clinician to determine if PrEP might be right for you.” “#PrEP2BeSafe like the #MenOfPrEP with your NEW HAPPY PILL”	93	37.2
How PrEP works	By taking Truvada/PrEP (a combination of two drugs, tenofovir and emtricitabine) daily, the presence of the medicine in the bloodstream can stop HIV from taking hold and spreading in the body.	“PrEP involves taking the combination drug emtricitabine-tenofovir (Truvada) or emtricitabine plus tenofovir alafenamide (Descovy) every day.” “Having PrEP medicine in your bloodstream can stop HIV from taking hold and spreading in your body.”	80	32.0
Effective in preventing HIV	Discusses the effectiveness of PrEP, such as successful prevention of HIV and statistics on prevention	“When taken daily, PrEP pills can reduce the risk of contracting HIV through sex by about 99 percent.”	70	28.0
Who can use PrEP	Claiming or mentioning who should receive PrEP (ie, MSM <sup>d</sup> , “high-risk” individuals)	“Pre-exposure prophylaxis (or PrEP) is a daily medication that allows people at very high risk for HIV to lower their chances of getting infected.” “PrEP has been shown to reduce risk of HIV infection through sex for gay and bisexual men, transgender women, and heterosexual men and women, as well as among people who inject drugs.”	68	27.2
How to obtain PrEP	Mentioning where or how to obtain PrEP	“The Ready, Set, PrEP program makes PrEP medication available at no cost for qualifying recipients. To receive PrEP medication through this program, you must: -Lack prescription drug coverage -Be tested for HIV with a negative result -Have a prescription for PrEP Talk to your health care provider or find a provider at HIV.gov Locator to find out if PrEP is right for you.”	55	22.0
Costs of PrEP	Mentioning the cost of PrEP or if insurance covers it	“The Ready, Set, PrEP program makes PrEP medication available at no cost for qualifying recipients.”	49	19.6
Antistigma	Owning the words, re-appropriate use, critique the use of stigma (eg, Truvada whore)	“The provision of the pill not only helps to reduce infections but allows for vulnerable populations and those often under a stigma, the opportunity to have access and be even more safe.”	20	8.0
Adverse effects	Mentioning the adverse effects of consuming PrEP/Truvada (ie, kidney damage)	“Have you or a loved one taken the antiretroviral, Truvada, and experienced osteoporosis, kidney failure, and broken or brittle bones?”	17	6.8
Stigmatization of PrEP users	Stigmatization of PrEP users, and negative attitudes or beliefs directed toward PrEP users	“Azealia Banks apologizes for ‘extremely insensitive’ comments on PrEP meds – ‘gay men should just be responsible, so they don’t have to take a f*cking pill’”	5	2.0

<sup>a</sup>PrEP: pre-exposure prophylaxis.

<sup>b</sup>HIV: human immune virus.

<sup>c</sup>FDA: Food and Drug Administration.

<sup>d</sup>MSM: men who have sex with men.

Based on previous research, we also coded for whether each post included any race-associated hashtags (eg, #blacklove), male-associated tags (eg, man, boy, male), female-associated

tags (eg, woman, girl, female), transgender-associated tags (eg, #trans, transgirl), and any other tags related to the PrEP guidelines of the Centers for Disease Control and Prevention

(CDC) (eg, #lgbt, #lgbtq, #gay, #queer, #heterosexual, #drugs) [3,41]. Previous research highlights the utility of hashtags in reaching a wider audience, so we determined that this would be an effective way to identify the target audiences of these posts [42,43]. The raters relied on hashtags and captions to identify the target audience and did not attempt to code the physical characteristics of those featured in photographs, as assuming one's ethnic background or gender identity did not

seem appropriate. Finally, we coded for the source that the post credited (eg, CDC, medical doctor), and whether the posts gave firsthand accounts of experiences with PrEP, whether the person writing the post was a child or an adult, and whether the firsthand experience with PrEP originated from an individual who belongs to the MSM group. We identified over 500 unique hashtags. Table 2 displays the top 10 most frequently used hashtags in this sample.

**Table 2.** Ten most common hashtags found in pre-exposure prophylaxis-related Instagram posts (N=541).

Hashtag	Count, n	%
#prep	41	7.6
#hiv	39	7.2
#hivprevention	32	5.9
#truvada	22	4.1
#gayman	16	3.0
#gaytheringhotel	13	2.4
#descovyprep	12	2.2
#truvadaprep	12	2.2
#sexualhealth	11	2.0
#aids	9	1.7

Two raters independently coded 57 characteristics of the posts in multiple stages. Specifically, the raters coded a set of Instagram posts after which the raters and a moderator discussed and resolved the discrepancies between ratings. This process was repeated for a total of four coding stages wherein each post was double coded on each variable. This was done to identify issues with the coding protocol and other errors. For instance, 1 post from the first set of 20 posts ultimately resulted in a broken link on the date which the raters met to reconcile their codes and was ultimately excluded from the final data set. After the first discussion on the coding differences and questions about the interpretation of the codebook definitions, the codebook was revisited with feedback from the raters to make the definitions more specific. Next, the two raters independently coded another 42 Instagram posts and repeated the process. Then, the raters independently coded another 40 Instagram posts and met to discuss coding differences. This process was then repeated with the remaining 149 Instagram posts. The codes from these four iterative rating stages were merged to form a data set of 250 posts. The lowest observed Cohen kappa ( $\kappa$ )

statistic found in this merged data set was -0.02, which occurred when the raters attempted to code for Stigma of PrEP Users present in the posts. Only 2 variables reported  $\kappa$  values that were less than 0.21, which indicates fair or higher levels of agreement across the various observations [44]. The median  $\kappa$  value was 0.76. After identifying the discrepant codes, the two raters and the moderator met to reconcile the codes for each variable in this data set and reached a perfect agreement over the final data set.

### Data Analysis

We conducted all analyses using IBM SPSS Statistics (Release 26.0.0.0) and R (Foundation for Statistical Computing, Vienna, Austria; <https://www.R-project.org/>). The measures of central tendency (eg, mean, median) and dispersion (eg, standard deviation [SD]) were calculated for the continuous variables. Frequencies and percentages were calculated for all the categorical variables. Lastly, a Chi-square test was conducted to determine the statistical differences between source types (ie, individual versus organization) regarding a variety of other characteristics, including those displayed in Tables 1 and 3.

**Table 3.** Source attribution in pre-exposure prophylaxis-related Instagram posts (N=250).

Source attribution: Posts citing information from the sources listed below	n	%
The Centers for Disease Control and Prevention or other another federal entity or foreign equivalent	25	10.0
PrEP <sup>a</sup> manufacturer (eg, Gilead, Merck)	22	8.8
Government officials (eg, senators, governors, representatives) or from political organizations	17	6.8
A celebrity	13	5.2
Another source such as WebMD, Mayo Clinic	11	4.4
A state or local health department	9	3.6
A medical doctor (physician)	8	3.2
A member of the research community (eg, researcher, scientist)	7	2.8
The World Health Organization	2	0.8

<sup>a</sup>PrEP: pre-exposure prophylaxis.

## Results

### Descriptive Statistics

For the 250 reviewed PrEP-related Instagram posts, the mean number of likes was 844.5 (SD 3,596.2; median 43.5; range 1-45,674); the mean number of comments was 16.96 (SD 77.1; median 1; range 0-1,086). For the 30 video posts, the mean number of views was 1033.4 (SD 6,523.3; median 0; range 0-64,866). Data for the number of followers were available for 193 posts; the mean number of followers was 114,487.6 at the time of posting (SD 598,905.8; median 11,705; range 459-7,780,734). As we focused on analyzing the textual content of the Instagram posts, we found that the text character counts varied. The mean number of characters in the posts was 645.4 (SD 467.1; median 515; range 1-2,196).

### Source Characteristics

More than 75% of all the posts (193/250, 77.2%) were posted by organizations, whereas 22% (55/250) were posted by individuals. The remaining posts came from Instagram accounts that were unclear in terms of whether they were representing an individual or organization. Among the 193 Instagram posts made by organizations, the largest identified types were health information providers (115/193, 59.6%), nonprofit/advocacy groups (95/193, 49.2%), and health care organizations (89/193, 46.1%). Among the 55 posts made by individuals, the largest identified types were those having a political affiliation (6/55, 10.9%), journalists (5/55, 9.1%), nurses or allied health workers (5/55, 9.1%), and physicians (4/55, 7.3%), as shown in [Tables 4 and 5](#).

**Table 4.** Source characteristics of Instagram posts among individuals regarding pre-exposure prophylaxis (N=55).

Information available in the posts, biographies, or profiles of individuals	n	%
Politics (eg, political affiliation)	6	10.9
Being a journalist or member of the press	5	9.1
Being a nurse or allied health worker	5	9.1
Being a physician (eg, medical doctor, MD <sup>a</sup> , DO <sup>b</sup> , resident)	4	7.3
Being a health educator	3	5.5
Religion or spirituality (eg, scripture, prayer, god)	2	3.6
Being a child (eg, son or daughter)	1	1.8
Being a parent (eg, mother or father)	0	0
Being an epidemiologist	0	0
A personal account from individuals with a firsthand experience with PrEP <sup>c</sup> , Truvada, or HIV <sup>d</sup> /AIDS <sup>e</sup>	10	18.2
A personal account originating from an individual who is a man who has sex with men (eg, GBM <sup>f</sup> )	15	27.3

<sup>a</sup>MD: Doctor of Medicine.

<sup>b</sup>DO: Doctor of Osteopathic Medicine.

<sup>c</sup>PrEP: pre-exposure prophylaxis.

<sup>d</sup>HIV: human immunodeficiency virus.

<sup>e</sup>AIDS: acquired immunodeficiency syndrome.

<sup>f</sup>GBM: gay and bisexual men.

**Table 5.** Source characteristics of Instagram posts among organizations regarding pre-exposure prophylaxis (N=193).

Information available in the posts, profiles, biographies, or affiliated websites of organizations	n	%
Health information provider	115	59.6
Nonprofit/advocacy group	95	49.2
Health care organization	89	46.1
Business (eg, company, franchise, store, product, or service)	44	22.8
Nonhealth-related advocacy group	42	21.8
News or media organization	40	20.7
City, state, or federal government	16	8.3
School or school district	1	0.5

It must be noted that the source characteristics in Table 4 do not add up to 100% owing to information being unavailable on the individual poster's Instagram bio. The source characteristics in Table 5 exceed 100% because an organization may be included under multiple categories (eg, a business and a health care organization).

### Media Type

Of the 250 Instagram posts reviewed, more than two-thirds (174/250, 69.6%) included images (ie, nonmoving photo image or snapshot), more than half (142/250, 56.8%) included infographics (ie, photos or images containing factual information/data/charts), and approximately one-tenth (30/250, 12%) included a video (eg, video, Boomerang, graphic interchange format [GIF] files).

### RQs

**RQ1:** *What is the textual content of PrEP-related Instagram posts?* Table 1 displays the results answering our first research question. We observed that more than half of all the reviewed Instagram posts defined PrEP (137/250, 54.8%). Fewer posts promoted PrEP use (93/250, 37.2%), explained how PrEP works (80/250, 32%), and included information on PrEP's effectiveness (70/250, 28%) or who can use PrEP (68/250, 27.2%). Less than one-quarter of all the reviewed posts provided information regarding how to obtain PrEP (55/250, 22%), costs related to PrEP (49/250, 19.6%), or adverse effects of PrEP (17/250, 6.8%). More posts were classified as battling stigma (20/250, 8%) than those classified as stigmatizing PrEP users (5/250, 2%). As displayed in Table 3, less than half of all the posts provided some source attribution (eg, citation) for the information posted. The most cited sources in the reviewed posts were the CDC (25/250, 10%) and PrEP manufacturers (eg, Gilead; 22/250, 8.8%).

**RQ2:** *Which priority populations are the focus of PrEP-related Instagram posts?* When answering our second research question, we found that PrEP-related Instagram posts were most commonly focused on GBM and MSM. The most commonly hashtagged priority population among posts included MSM but not necessarily bisexual men (eg, #gayman, #gaymen, #gayboy; 69/250, 27.6%). Few Instagram posts contained only male-associated hashtags (eg, #men; 39/250, 15.6%) or only female-associated tags (eg, #girlboss, #girlpower; 7/250, 2.8%). Very few posts contained race- or ethnicity-related hashtags (11/250, 4.4%), such as #BLACKPOWER and #latinxpower.

Even fewer posts contained transgender-associated tags (e.g., #trans, #transgirl, #translatina; 5/250, 2%). No posts contained tags related to heterosexuals or injection drug users.

**RQ3:** *How did the textual content of PrEP-related posts vary by source (eg, individual accounts versus organization accounts)?* Instagram posts from organizations were more likely to describe who can use PrEP, compared to those posted by individuals, with  $X^2_1=9.7$  (N=248) and  $P=.002$ . Almost one-third (62/193, 32.1%) of the posts from organizations described who can use PrEP, compared with 11% (6/55) of the posts from individuals. Conversely, Instagram posts from individuals were more likely to mention the adverse effects of PrEP, compared to the posts by organizations, with  $X^2_1=24.8$  (N=248) and  $P<.001$ . More than 20% (12/55, 21.8%) of the posts from individuals mentioned PrEP adverse effects, compared to 3% (6/193) of the posts from organizations.

Regarding source attribution differences, Instagram posts from organizations were more likely to cite information from the CDC or other federal sources, compared to those posted by individuals, with  $X^2_1=5.3$  (N=248) and  $P=.02$ . More than 10% (24/193, 12.4%) of the posts from organizations cited the CDC, compared to 2% (1/55) of the posts from individuals. Conversely, Instagram posts from individuals were more likely to cite information from a PrEP manufacturer compared to the posts by organizations, with  $X^2_1=9.8$  (N=248) and  $P=.002$ . Almost 20% (10/55, 18.2%) of the posts from individuals mentioned Gilead or Merck, or their representatives, compared to 5% (10/193) of the posts from organizations. No other statistical differences were found.

## Discussion

### Principal Results

This research study is among the first to review Instagram posts for textual content specifically related to PrEP. We found that PrEP awareness/information (eg, posts defining PrEP) comprised the largest content focus; gay men comprised the priority population most commonly represented in these posts; organizations and individuals differed somewhat in terms of their post contents. These findings are further contextualized below.

## Comparison With Prior Work

Our results, similar to those reported in studies involving PrEP-related content on Twitter [27] and on YouTube [23,24], suggest that a broad range of topics are mentioned in PrEP-related posts on Instagram. Our finding that awareness-related posts featuring PrEP definitions represent the largest content focus is consistent with previous research results regarding PrEP messages on Twitter. Schwartz and Grimm offered possible explanations for why this may be the most prominent form of PrEP content on social media [27] and identified potential reasons for low prescription rates [45] and promotion of PrEP [46], along with the uncertainty conveyed in online news stories about PrEP appearing in the United States [42].

Uncertainty about health-related topics can drive information seeking [47] and if clear PrEP-related information is unavailable through traditional information sources, the uncertainty surrounding PrEP may prompt some individuals to consult alternative sources (including social media) to fulfill their informational needs. This notion is supported by Breen et al [25] who found increases in Twitter posts focusing on PrEP (as well as other health-related topics) in recent years. Young adults (especially those aged 18-24 years) may turn to social media as most individuals in this age group are already daily users of platforms such as Instagram [20]. Owing to the sensitive nature of PrEP-related information, these young adults may seek this information and other sexual health-related information from online sources (including social media), rather than from health care providers, or family and friends [48]. This could be viewed positively as organizations seem to be utilizing Instagram to provide PrEP-related information; however, as described in the following paragraphs, PrEP-related posts from organizations differ significantly from those of individuals in ways that may influence their impact on PrEP uptake and adherence.

One of the interesting findings from this research is that the current population-related focus appears to be on gay men, as the MSM population is the most commonly hashtagged priority population in the posts (eg, #gayman, #gaysmen, #gayboy), although these individuals are not necessarily bisexual men. As rates of HIV diagnoses remain high among gay men, PrEP promotion messaging should focus on this important priority population. However, given that the rate of insertive condomless anal sex acts with casual partners is statistically significantly higher in bisexual men compared to gay men [49], a PrEP promotion focus on bisexual men is sorely lacking in this social media space. Such a focus is important, as research indicates that focusing on HIV in cities with high numbers of bisexual men (men who have sex with men and women) may have the dual effect of improving the health of the bisexual community and the health of MSM populations that have a high burden of HIV [50].

We also found that very few Instagram posts contained race- or ethnicity-related co-occurring hashtags, and even fewer posts contained transgender-associated co-occurring tags. Given that Black individuals account for a higher proportion of new HIV diagnoses and people living with HIV, compared to other races/ethnicities, people of color represent important priority

populations in terms of HIV risk and PrEP promotion [51]. This is especially true for men and women, as Black men represented 31% of the new HIV diagnoses in the United States, and Black women are approximately 16 times more likely to receive a diagnosis of HIV infection compared to White women [51,52]. In addition to a weak social media focus on PrEP for people of color, we found that only 2% (5/250) of the posts contained transgender-associated tags although transgender persons are at a very high risk for HIV infection [53]. Lastly, we found no PrEP promotion focusing on heterosexual adults or injection drug users in Instagram posts, both of which represent groups for which PrEP is indicated in the United States [8]. The lack of PrEP promotion on Instagram for bisexual men, people of color, transgender persons, heterosexual adults, and people who inject drugs remains a critical missed opportunity.

Our results also show some differences in the source characteristics of the PrEP-related posts as well as the textual content shared by organizations and individuals. Organizations comprised the vast majority (193/250, 77.2%) of PrEP-related Instagram posts, a stark difference compared to that reported by McLaughlin et al [26] who stated that organizations were less common creators of tweets focusing on PrEP. Organizations are significantly more likely to share information regarding who can use PrEP. This focus may coincide with the advocacy goals of HIV nonprofit organizations seeking to leverage social media as a platform for intervening and expanding organizational capacity to increase PrEP awareness, disseminating educational material, and enhancing engagement with members of the public [54,55]. If part of these public engagement and outreach efforts involves partnerships with federal agencies, it may help explain why organizations (compared to individuals) in our data are more likely to cite the CDC and other federal sources in their PrEP-related content. Another possibility is that these organizations are more likely to mention federal agencies as reliable sources for scientific research on PrEP in their content to offer clarity and reduce public uncertainty about PrEP [27,56,57]. In contrast, individuals were more likely to include PrEP-related content focusing on the adverse effects of PrEP and were also more likely to mention manufacturers such as Gilead or Merck, or their representatives. One explanation for this result may be that individuals are more likely to focus on the adverse effects of PrEP and cite PrEP manufacturers as part of larger public discussions surrounding consumer experiences with the pharmaceutical drug industry and drug safety. Prior research suggests that social media can be a rich source for these public discussions [58]. These findings highlight the perceived differences between posts made by PrEP users and those organizations that develop or promote PrEP. We recommend that organizations work with communities related to PrEP or sexual health influencers on Instagram to develop and maintain explicit partnerships for implementing communication strategies aimed at reducing barriers to PrEP uptake and adherence (eg, stigma, lack of knowledge about adverse effects, and costs). This strategy has been examined by previous research into Instagram suggesting that influencers provide several techniques for disseminating information that may be less possible for organizations (eg, word-of-mouth dissemination and celebrity status updates) [59-61].

## Limitations

This research should be considered within the context of its limitations. First, the Instagram posts included in this research are only those that are publicly available. Although this is consistent with the data showing that most social media users report setting all their social media accounts to private [62], it presented a limitation to our study, as we were unable to generalize our findings beyond publicly accessible posts. Second, although we followed a methodical process for identifying hashtags to use in our search for posts, we may have excluded some hashtags, resulting in missed posts pertaining to PrEP. Third, of the 193 posts coded as stemming from organizations, 49 originated from a single account (Prevention305), which represents 25% of that organizational subgroup. This may have generally skewed the findings for organizations toward Prevention305's PrEP awareness campaign goals. Fourth, as this study was our first attempt to apply a predefined codebook used in textual analyses of Twitter data to Instagram, we have not reported on the data extracted from photos or videos. The only exception was when we described the basic topology of the images (ie, whether the image was a photo, an infographic, or a video). This necessitates a slightly different application of our methodology than what was stated as the purpose of this study and may be utilized in future projects to expand on the findings highlighted in the current paper. Lastly, because we focused our review on English-only posts, we may have missed important Instagram content related to Spanish-speaking people (eg, Latinx).

Despite these limitations, our study exhibits some strengths. This study is among the first to examine PrEP-related textual content on Instagram. Despite its popularity, especially among young adults, Instagram remains an understudied platform, and the current study begins to address this gap in the extant literature. In addition, although our sample only includes publicly available posts, to the extent that as these posts are

indeed reflective of the public PrEP-related information environment on Instagram, they may point toward types of PrEP-related content that Instagram users seeking information about PrEP are likely to encounter. Future research may confirm this if interviews or surveys are conducted to determine the most common types of PrEP-related content Instagram users recall encountering during active searches for PrEP information. Public health professionals may also consider the publicly available user comments for PrEP-related content posted on Instagram to gauge user reactions. These user comments may offer insights into how priority populations may respond to PrEP-related content serving as the basis for future PrEP-related interventions. Future studies should also strive to examine such content across social media platforms, including Instagram, Facebook, and Twitter.

## Conclusions

The National AIDS Strategy's call to more clearly articulate the science surrounding HIV risk and prevention is more fully addressed by first understanding the current public information environment surrounding PrEP. The present study seeks to begin answering this call by offering a snapshot of how PrEP is being discussed (and by whom) on one of the most popular social media platforms. In addition, this study responds to the National AIDS Strategy's recommendation to develop campaign strategies that leverage the unique properties of emerging digital technologies by laying the foundation for big data approaches that may be applied to glean insights for messaging in PrEP campaigns and interventions. These findings highlight the additional efforts required to reach the National AIDS Strategy's goal of improved communication surrounding PrEP. The small number of Instagram posts that feature PrEP highlight a less than optimal level of engagement, and the current study should serve as a call for investigators to utilize emerging tools such as Instagram more effectively to engage priority populations in conversations around PrEP.

## Conflicts of Interest

None declared.

## References

1. Centers for Disease Control & Prevention. HIV Surveillance Report, 2018 (updated). Center for Disease Control and Prevention. URL: <https://www.cdc.gov/hiv/pdf/library/reports/surveillance/cdc-hiv-surveillance-report-2018-updated-vol-31.pdf> [accessed 2021-07-16]
2. Clark H, Babu AS, Wiewel EW, Opoku J, Crepaz N. Diagnosed HIV infection in transgender adults and adolescents: results from the national HIV surveillance system, 2009–2014. *AIDS Behav* 2017 Sep;21(9):2774-2783 [FREE Full text] [doi: [10.1007/s10461-016-1656-7](https://doi.org/10.1007/s10461-016-1656-7)] [Medline: [28035497](https://pubmed.ncbi.nlm.nih.gov/28035497/)]
3. Preexposure Prophylaxis for the Prevention of HIV Infection in the United States – 2017 Update: Clinical Practice Guideline. 2017. URL: <https://www.cdc.gov/hiv/pdf/risk/prep/cdc-hiv-prep-guidelines-2017.pdf> [accessed 2021-07-16]
4. Baeten JM, Donnell D, Ndase P, Mugo NR, Campbell JD, Wangisi J, Partners PrEP Study Team. Antiretroviral prophylaxis for HIV prevention in heterosexual men and women. *N Engl J Med* 2012 Aug;367(5):399-410 [FREE Full text] [doi: [10.1056/NEJMoa1108524](https://doi.org/10.1056/NEJMoa1108524)] [Medline: [22784037](https://pubmed.ncbi.nlm.nih.gov/22784037/)]
5. Choopanya K, Martin M, Suntharasamai P, Sangkum U, Mock PA, Leethochawalit M, et al. Antiretroviral prophylaxis for HIV infection in injecting drug users in Bangkok, Thailand (the Bangkok Tenofovir Study): a randomised, double-blind, placebo-controlled phase 3 trial. *The Lancet* 2013 Jun;381(9883):2083-2090. [doi: [10.1016/S0140-6736\(13\)61127-7](https://doi.org/10.1016/S0140-6736(13)61127-7)] [Medline: [23769234](https://pubmed.ncbi.nlm.nih.gov/23769234/)]

6. Grohskopf LA, Chillag KL, Gvetadze R, Liu AY, Thompson M, Mayer KH, et al. Randomized trial of clinical safety of daily oral tenofovir disoproxil fumarate among HIV-uninfected men who have sex with men in the United States. *J Acquir Immune Defic Syndr* 2013 Sep;64(1):79-86. [doi: [10.1097/QAI.0b013e31828ece33](https://doi.org/10.1097/QAI.0b013e31828ece33)] [Medline: [23466649](https://pubmed.ncbi.nlm.nih.gov/23466649/)]
7. Thigpen MC, Kebaabetswe PM, Paxton LA, Smith DK, Rose CE, Segolodi TM, TDF2 Study Group. Antiretroviral preexposure prophylaxis for heterosexual HIV transmission in Botswana. *N Engl J Med* 2012 Aug;367(5):423-434. [doi: [10.1056/NEJMoa1110711](https://doi.org/10.1056/NEJMoa1110711)] [Medline: [22784038](https://pubmed.ncbi.nlm.nih.gov/22784038/)]
8. Smith DK, Van Handel M, Grey J. Estimates of adults with indications for HIV pre-exposure prophylaxis by jurisdiction, transmission risk group, and race/ethnicity, United States, 2015. *Ann Epidemiol* 2018 Dec;28(12):850-857.e9. [doi: [10.1016/j.annepidem.2018.05.003](https://doi.org/10.1016/j.annepidem.2018.05.003)] [Medline: [29941379](https://pubmed.ncbi.nlm.nih.gov/29941379/)]
9. Golub SA, Fikslin RA, Starbuck L, Klein A. High rates of PrEP eligibility but low rates of PrEP access among a national sample of transmasculine individuals. *J Acquir Immune Defic Syndr* 2019 Sep;82(1):e1-e7 [FREE Full text] [doi: [10.1097/QAI.0000000000002116](https://doi.org/10.1097/QAI.0000000000002116)] [Medline: [31232834](https://pubmed.ncbi.nlm.nih.gov/31232834/)]
10. Ojikutu BO, Bogart LM, Higgins-Biddle M, Dale SK, Allen W, Dominique T, et al. Facilitators and barriers to pre-exposure prophylaxis (PrEP) use among black individuals in the United States: results from the National Survey on HIV in the Black Community (NSHBC). *AIDS Behav* 2018 Nov;22(11):3576-3587 [FREE Full text] [doi: [10.1007/s10461-018-2067-8](https://doi.org/10.1007/s10461-018-2067-8)] [Medline: [29468493](https://pubmed.ncbi.nlm.nih.gov/29468493/)]
11. Parsons JT, Rendina HJ, Lassiter JM, Whitfield THF, Starks TJ, Grov C. Uptake of HIV pre-exposure prophylaxis (PrEP) in a national cohort of gay and bisexual men in the United States. *J Acquir Immune Defic Syndr* 2017 Mar;74(3):285-292 [FREE Full text] [doi: [10.1097/QAI.0000000000001251](https://doi.org/10.1097/QAI.0000000000001251)] [Medline: [28187084](https://pubmed.ncbi.nlm.nih.gov/28187084/)]
12. Dodge B, Ford JV, Bo N, Tu W, Pachankis J, Herbenick D, et al. HIV risk and prevention outcomes in a probability-based sample of gay and bisexual men in the United States. *J Acquir Immune Defic Syndr* 2019 Dec;82(4):355-361 [FREE Full text] [doi: [10.1097/QAI.0000000000002151](https://doi.org/10.1097/QAI.0000000000002151)] [Medline: [31517706](https://pubmed.ncbi.nlm.nih.gov/31517706/)]
13. Finlayson T, Cha S, Xia M, Trujillo L, Denson D, Prejean J, et al. Changes in HIV Preexposure Prophylaxis Awareness and Use Among Men Who Have Sex with Men — 20 Urban Areas, 2014 and 2017. *MMWR Morb Mortal Wkly Rep* 2019 Jul;68(27):597-603 [FREE Full text] [doi: [10.15585/mmwr.mm6827a1](https://doi.org/10.15585/mmwr.mm6827a1)]
14. Allen ST, O'Rourke A, White RH, Smith KC, Weir B, Lucas GM, et al. Barriers and facilitators to PrEP use among people who inject drugs in rural Appalachia: a qualitative study. *AIDS Behav Internet* 2019 Dec 2020;24:1942-1950.
15. Jo Y, Bartholomew TS, Doblecki-Lewis S, Rodriguez A, Forrest DW, Tomita-Barber J, et al. Interest in linkage to PrEP among people who inject drugs accessing syringe services; Miami, Florida. *PLoS ONE* 2020 Apr 16;15(4):e0231424. [doi: [10.1371/journal.pone.0231424](https://doi.org/10.1371/journal.pone.0231424)]
16. McFarland W, Lin J, Santos G, Arayasirikul S, Raymond HF, Wilson E. Low PrEP awareness and use among people who inject drugs, San Francisco, 2018. *AIDS Behav* 2019 Sep;24(5):1290-1293. [doi: [10.1007/s10461-019-02682-7](https://doi.org/10.1007/s10461-019-02682-7)]
17. Sherman SG, Schneider KE, Park JN, Allen ST, Hunt D, Chaulk CP, et al. PrEP awareness, eligibility, and interest among people who inject drugs in Baltimore, Maryland. *Drug Alcohol Depend* 2019 Feb;195:148-155 [FREE Full text] [doi: [10.1016/j.drugalcdep.2018.08.014](https://doi.org/10.1016/j.drugalcdep.2018.08.014)] [Medline: [30639794](https://pubmed.ncbi.nlm.nih.gov/30639794/)]
18. National HIV/AIDS strategy for the United States: Updated to 2020. White House Office of National AIDS Policy. 2015. URL: <https://www.hiv.gov/sites/default/files/nhas-2020-action-plan.pdf> [accessed 2021-07-16]
19. Ahlqvist T, Bäck A, Heinonen S, Halonen M. Road-mapping the societal transformation potential of social media. *Foresight* 2010 Aug 31;12(5):3-26. [doi: [10.1108/14636681011075687](https://doi.org/10.1108/14636681011075687)]
20. Perrin A, Anderson M. Share of U.S. adults using social media, including Facebook, is mostly unchanged since 2018. Internet. Pew Research Center. URL: <https://www.pewresearch.org/fact-tank/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018/> [accessed 2021-07-16]
21. Demographics of social media users and adoption in the United States. Internet. Pew Research Center. URL: <https://www.pewresearch.org/internet/fact-sheet/social-media/> [accessed 2021-07-16]
22. Seidenberg AB, Jo CL, Ribisl KM, Lee JGL, Buchting FO, Kim Y, et al. A national study of social media, television, radio, and internet usage of adults by sexual orientation and smoking status: implications for campaign design. *Int J Environ Res Public Health* 2017 Apr;14(4):450 [FREE Full text] [doi: [10.3390/ijerph14040450](https://doi.org/10.3390/ijerph14040450)] [Medline: [28430161](https://pubmed.ncbi.nlm.nih.gov/28430161/)]
23. Kecojevic A, Meleo-Erwin ZC, Basch CH, Hammouda M. A thematic analysis of pre-exposure prophylaxis (PrEP) YouTube videos. *J Homosex* 2020 Jan;1-22. [doi: [10.1080/00918369.2020.1712142](https://doi.org/10.1080/00918369.2020.1712142)] [Medline: [31939710](https://pubmed.ncbi.nlm.nih.gov/31939710/)]
24. Kecojevic A, Basch C, Basch C, Kernan W. Pre-exposure prophylaxis YouTube videos: content evaluation. *JMIR Public Health Surveill* 2018 Feb;4(1):e19 [FREE Full text] [doi: [10.2196/publichealth.7733](https://doi.org/10.2196/publichealth.7733)] [Medline: [29467119](https://pubmed.ncbi.nlm.nih.gov/29467119/)]
25. Breen P, Kelly J, Heckman T, Quinn S. Mining pre-exposure prophylaxis trends in social media. 2016 Presented at: 2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA); October 17-19, 2016; Montreal, QC, Canada p. 214 URL: <https://ieeexplore.ieee.org/document/7796907> [doi: [10.1109/dsaa.2016.29](https://doi.org/10.1109/dsaa.2016.29)]
26. McLaughlin ML, Hou J, Meng J, Hu C, An Z, Park M, et al. Propagation of information about preexposure prophylaxis (PrEP) for HIV prevention through Twitter. *Health Commun* 2016 Aug;31(8):998-1007. [doi: [10.1080/10410236.2015.1027033](https://doi.org/10.1080/10410236.2015.1027033)] [Medline: [26756069](https://pubmed.ncbi.nlm.nih.gov/26756069/)]
27. Schwartz J, Grimm J. PrEP on Twitter: information, barriers, and stigma. *Health Commun* 2017 Apr;32(4):509-516. [doi: [10.1080/10410236.2016.1140271](https://doi.org/10.1080/10410236.2016.1140271)] [Medline: [27295507](https://pubmed.ncbi.nlm.nih.gov/27295507/)]



28. Hutchinson A. Instagram Now Has a Billion Monthly Active Users Internet. Social Media Today. URL: <https://www.socialmediatoday.com/news/instagram-now-has-a-billion-monthly-active-users/526188/> [accessed 2021-07-16]
29. Hutchinson A. Top 10 Instagram Stats for 2020 Infographic Internet. Social Media Today. URL: <https://www.socialmediatoday.com/news/top-10-instagram-stats-for-2020-infographic/569641/> [accessed 2021-07-16]
30. Kamel Boulos M, Giustini D, Wheeler S. Instagram and WhatsApp in health and healthcare: an overview. *Future Internet* 2016 Jul;8(3):37 [FREE Full text] [doi: [10.3390/fi8030037](https://doi.org/10.3390/fi8030037)]
31. Chung C, Agapie E, Schroeder J, Mishra S, Fogarty J, Munson SA. When personal tracking becomes social: examining the use of Instagram for healthy eating. *Proc SIGCHI Conf Hum Factor Comput Syst* 2017 May;2017:1674-1687 [FREE Full text] [doi: [10.1145/3025453.3025747](https://doi.org/10.1145/3025453.3025747)] [Medline: [28516174](https://pubmed.ncbi.nlm.nih.gov/28516174/)]
32. Ward J. A content analysis of celebrity Instagram posts and parasocial interaction. *Elon J Undergrad Res Commun Internet* 2016;7(1):1-1 [FREE Full text]
33. Nobles AL, Leas EC, Latkin CA, Dredze M, Strathdee SA, Ayers JW. #HIV: alignment of HIV-related visual content on Instagram with public health priorities in the US. *AIDS Behav* 2020 Jul;24(7):2045-2053. [doi: [10.1007/s10461-019-02765-5](https://doi.org/10.1007/s10461-019-02765-5)] [Medline: [31916098](https://pubmed.ncbi.nlm.nih.gov/31916098/)]
34. Simmons J. The InstaPrEP project: a social marketing intervention on Instagram to increase oral HIV prophylaxis in black men who have sex with men (Doctoral dissertation, UCLA). UCLA Electronic Theses and Dissertations. 2020. URL: <https://escholarship.org/uc/item/8kj814kg> [accessed 2021-07-16]
35. Whiteley L, Healy MG, Lowery A, Haubrick KK, Brown LK. An evaluation and review of English language pre-exposure prophylaxis websites and YouTube videos. *Int J STD AIDS* 2020 Mar;31(5):460-466. [doi: [10.1177/0956462420905271](https://doi.org/10.1177/0956462420905271)]
36. CrowdTangle. CrowdTangle | content discovery and social monitoring made easy Internet. CrowdTangle. URL: <https://www.crowdtangle.com/> [accessed 2021-07-16]
37. Chen N, Drouhard M, Kocielnik R, Suh J, Aragon CR. Using machine learning to support qualitative coding in social science. *ACM Trans Interact Intell Syst* 2018 Jul;8(2):1-20. [doi: [10.1145/3185515](https://doi.org/10.1145/3185515)]
38. Briones R, Nan X, Madden K, Waks L. When vaccines go viral: an analysis of HPV vaccine coverage on YouTube. *Health Commun* 2012;27(5):478-485. [doi: [10.1080/10410236.2011.610258](https://doi.org/10.1080/10410236.2011.610258)] [Medline: [22029723](https://pubmed.ncbi.nlm.nih.gov/22029723/)]
39. Madden K, Nan X, Briones R, Waks L. Sorting through search results: a content analysis of HPV vaccine information online. *Vaccine* 2012 May;30(25):3741-3746. [doi: [10.1016/j.vaccine.2011.10.025](https://doi.org/10.1016/j.vaccine.2011.10.025)] [Medline: [22019758](https://pubmed.ncbi.nlm.nih.gov/22019758/)]
40. Quintero Johnson J, Sionean C, Scott AM. Exploring the presentation of news information about the HPV vaccine: a content analysis of a representative sample of U.S. newspaper articles. *Health Commun* 2011 Sep;26(6):491-501. [doi: [10.1080/10410236.2011.556080](https://doi.org/10.1080/10410236.2011.556080)] [Medline: [21469005](https://pubmed.ncbi.nlm.nih.gov/21469005/)]
41. Muralidhara S, Paul MJ. #Healthy Selfies: exploration of health topics on Instagram. *JMIR Public Health Surveill* 2018 Jun;4(2):e10150. [doi: [10.2196/10150](https://doi.org/10.2196/10150)]
42. Bastos MT, Raimundo RLG, Travitzki R. Gatekeeping Twitter: message diffusion in political hashtags. *Media Cult Soc* 2013 Mar;35(2):260-270. [doi: [10.1177/0163443712467594](https://doi.org/10.1177/0163443712467594)]
43. Ma Z, Sun A, Cong G. On predicting the popularity of newly emerging hashtags in Twitter. *J Am Soc Inf Sci Technol* 2013 May;64(7):1399-1410 [FREE Full text] [doi: [10.1002/asi.22844](https://doi.org/10.1002/asi.22844)]
44. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977 Mar;33(1):159-174. [doi: [10.2307/2529310](https://doi.org/10.2307/2529310)]
45. Doblecki-Lewis S, Cohen S, Liu A. Clinical treatment options infectious diseases: update on PrEP implementation, adherence, and advances in delivery. *Curr Treat Options Infect Dis* 2015 Jun;7(2):101-112 [FREE Full text] [doi: [10.1007/s40506-015-0046-4](https://doi.org/10.1007/s40506-015-0046-4)] [Medline: [25931990](https://pubmed.ncbi.nlm.nih.gov/25931990/)]
46. McNeil JD, Tuller D. AIDS groups back anti-H.I.V. pill amid concerns doctors and patients will resist. *The New York Times*. 2014 May 16. URL: <https://www.nytimes.com/2014/05/16/health/aids-groups-back-anti-hiv-pill-amid-concerns-doctors-and-patients-will-resist.html> [accessed 2021-07-16]
47. Rains SA. Health information seeking and the World Wide Web: an uncertainty management perspective. *J Health Commun* 2014;19(11):1296-1307. [doi: [10.1080/10810730.2013.872731](https://doi.org/10.1080/10810730.2013.872731)] [Medline: [24730566](https://pubmed.ncbi.nlm.nih.gov/24730566/)]
48. Mitchell KJ, Ybarra ML, Korchmaros JD, Kosciw JG. Accessing sexual health information online: use, motivations and consequences for youth with different sexual orientations. *Health Educ Res* 2014 Feb;29(1):147-157 [FREE Full text] [doi: [10.1093/her/cyt071](https://doi.org/10.1093/her/cyt071)] [Medline: [23861481](https://pubmed.ncbi.nlm.nih.gov/23861481/)]
49. Feinstein BA, Moran KO, Newcomb ME, Mustanski B. Differences in HIV Risk behaviors between self-identified gay and bisexual young men who are HIV-negative. *Arch Sex Behav* 2019 Jan;48(1):261-275 [FREE Full text] [doi: [10.1007/s10508-018-1148-0](https://doi.org/10.1007/s10508-018-1148-0)] [Medline: [29508171](https://pubmed.ncbi.nlm.nih.gov/29508171/)]
50. Raymond HF, Al-Tayyib A, Neaigus A, Reilly KH, Braunstein S, Brady KA, et al. HIV among MSM and heterosexual women in the United States: an ecologic analysis. *J Acquir Immune Defic Syndr* 2017 Jul;75 Suppl 3:S276-S280. [doi: [10.1097/QAI.0000000000001422](https://doi.org/10.1097/QAI.0000000000001422)] [Medline: [28604428](https://pubmed.ncbi.nlm.nih.gov/28604428/)]
51. HIV and African Americans.: Division of HIV/AIDS Prevention, National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, Centers for Disease Control and Prevention; 2021 Jan 20. URL: <https://www.cdc.gov/hiv/group/raciaethnic/africanamericans/index.html> [accessed 2021-07-16]

52. McCree DH, Sutton M, Bradley E, Harris N. Changes in the disparity of HIV diagnosis rates among black women - United States, 2010-2014. *MMWR Morb Mortal Wkly Rep* 2017 Feb;66(4):104-106 [FREE Full text] [doi: [10.15585/mmwr.mm6604a3](https://doi.org/10.15585/mmwr.mm6604a3)] [Medline: [28151925](https://pubmed.ncbi.nlm.nih.gov/28151925/)]
53. Pitasi MA, Oraka E, Clark H, Town M, DiNenno EA. HIV testing among transgender women and men — 27 states and Guam, 2014–2015. *MMWR Morb Mortal Wkly Rep* 2017 Aug 25;66(33):883-887. [doi: [10.15585/mmwr.mm6633a3](https://doi.org/10.15585/mmwr.mm6633a3)]
54. Huang Y, Lin Y, Saxton GD. Give me a like: How HIV/AIDS nonprofit organizations can engage their audience on Facebook. *AIDS Educ Prev* 2016 Dec;28(6):539-556. [doi: [10.1521/aeap.2016.28.6.539](https://doi.org/10.1521/aeap.2016.28.6.539)] [Medline: [27925489](https://pubmed.ncbi.nlm.nih.gov/27925489/)]
55. Sophus AI, Mitchell JW. A review of approaches used to increase awareness of pre-exposure prophylaxis (PrEP) in the United States. *AIDS Behav* 2019 Jul;23(7):1749-1770. [doi: [10.1007/s10461-018-2305-0](https://doi.org/10.1007/s10461-018-2305-0)] [Medline: [30306434](https://pubmed.ncbi.nlm.nih.gov/30306434/)]
56. Saberi P, Gamarel KE, Neilands TB, Comfort M, Sheon N, Darbes LA, et al. Ambiguity, ambivalence, and apprehensions of taking HIV-1 pre-exposure prophylaxis among male couples in San Francisco: a Mixed methods study. *PLoS ONE* 2012 Nov;7(11):e50061. [doi: [10.1371/journal.pone.0050061](https://doi.org/10.1371/journal.pone.0050061)]
57. Jaspal R, Daramilas C. Perceptions of pre-exposure prophylaxis (PrEP) among HIV-negative and HIV-positive men who have sex with men (MSM). *Cogent Med* 2016 Nov;3(1):1256850. [doi: [10.1080/2331205x.2016.1256850](https://doi.org/10.1080/2331205x.2016.1256850)]
58. Pierce CE, Bourri K, Pamer C, Proestel S, Rodriguez HW, Van LH, et al. Evaluation of Facebook and Twitter monitoring to detect safety signals for medical products: an analysis of recent FDA safety alerts. *Drug Saf* 2017 Apr;40(4):317-331 [FREE Full text] [doi: [10.1007/s40264-016-0491-0](https://doi.org/10.1007/s40264-016-0491-0)] [Medline: [28044249](https://pubmed.ncbi.nlm.nih.gov/28044249/)]
59. Pilgrim K, Bohnet-Joschko S. Selling health and happiness how influencers communicate on Instagram about dieting and exercise: mixed methods research. *BMC Public Health* 2019 Aug;19:1054 [FREE Full text] [doi: [10.1186/s12889-019-7387-8](https://doi.org/10.1186/s12889-019-7387-8)] [Medline: [31387563](https://pubmed.ncbi.nlm.nih.gov/31387563/)]
60. Djafarova E, Rushworth C. Exploring the credibility of online celebrities' Instagram profiles in influencing the purchase decisions of young female users. *Comput Human Behav* 2017 Mar;68:1-7. [doi: [10.1016/j.chb.2016.11.009](https://doi.org/10.1016/j.chb.2016.11.009)]
61. Liu S, Jiang C, Lin Z, Ding Y, Duan R, Xu Z. Identifying effective influencers based on trust for electronic word-of-mouth marketing: A domain-aware approach. *Inform Sci* 2015 Jun;306:34-52. [doi: [10.1016/j.ins.2015.01.034](https://doi.org/10.1016/j.ins.2015.01.034)]
62. U.S. social media user account privacy 2018 | Statistic [Internet]. Statista. 2021 Jan 28. URL: <https://www.statista.com/statistics/934874/users-have-private-social-media-account-usa/> [accessed 2021-07-16]

## Abbreviations

**CDC:** Centers for Disease Control and Prevention

**FTC:** emtricitabine

**GBM:** gay and bisexual men

**HPV:** human papillomavirus

**MSM:** men who have sex with men

**PrEP:** pre-exposure prophylaxis

**RQ:** research question

**SD:** standard deviation

**STDs:** sexually transmitted diseases

**TDF:** tenofovir disoproxil fumarate

*Edited by G Eysenbach; submitted 27.08.20; peer-reviewed by A Kecojevic, JP Allem, L Laestadius; comments to author 18.09.20; revised version received 12.11.20; accepted 29.05.21; published 27.07.21.*

*Please cite as:*

Walsh-Buhi E, Houghton RF, Lange C, Hockensmith R, Ferrand J, Martinez L

Pre-exposure Prophylaxis (PrEP) Information on Instagram: Content Analysis

*JMIR Public Health Surveill* 2021;7(7):e23876

URL: <https://publichealth.jmir.org/2021/7/e23876>

doi: [10.2196/23876](https://doi.org/10.2196/23876)

PMID: [34061759](https://pubmed.ncbi.nlm.nih.gov/34061759/)

©Eric Walsh-Buhi, Rebecca Fagen Houghton, Claire Lange, Ryli Hockensmith, John Ferrand, Lourdes Martinez. Originally published in *JMIR Public Health and Surveillance* (<https://publichealth.jmir.org>), 27.07.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Public Health and Surveillance*, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

# The Effect of a Wordless, Animated, Social Media Video Intervention on COVID-19 Prevention: Online Randomized Controlled Trial

Alain Vandormael<sup>1</sup>, PhD; Maya Adam<sup>2</sup>, PhD; Merlin Greuel<sup>1</sup>, MGH; Jennifer Gates<sup>3</sup>, BA; Caterina Favaretti<sup>1</sup>, MSc; Violetta Hachaturyan<sup>1</sup>, MSc; Till Bärnighausen<sup>1,4,5</sup>, PhD

<sup>1</sup>Heidelberg Institute of Global Health, Heidelberg University, Heidelberg, Germany

<sup>2</sup>Department of Pediatrics, Stanford University School of Medicine, Stanford, CA, United States

<sup>3</sup>Icahn School of Medicine at Mount Sinai, New York, NY, United States

<sup>4</sup>Africa Health Research Institute, Durban, South Africa

<sup>5</sup>Department of Global Health and Population, Harvard TH Chan School of Public Health, Boston, MA, United States

**Corresponding Author:**

Alain Vandormael, PhD

Heidelberg Institute of Global Health

Heidelberg University

130.3 Im Neuenheimer Feld

Heidelberg, 69120

Germany

Phone: 49 163 481 3452

Email: [alain.vandormael@uni-heidelberg.de](mailto:alain.vandormael@uni-heidelberg.de)

## Abstract

**Background:** Innovative approaches to the dissemination of evidence-based COVID-19 health messages are urgently needed to counter social media misinformation about the pandemic. To this end, we designed a short, wordless, animated global health communication video (the CoVideo), which was rapidly distributed through social media channels to an international audience.

**Objective:** The objectives of this study were to (1) establish the CoVideo's effectiveness in improving COVID-19 prevention knowledge, and (2) establish the CoVideo's effectiveness in increasing behavioral intent toward COVID-19 prevention.

**Methods:** In May and June 2020, we enrolled 15,163 online participants from the United States, Mexico, the United Kingdom, Germany, and Spain. We randomized participants to (1) the CoVideo arm, (2) an attention placebo control (APC) arm, and (3) a do-nothing arm, and presented 18 knowledge questions about preventive COVID-19 behaviors, which was our first primary endpoint. To measure behavioral intent, our second primary endpoint, we randomized participants in each arm to five list experiments.

**Results:** Globally, the video intervention was viewed 1.2 million times within the first 10 days of its release and more than 15 million times within the first 4 months. Knowledge in the CoVideo arm was significantly higher (mean 16.95, 95% CI 16.91-16.99) than in the do-nothing (mean 16.86, 95% CI 16.83-16.90;  $P < .001$ ) arm. We observed high baseline levels of behavioral intent to perform many of the preventive behaviors featured in the video intervention. We were only able to detect a statistically significant impact of the CoVideo on one of the five preventive behaviors.

**Conclusions:** Despite high baseline levels, the intervention was effective at boosting knowledge of COVID-19 prevention. We were only able to capture a measurable change in behavioral intent toward one of the five COVID-19 preventive behaviors examined in this study. The global reach of this health communication intervention and the high voluntary engagement of trial participants highlight several innovative features that could inform the design and dissemination of public health messages. Short, wordless, animated videos, distributed by health authorities via social media, may be an effective pathway for rapid global health communication during health crises.

**Trial Registration:** German Clinical Trials Register DRKS00021582; <https://tinyurl.com/6r4zkbbn>

**International Registered Report Identifier (IRRID):** RR2-10.1186/s13063-020-04942-7

(*JMIR Public Health Surveill* 2021;7(7):e29060) doi:[10.2196/29060](https://doi.org/10.2196/29060)

## KEYWORDS

social media; cultural and social implications; randomized controlled trial; list experiment; information literacy; COVID-19; pandemic; digital health; infodemiology; global health; public health

## Introduction

Soon after the outbreak of the COVID-19 pandemic, health-related misinformation flooded the social media space [1,2]. Compelling, but often misleading, content captured the attention of a frightened global community [2]. The rapid and widespread dissemination of such misinformation on social media often overshadowed evidence-based recommendations released through more traditional public health communication channels. As a result, dangerous messages that increased the spread of COVID-19 and led to adverse health outcomes were allowed to spread to the estimated 3.8 billion people worldwide who use social media [3]. Tedros Ghebreyesus, Director-General of the World Health Organization warned, “We’re not just fighting an epidemic; we’re fighting an infodemic” [2].

There is a critical need to rapidly disseminate evidence-based informational videos on social media channels to counteract the epidemic of COVID-19 misinformation. To date, public health efforts have focused on correcting misinformation and debunking myths [4]. As such, these measures have almost exclusively been reactive rather than proactive. The corrective content itself has not been designed to incorporate the very characteristics that support the viral spread of content on social media [5]. For this reason, social media interventions designed to correct misinformation have unfortunately demonstrated far less impact than the content they aim to correct [4]. Researchers studying this emerging global health communication approach have urged health authorities to enter the social media arena more intentionally, with the aim of disseminating valid information, evaluating its impact, and reducing the knowledge translation gap [3]. Social media health messaging interventions need to do more than convey reliable information. They must be as emotionally compelling as they are evidence-based, if public health authorities are to reach broad, global audiences [5]. They also need to be accessible and tailored for cross-cultural acceptability [6].

In March 2020, we designed a short, wordless, and animated video to disseminate information about preventing the spread of COVID-19 [7,8]. The intervention video (the CoVideo) promotes evidence-based messages that focus on a set of preventive behaviors such as hand washing, social distancing, and the sanitation of kitchen surfaces, among others. Importantly, the CoVideo incorporates audience engagement characteristics that motivate widespread sharing on social media [5]. For example, it includes a compelling, familiar narrative and characters that are culturally agnostic; and the soundtrack is designed to evoke high-arousal emotions [9], which reflects the anxiety, altruism, and solidarity [10] of the global community. The CoVideo was released on Stanford Medicine’s YouTube channel on March 21, 2020, and within 10 days reached 332,000 views on YouTube, 220,000 views on Instagram, 294,000 views on Facebook, and 402,000 views on Twitter, with a cumulative count of 1.2 million [6]. It continued to spread organically across social media channels, due to

reposting by several global health authorities, including government departments of health, community health organizations, and media channels around the world [6]. Within 4 months, the CoVideo had reached more than 5.8 million people through their social media accounts.

In this study, we evaluate the effectiveness of the CoVideo to improve knowledge and behavioral intent toward COVID-19 prevention. According to the Theory of Planned Behavior, the intention to act is considered to be the immediate determinant of action [11]. Here, we frame behavioral intent as representing the participant’s commitment to undertake COVID-19 prevention behaviors in the next 7 days, which is the second outcome of our study [12,13]. As the primary outcome of our study, we aim to measure changes in knowledge about COVID-19 prevention. Knowledge is often considered to be a necessary but not sufficient condition for motivating a healthy behavior [14]. Specifically, the Theory of Planned Behavior posits that knowledge is more likely to be correlated with behavior if correct answers on the knowledge test support the practice of that behavior [15]. Results from this study, which incorporates several innovations in global health communication, can inform the development of future videos to disseminate evidence-based recommendations related to COVID-19 and other public health emergencies.

## Methods

### Trial Design

This is a multisite, parallel-group randomized controlled trial (RCT) comparing the effectiveness of a short informational video on COVID-19 prevention. To evaluate the effectiveness of the CoVideo, we enrolled participants from five countries into a large, online RCT. We randomly assigned participants to the CoVideo [7], an attention placebo control (APC) video [16], or no video (do-nothing arm), and measured change in knowledge of COVID-19 prevention behaviors (first endpoint) and change in self-reported behavioral intent toward COVID-19 prevention (second endpoint). Our RCT included two innovative experimental approaches. First, we used the APC to isolate the content effect of the CoVideo (the active component of the COVID-19–related health messaging and its delivery design) from the attention effect of watching a video (the inactive component of the intervention). Second, we nested a list experiment in each trial arm to reduce socially desirable responses to the behavioral intent questions. Both approaches were leveraged to improve the accuracy of our estimates. The study and its outcomes were registered with the German Clinical Trials Register [17] on May 12, 2020 (DRKS00021582). Ethical approval was obtained from the Stanford University Institutional Review Board on April 12, 2020 (#55820). There were no changes to the trial outcomes or methods after the trial commenced.

## Participants

We used an online platform called Prolific [18] to enroll participants from the United States, Mexico, the United Kingdom, Germany, and Spain into the RCT [8]. Participant eligibility included being 18 to 59 years of age (male, female, or other), being a resident of one of the five countries, and having proficiency in English, German, or Spanish. The trial was hosted and deployed on Gorilla [19], which is a cloud platform that provides versatile tools to undertake online, experimental, and behavioral research [20]. Participants were compensated an equivalent of £1 (US\$ 1.39) for a 10-minute completion time. To prevent duplicate participation, Prolific uses a number of tracking mechanisms, including IP and internet service provider address detection [21].

## Procedures

Participants began by answering basic demographic questions about their age, sex, primary language, country of residence, and highest education completed.

The Gorilla algorithm then randomly assigned participants 1:1:1 to the CoVideo, APC video, or do-nothing groups. Participants were required to watch the CoVideo or the APC video once from start to finish. The CoVideo is animated with sound effects but does not include any words, speech, or text. It explains how the novel coronavirus is spread (airborne, physical contact) and recommends best practices to prevent onward transmission (staying at home, not congregating in public spaces, and sanitizing hands/surfaces). It also covers the mass media coverage of the outbreak and the public's response to this media coverage, which includes a subplot on the stockpiling of essential goods, and the impact thereof on health care services and resources (eg, doctors being unable to access protective equipment). The total duration of the CoVideo is 2 minutes, 30 seconds.

The APC is also a wordless, animated video with the same duration as the CoVideo. Its content describes how small choices become actions, which become habits, which become a way of life. We included an APC to account for possible attention effects elicited by the video format. APC conditions should mimic the “inactive” components of an intervention—the effect of watching the video—while not containing any of the “active” intervention components—the content delivered by the video [22]. We did not make the assumption that the CoVideo is better than nothing (ie, no video). It is possible that the CoVideo could motivate reactance to our COVID-19 prevention message [23-25].

After completing the intervention (CoVideo, APC, do-nothing), participants answered 18 knowledge questions on preventive COVID-19 behaviors. All items required true or false responses, and all participants received the knowledge items. After completing the knowledge questions, participants then completed five list experiments. For each list experiment, we randomized participants 1:1 to a control list or a control list plus a sensitive item about behavioral intent toward social distancing, washing hands, cleaning dishes, cleaning kitchen surfaces, and the stockpiling of essential goods. The control group received a list of 5 items that were unrelated to COVID-19. For example,

in the first list experiment, we asked: “How many of the five statements do you agree with? We don't want to know which ones, just answer how many: 1. Spend time watching TV, 2. Do the vacuuming, 3. Pick a fight with my partner, 4. Eat a low sugar diet, 5. Rinse my nose with salt water daily.” The treatment group received the same 5 items and 1 additional “sensitive” item, “Go out with my friends,” which indicates behavioral intent to social distance (or not) during lockdown restrictions. We used the list experiments to reduce social desirability bias [26,27] and designed them in line with best practices [28].

## Statistical Analysis

We summarized the participant characteristics by obtaining mean (SD) values for age, gender, primary language, country of residence, and education status. Using the Gorilla platform, we identified and excluded participants from the analysis who were lost, defined as those who did not complete the survey from start to finish. Because we could not determine if participants watched some or all of the CoVideo or APC video, we used an intention-to-treat analysis.

For the first endpoint, we calculated a knowledge score for each participant by adding the correct responses (min=0, max=18). Participants had a time limit of 30 seconds to answer each knowledge item, preventing them from searching for answers on the internet. If the participant timed out, they received a missing value of 9. This missing value was recoded as an incorrect answer to the knowledge item, since the participant could not correctly answer the question in the allotted time. We used an analysis of variance (ANOVA) model with and the Tukey honestly significant difference to test for statistically significant differences (with  $\alpha=0.05$ ) in mean knowledge between the CoVideo, APC, and do-nothing arms. The ANOVA model is  $y=b_1\text{VideoArm}$ , where  $y$  is the number of knowledge statements that the participant correctly answered and VideoArm represents the treatment arm.

For the second endpoint, we calculated the prevalence of behavioral intent to perform COVID-19 preventive behaviors for each list experiment. Let  $C_j$  denote the number of items that the  $j$ th participant selected from the control list (min=0, max=5), and let  $T_j$  be the number of items that the  $j$ th participant selected from the treatment list (min=0, max=6). We calculated the mean score for the control list, denoted by  $\bar{C}$  and treatment list, denoted by  $\bar{T}$ , for the  $i$ th list experiment ( $i=1, \dots, 5$ ). Let the superscripts *cov* denote the CoVideo, *apc* denote the APC, and *no* denote the do-nothing arms, and let  $k$  denote the  $k$ th trial arm ( $k \in [\text{cov}, \text{apc}, \text{no}]$ ). For list experiment  $i$  and trial arm  $k$ , we then estimated the prevalence of behavioral intent, denoted by  $P_i^k$ , as the difference between the treatment and control, such that  $P_i^k = (T_i^k - C_i^k) \times 100$ . From these estimates, we calculated the total, content, and attention effect of the CoVideo. Let  $D_i^{\text{Tot}}$  denote the total effect, which is estimated by  $P_i^{\text{cov}} - P_i^{\text{no}}$ ; and let  $D_i^{\text{Att}}$  denote the attention effect, which is estimated by  $P_i^{\text{apc}} - P_i^{\text{no}}$ . These analyses are analogous to difference-in-difference analyses, which we implemented by specifying the main and

interaction terms in an ordinary least squares (OLS) regression model. The OLS equation for the  $i$ th list experiment is given as:

$$y = b_0 + b_1 \text{VideoArm} + b_2 \text{TreatList} + b_3 (\text{VideoArm} \times \text{TreatList}),$$

where  $y$  is the number of statements in the list that the participant agreed with, VideoArm indicates the  $k$ th arm, and TreatList indicates assignment to the treatment or control list. We calculated standard errors, 95% CIs, and  $P$  values (with  $\alpha=.05$ ) for linear combinations of coefficients from the OLS model.

### Informed Consent

All participants underwent a process of informed consent on the Prolific platform. The consent form explained the purpose of the study, the risks and benefits of the research, and how to contact the study investigators (or the Stanford University ethics review board). By clicking the link, participants agreed to participate in our study, and were redirected to the Gorilla platform, where additional information was given. Participants could withdraw from the study at recruitment or at any point during the experiment.

### Confidentiality

Each participant was assigned a unique, anonymized ID on Prolific and had no identifying information associated with it. We informed participants that their names could be revealed to us if they emailed the study investigators. The study investigators kept this information confidential.

### Blinding

Because Prolific handled the interaction between the study investigators and participants, the participants were completely anonymous to the study investigators. Participants self-responded to the survey questions and self-submitted their responses anonymously on the Gorilla platform. Only the

participant's unique, anonymized ID was used to manage the linking between the Prolific and Gorilla platforms. The study investigators were blinded to the group allocation [8].

### Adverse Event Reporting and Harms

No adverse events or harms were observed given the online format of the trial.

### Data Availability

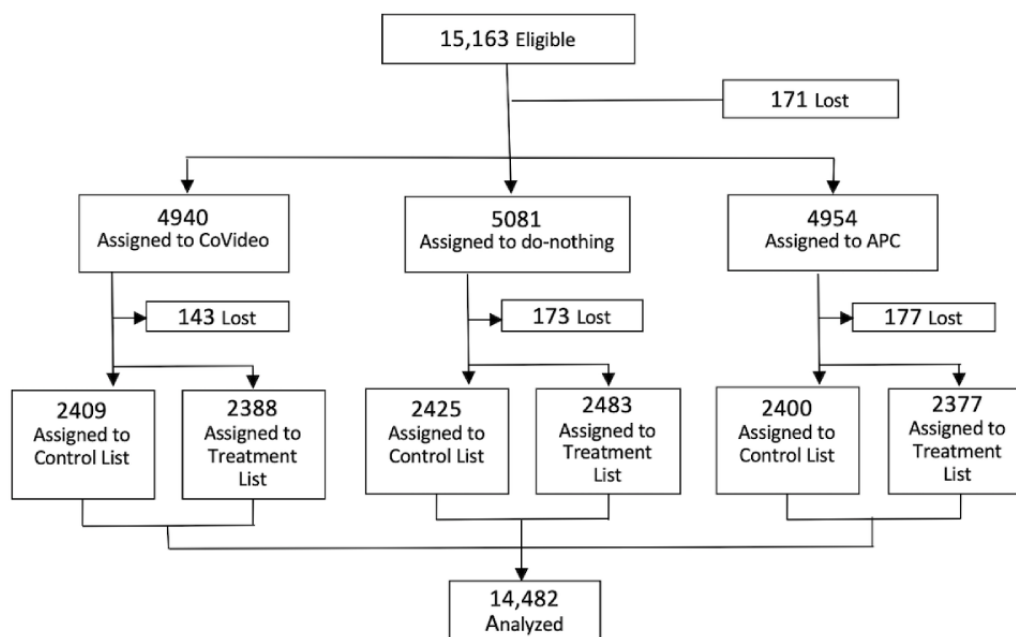
The data that support the findings of this study are available from the corresponding author upon request.

## Results

Between May 13, 2020, and June 23, 2020, 15,163 participants from the United States, Mexico, the United Kingdom, Germany, and Spain were enrolled in our RCT. Between recruitment and randomization, 171 participants were lost and 14,992 participants were randomly assigned to the CoVideo ( $n=4940$ ), APC ( $n=4954$ ), and do-nothing ( $n=5081$ ) arms (Figure 1). After randomization, another 173 (do-nothing), 177 (APC), and 143 (CoVideo) participants were lost for unknown reasons (possibly due technical issues like lost internet connection; difficulties linking to the video host, YouTube; server complications, etc). A total of 14,482 participants completed the trial and contributed data to the final analysis.

The majority of participants reported their residence in the United Kingdom ( $n=8519$ , 58.8%) or the United States ( $n=3765$ , 26%), and 84.9% ( $n=12,288$ ) of participants reported English as their first language. The sample was relatively well educated, with 81.6% ( $n=11,812$ ) having some college education or higher (bachelor's, master's/equivalent, or PhD). Table 1 shows the percentage of participants in each arm and treatment list by age, gender, country of residence, educational status, and primary language.

**Figure 1.** Trial design. After recruitment, participants were randomly assigned (1:1:1) to the CoVideo, attention placebo control (APC), or do-nothing arms. Participants in each trial arm were also randomized (1:1) to a control list (5 items; no sensitive item) or a treatment list (6 items; with 1 sensitive item) about behavioral intent toward social distancing, washing hands, cleaning dishes, cleaning kitchen surfaces, and the stockpiling of essential goods.



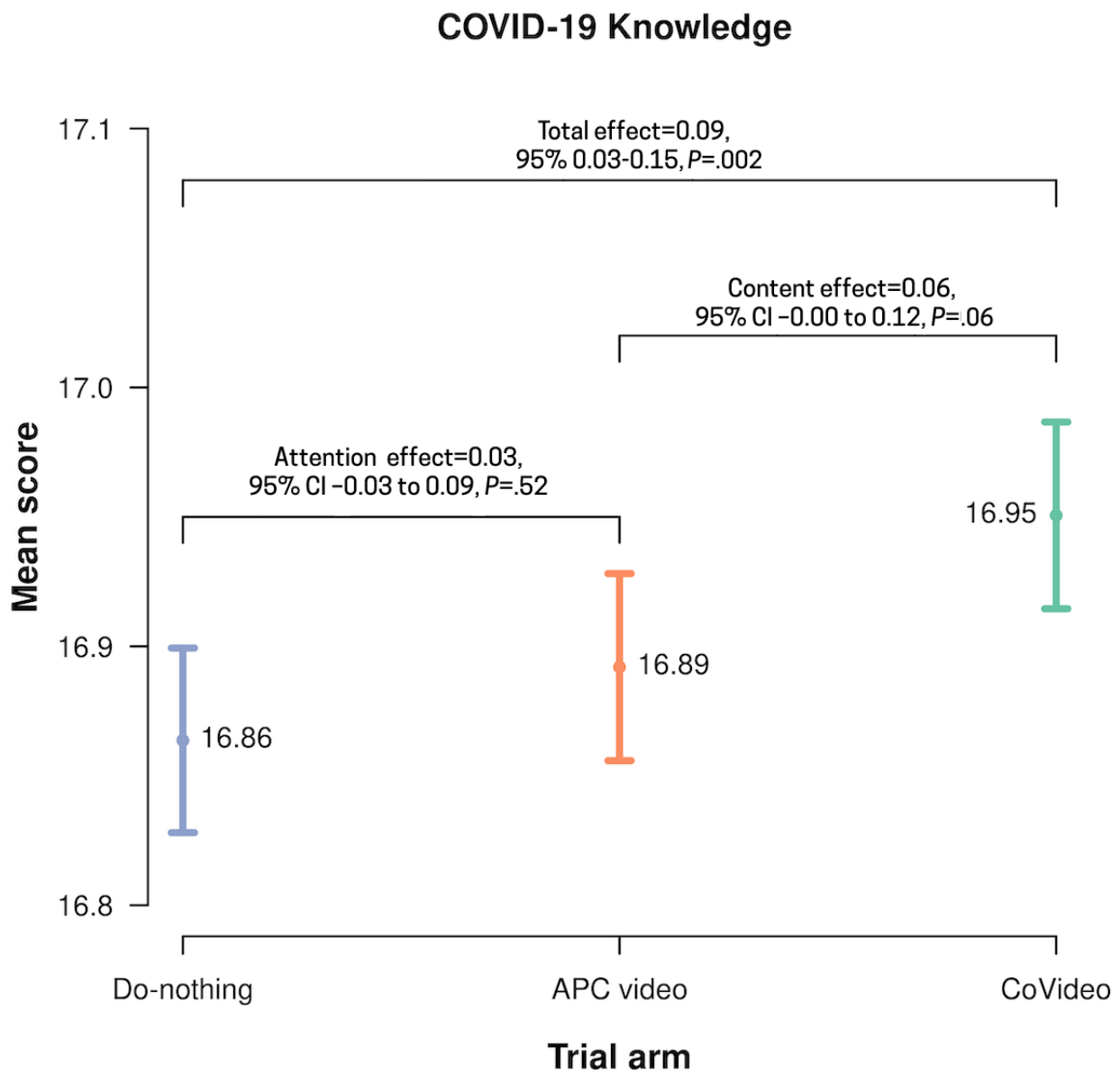
**Table 1.** Baseline demographic characteristics of participants by trial and list experiment arms (collected from 14,482 participants between May 2020 and June 2020).

Characteristic	Do-nothing		APC <sup>a</sup>		CoVideo		P value
	Control list, n (%)	Treatment list, n (%)	Control list, n (%)	Treatment list, n (%)	Control list, n (%)	Treatment list, n (%)	
<b>Age</b>							.98
18-24 years	672 (27.7)	691 (27.8)	649 (27.0)	640 (26.9)	656 (27.2)	667 (27.9)	
25-34 years	877 (36.2)	902 (36.3)	866 (36.1)	880 (37.0)	884 (36.7)	848 (35.5)	
35-44 years	475 (19.6)	502 (20.2)	484 (20.2)	456 (19.2)	479 (19.9)	470 (19.7)	
45-54 years	285 (11.8)	295 (11.9)	297 (12.4)	279 (11.7)	280 (11.6)	299 (12.5)	
55-59 years	116 (4.8)	93 (3.7)	104 (4.3)	122 (5.1)	110 (4.6)	104 (4.4)	
<b>Gender</b>							.38
Female	1316 (54.3)	1298 (52.3)	1353 (56.4)	1269 (53.4)	1306 (54.2)	1310 (54.9)	
Male	1090 (44.9)	1167 (47.0)	1037 (43.2)	1092 (45.9)	1088 (45.2)	1063 (44.5)	
Other	19 (0.8)	18 (0.7)	10 (0.4)	16 (0.7)	15 (0.6)	15 (0.6)	
<b>Country of residence</b>							>.99
Germany	118 (4.9)	135 (5.4)	132 (5.5)	116 (4.9)	130 (5.4)	124 (5.2)	
Mexico	116 (4.8)	119 (4.8)	119 (5.0)	117 (4.9)	114 (4.7)	117 (4.9)	
Spain	124 (5.1)	126 (5.1)	125 (5.2)	121 (5.1)	123 (5.1)	122 (5.1)	
United Kingdom	1418 (58.5)	1453 (58.5)	1384 (57.7)	1437 (60.5)	1429 (59.3)	1398 (58.5)	
United States	649 (26.8)	650 (26.2)	640 (26.7)	586 (24.7)	613 (25.4)	627 (26.3)	
<b>Education status</b>							.35
Primary school	66 (2.7)	90 (3.6)	66 (2.8)	61 (2.6)	83 (3.4)	87 (3.6)	
High school	360 (14.8)	377 (15.2)	383 (16.0)	360 (15.1)	364 (15.1)	373 (15.6)	
Bachelor's, some college	1551 (64.0)	1570 (63.2)	1529 (63.7)	1507 (63.4)	1526 (63.3)	1497 (62.7)	
Master's/PhD	448 (18.5)	446 (18.0)	422 (17.6)	449 (18.9)	436 (18.1)	431 (18.0)	
<b>First language</b>							>.99
German	117 (4.8)	135 (5.4)	130 (5.4)	116 (4.9)	128 (5.3)	124 (5.2)	
English	2068 (85.3)	2103 (84.7)	2026 (84.4)	2022 (85.1)	2044 (84.8)	2025 (84.8)	
Spanish (Mexico)	124 (5.1)	126 (5.1)	125 (5.2)	123 (5.2)	123 (5.1)	122 (5.1)	
Spanish	116 (4.8)	119 (4.8)	119 (5.0)	116 (4.9)	114 (4.7)	117 (4.9)	

The knowledge questionnaire had an acceptable reliability correlation coefficient of 0.65 (split-half). Overall, there was extraordinarily high attainment of COVID-19 knowledge. In the do-nothing arm, participants correctly answered 16.86 (95% CI 16.83-16.90) out of 18 items, which is a 93.7% correct response rate (Figure 2). With this high baseline score, the CoVideo could therefore only increase knowledge by a maximum of 1.14 points. Relative to the do-nothing arm, the CoVideo increased knowledge by 0.09 points (mean 16.95, 95% CI 16.91-16.99;  $P=.002$ ), which represents an increase of 7.6% (0.09/1.14) (Figure 2). The average score for the APC arm was

16.89 (95% CI 16.86-16.93), a correct response rate of 93.8%. When we removed the attention effect of the video format, the CoVideo increased overall knowledge by 0.06 points ( $P=.06$ ), which represents an increase of 5.3% (0.06/1.11). Figure 3 shows the proportion of correct responses to each of the 18 knowledge items (see also Table S1 in Multimedia Appendix 1). The highest correctly answered item ("An effective way to prevent COVID-19 spread is to wash your hands frequently with soap and water") had a correct response rate of 99.4%; most items had a >90% correct response rate.

**Figure 2.** Mean scores for the COVID-19 knowledge questions by trial arm (N=14,482). Differences between the CoVideo, attention placebo control (APC), and do-nothing arms are reported with *P* values. Total effect represents the difference in means between the CoVideo and do-nothing arms, attention effect represents the difference in means between the APC and do-nothing arms, and content effect represents the difference in means between the CoVideo and APC arms.





**Figure 3.** The proportion of correct answers for each knowledge item in the CoVideo, attention placebo control (APC), and do-nothing arms (N=14,482).

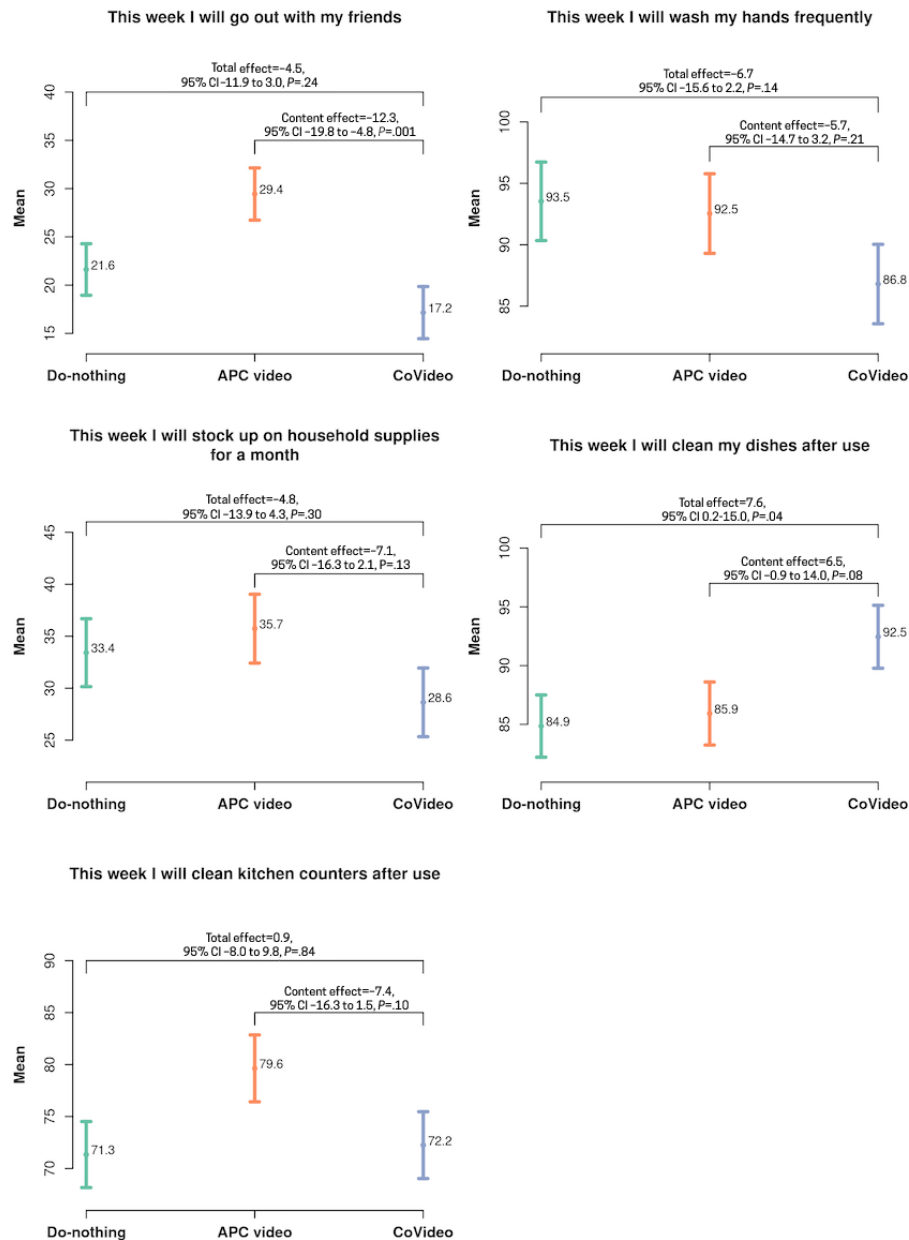


Figure S1 in Multimedia Appendix 1 shows the mean scores for the five list experiments by trial arm and list group. These mean scores were used to calculate the prevalence of behavioral intent for each preventive COVID-19 behavior, including the total and content effects with 95% CIs and *P* values (Figure 4). Scores for the treatment list are higher because the treatment list has 6 items and the control list has 5 items. For a given trial arm, the difference between the treatment and control means represents the prevalence of intent to undertake the preventive

COVID-19 behavior. For example, for the first list experiment in the CoVideo arm (“this week I will go out with friends”),  $\bar{x}_T=2.20$  is the treatment mean and  $\bar{x}_C=2.03$  is the control mean. The prevalence is then  $\frac{\bar{x}_T - \bar{x}_C}{\bar{x}_C} \times 100 = \frac{2.20 - 2.03}{2.03} \times 100 = 17.2$ , as shown in Figure 2. Similarly, the prevalence for the APC arm is  $\frac{\bar{x}_T - \bar{x}_C}{\bar{x}_C} \times 100 = \frac{2.33 - 2.03}{2.03} \times 100 = 29.4$ . For our secondary outcome, we report that participants in the CoVideo arm had lower behavioral intent to go out with friends when compared

with the APC (content effect=-0.123,  $P<.001$ ) and do-nothing (total effect=-0.045,  $P=.24$ ) arms.

**Figure 4.** The prevalence of behavioral intent for each of the five list experiments with 95% CIs. Differences between the CoVideo, attention placebo control (APC), and do-nothing arms are reported with  $P$  values. Total effect represents the difference in means between the CoVideo and do-nothing arms, attention effect represents the difference in means between the APC and do-nothing arms, and content effect represents the difference in means between the CoVideo and APC arms.



## Discussion

In this study, we tested an intervention with several innovations in global health communication that catalyzed a broad, organic global reach on social media [6,8-10]. The intervention, called the CoVideo, packaged critical health messages about COVID-19 prevention within a compelling, familiar narrative, using characters that were free of cultural identifiers and a soundtrack designed to evoke high-arousal emotions. Our results showed that baseline levels of COVID-19 prevention were high, and that the CoVideo intervention increased this prevention knowledge by another 7.6% and 5.3% relative to the do-nothing and APC arms, respectively. It was also found that the CoVideo

intervention improved behavioral intent toward COVID-19 prevention when compared with the APC and do-nothing arms.

To evaluate the effectiveness of the CoVideo on knowledge and behavioral intent toward COVID-19 prevention, we used a large, online RCT to enroll 15,163 participants from the United States, Spain, Germany, the United Kingdom, and Mexico. The results for our first endpoint showed high knowledge of COVID-19 prevention behaviors across the five countries. For the three trial arms, the average number of correct answers was nearly 17 out of 18 items, a correct response rate of approximately 94%. Moderate to high knowledge levels about COVID-19 prevention measures among the general public were also

observed earlier in Ecuador [29] and the United States [30-32]. On the other hand, a recent systematic review on knowledge, attitude, and practices toward the COVID-19 pandemic on the American continent concluded that many people have insufficient knowledge about the virus, highlighting the need to develop effective educational tools and materials on COVID-19 prevention [33]. The high baseline levels of COVID-19 knowledge in our study could be due to the delay of several weeks that occurred between the original release of the CoVideo and the launch of our online trial, as we awaited ethics approval, and designed and registered the trial. This lag likely facilitated exposure of our participants to COVID-19 prevention messages from other sources. Our results suggest, as we drift deeper into the pandemic, it may be unnecessary to spend more money on public health campaigns to improve COVID-19 prevention knowledge in the five countries from which we enrolled participants.

An important study finding was that the CoVideo improved already high levels of COVID-19 prevention knowledge. In the do-nothing and APC arms, only 1.14 and 1.11 additional correct items were needed to reach a perfect (100%) score, respectively. Our results showed that the CoVideo boosted COVID-19 prevention knowledge by another 7.6% relative to the do-nothing ceiling and by 5.3% relative to the APC ceiling. It seems plausible, therefore, that the CoVideo could significantly improve COVID-19 prevention knowledge in countries where baseline knowledge levels are currently low or moderate.

For our second endpoint, we nested a list experiment in each trial arm to evaluate the effect of the CoVideo on self-reported behavioral intent toward COVID-19 prevention. We used this experimental approach because it is likely that participants (at the time of enrollment) were already primed to give socially desirable responses to questions about COVID-19 prevention. The indirect questions (ie, how many statements do you agree with) provide protection to participants who have no behavioral intent toward COVID-19 prevention, without revealing this intention directly [27]. Our results showed that behavioral intent to go out with friends during stay-at-home recommendations and to stockpile household goods was lower in the CoVideo arm when compared with the APC and do-nothing arms, but not significantly so. We also observed that participants had higher behavioral intent to prevent COVID-19 spread by cleaning dishes after use when compared with the do-nothing arm (significantly different) and APC arm (not significantly different). Several studies have used the list experiment technique in the context of COVID-19 and found that list experiments were less favorable than simpler, traditional measurements, concluding that social desirability had no impact on the reported compliance with COVID-19 regulations [34,35]. On the contrary, other scholars have argued that the list experiment approach counters social desirability and is, therefore, less likely to introduce measurement errors presented by direct questions that measure self-reported compliance with COVID-19 guidance [36,37].

Our study is innovative in its use of both a list experiment and an APC video. Our APC video was selected to account for the possible attention effects elicited by the CoVideo intervention. The APC was designed to mimic the inactive components of

the CoVideo intervention (the effect of watching a video of the same length), while not containing the active intervention component (the content of the COVID-19-related health messaging and its delivery design) [22]. The APC, therefore, enabled us to decompose the total intervention effect, which is the difference in knowledge means between the CoVideo and do-nothing groups, into the sum of the content and attention effects. We are not aware of any study that has used this approach to isolate the active component (the content effect) of the intervention video itself. For this purpose, we advise researchers using APCs to choose their APC topic carefully, and to avoid any potential effect of the placebo content on the outcomes being studied.

Our study had several limitations. At the time of our study, no validated scale on COVID-19 knowledge prevention existed. Nevertheless, we used best practices from the survey methods field to inform the design and development of the knowledge questions [38]. Another limitation is that we could not determine if participants watched some or all of the CoVideo or APC video. Once participants were randomized to a video, they could not skip to the end or fast-forward without ending the study. However, it is possible in some cases that the participants could have been engaged in other activities while the video was playing. Because of potential noncompliance, we used an intention-to-treat analysis. One possible limitation is that high baseline knowledge likely reflects the high educational status of our online sample, with 81.6% having some college education or higher (bachelor's/equivalent, master's, PhD). Our sample was likely more educated than the general populations of the United States, the United Kingdom, Germany, Spain, and Mexico. A similar educational distribution has been reported in a recent web-based study on COVID-19 knowledge in the United States and United Kingdom [39].

Together, the findings of this study present innovative propositions for content design, dissemination, and evaluation of rapid global health communication interventions. Content designs that emphasize cultural accessibility, convey a compelling narrative, and elicit high-arousal emotions could fuel rapid dissemination across the 3.8 billion global citizens currently using social media. The wordless, animated approach also minimizes barriers traditionally associated with underlying differences in language and literacy levels. Given the massive global penetration of social media, short, animated, wordless video messages, designed to spread organically, may help public health authorities reach people where they are (ie, social media). Evaluating these interventions using online trials, APCs and list experiments can help expedite results and strengthen our efficacy evaluations. The value of such an approach becomes especially apparent during global crises in which lost weeks translate into lost lives. Accessible and compelling video health messages that lean on the shared characteristics of our global community could facilitate the spread of time-sensitive health messages. Public health authorities poised to implement these innovative health communication solutions could better support a global community facing unprecedented, shared challenges.

## Authors' Contributors

AV, MA, and VH wrote the paper. AV and CF undertook the statistical analysis. MA designed, produced, and created the

CoVideo. TB, AV, and MA designed the trial. AV, TB, MA, and MG contributed to the questionnaire development. All authors provided comments and feedback.

## Acknowledgments

TB is funded by the Alexander von Humboldt University Professor Prize.

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. All authors had full access to all the data in the study and accepted responsibility to submit for publication.

## Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary material.

[DOCX File, 83 KB - [publichealth\\_v7i7e29060\\_app1.docx](#) ]

Multimedia Appendix 2

CONSORT-EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 617 KB - [publichealth\\_v7i7e29060\\_app2.pdf](#) ]

## References

1. Sharma K, Seo S, Meng C, Rambhatla S, Liu Y. COVID-19 on Social Media: Analyzing Misinformation in Twitter Conversations. arXiv. Preprint posted online Mar 26, 2020 [FREE Full text]
2. Pennycook G, McPhetres J, Zhang Y, Lu JG, Rand DG. Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention. *Psychol Sci* 2020 Jul 30;31(7):770-780 [FREE Full text] [doi: [10.1177/0956797620939054](#)] [Medline: [32603243](#)]
3. Cuello-Garcia C, Pérez-Gaxiola G, van Amelsvoort L. Social media can have an impact on how we manage and investigate the COVID-19 pandemic. *J Clin Epidemiol* 2020 Nov 27;127:198-201 [FREE Full text] [doi: [10.1016/j.jclinepi.2020.06.028](#)] [Medline: [32603686](#)]
4. Walter N, Brooks JJ, Saucier CJ, Suresh S. Evaluating the Impact of Attempts to Correct Health Misinformation on Social Media: A Meta-Analysis. *Health Commun* 2020 Aug 06:1-9. [doi: [10.1080/10410236.2020.1794553](#)] [Medline: [32762260](#)]
5. Guadagno RE, Rempala DM, Murphy S, Okdie BM. What makes a video go viral? An analysis of emotional contagion and Internet memes. *Computers in Human Behavior* 2013 Nov;29(6):2312-2319. [doi: [10.1016/j.chb.2013.04.016](#)]
6. Adam M, Bärnighausen T, McMahon S. Design for extreme scalability: A wordless, globally scalable COVID-19 prevention animation for rapid public health communication. *J Glob Health* 2020 Jun;10(1):010343 [FREE Full text] [doi: [10.7189/jogh.10.010343](#)] [Medline: [32411360](#)]
7. Global COVID-19 Prevention. Stanford Medicine. 2020. URL: <https://www.youtube.com/watch?v=rAj38E7vrS8> [accessed 2021-06-07]
8. Vandormael A, Adam M, Greuel M, Bärnighausen T. A short, animated video to improve good COVID-19 hygiene practices: a structured summary of a study protocol for a randomized controlled trial. *Trials* 2020 Jun 03;21(1):469 [FREE Full text] [doi: [10.1186/s13063-020-04449-1](#)] [Medline: [32493460](#)]
9. Berger J, Milkman KL. What Makes Online Content Viral? *Journal of Marketing Research* 2012 Apr 01;49(2):192-205. [doi: [10.1509/jmr.10.0353](#)]
10. Cheng KK, Lam TH, Leung CC. Wearing face masks in the community during the COVID-19 pandemic: altruism and solidarity. *The Lancet* 2020 Apr. [doi: [10.1016/s0140-6736\(20\)30918-1](#)]
11. Ajzen I. The theory of planned behavior. *Organizational Behavior and Human Decision Processes* 1991 Dec;50(2):179-211. [doi: [10.1016/0749-5978\(91\)90020-T](#)]
12. Lin N, Roberts KR. Predicting and explaining behavioral intention and hand sanitizer use among US Army soldiers. *Am J Infect Control* 2017 Apr 01;45(4):396-400. [doi: [10.1016/j.ajic.2016.11.008](#)] [Medline: [27979387](#)]
13. Conner M, Sparks P. Theory of planned behaviour and the reasoned action approach. In: Conner M, Norman P, editors. *Predicting and Changing Health Behaviour: Research and Practice With Social Cognition Models*. United Kingdom: Open University Press; 2015:142-188.
14. Helweg-Larsen M, Collins B. A Social Psychological Perspective on the Role of Knowledge About AIDS in AIDS Prevention. *Curr Dir Psychol Sci* 2016 Jun 22;6(2):23-25. [doi: [10.1111/1467-8721.ep11512614](#)]

15. Ajzen I, Joyce N, Sheikh S, Cote NG. Knowledge and the Prediction of Behavior: The Role of Information Accuracy in the Theory of Planned Behavior. *Basic and Applied Social Psychology* 2011 May 16;33(2):101-117. [doi: [10.1080/01973533.2011.568834](https://doi.org/10.1080/01973533.2011.568834)]
16. THE CHOICE (Short Animated Movie). Project Better Self. 2018. URL: <https://www.youtube.com/watch?v=HEnohs6yYw> [accessed 2021-06-07]
17. German Clinical Trials Register. 2021. URL: <http://www.drks.de/> [accessed 2021-04-15]
18. Prolific. 2021. URL: <https://www.prolific.co/> [accessed 2021-04-14]
19. Gorilla. 2021. URL: <https://app.gorilla.sc/> [accessed 2021-04-14]
20. Anwyl-Irvine AL, Massonnié J, Flitton A, Kirkham N, Evershed JK. Gorilla in our midst: An online behavioral experiment builder. *Behav Res Methods* 2020 Feb 23;52(1):388-407 [FREE Full text] [doi: [10.3758/s13428-019-01237-x](https://doi.org/10.3758/s13428-019-01237-x)] [Medline: [31016684](https://pubmed.ncbi.nlm.nih.gov/31016684/)]
21. Bradley P. Bots and data quality on crowdsourcing platforms. Prolific Blog. URL: <https://blog.prolific.co/bots-and-data-quality-on-crowdsourcing-platforms/> [accessed 2021-06-07]
22. Freedland KE, Mohr DC, Davidson KW, Schwartz JE. Usual and unusual care: existing practice control groups in randomized controlled trials of behavioral interventions. *Psychosom Med* 2011 May;73(4):323-335 [FREE Full text] [doi: [10.1097/PSY.0b013e318218e1fb](https://doi.org/10.1097/PSY.0b013e318218e1fb)] [Medline: [21536837](https://pubmed.ncbi.nlm.nih.gov/21536837/)]
23. Dillard JP, Shen L. On the Nature of Reactance and its Role in Persuasive Health Communication. *Communication Monographs* 2005 Jun;72(2):144-168. [doi: [10.1080/03637750500111815](https://doi.org/10.1080/03637750500111815)]
24. Miller CH, Lane LT, Deatrick LM, Young AM, Potts KA. Psychological Reactance and Promotional Health Messages: The Effects of Controlling Language, Lexical Concreteness, and the Restoration of Freedom. *Human Comm Res* 2007 Apr;33(2):219-240 [FREE Full text] [doi: [10.1111/j.1468-2958.2007.00297.x](https://doi.org/10.1111/j.1468-2958.2007.00297.x)]
25. Richards AS, Bessarabova E, Banas JA, Bernard DR. Reducing Psychological Reactance to Health Promotion Messages: Comparing Preemptive and Postscript Mitigation Strategies. *Health Commun* 2020 Oct 27:1-9. [doi: [10.1080/10410236.2020.1839203](https://doi.org/10.1080/10410236.2020.1839203)] [Medline: [33106046](https://pubmed.ncbi.nlm.nih.gov/33106046/)]
26. Kramon E, Weghorst K. (Mis)Measuring Sensitive Attitudes with the List Experiment. *Public Opin Q* 2019;83(S1):236-263 [FREE Full text] [doi: [10.1093/poq/nfz009](https://doi.org/10.1093/poq/nfz009)]
27. Corstange D. Sensitive Questions, Truthful Answers? Modeling the List Experiment with LISTIT. *Polit anal* 2017 Jan 04;17(1):45-63 [FREE Full text] [doi: [10.1093/pan/mpn013](https://doi.org/10.1093/pan/mpn013)]
28. Glynn AN. What Can We Learn with Statistical Truth Serum? Design and Analysis of the List Experiment. *Public Opinion Quarterly* 2013;77(S1):159-172. [doi: [10.1093/poq/nfs070](https://doi.org/10.1093/poq/nfs070)]
29. Bates BR, Moncayo AL, Costales JA, Herrera-Cespedes CA, Grijalva MJ. Knowledge, Attitudes, and Practices Towards COVID-19 Among Ecuadorians During the Outbreak: An Online Cross-Sectional Survey. *J Community Health* 2020 Dec 11;45(6):1158-1167 [FREE Full text] [doi: [10.1007/s10900-020-00916-7](https://doi.org/10.1007/s10900-020-00916-7)] [Medline: [32915380](https://pubmed.ncbi.nlm.nih.gov/32915380/)]
30. O'Connor R, Opsasnick L, Benavente JY, Russell AM, Wismer G, Eifler M, et al. Knowledge and Behaviors of Adults with Underlying Health Conditions During the Onset of the COVID-19 U.S. Outbreak: The Chicago COVID-19 Comorbidities Survey. *J Community Health* 2020 Dec 11;45(6):1149-1157 [FREE Full text] [doi: [10.1007/s10900-020-00906-9](https://doi.org/10.1007/s10900-020-00906-9)] [Medline: [32780294](https://pubmed.ncbi.nlm.nih.gov/32780294/)]
31. Geana MV. Kansans in the Middle of the Pandemic: Risk Perception, Knowledge, Compliance with Preventive Measures, and Primary Sources of Information about COVID-19. *Kans J Med* 2020;13:160-164 [FREE Full text] [Medline: [32612749](https://pubmed.ncbi.nlm.nih.gov/32612749/)]
32. Quandt SA, LaMonto NJ, Mora DC, Talton JW, Laurienti PJ, Arcury TA. COVID-19 Pandemic among Latinx Farmworker and Nonfarmworker Families in North Carolina: Knowledge, Risk Perceptions, and Preventive Behaviors. *Int J Environ Res Public Health* 2020 Aug 10;17(16):5786 [FREE Full text] [doi: [10.3390/ijerph17165786](https://doi.org/10.3390/ijerph17165786)] [Medline: [32785108](https://pubmed.ncbi.nlm.nih.gov/32785108/)]
33. Sarria-Guzmán Y, Fusaro C, Bernal JE, Mosso-González C, González-Jiménez FE, Serrano-Silva N. Knowledge, Attitude and Practices (KAP) towards COVID-19 pandemic in America: A preliminary systematic review. *J Infect Dev Ctries* 2021 Jan 31;15(1):9-21 [FREE Full text] [doi: [10.3855/jidc.14388](https://doi.org/10.3855/jidc.14388)] [Medline: [33571141](https://pubmed.ncbi.nlm.nih.gov/33571141/)]
34. Larsen M, Nyrop J, Petersen MB. Do Survey Estimates of the Public's Compliance with COVID-19 Regulations Suffer from Social Desirability Bias? *JBPA* 2020 Aug 24;3(2). [doi: [10.30636/jbpa.32.164](https://doi.org/10.30636/jbpa.32.164)]
35. Selb P, Munzert S. Can we directly survey adherence to non-pharmaceutical interventions? *Surv Res Methods* 2020;14(2):205-209 [FREE Full text] [doi: [10.18148/srm/2020.v14i2.7759](https://doi.org/10.18148/srm/2020.v14i2.7759)]
36. Becher M, Stegmueller D, Brouard S, Kerrouche E. Comparative Experimental Evidence on Compliance with Social Distancing During the Covid-19 Pandemic. *SSRN Journal*. Preprint posted online Jul 24, 2020 [FREE Full text] [doi: [10.2139/ssrn.3652543](https://doi.org/10.2139/ssrn.3652543)]
37. Timmons S, McGinnity F, Belton C, Barjaková M, Lunn P. It depends on how you ask: measuring bias in population surveys of compliance with COVID-19 public health guidance. *J Epidemiol Community Health* 2020 Oct 16;75(4):387-389. [doi: [10.1136/jech-2020-215256](https://doi.org/10.1136/jech-2020-215256)] [Medline: [33067252](https://pubmed.ncbi.nlm.nih.gov/33067252/)]
38. Saris W. Design, Evaluation, and Analysis of Questionnaires for Survey Research, Second Edition. Hoboken, NJ: Wiley; 2014.

39. Geldsetzer P. Knowledge and Perceptions of COVID-19 Among the General Public in the United States and the United Kingdom: A Cross-sectional Online Survey. *Ann Intern Med* 2020 Jul 21;173(2):157-160 [[FREE Full text](#)] [doi: [10.7326/M20-0912](https://doi.org/10.7326/M20-0912)] [Medline: [32196071](https://pubmed.ncbi.nlm.nih.gov/32196071/)]

## Abbreviations

**ANOVA:** analysis of variance

**APC:** attention placebo control

**OLS:** ordinary least squares

**RCT:** randomized controlled trial

*Edited by T Sanchez; submitted 24.03.21; peer-reviewed by H Oyama, S Doan; comments to author 29.04.21; revised version received 11.05.21; accepted 17.06.21; published 27.07.21.*

*Please cite as:*

*Vandormael A, Adam M, Greuel M, Gates J, Favaretti C, Hachaturyan V, Bärnighausen T*

*The Effect of a Wordless, Animated, Social Media Video Intervention on COVID-19 Prevention: Online Randomized Controlled Trial*  
*JMIR Public Health Surveill* 2021;7(7):e29060

*URL:* <https://publichealth.jmir.org/2021/7/e29060>

*doi:* [10.2196/29060](https://doi.org/10.2196/29060)

*PMID:* [34174778](https://pubmed.ncbi.nlm.nih.gov/34174778/)

©Alain Vandormael, Maya Adam, Merlin Greuel, Jennifer Gates, Caterina Favaretti, Violetta Hachaturyan, Till Bärnighausen. Originally published in *JMIR Public Health and Surveillance* (<https://publichealth.jmir.org>), 27.07.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Public Health and Surveillance*, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

# Building an Interactive Geospatial Visualization Application for National Health Care–Associated Infection Surveillance: Development Study

Shuai Zheng<sup>1</sup>, PhD; Jonathan R Edwards<sup>1</sup>, MStat; Margaret A Dudeck<sup>1</sup>, MPH; Prachi R Patel<sup>1</sup>, MPH; Lauren Wattenmaker<sup>1</sup>, MPH; Muzna Mirza<sup>1</sup>, MD, MS; Sheri Chernetsky Tejedor<sup>1,2</sup>, MD; Kent Lemoine<sup>1</sup>, MBA; Andrea L Benin<sup>1</sup>, MD; Daniel A Pollock<sup>1</sup>, MD

<sup>1</sup>Division of Healthcare Quality Promotion, Centers for Disease Control and Prevention, Atlanta, GA, United States

<sup>2</sup>Department of Medicine, Division of Hospital Medicine, Emory University School of Medicine, Atlanta, GA, United States

**Corresponding Author:**

Shuai Zheng, PhD

Division of Healthcare Quality Promotion

Centers for Disease Control and Prevention

1600 Clifton Road

Atlanta, GA

United States

Phone: 1 4045614316

Email: [lgu3@cdc.gov](mailto:lgu3@cdc.gov)

## Abstract

**Background:** The Centers for Disease Control and Prevention's (CDC's) National Healthcare Safety Network (NHSN) is the most widely used health care–associated infection (HAI) and antimicrobial use and resistance surveillance program in the United States. Over 37,000 health care facilities participate in the program and submit a large volume of surveillance data. These data are used by the facilities themselves, the CDC, and other agencies and organizations for a variety of purposes, including infection prevention, antimicrobial stewardship, and clinical quality measurement. Among the summary metrics made available by the NHSN are standardized infection ratios, which are used to identify HAI prevention needs and measure progress at the national, regional, state, and local levels.

**Objective:** To extend the use of geospatial methods and tools to NHSN data, and in turn to promote and inspire new uses of the rendered data for analysis and prevention purposes, we developed a web-enabled system that enables integrated visualization of HAI metrics and supporting data.

**Methods:** We leveraged geocoding and visualization technologies that are readily available and in current use to develop a web-enabled system designed to support visualization and interpretation of data submitted to the NHSN from geographically dispersed sites. The server–client model–based system enables users to access the application via a web browser.

**Results:** We integrated multiple data sets into a single-page dashboard designed to enable users to navigate across different HAI event types, choose specific health care facility or geographic locations for data displays, and scale across time units within identified periods. We launched the system for internal CDC use in January 2019.

**Conclusions:** CDC NHSN statisticians, data analysts, and subject matter experts identified opportunities to extend the use of geospatial methods and tools to NHSN data and provided the impetus to develop NHSNViz. The development effort proceeded iteratively, with the developer adding or enhancing functionality and including additional data sets in a series of prototype versions, each of which incorporated user feedback. The initial production version of NHSNViz provides a new geospatial analytic resource built in accordance with CDC user requirements and extensible to additional users and uses in subsequent versions.

(*JMIR Public Health Surveill* 2021;7(7):e23528) doi:[10.2196/23528](https://doi.org/10.2196/23528)

**KEYWORDS**

data visualization; geospatial information system; health care–associated infection

## Introduction

The National Healthcare Safety Network (NHSN) is a web-based system developed and maintained by the Centers for Disease Control and Prevention (CDC), and used by health care facilities, the CDC, and other agencies and organizations for surveillance of health care–associated infections (HAIs), antimicrobial use and resistance, and other health care events and processes. NHSN data serve a variety of purposes that include infection prevention, antimicrobial stewardship, and clinical quality measurement. The CDC launched the system in 2005 and its current surveillance coverage extends to over 37,000 health care facilities nationwide. Because of state and federal HAI reporting requirements, NHSN’s geographic and health care facility coverage is extensive and the granular HAI data and summary statistics available for use are voluminous.

Health care facilities submit data to the system in accordance with NHSN’s surveillance protocols, and the facilities access the system’s analytic features to produce summary statistics for their own facility, including the cumulative attributable difference (CAD) [1]—a benchmark of the difference between observed and predicted HAIs—and the standardized infection ratio (SIR) [2]—a measure of the ratio of observed to predicted HAIs. Summary statistics enable intrafacility and interfacility comparisons for individual health care facilities and groups of facilities, and also provide a means to assess HAI prevention needs or progress at local, state, and national levels.

All data that health care facilities submit to NHSN are facility identifiable and can be linked to a street address, which enables analysis and display of NHSN data by the geographic location of each reporting facility. Most of the data that hospitals report also specify the patient care location within the hospital where the data originated (eg, intensive care unit [ICU], and, more specifically, by type of unit, such as surgical or medical ICU). For analytic purposes, NHSN data may be aggregated at a variety of geographic levels, from patient care locations within individual facilities to facility, regional, state, and national levels. NHSN analyses also may call for data to be aggregated according to predetermined time intervals, such as calendar quarter and year, or by customized data ranges.

The wide variety of granular data submitted to NHSN, the increasing volume of NHSN data and complementary data available from external sources, and the broad range of NHSN data analyses that are produced, planned, or could be made possible present an opportunity to leverage the large amount of data on a “big data” platform for NHSN data users. Among the basic requirements are systematically organizing the massive amount of data available because of NHSN’s surveillance coverage and making those data readily available to NHSN data analysts in ways that help maximize the usability and usefulness of the data for analytic purposes.

The importance of data visualization for exploratory data analysis is axiomatic across many disciplines. In epidemiology, visualization of geospatial data can reveal meaningful associations between a local site, such as a health care facility, and its surrounding environment [3]. Compared with traditional data representations, such as tables or static figures, and

particularly when the data can be characterized as large volume, high complexity, and highly disparate [4], data visualization based on a geographic information system (GIS) display offers enhanced opportunities to use them for public health purposes, such as for infectious diseases prevention and control. In practical terms, data visualization bolsters public health efforts, including outbreak detection and response activities [5,6].

At the CDC and throughout public health, data visualization is an important means of enhancing efficiencies and scalability of data analysis and dissemination [7]. Data visualization tools and platforms developed at the CDC’s enterprise and division levels typically serve domain-specific programmatic needs, such as environmental and behavioral health. These programmatic systems tend to focus on data presentations at geospatial levels—municipal, county, state, and national—located above the health care facility level. The paramount importance of facility-level data for HAI surveillance and prevention places a premium on aggregating, processing, and exhibiting data at a more granular level than is typical for other public health domains, while also enabling HAI data to be visualized at higher geospatial levels. Commercially available, ready-to-use visualization tools provide an array of important capabilities, such as rendering simple visual displays of bivariate analytic results [8]. However, the shortcomings of these systems typically include limited capacity to represent analytic results that involve multiple variables, aspects of data in parallel, or complex relationships between various data sources [9]. These limitations led the NHSN program to make the decision to develop a system for visualization of NHSN data sets—NHSNViz. Among the most important requirements identified and prioritized for the NHSN platform are the capacity to capture the logic embedded in data and to stimulate knowledge discovery. Therefore, consolidating multiple data sources and layers into the same space is a prerequisite. Additionally, the underlying connection between multiple variables of data needs to be explicitly expressed via proper user interaction design, which can streamline otherwise complicated surveillance inquiry tasks. All these requirements would be better satisfied with an ad hoc solution which is not limited by predefined visualization templates.

In summary, this project is driven by the following research questions: in terms of HAI prevention analyses, “What are the critical data sources and elements,” “What is the most intuitive and user-friendly visualization representation of each variable,” and “How could multiple facets of the various data sets be integrated into the same display systematically?”. By leveraging the expertise of infection preventionists, these topics have been explored in depth. In the following sections, we introduce our preliminary responses for these research questions as well as the methodology of our analyses process.

## Methods

### Overview of Development of NHSNViz

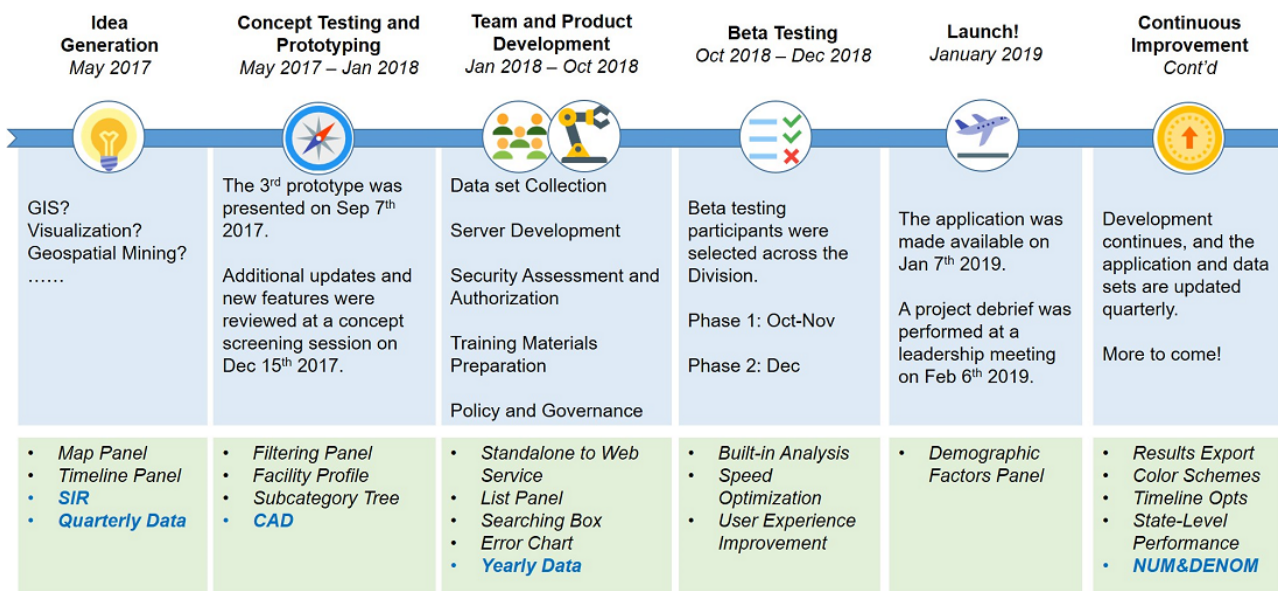
Statisticians, epidemiologists, data scientists, and other stakeholders, including policy and communications staff, spurred NHSNViz development (Figure 1). Their input was translated into technical requirements for the system’s design and



functionality, which in turn were applied in an agile software engineering life cycle characterized by a short feedback loop and sequential improvements that incorporated stakeholder feedback [10]. In particular, all requirements proposed by stakeholders are decomposed, itemized, and documented

systematically. Each feature is then prioritized based on its value and the overall business strategy. Finally, the design and implementation are reconfirmed via sketch, paper interface, or prototype to guarantee transparent communication between stakeholders and developers.

**Figure 1.** NHSNViz development timeline. CAD: cumulative attributable difference, GIS: geographic information system, NUM&DENOM: numerator and denominator, SIR: standardized infection ratio.



The project began in earnest in mid-2017 following an increase in stakeholder requests for a ready-to-use set of tools that would facilitate geospatial analysis and visualization of NHSN data. These requests prompted an initial round of brainstorming conducted by nominal group techniques, formal requirements gathering based on group interviews, rapid prototype development followed by a user observation test, approval for a proof-of-concept project, organization of a development team, and initiation of a project aimed at producing a beta test version. In October 2018, the beta test version was released and tested by over 20 individuals, drawn from diverse disciplinary backgrounds and a variety of programmatic units within CDC’s Division of Healthcare Quality Promotion. All beta testers completed training with a 1-hour online learning course. Most users easily comprehended the operations of the system. Afterward, user feedback was collected via either a survey form or a structure interview. These suggestions are primarily related to the choice of visualization representation, for example, how to represent “null” value on the map and how to highlight an increase or decrease on the timeline.

Beta test user experience and recommendations guided further development of the NHSNViz application, which was launched for initial production use in January 2019. Since then, user input has led to additional NHSNViz enhancements, specifically the addition of new analysis and visualization features and new NHSN data sets.

NHSNViz’s continuous development and release cycle provides opportunities for incremental improvements while retaining a core set of maintenance processes and system features:

- Spatial locations of individual health care facilities are gathered via geocoding, providing an aid to identifying

facilities and geographic areas where HAI prevention progress has been made or has yet to be achieved.

- Hospital characteristics, including hospital type (eg, major teaching hospital), number of hospital beds, and hospital network affiliation, enabling rapid assessments of individual hospitals with high or low SIRs.
- Quarterly and yearly aggregations of granular HAI data and summary statistics, such as numerator and denominator data, SIRs, and 95% confidence intervals are displayed. The availability of SIRs for each facility and for individual patient care locations/units within facilities promotes in-depth understanding of where HAI prevention resources are most needed.
- Demographic data acquired from external sources on income and census statistics provide additional data layers to support policy- and population-related analysis.

All these requirements eventually spawned a high-dimensional, high-complexity schema. To build such a big data application, modeling the relation among multiple data elements and constricting response time are the 2 major technical challenges to be addressed.

### Geocoding

Geocoding, which transforms physical locations such as mailing addresses to latitude- and longitude-based geographic coordinates, is an essential component for GIS applications. We used the Geocoding web service of Google [11] to obtain coordinates information for the NHSNViz project. As a part of the data collection process, the health care facility address that facilities report to NHSN in NHSN’s annual survey, including street, city, and state and zip code, was used to invoke the

Geocoding service, and then the coordinates were parsed out from the JSON format response.

### Data Format and Preparation

Indexing and merging multidimensional tabular data into a single entity helps to optimize querying performance. The JSON format provides a light-weight semistructured norm for hosting multidimensional data sources. The particular JSON format specification, GeoJSON, is designed for encoding geographic data, with flexible customization of attributes as needed. We used the Point geometry type of GeoJSON to represent health care facilities, which enables transformation of each facility's profile and measurement to subattributes under its "property" field. Coordinates collected from the geocoding process are inserted into the "geometry" field. This data transformation is completed offline as a part of a prior processing step to optimize the client-side performance.

### Interactive Map

An interactive map serves as the background canvas for a GIS system. Leaflet is a prevalent open-source JavaScript library for visualizing interactive maps. It is lightweight and yet provides the most commonly used map features such as layers, markers, and choropleth maps. Leaflet supports GeoJSON format directly. We apply this framework to NHSNViz by converting each JSON file entry (ie, a health care facility) to a "dot" marker on the map. The color of each "dot" and the corresponding legend are used to depict the level of a value, such as SIR or CAD.

### Visualization

The graphical form that users want to see when they visualize data depends on specific analytic interests (eg, basic frequency distributions, temporal trends, geographic comparisons, or other data relationships). AnyChart is a JavaScript library for rendering charts that takes into account a wide range of user needs and preferences. It supports more than 70 flexible and interactive chart types. We leveraged this versatility, in conjunction with an analysis of NHSN data users' requirements, to enable NHSNViz users to render a variety of charts:

- Stacked column chart: This is used to represent data with multiple segments. It demonstrates how each segment contributes to the overall number and displays the changes overtime.
- Mosaic chart: Also known as the heatmap chart. It represents a data matrix while each cell uses different colors to represent different values.
- Waterfall chart versus Column chart: Both represent time-based data series. However, waterfall chart highlights the change between data points using floating columns while the column chart represents absolute value with the length of columns. With markers showing the upper and lower bounds, a column chart works as an error chart and could be used to display variability details such as the confidence interval of SIR.
- Treemap chart: This renders categorical data in a hierarchical way. For comparison purpose, it displays the proportion of every subcomponent and exhibits the value of each simultaneously, in other words, 2 facets of the entities could be visualized in parallel.

Use of these chart types in NHSNViz is illustrated in the "User Interface" section.

## Results

### System Architecture

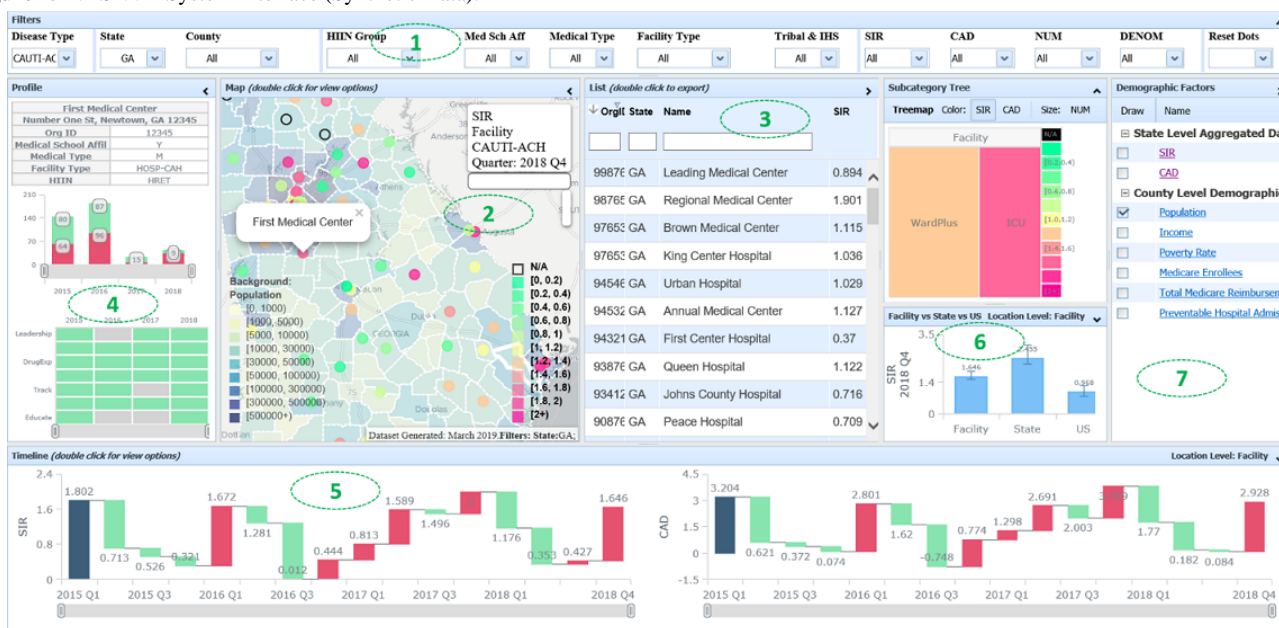
NHSNViz's underlying architecture is a classic client-server design with an HTML/JavaScript-based front end. When the server is started, all data and configuration files are loaded from a Windows shared drive with a JCIFS SMB connection. This design isolates the relatively stable source code from the regularly updated configuration parameters and data sets. To optimize NHSNViz's client-side runtime performance, all required data are downloaded in batch when the application is opened in a web browser. Next, all querying and filtering interactions with the application are completed on the client side, which minimizes the hardware and bandwidth requirements for the server. Technical specification of the server are as follows:

- OS System: Linux CentOS 7
- Server: Tomcat 9
- JavaScript Libraries: Leaflet 1.5.1, AnyChart 7.4, DHTMLX 5.1

### User Interface

The resulting user interface can represent complex analytic results and consolidate multiple variables as well as multifaceted relationships between them. NHSNViz's JavaScript-based interface leverages state-of-the-art visualization libraries, including Leaflet, AnyChart, and DHTMLX, which enables integration of NHSN data sets and the supplementary demographic data sets into a single-page dashboard (Figure 2).

Figure 2. NHSNViz System Interface (Synthetic Data).



This 1-dashboard, multiple-view design enables both experienced users and beginners to navigate easily across different geographic locations and periods. Users can conveniently browse demographic characteristics for selected geographic areas, clinical quality performance measures previously calculated and made available through NHSNViz at the health care facility, state, and national levels, and health care facility attributes such as bed size and medical school affiliation. These interactive features enable dynamic queries based on user’s operations.

A user may choose panels of interest to customize the layout of the interface for a specific task, and then navigate through different HAI types, time points, locations, and facility aspects.

To configure the Filters Panel (1), the user can focus on facilities displayed via Map (2) and List (3). Selecting either a dot on the map or an entry on the list will render detailed profile information and performance measures of the facility. One can also export the data set displayed on the List directly into a CSV file.

The Profile Panel (4) displays hospital bed sizes, including both ICU and non-ICU locations, using a stacked column chart, enabling data visualization for each subcategory segment. Additionally, the status of hospitals’ antimicrobial stewardship programs is represented by a mosaic chart, which is usually used to illustrate a categorical data matrix.

The Timeline Panel (5) at the bottom of the NHSNViz display enables access to clinical quality measures for calendar quarters or years, presented via a column chart or waterfall chart, highlighting the absolute or relative values, respectively. The Right Panel (6) displays additional performance details such as the Treemap chart representing patient care location groups within a facility that has a tree structure hierarchy and allows

drill-down; the color of each unit represents a performance measure such as SIR while the size represents either the denominator or numerator value. In addition, both the Timeline and Treemap are interactive. The user may update performance measures displayed on the map and list to certain time or clinical unit by simply clicking on the corresponding data point.

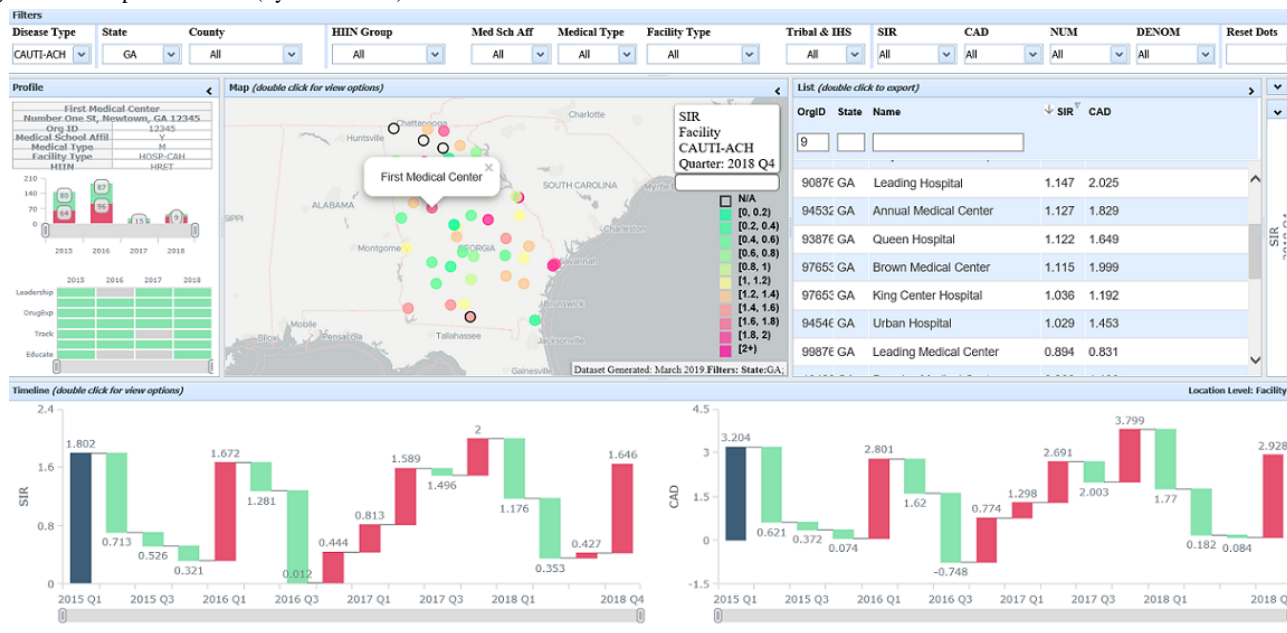
The Demographic Factors Panel (7) supports aggregated measures and external complementary data sets. These data could be rendered as a Choropleth map to provide a resource for advanced data analysis.

According to the feedback collected during the beta test, we have polished the application further to optimize the user experience. The enhancements include but are not limited to adding different color themes to ease the interpretation of data, enabling direct data and image export through the interface, as well as providing extra tooltips and messages to smooth the learning curve. All these updates have been reviewed by the proposers and well received.

### Example Use Cases

A typical NHSNViz use case is identification and prioritization of health care facilities in a state health department’s jurisdiction for purposes of initiating or intensifying targeted HAI prevention efforts (Figure 3). First, a user can select a single state and specify a range of health care facility clinical performance scores from the Filters panel, which yields a subsetting of the facilities that meet that criteria for targeted prevention. Next, the user can open the List panel, sort facilities in descending order by SIR or CAD, and inspect them individually. In addition, the user may browse the filtered results using the map, probing the distribution of high incidence or regions of high concentration.

Figure 3. Example Use Cases (Synthetic Data).



An additional use of NHSNViz is reviewing the performance score distribution of a facility over a specific period. After choosing a facility via either the Map or the List, the user may configure view options of the Timeline panel to browse SIR and CAD measures by quarter-year or year. To complement this comparison, a user can view state- and national-level performance over the same time units and period. Lastly, to enrich the context for data analysis, demographic information on the surrounding area, such as population, income, and nearby airports, could be rendered on the map as well.

### Discussion

NHSNViz achieved its primary goals of making NHSN data more readily accessible for users, enabling easy data visualization across an array of user-selected, graphical displays, and providing a means for users to gain new insights from data displayed in temporal and geographical forms. In particular, the application facilitates access to an extensive amount of high-dimensional data across multiple complex metrics which can be used for analysis and action at all geographical levels of public health. In this study, we have explored a large amount of data sources related to HAI prevention and selected the most useful ones. For each data set, we picked the proper visualization method to exhibit the pattern or highlight the trend, facilitating user’s interpretation of data. In addition, all the facets of the various data sets have been indexed in a systematic way based on the geographic and temporal information, which allows users to navigate through different facilities and periods conveniently. These experiences could serve as a reference for other visualization applications and systems addressing the same topic. HAI prevention is a joint effort made by the facilities, regions, and nation. Visualizing HAI data expedites data analyses and benchmarks the prevention advancement, therefore drives infectious preventionists toward conquering HAI.

Launched in January 2019, NHSNViz serves as a key resource for CDC data users in assessing HAI prevention and a platform

for further enhancements and extension of the system’s features to additional external users. By the end of 2019, the users of NHSNViz had invoked about 20,000 map tile requests in total. The 2-year development and maintenance experience also established guidelines for future projects. An external facing version of the application has been discussed and is in the planning stages, so the data could be shared with NHSN user groups with access control. Although adding supplementary data sets on the fly is not yet supported, additional requests for application features and data sets are addressed carefully in an agile way to enable continuous improvement. Currently the NHSN data sets used by the application are updated on a quarterly basis. Formal quality assurance process is performed by a combination of validation script and manual review. During each data preparation process, the raw data sets are scanned by a program to ensure the completeness and consistency of data. Whenever the application is populated with an updated data set, the quality assurance team will manually review and test the rendering of data to guarantee that the new data elements are properly represented.

Careful evaluation of NHSN’s different data types, in terms of their specific characteristics and their uses for various surveillance and prevention purposes, was a key consideration and a vitally important design determinant throughout the NHSNViz development process. This user-centered approach governed decisions about which graphical displays and interface features would help reveal data in the most meaningful way and optimize the application’s value [12].

A second overarching influence on the application’s design was the goal of enabling NHSNViz to serve as a resource to inspire knowledge discovery, hypothesis generation, analytic exploration, and data presentations. The 1-page dashboard design helps users to view and explore different data sources simultaneously in an integrated way [13]. Promoting ease of learning and inspiring user insights were top priorities and are reflected in the intricately engineered yet easy to operate system interface.

Further, we have identified several additional directions for future enhancements. GIS information resources can activate access to meaningful background data, prompt linkages between the health care facility's geographic location and contextual social and economic factors, and support advanced analysis such as cooccurrence and network [3]. Enriching built-in geospatial mining functionalities could significantly automate and reinforce data analysis. For example, a thematic map, which is also known as a heatmap, could be created based on density and severity of data points. Another example of this direction is the analyses related to COVID-19. Rendering the distribution of COVID-19 cases as the background information may support the assessment of pandemic influences on HAI prevention needs or efforts. Such an artificial layer could expose geographic

clustering of unexpected instances, which may be adopted to target problem areas of HAI. Another potential enhancement for analytics is to enable outbreak detection, such as the timely hotspot analyses for COVID-19. This feature requires near real-time data feed, and then spatial-temporal information would be assessed collectively using diagnostic algorithms [14]. Such a predictive early warning model may portray threats even before human's cognition, which promotes timely information dissemination for epidemiology surveillance.

In summary, as a ready-to-use application, NHSNViz enables NHSN data users to gain access to a wide array of surveillance data with complex relationships beyond what is typically available for purposes of analysis, visualization, and reporting.

## Conflicts of Interest

None declared.

## References

1. Soe MM, Gould CV, Pollock D, Edwards J. Targeted Assessment for Prevention of Healthcare-Associated Infections: A New Prioritization Metric. *Infect Control Hosp Epidemiol* 2015 Dec;36(12):1379-1384 [FREE Full text] [doi: [10.1017/ice.2015.201](https://doi.org/10.1017/ice.2015.201)] [Medline: [26310913](https://pubmed.ncbi.nlm.nih.gov/26310913/)]
2. The NHSN Standardized Infection Ratio. URL: <https://www.cdc.gov/nhsn/pdfs/ps-analysis-resources/nhsn-sir-guide.pdf> [accessed 2021-07-05]
3. Najafabadi AT. Applications of GIS in health sciences. *Shiraz E-Medical Journal* 2009 Oct;10(4):221-230 [FREE Full text]
4. Carroll LN, Au AP, Detwiler LT, Fu T, Painter IS, Abernethy NF. Visualization and analytics tools for infectious disease epidemiology: a systematic review. *J Biomed Inform* 2014 Oct;51:287-298 [FREE Full text] [doi: [10.1016/j.jbi.2014.04.006](https://doi.org/10.1016/j.jbi.2014.04.006)] [Medline: [24747356](https://pubmed.ncbi.nlm.nih.gov/24747356/)]
5. Choi J, Cho Y, Shim E, Woo H. Web-based infectious disease surveillance systems and public health perspectives: a systematic review. *BMC Public Health* 2016 Dec 08;16(1):1238 [FREE Full text] [doi: [10.1186/s12889-016-3893-0](https://doi.org/10.1186/s12889-016-3893-0)] [Medline: [27931204](https://pubmed.ncbi.nlm.nih.gov/27931204/)]
6. Shneiderman B, Plaisant C, Hesse B. Improving Healthcare with Interactive Visualization. *Computer* 2013 May;46(5):58-66. [doi: [10.1109/MC.2013.38](https://doi.org/10.1109/MC.2013.38)]
7. Richards CL, Iademarco MF, Atkinson D, Pinner RW, Yoon P, Mac Kenzie WR, et al. Advances in Public Health Surveillance and Information Dissemination at the Centers for Disease Control and Prevention. *Public Health Rep* 2017;132(4):403-410 [FREE Full text] [doi: [10.1177/0033354917709542](https://doi.org/10.1177/0033354917709542)] [Medline: [28609194](https://pubmed.ncbi.nlm.nih.gov/28609194/)]
8. Zhang L, Stoffel A, Behrisch M, Mittelstadt S, Schreck T, Pompl R, et al. Visual analytics for the big data era — A comparative review of state-of-the-art commercial systems. New York, NY: IEEE Computer Society; 2012 Presented at: 2012 IEEE Conference on Visual Analytics Science and Technology (VAST); October 14-19, 2012; Washington, DC p. 173-182. [doi: [10.1109/VAST.2012.6400554](https://doi.org/10.1109/VAST.2012.6400554)]
9. Ola O, Sedig K. Beyond simple charts: Design of visualizations for big health data. *Online J Public Health Inform* 2016;8(3):e195 [FREE Full text] [doi: [10.5210/ojphi.v8i3.7100](https://doi.org/10.5210/ojphi.v8i3.7100)] [Medline: [28210416](https://pubmed.ncbi.nlm.nih.gov/28210416/)]
10. Highsmith J, Cockburn A. Agile software development: the business of innovation. *Computer* 2001;34(9):120-127. [doi: [10.1109/2.947100](https://doi.org/10.1109/2.947100)]
11. Google Geocoding Web Service. URL: <https://developers.google.com/maps/documentation/geocoding/start> [accessed 2021-07-05]
12. Abras C, Maloney-Krichmar D, Preece J. User-centered design. In: Bainbridge W, editor. *Encyclopedia of Human-Computer Interaction*. Thousand Oaks, CA: Sage Publications; 2004:445-456.
13. Wang Baldonado MQ, Woodruff A, Kuchinsky A. Guidelines for Using Multiple Views in Information Visualization. In: *AVI '00: Proceedings of the Working Conference on Advanced Visual Interfaces*. New York, NY: Association for Computing Machinery; 2000 Presented at: *Advanced Visual Interfaces (AVI '00)*; May 2000; Palermo Italy p. 110-119. [doi: [10.1145/345513.345271](https://doi.org/10.1145/345513.345271)]
14. Watkins RE, Eagleson S, Beckett S, Garner G, Veenendaal B, Wright G, et al. Using GIS to create synthetic disease outbreaks. *BMC Med Inform Decis Mak* 2007 Feb 14;7:4 [FREE Full text] [doi: [10.1186/1472-6947-7-4](https://doi.org/10.1186/1472-6947-7-4)] [Medline: [17300714](https://pubmed.ncbi.nlm.nih.gov/17300714/)]

## Abbreviations

**CAD:** cumulative attributable difference  
**CDC:** Centers for Disease Control and Prevention  
**GIS:** geographic information system  
**HAI:** health care–associated infection  
**ICU:** intensive care unit  
**NHSN:** National Healthcare Safety Network  
**SIR:** standardized infection ratio

*Edited by T Sanchez; submitted 15.08.20; peer-reviewed by M Mars, L Simonse, Z Lu; comments to author 02.11.20; revised version received 05.12.20; accepted 10.05.21; published 30.07.21.*

*Please cite as:*

*Zheng S, Edwards JR, Dudeck MA, Patel PR, Wattenmaker L, Mirza M, Tejedor SC, Lemoine K, Benin AL, Pollock DA  
Building an Interactive Geospatial Visualization Application for National Health Care–Associated Infection Surveillance: Development Study*

*JMIR Public Health Surveill 2021;7(7):e23528*

*URL: <https://publichealth.jmir.org/2021/7/e23528>*

*doi: [10.2196/23528](https://doi.org/10.2196/23528)*

*PMID: [34328436](https://pubmed.ncbi.nlm.nih.gov/34328436/)*

©Shuai Zheng, Jonathan R Edwards, Margaret A Dudeck, Prachi R Patel, Lauren Wattenmaker, Muzna Mirza, Sheri Chernetsky Tejedor, Kent Lemoine, Andrea L Benin, Daniel A Pollock. Originally published in JMIR Public Health and Surveillance (<https://publichealth.jmir.org>), 30.07.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.

Corrigenda and Addenda

# Correction: Global Changes and Factors of Increase in Caloric/Salty Food Intake, Screen Use, and Substance Use During the Early COVID-19 Containment Phase in the General Population in France: Survey Study

Benjamin Rolland<sup>1,2</sup>, MD, PhD; Frédéric Haesebaert<sup>3</sup>, MD, PhD; Elodie Zante<sup>3</sup>, MD; Amine Benyamina<sup>4,5</sup>, MD, PhD; Julie Haesebaert<sup>6</sup>, MD, PhD; Nicolas Franck<sup>5</sup>, MD, PhD

<sup>1</sup>Service Universitaire d'Addictologie de Lyon, CH Le Vinatier, Hospices Civils de Lyon, Bron, France

<sup>2</sup>Inserm U1028, CNRS UMR5292, Université Claude Bernard Lyon 1, Bron, France

<sup>3</sup>Centre Ressource de Réhabilitation Psychosociale, CH Le Vinatier, Lyon, France

<sup>4</sup>PSYCOMADD 4872, Université Paris-Saclay, Paris, France

<sup>5</sup>CH Le Vinatier, Pôle Centre Rive Gauche, Bron, France

<sup>6</sup>Service de recherche et d'épidémiologie, Pôle de santé publique, Hospices Civils de Lyon, Lyon, France

**Corresponding Author:**

Benjamin Rolland, MD, PhD

Service Universitaire d'Addictologie de Lyon

CH Le Vinatier

Hospices Civils de Lyon

Bâtiment 502

Bron, 69500

France

Phone: 33 0437915075

Email: [benjrolland@gmail.com](mailto:benjrolland@gmail.com)

**Related Article:**

Correction of: <http://mhealth.jmir.org/2020/3/e19630/>

(*JMIR Public Health Surveill* 2021;7(7):e31906) doi:[10.2196/31906](https://doi.org/10.2196/31906)

In “Global Changes and Factors of Increase in Caloric/Salty Food Intake, Screen Use, and Substance Use During the Early COVID-19 Containment Phase in the General Population in France: Survey Study” (*JMIR Public Health Surveill* 2020;6(3):e19630), one error was noted.

In the originally published article, one paragraph in the *Results* section incorrectly referred to “alcohol use” instead of “cannabis use.” The full paragraph was published as follows:

*Finally, regarding cannabis use, 620/11,391 (5.44%) participants reported using cannabis. Among the, 263/620 (39.49%) reported that they had not changed their average daily use of alcohol, whereas 162 (24.32%) declared having moderately increased their alcohol use, 46 (6.91%) increased their alcohol use in a difficult-to-control manner, 150 (22.52%) reduced or stopped their alcohol use without craving/withdrawal, and 45 (6.76%) reduced their alcohol use with craving/withdrawal.*

This paragraph has been corrected as follows:

*Finally, regarding cannabis use, 620/11,391 (5.44%) participants reported using cannabis. Among the, 263/620 (39.49%) reported that they had not changed their average daily use of cannabis, whereas 162 (24.32%) declared having moderately increased their cannabis use, 46 (6.91%) increased their cannabis use in a difficult-to-control manner, 150 (22.52%) reduced or stopped their cannabis use without craving/withdrawal, and 45 (6.76%) reduced their cannabis use with craving/withdrawal.*

The correction will appear in the online version of the paper on the JMIR Publications website on July 20, 2021, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 08.07.21; this is a non-peer-reviewed article; accepted 08.07.21; published 20.07.21.

*Please cite as:*

Rolland B, Haesebaert F, Zante E, Benyamina A, Haesebaert J, Franck N

Correction: Global Changes and Factors of Increase in Caloric/Salty Food Intake, Screen Use, and Substance Use During the Early COVID-19 Containment Phase in the General Population in France: Survey Study

*JMIR Public Health Surveill* 2021;7(7):e31906

URL: <https://publichealth.jmir.org/2021/7/e31906>

doi: [10.2196/31906](https://doi.org/10.2196/31906)

PMID: [34283724](https://pubmed.ncbi.nlm.nih.gov/34283724/)

©Benjamin Rolland, Frédéric Haesebaert, Elodie Zante, Amine Benyamina, Julie Haesebaert, Nicolas Franck. Originally published in *JMIR Public Health and Surveillance* (<https://publichealth.jmir.org>), 20.07.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Public Health and Surveillance*, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.



Original Paper

# The Psychosocial Impacts of COVID-19 on a Sample of Australian Adults: Cross-sectional Survey and Sentiment Analysis

Jillian Ryan<sup>1,2</sup>, PhD; Hamza Sellak<sup>1,3</sup>, PhD; Emily Brindal<sup>2</sup>, PhD

<sup>1</sup>Precision Health Future Science Platform, Commonwealth Scientific and Industrial Research Organisation, Adelaide, Australia

<sup>2</sup>Nutrition and Health Program, Health and Biosecurity, Commonwealth Scientific and Industrial Research Organisation, Adelaide, Australia

<sup>3</sup>Data61, Commonwealth Scientific and Industrial Research Organisation, Melbourne, Australia

**Corresponding Author:**

Jillian Ryan, PhD

Precision Health Future Science Platform

Commonwealth Scientific and Industrial Research Organisation

13 Kintore Avenue

Adelaide, 5000

Australia

Phone: 61 8 8305 0644

Email: [jillian.ryan@csiro.au](mailto:jillian.ryan@csiro.au)

## Abstract

**Background:** The COVID-19 pandemic has had enormous impacts on people's lives, including disruptions to their normal ways of behaving, working, and interacting with others. Understanding and documenting these experiences is important to inform the ongoing response to COVID-19 and disaster preparedness efforts.

**Objective:** The aim of this study was to examine the psychosocial impacts of COVID-19 on a sample of Australian adults.

**Methods:** The data analyzed were derived from a larger cross-sectional survey of Australian adults that was administered during the month of May 2020. Participants (N=3483) were asked in which ways COVID-19 had most greatly impacted them; the responses produced a text data set containing 1 COVID-19 impact story for each participant, totaling 86,642 words. Participants also completed assessments of their sociodemographic characteristics (sex, age, financial stress), level of concern related to COVID-19, personality trait profile, and satisfaction with life. Impact stories were analyzed using sentiment analysis and compared against the Theoretical Domains Framework to determine the most frequently impacted life domains. Finally, a multinomial regression analysis, stratified by participant sex, was conducted to identify the associations of psychological and demographic socializations with sentiment toward COVID-19.

**Results:** In total, 3483 participants completed the survey, the majority of whom were female (n=2793, 80.2%). Participants' impact stories were most commonly categorized as neutral (1544/3483, 44.3%), followed by negative (1136/3483, 32.6%) and positive (802/3483, 23.1%). The most frequently impacted life domains included behavioral regulation, environmental context and resources, social influences, and emotions, suggesting that the COVID-19 pandemic was impacting these areas of participants' lives the most. Finally, the regression results suggested that for women, lower satisfaction with life and higher financial stress were associated with increased likelihood of negative, rather than positive, sentiment ( $P<.001$ ); however, the proportion of variance in the sentiment that was explained was very small (<5%).

**Conclusions:** Participant sentiment toward COVID-19 varied. High rates of neutral and negative sentiment were identified. Positive sentiment was identified but was not as common. Impacts to different areas of people's lives were identified, with a major emphasis on behavioral regulation and related domains such as social influences, environmental context and resources, and emotions. Findings may inform the development of mental health and social support resources and interventions to help alleviate the psychosocial consequences of disaster response measures.

(*JMIR Public Health Surveill* 2021;7(7):e29213) doi:[10.2196/29213](https://doi.org/10.2196/29213)

**KEYWORDS**

COVID-19; natural language processing; behavioural science; NLP; behavior; impact; Australia; community; sentiment; data set; machine learning

## Introduction

The COVID-19 pandemic is among the most disruptive and significant public health crises in recent human history. At the time of writing, nearly 3 million people had died from COVID-19 [1], while containment efforts had caused enormous upheaval to lives and livelihoods worldwide. In countries such as Australia, where viral spread has been relatively well controlled, the largest effects of the pandemic have been related not to the virus itself and its high mortality and infection rates, but to the containment measures needed to control outbreaks and how these measures have affected the community. People's lives have changed dramatically, and adapting to the next phase of the pandemic and postpandemic life will require adjustment to a "new normal." This means accepting a change of lifestyle in which alternative working arrangements, social distancing, and travel restrictions become the norm [2]. Adapting behavior can be challenging at the best of times, and in this instance, it is further complicated by feelings of loss related to letting go of previous ways of life [3]. The process of grief has been well documented, including nonlinear stages of disbelief, yearning, anger, depression, and acceptance [4]; however, more needs to be understood about the process and implications of the psychosocial impacts of COVID-19.

The virus that causes COVID-19, SARS-CoV-2, was first detected in Australia in January 2020. The Australian government's approach to the mitigation of disease spread was strict and proactive, including significant restrictions on travel and public gatherings, quarantine protocols, and social distancing measures [5]. At the core of Australia's public health response was a national shutdown that occurred during the months of March, April, and May 2020. This shutdown had the objectives of (1) delaying the impending epidemic to allow for resource planning to occur and (2) "flattening the curve" to reduce case numbers in Australia by minimizing the opportunity for disease transmission [5]. Shutdown measures consisted of a broad suite of restrictions, including the closure of gyms, pools, cinemas, and other health and entertainment facilities, a change to remote learning for universities and other higher education venues, restrictions on freedom to leave the house for nonessential reasons, restrictions on visitation to residential aged care facilities and hospitals, hygiene and social distancing measures, and the cancellation of events such as Australian and New Zealand Army Corp Day celebrations, arts, and sporting events [6].

The Australian government's response to the pandemic was particularly strong compared to that of governments of other countries [7], and containment measures were enforced by law. In Western Australia, where the conditions are most stringent, individuals can face fines of up to Aus \$50,000 (approximately US \$38,000) or 12 months of imprisonment for breaching COVID mitigation rules. In other states, fines of between Aus \$200 and \$4000 (approximately US \$152 and \$3052) were issued [8]. Economic stimulus was provided to balance the effects of the containment measures. Stimulus efforts included the JobSeeker and JobKeeper programs, which effectively doubled income support payments and supplemented employees' wages for businesses that were affected [9]. An additional Aus

\$1.1 billion (approximately US \$841,313,000) was spent expanding mental health and telehealth services, increasing domestic violence services, and increasing food relief services [9].

Multicountry analyses suggest that Australia's strict and proactive COVID-19 containment measures contributed to the delay and prevention of large infection outbreaks and to overall optimal outcomes [7]. However, they also imposed substantial limitations on Australians' personal freedoms, livelihoods, and ability to maintain social connectivity. International travel effectively ceased due to the pandemic, with the number of citizens travelling outside of Australia plummeting to just 5050 individuals in May 2020, a >99% drop from the 798,700 departures recorded in May 2019 [10]. Similarly, domestic travel dropped from around 5 million trips in May 2019 to about 200,000 in May 2020 [11]. The shutdown also transformed ways of working, with only workers who were providing essential services (eg, emergency services, utilities), food and groceries, and health care permitted to travel to work. Australian research has documented significant declines in mental health, psychological distress, and contagion anxiety during the pandemic [12]; negative shifts in behavioral indicators of health, including physical activity, diet, sleep, alcohol consumption, and tobacco smoking [13]; and increases in food insecurity, with the strongest impacts felt by Australians with disabilities, those in rural areas, and others who face economic disadvantage or vulnerability [14,15]. COVID-19 is therefore a social crisis in addition to a public health crisis, and a greater understanding of the nature and extent of its impacts is needed [16].

Natural language processing offers a promising strategy to help understand the nature and impacts of complex health problems. Natural language processing uses intelligent computer algorithms to detect patterns and themes in unstructured data sets commonly containing text data [17]. A key advantage of this approach is the ability to automatically monitor or rapidly analyze unstructured data to identify and comprehend unanticipated or unforeseen health- and medical-related needs in the community (sometimes called *inforeveillance*), monitor community sentiment toward health, and identify key geopolitical and psychosocial drivers for health-related behaviors [18]. Natural language processing can aid understanding of disease transmission and epidemic and pandemic trajectories by shedding light on community attitudes, behaviors, and experiences related to specific diseases. In particular, a substantial body of research has used sentiment analysis, which seeks to detect positive, negative, and neutral sentiments within text data, to understand aspects of the COVID-19 pandemic and in the context of other infectious diseases [19]. Broadly speaking, the aim of sentiment analysis is to classify words, sentences, or other units of text data according to the positive or negative polarity of that text [20]. These sentiment scores can then be interpreted in terms of the extent to which participant sentiment is favorable toward a focal issue, which can then inform more relevant and appealing messaging or identify intervention needs.

Natural language processing has a key advantage of being able to process large qualitative data sets rapidly and with greater objectivity than manual analysis. For example, it can help to

identify community emotions toward different COVID-19 containment measures, identify geographical or sociodemographic correlates of vaccine hesitancy, and even detect outbreaks based on rapid analysis of social media data. In one example, researchers used machine learning to collect and analyze 86 million tweets published on the web-based social media platform Twitter in the United States to understand public sentiment toward COVID-19 and how it changed as the pandemic continued [21]. The results indicated an increasing volume of tweets over time and overall negative sentiment on the Twitter platform throughout the year.

Less explored is the potential application of natural language processing and sentiment analysis of research data that include populations representing a variety of community groups and that can be linked with demographic or other characteristics to better inform data-driven public health decision-making. Few studies have used this approach with purpose-collected research data sets and with research samples other than Twitter users [19]. Further efforts that use robust natural language processing techniques to identify and document both the expected (eg, shifts in health and travel behaviors, income streams) and unexpected impacts of containment measures are needed, as observational inferences are critical to optimizing public health response during the COVID-19 pandemic and in preparation for future public health and natural disasters. The purpose of this study is to use natural language processing to examine the breadth and nature of the impacts of the national shutdown on Australians during May 2020. More specifically, the study aimed to assess community sentiment toward the impacts of COVID-19, explore the nature of the impacts of COVID-19 and which life domains are most commonly affected, and identify psychological (personality traits, COVID-19-related concerns, and satisfaction with life) and sociodemographic characteristics (sex, age, and financial stress) that predict sentiment toward the impacts of the COVID-19 pandemic.

## Methods

The data analyzed in this study were collected as part of a larger cross-sectional survey of Australian adults that was conducted in the month of May 2020. Full details of the study methodology and primary findings have been published elsewhere [22]. The study received ethics approval from the Commonwealth Scientific and Industrial Research Organization Human Research Ethics Low-Risk Committee (LR2020/026), and all participants provided informed consent prior to completing the survey.

### Procedure

Convenience sampling methods were used in which a web-based survey was distributed to an email list of participants in a health and well-being program who had consented to being contacted about future studies or other tools relating to health and diet. In general, the list of contacts contained a higher proportion of women, and the members were slightly older and more educated in comparison with the general Australian population. Each member on this list was sent an email that included an invitation to participate and a link to the web-based survey and informed consent process.

## Materials

The web-based survey assessed the participants' sociodemographic characteristics (sex and age), COVID-19 impacts, personality traits, and subjective well-being. COVID-19 impact stories, which formed the basis for the machine learning analysis, were collected via a single open-ended survey item. This item asked participants to finish the following sentence: "The COVID-19 outbreak has most greatly impacted..."

Participants' psychosocial and demographic characteristics were also captured via self-report survey items. COVID-19-related financial stress was captured by a single item that asked participants to consider any financial stress they might have experienced during the COVID-19 pandemic and rate the extent to which they were unsure how they would pay upcoming bills on time. Responses were recorded on a 7-point Likert-type scale and ranged from 1, strongly disagree, to 7, strongly agree. The 14-item COVID-19 Concerns Scale was included to assess the participants' concerns related to COVID-19. This scale asks respondents to indicate the extent to which they are concerned about different aspects of the pandemic, such as becoming infected with COVID-19, losing a job, or isolation from friends and family members [23]. Responses are recorded on a 4-point scale ranging from 1, to a great extent, to 4, not at all, with "I don't know" also included as an option. The participants' personality traits were captured by the validated Big Five Inventory-2-S [24], which assesses the "big five" personality traits: openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism. This scale consists of 30 items, with responses captured on a 5-point Likert scale ranging from 1, strongly disagree, to 5, strongly agree. Finally, life satisfaction was assessed by the Satisfaction With Life Scale [25], which consists of 5 items that are measured on a 7-point Likert scale ranging from 1, strongly disagree, to 7, strongly agree. The scale is a validated measure that captures global life satisfaction and has adequate psychometric properties [25].

## Data Analysis

The survey was attempted by 4313 individuals; of these, 3483 answered the COVID-19 impact open-ended question (80.3% completion rate) required for the current study. Qualitative data were analyzed to detect sentiment and key themes using advanced natural language processing tools. We used the Stanford CoreNLP sentiment annotator [26], a machine learning model that uses recursive neural networks to perform sentiment analysis and classify input text on a 5-point scale of very negative, negative, neutral, positive, and very positive (higher scores indicate more positive sentiment). In addition, we analyzed data using the Valence Aware Dictionary and Sentiment Reasoner (VADER), a lexicon and rule-based sentiment analysis tool that produces a unidimensional measure of sentiment for a given sentence that reflects a summed score for each word in that text [27]. VADER and Stanford CoreNLP rely on distinct mechanisms that enable each tool to provide a different perspective on the data, with a core difference being that Stanford CoreNLP uses discrete but more detailed sentiment categories (5 sentiment levels) and VADER uses continuous numerical sentiment score values (continuous values that are more suitable to plotting) [28]. VADER produces summed

sentiment scores that are normalized between -1 (most extreme negative) and +1 (most extreme positive). We used VADER sentiment scores to produce a plot depicting the distribution of sentiment and Stanford CoreNLP scores in subsequent analysis as a way to categorize the participants' impact scores from very positive to very negative. The sentiment analysis results were manually inspected, and overall, the classifications appeared to be consistent with our expectations (see [Multimedia Appendix 1](#)).

Following text preprocessing, we automatically extracted the most frequent terms to identify themes in the data set, consistent with previous research [20]. The most frequently occurring words impacted the themes used in further analysis, as content themes and were manually classified according to the Theoretical Domains Framework [29]. This model acknowledges 14 life domains that are relevant to understanding and changing human behavior: knowledge, skills, social/professional role and identity, beliefs about capabilities, optimism, beliefs about consequences, reinforcement, intentions, goals, memory, attention and decision processes, environmental context and resources, social influences, emotions, and behavioral regulation. The Theoretical Domains Framework comprises established standardized terminology and organizing constructs for use in exploration, prediction, and intervention of human behavior. This framework is of particular value as a comprehensive framework that allows for the contextualization of different personal, interpersonal, and environmental influences on behavior and mood. COVID-19 impacts are therefore presented as frequencies with quotes provided as examples.

To test associations between participant characteristics and sentiment scores, an ordinal regression procedure was commenced. However, during the process, it became evident that our data did not meet the assumption of proportional odds. Therefore, multinomial logistic regression was undertaken to examine whether age, financial stress, life satisfaction, and personality traits (openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism) were predictive of sentiment score. Analyses were stratified by gender. Because the very positive and very negative cells comprised fewer than 25 cases each, we merged them with the positive and negative cells, respectively, to arrive at 3 categories of the dependent variable: positive, neutral, and negative sentiment. The data

were first examined to check that they met the model assumptions of an absence of multicollinearity based on variance inflation factors and tolerance, linearity to the logit, and absence of outliers [30,31]. These assumptions were met, and the multinomial regression could be conducted.

Variables were selected for inclusion in the model based on theoretical background knowledge [32]. There is a large body of research that demonstrates significant associations between individual differences in factors such as sociodemographic characteristics, personality trait profiles, and subjective well-being with psychosocial well-being outcomes [33]. Furthermore, insights into these associations have potential health communication applications, as they can inform the development of tailored public health interventions or identify in-need target audience segments within larger populations [34]. As this was an exploratory analysis with a primary aim to deduce novel insights from the data rather than to test a prespecified hypothesis, we included all variables of interest in the final regression model to determine each variable's strength of association relative to the other variables included in the model. Continuous and ordinal variables of age, openness, conscientiousness, extraversion, agreeableness, neuroticism, satisfaction with life, and financial stress were entered as covariates [30]. Positive sentiment scores were set as the reference category such that the results reflect the likelihood of obtaining a neutral or negative sentiment score in comparison with a positive score. Statistical significance was deemed to be reached when  $\alpha < .05$ , and standardized beta weights with 95% confidence intervals have been reported.

## Results

The majority of the participants were female (see [Table 1](#)). The participants' mean age was 57.1 years (SD 12.2), and participants reported an average COVID-19-related financial stress score of 1.9 (SD 1.5) out of 7, with lower scores equating to less financial distress.

The qualitative data set contained 3483 COVID-19 impact stories, comprising a total of 86,642 words. The average length of the impact stories was 25 words; however, men provided shorter stories on average (mean 18.4, SD 24.8) compared to women (mean 26.4, SD 27.3).

**Table 1.** Participant characteristics among a sample of Australian adults in a cross-sectional study investigating the psychosocial impacts of COVID-19 in May 2020 (N=3483). For all variables, a higher score indicates a higher level of that variable.

Characteristic	Value, mean (SD)		
	Male participants (n=417)	Female participants (n=2793)	Overall sample (N=3483)
Age	61.8 (12.4)	56.2 (12.0)	57.1 (12.2)
Sentiment <sup>a</sup>	2.9 (0.7)	2.9 (0.8)	2.9 (0.8)
Financial stress <sup>b</sup>	1.9 (1.5)	1.9 (1.5)	1.9 (1.5)
Satisfaction with life <sup>b</sup>	4.4 (1.4)	4.1 (1.5)	4.2 (1.5)
COVID-19-related concerns <sup>c</sup>	2.5 (0.5)	2.6 (0.5)	2.6 (0.5)
Openness to experience <sup>a</sup>	3.5 (0.6)	3.6 (0.7)	3.6 (0.7)
Conscientiousness <sup>a</sup>	3.7 (0.7)	3.8 (0.7)	3.8 (0.7)
Extraversion <sup>a</sup>	3.1 (0.7)	3.0 (0.7)	3.1 (0.7)
Agreeableness <sup>a</sup>	3.7 (0.6)	4.0 (0.6)	4.0 (0.6)
Neuroticism <sup>a</sup>	2.5 (0.8)	2.7 (0.9)	2.7 (0.9)

<sup>a</sup>Scores range from 1 to 5.

<sup>b</sup>Scores range from 1 to 7.

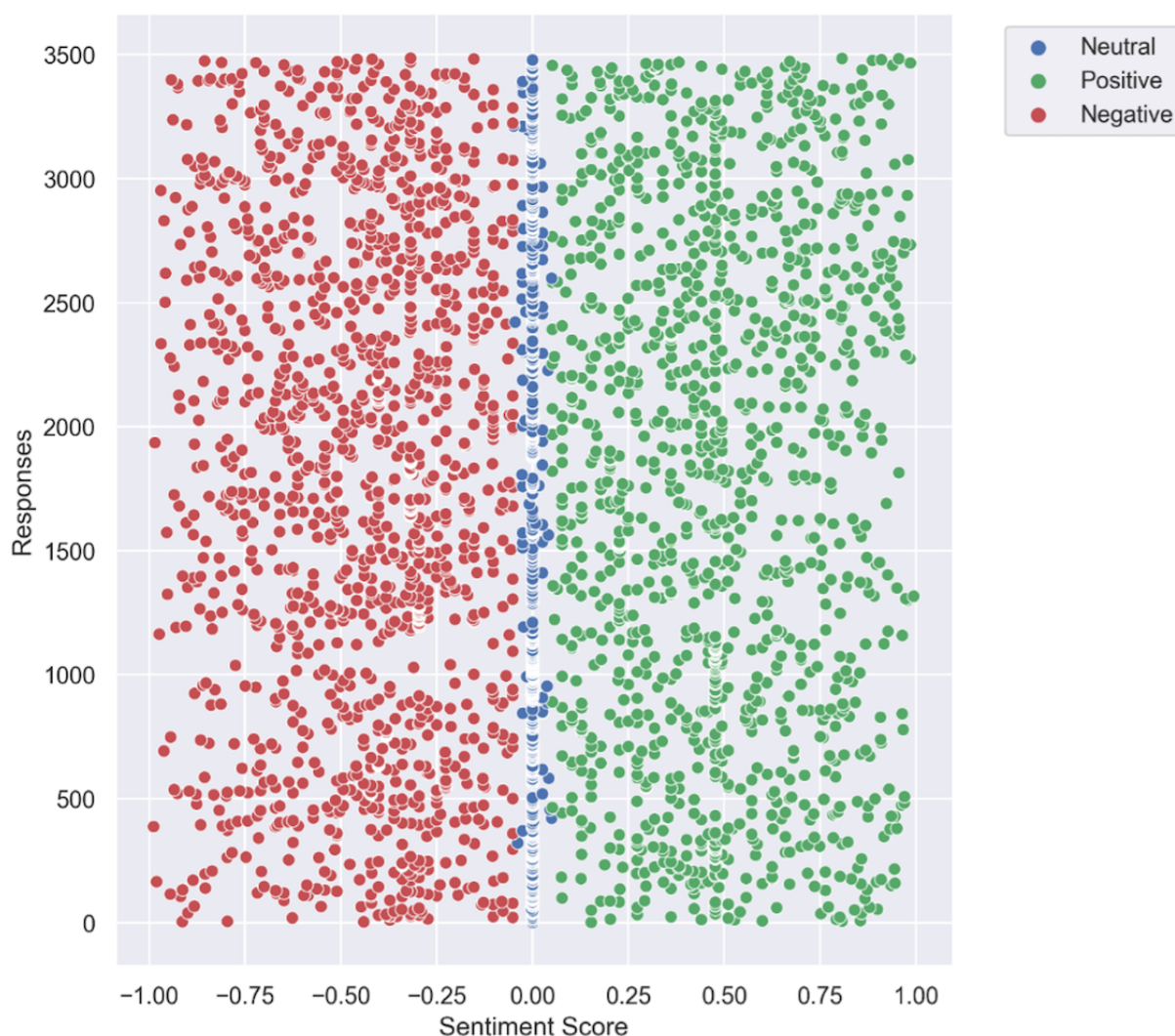
<sup>c</sup>Scores range from 1 to 4.

### Sentiment Toward COVID-19 Impacts

The average Stanford CoreNLP sentiment score fell into the neutral category (mean 2.91, SD 0.76). Neutral classifications were most common, representing 44.3% of impact stories (n=1544/3483). There was also a relatively balanced prevalence of negative (n=1136/3483, 32.6%) and positive (n=802/3483,

23.1%) sentiment scores. Approximately 1% (n=38/3483) of all COVID-19 impact stories were categorized at the extreme ends of the sentiment distribution (very negative or very positive). The distribution of the VADER sentiment scores is depicted in [Figure 1](#). Representative quotes from the data that exemplify each of the sentiment score categories are provided in [Multimedia Appendix 1](#).

**Figure 1.** Distribution of standardized Valence Aware Dictionary and Sentiment Reasoner (VADER) sentiment scores in COVID-19 impact stories reported by a sample of Australian adults during May 2020.



**Nature of COVID-19 Impacts Across Different Life Domains Based on Word Frequency Analysis**

The numbers and proportions of most frequently appearing words and their respective life domains are displayed in [Table 2](#)

2. The impacts were most commonly located within the four life domains of behavioral regulation, environmental context and resources, emotion, and social influences.

**Table 2.** Categorization of most frequently appearing words by theoretical life domain among impact stories from Australian adults in a cross-sectional study investigating the psychosocial impacts of COVID-19 in May 2020 (N=3483). Note: the theoretical life domains of skills, optimism, reinforcement, intentions, goals, and memory, attention, and decision processes were not identified in this data set.

Theoretical life domain	Top 50 words categorized into each domain, n (%)
Behavioral regulation	16 (32)
Environmental context and resources	13 (26)
Social influences	6 (12)
Emotion	6 (12)
Social/professional role and identity	4 (8)
Beliefs about consequences	3 (6)
Knowledge	1 (2)
Beliefs about capabilities	1 (2)

### **Behavioral Regulation**

Impacts occurred primarily in the life domain of behavioral regulation, which considers ability and changes in ability to perform different behaviors (see [Table 2](#)), including health behaviors (eg, physical activity, eating vegetables), social behaviors (eg, spending time with friends and family), and risk behaviors (eg, drinking alcohol). Key themes within this category related to how the COVID-19 pandemic influenced participants' abilities to fulfil their normal routines, with a focus on health and social behaviors, visiting others and travel, and use-of-time impacts. The closure of gyms and other health facilities was a particularly prevalent theme impacting participants' ability to maintain physical and social health behaviors. This impact was compounded among respondents who reported special needs for supervised exercise, pools, or other equipment that was not easily substitutable with virtual or home-based exercise, including older adults and people with disabilities, injuries, and postoperative rehabilitation needs. More generally, the health and fitness impacts of gym closures were closely intertwined with loss of social opportunities and isolation. Restrictions on travel was also a key theme within this life domain; it was related to the ability to travel for work and to visit and support friends and family. Positive impacts related to increased flexibility for employees, such as ability to work from home and reduced time commuting. Example quotes from the data are provided below.

*We have been unable to continue activities that are important to our health and fitness because of the closure of heated pools, massage clinics and other facilities for exercise suitable for aged persons with joint disabilities.* [Female, 72 years old, 3/5 (neutral) sentiment score]

*It's massively affected my relationship with food. The relationship was already bad, but social isolation and gyms closing were just the LAST thing I needed. it couldn't have come at a worse time.* [Female, 21 years old, 2/5 (negative) sentiment score]

*The COVID-19 outbreak has meant I am now working from home, which has been incredible for my mental health and wellbeing. I normally commute 2 hours per day, 3 times a week, and this was really taxing on me physically and emotionally. Working from home is so much more peaceful and stress-free. I am also the kind of person who benefits from working solo, so not being in the office environment has also been really positive for me. I was unable to get this opportunity before COVID-19, and only had 2 days a week approved to work closer to home. I now feel more confident to ask my employer for more days working from home.* [Female, 34 years old, 4/5 (positive) sentiment score]

### **Environmental Context and Resources**

The environmental context and resources life domain considers external factors that facilitate or inhibit the development and expression of abilities, independence, social competence, and adaptive behavior. Common impacts falling under this theme are related to changes in the home environment, lost

opportunities to have contact with or care for loved ones and friends, and difficulties acquiring groceries. The shutdown restrictions served as a major source of psychological distress and boredom for participants, particularly in relation to a lack of variation in environment, with work, socializing, and education of children now occurring primarily within the home location. Environmental constraints were also linked to behavioral outcomes (eg, changes in ability to eat healthily and exercise) as well as emotional outcomes such as stress, boredom, and anxiety. Some positive impacts were also noted as a result of fewer social obligations and reduced face-to-face contact.

*Mental health has been an issue as I feel we at home are just getting through the weekly routine of trying to get some work done while helping our son to home school and then planning the next grocery shop. No real fun as we live in a dull routine.* [Male, 43 years old, 2/5 (negative) sentiment score]

*My husband has dementia and is in a Nursing Home which is in total lockdown and I have not been able to have any contact with him since March 20th. I am devastated by this. Prior to lockdown I visited him daily. He continues to ask for me each morning and has no understanding as to why I no longer visit him.* [Female, 77 years old, 2/5 (negative) sentiment score]

*I enjoy people social distancing from me when I am out and like that shops are clean. I prefer catching up with people online rather than in person.* [Female, 41 years old, 4/5 (positive) sentiment score]

### **Emotion**

The third most commonly impacted life domain was emotion, which relates to individuals' cognitive and psychological reactions and ways of coping with events and circumstances. Salient themes within this domain included stress, mental health, and anxiety, with the descriptors of "loss" and "hard" featuring prominently. Positive impacts included increased feelings of neighborhood cohesion and reductions in stress related to shorter commutes and greater hygiene and social distancing. Positive impacts tended to relate to reduced life "busyness" and a more relaxed pace.

*My aunty past [sic] away. She was 84 and in a nursing home. Even though we (and the nursing home) knew she would pass away soon we were not allowed to visit in the last 2 weeks as the nursing home was in lock down. The funeral was limited to 10 people and we had to social distance. It was not the way we wanted to celebrate her life. I have had a 20% pay cut which is nothing compared to the sadness I feel with regard to not being able to see my aunty and say goodbye.* [Female, 47 years old, 2/5 (negative) sentiment score]

*It has greatly reduced the traffic in my street. So many more neighbours are out and about, riding their bikes, walking with family, working in their gardens, with time to stop and chat (properly socially distanced of course). We all make way for others on the footpaths, and say hello or wave as we pass. I feel we have a*

*more cohesive community spirit in our neighbourhood.* [Female, 68 years old, 4/5 (positive) sentiment score]

*Cancelled holidays. Limited socialisation with friends. Worry about maintaining health and well-being (i.e. not getting infected with Covid-19).* [Male, 60 years old, 2/5 (negative) sentiment score]

*Life has been less hectic and more relaxed. I have enjoyed not having any outside of my house responsibilities.* [Female, 61 years old, 3/5 (neutral) sentiment score]

### **Social Influences**

The fourth life domain that was commonly impacted for the participants was social influences, referring to how people's relationships with others affect their own thoughts, feelings, and behaviors. Themes within this category primarily related to friends and family members, particularly children.

*Life is unrecognisable. Husband stood down from Qantas, children bored and under stimulated by home schooling, children's sport was a big part of all our lives. Depressed.* [Female, 47 years old, sentiment score 2/5 (negative)]

*The pandemic had had more effect on my children and husband than me. My five year old has started experiencing anxiety in the night and sometimes cannot sleep. My other child has also become much more emotional.* [Female, 42 years old, 2/5 (negative) sentiment score]

*With children aged 4 & 2, I have enjoyed being home with them. No pressure for play dates, outings etc.* [Female, 37 years old, 4/5 (positive) sentiment score]

### **Psychosocial and Demographic Characteristics That Predict Sentiment Toward COVID-19**

Associations between participant characteristics and COVID-19 sentiment were analyzed using regression models (stratified by sex). Model 1, which tested the association between female participants' characteristics and sentiment scores, was statistically significant ( $\chi^2_{18}=75.8, P<.001$ ) (see Table 3). The model explained 3.2% (Nagelkerke  $R^2$ ) of the variance in sentiment (very small effect size). Of the variables included in the model, satisfaction with life, openness to experience, and financial stress were statistically significant predictors of sentiment. More specifically, for each 1-unit increase in satisfaction with life, there was a decrease in the odds of having a negative sentiment score rather than a positive sentiment score. Furthermore, for each 1-unit increase in openness to experience, the risk of reporting a negative relative to a positive sentiment score decreased by a factor of 0.832. Finally, for each 1-unit increase in financial stress, the odds of having a neutral rather than positive sentiment score increased by 1.128, and the odds of have a negative rather than positive sentiment score increased by 1.194.

Model 2, evaluating the association between participant characteristics and sentiment scores for men, was not statistically significant ( $\chi^2_{18}=19.5, P=.36$ ).



**Table 3.** Multinomial logistic regression identifying significant associations with COVID-19 sentiment among a sample of Australian adults reported in May 2020.

Characteristic	Positive sentiment, mean (SD)	Neutral sentiment				Negative sentiment			
		Mean (SD)	$\beta$	95% CI (lower)	95% CI (upper)	Mean (SD)	$\beta$	95% CI (lower)	95% CI (upper)
<b>Model 1 (female participants)</b>									
Age	56.8 (11.7)	56.1 (12.2)	0.998	0.990	1.007	56.0 (11.9)	0.998	0.989	1.007
Financial stress	1.65 (1.2)	1.9 (1.5)	<i>1.128</i> <sup>a</sup>	1.039	1.224	2.1 (1.6)	<i>1.194</i>	1.100	1.297
Satisfaction with life	3.2 (1.1)	3.0 (1.0)	0.923	0.832	1.024	2.9 (1.1)	<i>0.832</i>	0.747	0.927
COVID-19-related concerns	2.6 (0.5)	2.6 (0.5)	0.931	0.753	1.152	2.6 (0.5)	1.019	0.815	1.274
Openness to experience	3.6 (0.7)	3.5 (0.7)	<i>0.823</i>	0.710	0.953	3.7 (0.7)	1.039	0.889	1.213
Conscientiousness	3.9 (0.7)	3.8 (0.7)	0.977	0.838	1.139	3.8 (0.7)	0.939	0.800	1.103
Extraversion	3.1 (0.7)	3.0 (0.7)	0.915	0.791	1.059	3.0 (0.8)	0.902	0.774	1.051
Agreeableness	4.0 (0.6)	4.0 (0.6)	0.981	0.824	1.168	4.0 (0.6)	1.101	0.915	1.324
Neuroticism	2.6 (0.9)	2.7 (0.9)	0.982	0.856	1.127	2.7 (0.9)	0.991	0.858	1.145
<b>Model 2 (male participants)</b>									
Age	61.8 (14.5)	63.1 (11.0)	1.007	0.983	1.032	59.6 (13.1)	0.985	0.960	1.011
Financial stress	1.8 (1.4)	1.8 (1.5)	1.015	0.821	1.254	2.1 (1.6)	1.044	0.832	1.311
Satisfaction with life	3.2 (1.0)	3.2 (1.0)	0.933	0.677	1.287	3.0 (1.1)	0.760	0.533	1.084
COVID-19-related concerns	2.6 (0.5)	2.5 (0.5)	0.680	0.388	1.189	2.6 (0.5)	0.975	0.518	1.833
Openness to experience	3.4 (0.6)	3.5 (0.6)	1.244	0.781	1.983	3.4 (0.6)	1.141	0.674	1.932
Conscientiousness	3.7 (0.7)	3.7 (0.7)	0.915	0.572	1.461	3.6 (0.7)	0.852	0.504	1.441
Extraversion	3.0 (0.8)	3.1 (0.7)	1.281	0.825	1.988	3.1 (0.6)	1.072	0.653	1.757
Agreeableness	3.6 (0.7)	3.7 (0.6)	1.227	0.755	1.993	3.8 (0.6)	1.387	0.803	2.396
Neuroticism	2.5 (0.8)	2.4 (0.8)	1.055	0.678	1.641	2.5 (0.9)	0.825	0.505	1.346

<sup>a</sup>Italic text indicates a statistically significant variable.

## Discussion

### Principal Findings

This study sought to examine the psychosocial impact of COVID-19 on a sample of Australian adults. Natural language processing was applied to COVID-19 impact stories from more than 3000 Australian adults during the height of a national shutdown that was a core tenet of Australia's COVID-19

containment strategy. The majority of participants' impact stories (76.9%) were classified as having either neutral or negative sentiment. More than 70% of the main impacts detected through word frequency analysis fell into the life domains of behavioral regulation, environmental context and resources, and social influences. Statistically significant but small-magnitude associations of negative sentiment scores relative to positive scores with greater financial stress and lower

satisfaction with life were identified for female, but not male, participants.

Knowledge of community sentiment during public health events can inform better health outcomes by providing insights into the psychosocial and environmental drivers of human behavior and in detecting unanticipated events or trends within the community [35]. Similar to previous research [36-38], our findings suggest an overall trend toward neutral and negative sentiment toward COVID-19. Neutral sentiment tends to indicate lower levels of concern about COVID-19 or minimal experiences of personal impact. It can also be the case that negative impacts are balanced out by positive ones to arrive at a net neutral result, which is common during the course of disruptive events [36]. This principle of sentiment balance was reflected in the data, whereby many participants reported that some aspects of their lives, such as homeschooling children, had become more stressful but that other life pressures, such as work commutes and social obligations or overall life busyness, had eased.

Approximately one-third of the impact stories were classified as reflecting negative sentiment. Among female participants, negative sentiment was associated with financial stress and subjective well-being. This finding is consistent with previous research, particularly evidence that COVID-19 is an independent source of psychological distress among Australian adults [12] and demonstration of the association between economic indicators (eg, financial distress, unemployment) and negative psychological outcomes, including increased prevalence of psychological disorders and suicide rates [39]. These findings bolster previous calls to increase mental health support, including evidence-based prevention programs, during national disasters [39]. Our findings suggest that initiatives are particularly needed that target individuals who are experiencing financial stress (eg, who have lost their job or are seeking income support) as a targetable correlate of negative sentiment and, potentially, mental health. However, it should be noted that the magnitude of these associations was small, which suggests that another confounder that was not measured may be influencing sentiment levels.

A further 23% of participants expressed a positive sentiment toward COVID-19. Local factors such as relatively low numbers of cases in Australia as well as the composition of the current research sample are potential explanations. From a global perspective, Australia's COVID-19 response bore similarities with that of Thailand, South Korea, and Japan, including early and widespread intervention and very low rates of infection (<1000 cases per million) during 2020 [7]. This differs from countries in which responses were more anticipatory or reactive and that experienced high rates of infection (>4500 cases per million) as well as worse mortality outcomes during a similar time period [7]. It is likely that the proactiveness of Australia's response, combined with a general tendency of Australians to trust health and science authorities [40], contributed to higher confidence and optimism for Australia's recovery and therefore to the prevalence of neutral and positive sentiment. Epidemiological outcomes related to infection or mortality did not emerge as significant themes in the current data set, providing further indication that the main impacts were tied to

the consequences of containment measures rather than the virus itself.

In terms of the nature of COVID-19 impacts, the life domain that was by far the most affected was behavioral regulation. Behavioral regulation is a critical psychological determinant of positive adaptation to change [39,41,42]; however, during the COVID-19 pandemic, these abilities are inhibited. Gym closures and the subsequent lost opportunities to exercise and socialize was a salient example in this data set, particularly for older respondents or those with injury or disability who require specialized equipment (eg, pools) or supervision. Similarly, our data suggest profound social and emotional implications associated with restrictions on movement within and between communities. For example, many individuals were restricted from providing their usual informal support and care to members of their social and familial networks. In some cases, the content of the COVID-19 impact stories reflects experiences of highly emotional and potentially traumatic individual experiences that reflect isolation, boredom, and frustration with the situation. Some pertinent examples included individuals who were unable to visit patients with dementia, individuals unable to support loved ones who were dying, and ongoing worry about the safety, security, and health of others.

Public health measures designed to counteract limitations on behavioral regulation during health crises are key to minimizing the psychosocial burdens of disaster response strategies. Across the board, COVID-19 has accelerated the digitization of services, with examples including the transition toward mixed-modality health care provision including telehealth appointments and other forms of remote communication and care and the growth of virtual gyms and coaching. These shifts have not only served as integral measures to cope with the COVID-19 pandemic but also increased access to services for all in the Australian community [43]. However, the digital divide should also be acknowledged, and face-to-face and assisted services should be maintained for those with limited access or capacity to use web-based services [44,45]. Evidence suggests that multisystem-level support programs are effective, including interventions targeted at individuals in the community, health care providers, and indirect interventions that address predictors of mental health programs [39].

Looking to the future, public sentiment toward the impact of COVID-19 will continue to play an important role in how we adapt to changing conditions and new public health measures, including, for example, the rollout of vaccines, localized self-isolation orders and outbreak containment measures, and digital contact tracing. The value of natural language processing is well known; however, the sources of data used in these applications have traditionally been limited to existing data sources, particularly social media (eg, Twitter tweets) and electronic health record data. Our findings highlight the feasibility and value of natural language processing with purpose-collected research data to answer unique research questions. Future studies are also needed that apply and compare different natural language processing techniques with research data. Topic modeling is one example that identifies latent topics based on the clustering of similar terms within a data set. These more advanced techniques, which take into account not only

word frequency but also indications of proximity and other latent semantic patterns, may help to provide richer insights into the texts at hand.

### Strengths and Limitations

This is the first study to our knowledge to apply natural language processing to understand the psychosocial impacts of Australians during the COVID-19 pandemic using linked research data. Where possible, we used validated outcome measurement and analytical tools, including the robust Stanford CoreNLP analysis tool [26], to strengthen findings and allow comparisons with other studies. Limitations of the study included that the sample was biased toward women and relied upon self-report measures, which can be subject to social desirability bias. However, this can also be viewed as an advantage because the vast majority of natural language processing research draws upon social media data, such as tweets (from Twitter), as this is a common application of natural language processing. The demography of Twitter (and other social media platforms) is also biased and tends to overrepresent demographics such as urban residents, men, and people with specific interests in sports and politics [46,47]. Our sample, in comparison, is largely characterized by older women, who are not well represented in these platforms;

this enabled us to capture a different angle on the same topic. Still, findings may have limited generalizability outside of this sample, and they therefore provide a snapshot of the experiences, and their potential implications, of one moderately sized sample of Australian adults. In future, it would be pertinent to repeat this process with different population segments to build a more complete picture of the psychosocial impacts of COVID-19.

### Conclusions

Our study identified a large prevalence of negative and neutral sentiment toward COVID-19 among a sample of Australian adults. The most common impacts of COVID-19 were in the life domains of behavioral regulation, environmental contexts and resources, social influences, and emotions. Our findings shed some light on the profound disruptions that the COVID-19 pandemic has created and continues to create in people's routines and relationships, as well as some of the potential social and emotional consequences of these disruptions. Sentiment analysis as a deductive way to understand people's experiences is of critical importance during significant public health events. Ongoing analysis of community sentiment is needed to inform optimum disaster response and preparedness measures.

### Conflicts of Interest

None declared.

#### Multimedia Appendix 1

Example quotations from COVID-19 impact stories.

[DOCX File, 14 KB - [publichealth\\_v7i7e29213\\_app1.docx](#)]

### References

1. Coronavirus (COVID-19) dashboard. World Health Organization. URL: <https://covid19.who.int/> [accessed 2021-03-03]
2. Kirk C, Rifkin L. I'll trade you diamonds for toilet paper: Consumer reacting, coping and adapting behaviors in the COVID-19 pandemic. *J Bus Res* 2020 Sep;117:124-131 [FREE Full text] [doi: [10.1016/j.jbusres.2020.05.028](https://doi.org/10.1016/j.jbusres.2020.05.028)] [Medline: [32834208](https://pubmed.ncbi.nlm.nih.gov/32834208/)]
3. Ferguson N. Capturing human behaviour. *Nature* 2007 Apr 12;446(7137):733. [doi: [10.1038/446733a](https://doi.org/10.1038/446733a)] [Medline: [17429381](https://pubmed.ncbi.nlm.nih.gov/17429381/)]
4. Maciejewski PK, Zhang B, Block SD, Prigerson HG. An empirical examination of the stage theory of grief. *JAMA* 2007 Feb 21;297(7):716-723. [doi: [10.1001/jama.297.7.716](https://doi.org/10.1001/jama.297.7.716)] [Medline: [17312291](https://pubmed.ncbi.nlm.nih.gov/17312291/)]
5. Price D, Shearer F, Meehan M, McBryde E, Moss R, Golding N. Early analysis of the Australian COVID-19 epidemic. *Elife* 2020 Aug 13;9:e58785. [doi: [10.7554/elife.58785](https://doi.org/10.7554/elife.58785)]
6. A statement from the Australian Health Protection Principal Committee (AHPPC) about coronavirus (COVID-19). Australian Government Department of Health. 2020 Mar 17. URL: <https://www.health.gov.au/news/australian-health-protection-principal-committee-ahppc-coronavirus-covid-19-statement-on-17-march-2020-0> [accessed 2021-06-16]
7. Migone A. The influence of national policy characteristics on COVID-19 containment policies: a comparative analysis. *PDP* 2020 Aug 18;3(3):259-276. [doi: [10.1080/25741292.2020.1804660](https://doi.org/10.1080/25741292.2020.1804660)]
8. COVID-19 coronavirus: what you can and can't do. Western Australian Government: Department of the Premier and Cabinet. URL: <https://www.wa.gov.au/organisation/department-of-the-premier-and-cabinet/covid-19-coronavirus-what-you-can-and-cant-do> [accessed 2021-06-16]
9. Household impacts of COVID-19 survey. Australian Bureau of Statistics. URL: <https://www.abs.gov.au/statistics/people/people-and-communities/household-impacts-covid-19-survey> [accessed 2021-06-16]
10. Overseas arrivals and departures, Australia. Australian Bureau of Statistics. URL: <https://www.abs.gov.au/statistics/industry/tourism-and-transport/overseas-arrivals-and-departures-australia/latest-release> [accessed 2021-06-16]
11. Domestic aviation activity. Australian Bureau of Infrastructure and Transport Research Economics. URL: <https://www.bitre.gov.au/statistics/aviation/domestic> [accessed 2021-06-16]
12. Newby JM, O'Moore K, Tang S, Christensen H, Faasse K. Acute mental health responses during the COVID-19 pandemic in Australia. *PLoS One* 2020 Jul 28;15(7):e0236562 [FREE Full text] [doi: [10.1371/journal.pone.0236562](https://doi.org/10.1371/journal.pone.0236562)] [Medline: [32722711](https://pubmed.ncbi.nlm.nih.gov/32722711/)]

13. Stanton R, To QG, Khalesi S, Williams SL, Alley SJ, Thwaite TL, et al. Depression, anxiety and stress during COVID-19: associations with changes in physical activity, sleep, tobacco and alcohol use in Australian adults. *Int J Environ Res Public Health* 2020 Jun 07;17(11):4065 [FREE Full text] [doi: [10.3390/ijerph17114065](https://doi.org/10.3390/ijerph17114065)] [Medline: [32517294](https://pubmed.ncbi.nlm.nih.gov/32517294/)]
14. Kent K, Murray S, Penrose B, Auckland S, Visentin D, Godrich S, et al. Prevalence and socio-demographic predictors of food insecurity in Australia during the COVID-19 pandemic. *Nutrients* 2020 Sep 02;12(9):2682 [FREE Full text] [doi: [10.3390/nu12092682](https://doi.org/10.3390/nu12092682)] [Medline: [32887422](https://pubmed.ncbi.nlm.nih.gov/32887422/)]
15. O'Sullivan D, Rahamathulla M, Pawar M. The impact and implications of COVID-19: an Australian perspective. *Int J Community Soc Dev* 2020 Jul 01;2(2):134-151. [doi: [10.1177/2516602620937922](https://doi.org/10.1177/2516602620937922)]
16. Clark H, Gruending A. Invest in health and uphold rights to "build back better" after COVID-19. *Sex Reprod Health Matters* 2020 Dec 10;28(2):1781583 [FREE Full text] [doi: [10.1080/26410397.2020.1781583](https://doi.org/10.1080/26410397.2020.1781583)] [Medline: [32543338](https://pubmed.ncbi.nlm.nih.gov/32543338/)]
17. Jordan MI, Mitchell TM. Machine learning: trends, perspectives, and prospects. *Science* 2015 Jul 17;349(6245):255-260. [doi: [10.1126/science.aaa8415](https://doi.org/10.1126/science.aaa8415)] [Medline: [26185243](https://pubmed.ncbi.nlm.nih.gov/26185243/)]
18. Zunic A, Corcoran P, Spasic I. Sentiment analysis in health and well-being: systematic review. *JMIR Med Inform* 2020 Jan 28;8(1):e16023 [FREE Full text] [doi: [10.2196/16023](https://doi.org/10.2196/16023)] [Medline: [32012057](https://pubmed.ncbi.nlm.nih.gov/32012057/)]
19. Alamoodi A, Zaidan B, Zaidan A, Albahri O, Mohammed K, Malik R, et al. Sentiment analysis and its applications in fighting COVID-19 and infectious diseases: a systematic review. *Expert Syst Appl* 2021 Apr 01;167:114155 [FREE Full text] [doi: [10.1016/j.eswa.2020.114155](https://doi.org/10.1016/j.eswa.2020.114155)] [Medline: [33139966](https://pubmed.ncbi.nlm.nih.gov/33139966/)]
20. Aslam F, Awan TM, Syed JH, Kashif A, Parveen M. Sentiments and emotions evoked by news headlines of coronavirus disease (COVID-19) outbreak. *Humanit Soc Sci Commun* 2020 Jul 08;7(1). [doi: [10.1057/s41599-020-0523-3](https://doi.org/10.1057/s41599-020-0523-3)]
21. Valdez D, Ten Thij M, Bathina K, Rutter L, Bollen J. Social media insights into US mental health during the COVID-19 pandemic: longitudinal analysis of Twitter data. *J Med Internet Res* 2020 Dec 14;22(12):e21418 [FREE Full text] [doi: [10.2196/21418](https://doi.org/10.2196/21418)] [Medline: [33284783](https://pubmed.ncbi.nlm.nih.gov/33284783/)]
22. Brindal E, Ryan J, Kakoschke N, Golley S, Zajac I, Wiggins B. Individual differences and changes in lifestyle behaviours predict decreased subjective well-being during COVID-19 restrictions in an Australian sample. *J Public Health (Oxf)* 2021 Mar 04 [FREE Full text] [doi: [10.1093/pubmed/fdab040](https://doi.org/10.1093/pubmed/fdab040)] [Medline: [33683320](https://pubmed.ncbi.nlm.nih.gov/33683320/)]
23. Lavoie K, Bacon S. iCARE collaborator documents. OSF. URL: <https://osf.io/nswcm/> [accessed 2020-07-17]
24. Soto CJ, John OP. Short and extra-short forms of the Big Five Inventory-2: the BFI-2-S and BFI-2-XS. *J Res Pers* 2017 Jun;68:69-81. [doi: [10.1016/j.jrp.2017.02.004](https://doi.org/10.1016/j.jrp.2017.02.004)]
25. Diener E, Emmons RA, Larsen RJ, Griffin S. The Satisfaction With Life scale. *J Pers Assess* 1985 Feb;49(1):71-75. [doi: [10.1207/s15327752jpa4901\\_13](https://doi.org/10.1207/s15327752jpa4901_13)] [Medline: [16367493](https://pubmed.ncbi.nlm.nih.gov/16367493/)]
26. Manning C, Surdeanu M, Bauer J, Finkel J, Bethard S, McClosky D. The Stanford CoreNLP natural language processing toolkit. In: *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*. 2014 Presented at: 52nd Annual Meeting of the Association for Computational Linguistics; June 2014; Baltimore, MD. [doi: [10.3115/v1/p14-5010](https://doi.org/10.3115/v1/p14-5010)]
27. Hutto CJ, Gilbert E. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In: *Proceedings of the International AAAI Conference on Web and Social Media*. 2014 Presented at: International AAAI Conference on Web and Social Media; June 1-4, 2014; Ann Arbor, MI.
28. Lin B, Zampetti F, Bavota G, Di Penta M, Lanza M, Oliveto R. Sentiment analysis for software engineering: how far can we go? In: *ICSE '18: Proceedings of the 40th International Conference on Software Engineering*. 2018 May Presented at: 40th International Conference on Software Engineering; May 27-June 3, 2018; Gothenburg, Sweden p. 94-104. [doi: [10.1145/3180155.3180195](https://doi.org/10.1145/3180155.3180195)]
29. Cane J, O'Connor D, Michie S. Validation of the theoretical domains framework for use in behaviour change and implementation research. *Implement Sci* 2012 Apr 24;7(1):37 [FREE Full text] [doi: [10.1186/1748-5908-7-37](https://doi.org/10.1186/1748-5908-7-37)] [Medline: [22530986](https://pubmed.ncbi.nlm.nih.gov/22530986/)]
30. Ranganathan P, Pramesh CS, Aggarwal R. Common pitfalls in statistical analysis: logistic regression. *Perspect Clin Res* 2017;8(3):148-151 [FREE Full text] [doi: [10.4103/picr.PICR\\_87\\_17](https://doi.org/10.4103/picr.PICR_87_17)] [Medline: [28828311](https://pubmed.ncbi.nlm.nih.gov/28828311/)]
31. Starkweather J, Moske AK. Multinomial logistic regression. *University of North Texas Information Technology*. 2011 Aug. URL: [https://it.unt.edu/sites/default/files/mlr\\_jds\\_aug2011.pdf](https://it.unt.edu/sites/default/files/mlr_jds_aug2011.pdf) [accessed 2021-06-17]
32. Heinze G, Wallisch C, Dunkler D. Variable selection - a review and recommendations for the practicing statistician. *Biom J* 2018 May;60(3):431-449 [FREE Full text] [doi: [10.1002/bimj.201700067](https://doi.org/10.1002/bimj.201700067)] [Medline: [29292533](https://pubmed.ncbi.nlm.nih.gov/29292533/)]
33. Steptoe A, Deaton A, Stone AA. Subjective wellbeing, health, and ageing. *Lancet* 2015 Feb;385(9968):640-648. [doi: [10.1016/s0140-6736\(13\)61489-0](https://doi.org/10.1016/s0140-6736(13)61489-0)]
34. Slater MD. Theory and method in health audience segmentation. *J Health Commun* 1996 Jul;1(3):267-283. [doi: [10.1080/108107396128059](https://doi.org/10.1080/108107396128059)] [Medline: [10947364](https://pubmed.ncbi.nlm.nih.gov/10947364/)]
35. Schoutman M, Nelson EJ, Werner K, Shacham E, Elliott M, Ratnapradipa K, et al. Emerging technologies to measure neighborhood conditions in public health: implications for interventions and next steps. *Int J Health Geogr* 2016 Jun 23;15(1):20 [FREE Full text] [doi: [10.1186/s12942-016-0050-z](https://doi.org/10.1186/s12942-016-0050-z)] [Medline: [27339260](https://pubmed.ncbi.nlm.nih.gov/27339260/)]
36. Bhat M, Qadri M, Beg N, Kundroo M, Ahanger N, Agarwal B. Sentiment analysis of social media response on the Covid19 outbreak. *Brain Behav Immun* 2020 Jul;87:136-137 [FREE Full text] [doi: [10.1016/j.bbi.2020.05.006](https://doi.org/10.1016/j.bbi.2020.05.006)] [Medline: [32418721](https://pubmed.ncbi.nlm.nih.gov/32418721/)]

37. Duong V, Pham P, Yang T, Wang Y, Luo J. The ivory tower lost: how college students respond differently than the general public to the COVID-19 pandemic. ArXiv. Preprint posted online on April 21, 2020. [doi: [10.1109/asonam49781.2020.9381379](https://doi.org/10.1109/asonam49781.2020.9381379)]
38. Kumar A, Khan S, Kalra A. COVID-19 pandemic: a sentiment analysis. Eur Heart J 2020 Oct 14;41(39):3782-3783 [FREE Full text] [doi: [10.1093/eurheartj/ehaa597](https://doi.org/10.1093/eurheartj/ehaa597)] [Medline: [32678890](https://pubmed.ncbi.nlm.nih.gov/32678890/)]
39. Deady M, Tan L, Kugenthiran N, Collins D, Christensen H, Harvey SB. Unemployment, suicide and COVID-19: using the evidence to plan for prevention. Med J Aust 2020 Aug;213(4):153-154.e1. [doi: [10.5694/mja2.50715](https://doi.org/10.5694/mja2.50715)] [Medline: [32729145](https://pubmed.ncbi.nlm.nih.gov/32729145/)]
40. Pagliaro S, Sacchi S, Pacilli MG, Brambilla M, Lionetti F, Bettache K, et al. Trust predicts COVID-19 prescribed and discretionary behavioral intentions in 23 countries. PLoS One 2021 Mar 10;16(3):e0248334 [FREE Full text] [doi: [10.1371/journal.pone.0248334](https://doi.org/10.1371/journal.pone.0248334)] [Medline: [33690672](https://pubmed.ncbi.nlm.nih.gov/33690672/)]
41. Dunn DS, Brody C. Defining the good life following acquired physical disability. Rehabil Psychol 2008;53(4):413-425. [doi: [10.1037/a0013749](https://doi.org/10.1037/a0013749)]
42. Bilsker D, Fogarty AS, Wakefield MA. Critical issues in men's mental health. Can J Psychiatry 2018 Sep 19;63(9):590-596 [FREE Full text] [doi: [10.1177/0706743718766052](https://doi.org/10.1177/0706743718766052)] [Medline: [29673272](https://pubmed.ncbi.nlm.nih.gov/29673272/)]
43. Pérez Sust P, Solans O, Fajardo JC, Medina Peralta M, Rodenas P, Gabaldà J, et al. Turning the crisis into an opportunity: digital health strategies deployed during the COVID-19 outbreak. JMIR Public Health Surveill 2020 May 04;6(2):e19106 [FREE Full text] [doi: [10.2196/19106](https://doi.org/10.2196/19106)] [Medline: [32339998](https://pubmed.ncbi.nlm.nih.gov/32339998/)]
44. Newman L, Biedrzycki K, Baum F. Digital technology use among disadvantaged Australians: implications for equitable consumer participation in digitally-mediated communication and information exchange with health services. Aust. Health Review 2012;36(2):125. [doi: [10.1071/ah11042](https://doi.org/10.1071/ah11042)]
45. Hantrais L, Allin P, Kritikos M, Sogomonjan M, Anand PB, Livingstone S, et al. Covid-19 and the digital revolution. Contemp Soc Sci 2020 Oct 30;16(2):256-270. [doi: [10.1080/21582041.2020.1833234](https://doi.org/10.1080/21582041.2020.1833234)]
46. Mislove A, Lehmann S, Ahn Y, Onnela J, Rosenquist J. Understanding the demographics of Twitter users. In: Proceedings of the International AAAI Conference on Web and Social Media. 2011 Jul 05 Presented at: International AAAI Conference on Web and Social Media; July 17–21, 2011; Barcelona, Spain.
47. Bruns A, Moon B, Münch F, Sadkowsky T. The Australian Twittersphere in 2016: mapping the follower/followee network. SM+S 2017 Dec 13;3(4):205630511774816. [doi: [10.1177/2056305117748162](https://doi.org/10.1177/2056305117748162)]

## Abbreviations

**VADER:** Valence Aware Dictionary and Sentiment Reasoner

*Edited by T Sanchez, G Eysenbach; submitted 30.03.21; peer-reviewed by L Guo, D Carvalho; comments to author 26.04.21; revised version received 19.05.21; accepted 29.05.21; published 02.07.21.*

*Please cite as:*

*Ryan J, Sellak H, Brindal E*

*The Psychosocial Impacts of COVID-19 on a Sample of Australian Adults: Cross-sectional Survey and Sentiment Analysis*

*JMIR Public Health Surveill 2021;7(7):e29213*

*URL: <https://publichealth.jmir.org/2021/7/e29213>*

*doi: [10.2196/29213](https://doi.org/10.2196/29213)*

*PMID: [34081594](https://pubmed.ncbi.nlm.nih.gov/34081594/)*

©Jillian Ryan, Hamza Sellak, Emily Brindal. Originally published in JMIR Public Health and Surveillance (<https://publichealth.jmir.org>), 02.07.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

# Risk Factors of Cholera Transmission in Al Hudaydah, Yemen: Case-Control Study

Abdulqawi Mohammed Qaserah<sup>1</sup>, BSN; Mohammed Abdullah Al Amad<sup>2</sup>, MPH; Abdulwahed Abduljabbar Al Serouri<sup>2</sup>, PhD; Yousef Saleh Khader<sup>3</sup>, ScD

<sup>1</sup>Rapid Response Team, Yemen Ministry of Public Health and Population, Sana'a, Yemen

<sup>2</sup>Yemen Field Epidemiology Training Program, Yemen Ministry of Public Health and Population, Sana'a, Yemen

<sup>3</sup>Department of Community Medicine, Public Health, and Family Medicine, Faculty of Medicine, Jordan University of Science & Technology, Amman, Jordan

**Corresponding Author:**

Abdulqawi Mohammed Qaserah, BSN  
Rapid Response Team  
Yemen Ministry of Public Health and Population  
P.O. Box 299  
Al-Hasabah  
Sana'a  
Yemen  
Phone: 967 772435532  
Email: [aabuhosam33@gmail.com](mailto:aabuhosam33@gmail.com)

## Abstract

**Background:** Yemen has recently faced the largest cholera outbreak in the world, which started at the end of 2016. By the end of 2017, the cumulative reported cases from all governorates reached 777,229 with 2134 deaths. Al Hudaydah was one of the most strongly affected areas, with 88,741 (18%) cases and 244 (12%) deaths reported.

**Objective:** The aim of this study was to determine the risk factors associated with cholera transmission in Al Hudaydah city, Yemen.

**Methods:** From December 1, 2017 to January 10, 2018, a total of 104 patients with cholera (57 women and 47 men) who presented at cholera treatment centers in Al Hudaydah city with three or more watery stools in a 24-hour period and with moderate or severe dehydration were identified for inclusion in this study. Each case was matched by age and gender with two controls who were living in the neighboring house. A semistructured questionnaire was used to collect data on behavioral and environmental risk factors such as drinking water from public wells, storing water in containers, consumption of unwashed vegetables or fruits, and sharing a toilet.

**Results:** The median age of the cases and controls was 20 years (range 5-80) and 23 years (range 5-85), respectively. Only 6% of cases and 4% of controls were employed. Multivariate analysis showed that eating unwashed vegetables or fruits (odds ratio [OR] 7.0, 95% CI 1.6-30.6,  $P=.01$ ), storing water in containers (OR 3.0, 95% CI 1.3-7.3,  $P=.01$ ), drinking water from a public well (OR 2.5, 95% CI 1.1-5.7,  $P=.02$ ), and using a public toilet (OR 5.2, 95% CI 1.1-24.4,  $P=.04$ ) were significantly associated with cholera infection risk.

**Conclusions:** The cholera transmission risk factors in Al Hudaydah city were related to water and sanitation hygiene. Therefore, increasing awareness of the population on the importance of water chlorination, and washing fruits and vegetables through a health education campaign is strongly recommended.

(*JMIR Public Health Surveill* 2021;7(7):e27627) doi:[10.2196/27627](https://doi.org/10.2196/27627)

**KEYWORDS**

cholera; outbreak; risk factors, Yemen; Field Epidemiology Training Program

## Introduction

Cholera is an extremely virulent disease that can cause severe acute watery diarrhea. Cholera is mainly transmitted by ingestion of food or water contaminated with the bacterium *Vibrio cholera* O1 or O139, and is considered to be a serious threat to global public health. An estimated 3-5 million cases and over 100,000 deaths of cholera occur each year worldwide [1]. The infection is often mild or without symptoms but can be severe and can kill the host within hours in some cases if left untreated [2]. This disease particularly spreads in countries where people have unfavorable living conditions such as poor access to safe water and sanitary toilets [1,2].

Drinking contaminated water and poor food preservation methods are the major risk factors for cholera transmission [3,4]. Different risk factors have been reported in previous research, such as bathing in the river, eating dried fish, not boiling drinking water, living with people who had acute diarrhea, travel and eating outside the home, and consumption of unrefrigerated leftover food [5-9].

Yemen has faced the largest cholera outbreak in the world in recent years [10,11]. The first cases appeared in late September 2016, and by the end of 2017, the cumulative reported cases from all governorates reached 777,229 with 2134 deaths. Al Hudaydah was one of the most strongly affected areas, with 88,741 (18%) cases and 244 (12%) deaths reported [12]. A recent study performed in Aden identified the following risk factors of cholera transmission: a history of travelling and having visitors from outside the Aden governorate; eating outside the house; not washing fruits, vegetables, and khat (a local herbal stimulant) before consumption; using common-source water; and not using chlorine or soap in the household [10]. Despite the fact that Al Hudaydah was one of the most strongly affected areas, the possible risk factors of transmission in this city are still unclear. Therefore, the aim of this study was to determine the risk factors associated with cholera transmission in Al Hudaydah city, Yemen.

## Methods

### Study Area and Setting

Hudaydah city is the center of Al Hudaydah governorate located on the coast of the Red Sea, which is the fourth largest city in Yemen with a population of 400,000 habitants. Al Hudaydah is also well known as a city in which the majority of the population live below the poverty line, and lacks public services such as water and sanitation. There are three main districts in Al Hudaydah: Al Meina, Al Hawak, and Al Hali. In response to the 2017 cholera outbreak, the World Health Organization launched five cholera treatment centers (CTCs), including three centers in the three general hospitals: Al Thawrah, Al Salakhanh, and Al Aulfi hospitals.

We performed a case-control study matched by age and gender. Cases included individuals who presented to CTCs in Al Hudaydah city with acute watery diarrhea (ie, three or more watery stools in a 24-hour period) and moderate or severe dehydration during the study period. Cases were included if

they were at least 5 years of age, had a positive result to a rapid diagnostic test for cholera, and agreed to participate in the study. Cases were excluded if they lived outside of Al Hudaydah city after the start of the outbreak. Controls were selected from the same neighborhood of the cases among houses that had not reported any cases of cholera since the start of the outbreak. Individuals who were at least 5 years of age; had lived in the same neighborhood as the cases since the start of the outbreak up to April 27, 2017; did not have three liquid watery stools within 24 hours at any time since the start of outbreak; and agreed to participate in the study were included. Controls were specifically selected among individuals living in the house to the direct left of a case's house. If a control was not found in that house, the data collectors moved to the next house on the left. Controls were excluded if they lived outside of Al Hudaydah city at any time after the start of the outbreak.

Cholera risk factors were defined as any event or behavior related to water and food consumption, and hygiene practices of peoples living in Al Hudaydah city that could potentially increase the chance of becoming infected with cholera.

### Sample Size

The sample size was calculated assuming that 40% of the controls had been exposed. To detect an odds ratio (OR) of 2 between any of the studied exposure factors and the disease with a margin of error of 5%, the minimum sample size was estimated as 104 cases and 208 controls (using a ratio of cases to controls of 1:2) at a level of significance of .05 and power of 80%. A total of 104 cases were recruited from the five CTCs in Al Hudaydah city from December 1, 2017 to January 10, 2018.

### Data Collection

Data were collected during the period from December 1, 2017 to January 10, 2018. Well-trained health workers collected data using a semistructured questionnaire. The questionnaire was translated to Arabic and distributed to 10 health care providers who were included in our pilot test. The questionnaire was modified accordingly, and the final version was used to collect responses through face-to face-interviews with the cases and controls. The interviewers collected the data from cases or their caretakers at CTCs after they reviewed the registers for admitted patients. Houses of cases were also visited to collect data related to water, sanitation, hygiene, and food consumed. The interviewers searched for controls at neighboring houses and selected two controls for each case.

The questionnaires were used to collect demographic characteristics such as age, gender, address, neighborhood, street, and occupation. Clinical details, including the date of diarrhea onset, symptoms, and diagnosis, were recorded. Information on travel history, contact with infected persons, hygiene practices, eating outside the home, and attending gatherings were also collected. Source of water in the home (eg, public well, truck water, private well/borehole water, water containers); water used for drinking, preparing food, and washing; as well as the source of food were assessed.

## Ethics

This study was performed as one of the requirements for graduation from the Yemen Field Epidemiology Training Program. Ethical approval was obtained from the ethical committee at Yemen Ministry of Public Health and Population. Verbal consent was obtained from each participant. Participation was strictly voluntary, and confidentiality of participants was maintained throughout the study.

## Data Analysis

Data were entered into Epi Info version 7.2. Data were summarized using frequency distributions. Percentages were

compared using the OR. Univariate and multivariate binary logistic regression were used to determine factors associated with cholera. ORs with 95% CIs were calculated; a *P* value <.05 was considered statistically significant.

## Results

### Participant Characteristics

A total of 104 cases and 208 controls were included in this study. The median age was 20 and 23 years in the cases and controls, respectively. Only 6% of cases and 4% of controls were employed. Table 1 shows the demographic characteristics of the study participants.

**Table 1.** Demographic characteristics of study participants in Al Hudaydah, Yemen.

Characteristics	Cases (n=104)	Controls (n=208)
<b>Area of residence, n (%)</b>		
Al Hali	81 (77.9)	162 (77.9)
Al Hawak	15 (14.4)	30 (14.4)
Al Meina	8 (7.7)	16 (7.7)
<b>Gender, n (%)</b>		
Female	57 (54.8)	114 (54.8)
Male	47 (45.2)	94 (45.2)
Age (years), median (range)	20 (5-80)	23 (5-85)
<b>Occupation, n (%)</b>		
Employed	6 (5.8)	9 (4.3)
Unemployed	53 (51.0)	117 (56.3)
Student	45 (43.3)	82 (39.4)

### Factors Associated with Transmission

Table 2 shows the distribution of cholera risk factors among cases and controls in Al Hudaydah city, Yemen in 2017. A public well was more likely to be reported as a source of drinking water by cases compared to the controls. Cases were more likely than controls to use water containers to store water. Eating unwashed vegetables or fruits 1 week before the onset of symptoms was reported by significantly more cases than

controls. Buying water products in sachets in the street in the last 7 days was also reported by significantly more cases than controls.

Compared with controls, cases also significantly reported more contact with another infected family member, traveling 1 week before the onset of symptoms, using a public toilet or pit latrine, and lack of tap water in the toilet. None of the other factors investigated was significantly associated with cholera infection.



**Table 2.** Univariate analyses of cholera risk factors reported among cases and controls.

Risk factors	Cases (n=104), n (%)	Controls (n=208), n (%)	OR <sup>a</sup> (95% CI)	P value
<b>Source of home drinking water</b>			2.1 (1.2-3.7)	.008
Public well	84 (80.8)	138 (66.3)		
Truck	20 (18.3)	70 (33.7)		
Storing water in containers	92 (88.5)	149 (71.6)	3.0 (1.5-6.0)	.006
Eating unwashed vegetables or fruits	11 (10.6)	6 (2.9)	3.9 (1.4-11.0)	.007
Bought water or water products in sachets from the street in the last 7 days	58 (55.8)	83 (39.9)	1.7 (1.2-2.8)	.008
Contact with an infected person in the family	54 (51.9)	67 (32.2)	3.0 (0.5-18.6)	<.001
Travel history in the last 7 days	6 (5.8)	1 (0.5)	12.6 (1.5-106.7)	.002
<b>Type of toilet used</b>			2.7 (1.6-6.3)	.02
Public	97 (93.3)	174 (83.7)		
Private	7 (6.7)	34 (16.3)		
<b>Type of latrine used</b>			1.9 (1.2-3.2)	.005
Pit latrine	63 (60.6)	91 (43.8)		
Clean indoor latrine	41 (39.4)	117 (56.3)		
<b>Tap water in toilet</b>			1.9 (1.2-3.1)	.006
No	61 (58.7)	88 (42.3)		
Yes	43 (41.3)	120 (57.7)		

<sup>a</sup>OR: odds ratio.

### Multivariate Analysis of Risk Factors

The only factors that remained significantly associated with cholera in the multivariate analysis included eating unwashed

vegetables or fruits, storing water in containers, drinking water from a public well, and using a public toilet (Table 3).

**Table 3.** Multivariate analysis for cholera transmission risk factors.

Risk factor	aOR <sup>a</sup> (95% CI)	P value
Eating unwashed vegetables or fruits (yes vs no)	7.0 (1.6-30.6)	.01
Storing water in containers (yes vs no)	3.0 (1.3-7.3)	.01
Source of drinking water (public well vs truck)	2.5 (1.1-5.7)	.02
Type of toilet (public vs private)	5.2 (1.1-24.4)	.04

<sup>a</sup>aOR: adjusted odds ratio.

## Discussion

Cholera remains a global threat to public health, and an indicator of inequity and lack of social development [1]. Our study revealed that most of the cholera cases in Al Hudaydah, Yemen were in the Al Hali district, which is considered to be the poorest district in the city. This finding is in agreement with a study performed in an urban north-central Nigerian community in 2014 [3].

Drinking water from public wells was significantly associated with an increased odds of becoming infected with cholera. Two studies in Iran and Nigeria reported similar findings [9,13]. Storing water in containers is usually associated with a higher level of bacterial contamination if the water is not treated [13]. In our study, storing water in containers was associated with a

three-fold increase in the odds of catching cholera. This finding is in agreement with two previous studies in Vietnam and Tanzania, respectively [5,7]. However, a study performed in urban slums showed that storing water in narrow-necked earthenware vessels (called a "sorai") was effective in reducing the transmission of infection [14]. The differences between studies could be explained by the type of container used and likely also differences in the practice of treating water.

Eating unwashed vegetables or fruits was also significantly associated with an increased risk of cholera transmission, as reported in a previous study performed in Yemen [15]. Studies in Aden, South Sudan, and Nigeria reported the same finding [5,10,13]. Some studies reported that contact with a person having diarrhea and the presence of a cholera case at home are significantly associated with cholera transmission [13,14,16,17]. However, we did not find evidence to support these associations.

Sharing a toilet was also a significant predictor of cholera transmission in this study and other studies [5].

The main limitation of this study is that it was performed in Al Hudaydah city and not in all governorates, and therefore the results cannot be generalized for the whole country.

In conclusion, the cholera transmission risk factors in Al Hudaydah city, Yemen were mainly related to water and sanitation hygiene. Increasing public awareness on the importance of daily water chlorination, and washing fruits and vegetables prior to consumption through a health education campaign is strongly recommended.

## Acknowledgments

The authors would like to acknowledge the Training Programs in Epidemiology and Public Health Interventions Network (TEPHINET) and Global Health Development| Eastern Mediterranean Public Health Network (GHD|EMPHNET) for their support.

## Conflicts of Interest

None declared.

## References

1. Cholera fact sheet. World Health Organization. URL: <https://tinyurl.com/2t7c25ma> [accessed 2021-03-12]
2. Cholera -Vibrio cholerae infection. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/cholera/general/index.html> [accessed 2021-03-12]
3. Ishaku A, Shadrack BE, Ajumobi O, Olayinka A, Nguku P. Investigation of cholera outbreak in an urban North Central Nigerian community-the Akwanga Experience. *Public Health Res* 2014;4(1):7-12. [doi: [10.5923/j.phr.20140401.02](https://doi.org/10.5923/j.phr.20140401.02)]
4. Nsagha DS, Atashili J, Fon PN, Tanue EA, Ayima CW, Kibu OD. Assessing the risk factors of cholera epidemic in the Buea Health District of Cameroon. *BMC Public Health* 2015 Nov 14;15:1128 [FREE Full text] [doi: [10.1186/s12889-015-2485-8](https://doi.org/10.1186/s12889-015-2485-8)] [Medline: [26577770](https://pubmed.ncbi.nlm.nih.gov/26577770/)]
5. Acosta C, Galindo C, Kimario J, Senkoro K, Urassa H, Casals C, et al. Cholera outbreak in southern Tanzania: risk factors and patterns of transmission. *Emerg Infect Dis* 2001;7(3 Suppl):583-587 [FREE Full text] [doi: [10.3201/eid0707.010741](https://doi.org/10.3201/eid0707.010741)] [Medline: [11485679](https://pubmed.ncbi.nlm.nih.gov/11485679/)]
6. Amaah P. Quantitative and qualitative analysis of the knowledge, attitudes and social representations of cholera in the extreme northern region of Cameroon: the case of Maroua I, Maroua II and Mokolo. *Pan Afr Med J* 2014;17:253 [FREE Full text] [doi: [10.11604/pamj.2014.17.253.2799](https://doi.org/10.11604/pamj.2014.17.253.2799)] [Medline: [25309653](https://pubmed.ncbi.nlm.nih.gov/25309653/)]
7. Nguyen T, Pham Q, Do Q, Diep T, Phan H, Ho T, et al. Cholera returns to southern Vietnam in an outbreak associated with consuming unsafe water through iced tea: A matched case-control study. *PLoS Negl Trop Dis* 2017 Apr;11(4):e0005490 [FREE Full text] [doi: [10.1371/journal.pntd.0005490](https://doi.org/10.1371/journal.pntd.0005490)] [Medline: [28406946](https://pubmed.ncbi.nlm.nih.gov/28406946/)]
8. Ujjiga T, Wamala J, Mogga J, Othwonn T, Mutonga D, Kone-Coulibaly A, et al. Risk factors for sustained cholera transmission, Juba County, South Sudan, 2014. *Emerg Infect Dis* 2015 Oct;21(10):1849-1852. [doi: [10.3201/eid2110.142051](https://doi.org/10.3201/eid2110.142051)] [Medline: [26402715](https://pubmed.ncbi.nlm.nih.gov/26402715/)]
9. Moradi G, Rasouli MA, Mohammadi P, Elahi E, Barati H. A cholera outbreak in Alborz Province, Iran: a matched case-control study. *Epidemiol Health* 2016;38:e2016018. [doi: [10.4178/epih.e2016018](https://doi.org/10.4178/epih.e2016018)] [Medline: [27188308](https://pubmed.ncbi.nlm.nih.gov/27188308/)]
10. Camacho A, Bouhenia M, Alyusfi R, Alkohlani A, Naji MAM, de Radiguès X, et al. Cholera epidemic in Yemen, 2016-18: an analysis of surveillance data. *Lancet Glob Health* 2018 Jun;6(6):e680-e690 [FREE Full text] [doi: [10.1016/S2214-109X\(18\)30230-4](https://doi.org/10.1016/S2214-109X(18)30230-4)] [Medline: [29731398](https://pubmed.ncbi.nlm.nih.gov/29731398/)]
11. Federspiel F, Ali M. The cholera outbreak in Yemen: lessons learned and way forward. *BMC Public Health* 2018 Dec 04;18(1):1338 [FREE Full text] [doi: [10.1186/s12889-018-6227-6](https://doi.org/10.1186/s12889-018-6227-6)] [Medline: [30514336](https://pubmed.ncbi.nlm.nih.gov/30514336/)]
12. Yemen: Cholera Response Emergency Operations Center Situation Report No.6. reliefweb: UN Children's Fund. 2017 Oct 02. URL: <https://reliefweb.int/report/yemen/yemen-cholera-response-emergency-operations-center-situation-report-no-6-2-october-2017> [accessed 2017-12-10]
13. Hutin Y, Luby S, Paquet C. A large cholera outbreak in Kano City, Nigeria: the importance of hand washing with soap and the danger of street-vended water. *J Water Health* 2003 Mar;1(1):45-52. [Medline: [15384272](https://pubmed.ncbi.nlm.nih.gov/15384272/)]
14. Deb B, Sircar B, Sengupta P, De S, Mondal S, Gupta D, et al. Studies on interventions to prevent eltor cholera transmission in urban slums. *Bull World Health Organ* 1986;64(1):127-131 [FREE Full text] [Medline: [3488134](https://pubmed.ncbi.nlm.nih.gov/3488134/)]
15. Dureab F, Jahn A, Krisam J, Dureab A, Zain O, Al-Awlaqi S, et al. Risk factors associated with the recent cholera outbreak in Yemen: a case-control study. *Epidemiol Health* 2019 Apr 21;41:e2019015. [doi: [10.4178/epih.e2019015](https://doi.org/10.4178/epih.e2019015)] [Medline: [31010279](https://pubmed.ncbi.nlm.nih.gov/31010279/)]
16. Kone-Coulibaly A, Tshimanga M, Shambira G, Gombe N, Chadambuka A, Chonzi P, et al. Risk factors associated with cholera in Harare City, Zimbabwe, 2008. *East Afr J Public Health* 2010 Dec;7(4):311-317. [doi: [10.4314/eajph.v7i4.64754](https://doi.org/10.4314/eajph.v7i4.64754)] [Medline: [22066327](https://pubmed.ncbi.nlm.nih.gov/22066327/)]

17. Siddiqui F, Bhutto N, von Seidlein L, Khurram I, Rasool S, Ali M, et al. Consecutive outbreaks of *Vibrio cholerae* O139 and *V. cholerae* O1 cholera in a fishing village near Karachi, Pakistan. *Trans R Soc Trop Med Hyg* 2006 May;100(5):476-482. [doi: [10.1016/j.trstmh.2005.07.019](https://doi.org/10.1016/j.trstmh.2005.07.019)] [Medline: [16443247](https://pubmed.ncbi.nlm.nih.gov/16443247/)]

## Abbreviations

**CTC:** cholera treatment center

**OR:** odds ratio

*Edited by H Abbas, T Sanchez; submitted 31.01.21; peer-reviewed by F Lami, A Al-Shahethi; comments to author 27.02.21; revised version received 10.03.21; accepted 18.03.21; published 05.07.21.*

*Please cite as:*

*Qaserah AM, Al Amad MA, Al Serouri AA, Khader YS*

*Risk Factors of Cholera Transmission in Al Hudaydah, Yemen: Case-Control Study*

*JMIR Public Health Surveill* 2021;7(7):e27627

URL: <https://publichealth.jmir.org/2021/7/e27627>

doi: [10.2196/27627](https://doi.org/10.2196/27627)

PMID: [36260393](https://pubmed.ncbi.nlm.nih.gov/36260393/)

©Abdulqawi Mohammed Qaserah, Mohammed Abdullah Al Amad, Abdulwahed Abduljabbar Al Serouri, Yousef Saleh Khader. Originally published in *JMIR Public Health and Surveillance* (<https://publichealth.jmir.org>), 05.07.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Public Health and Surveillance*, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

# Performance of the Severe Acute Respiratory Illness Sentinel Surveillance System in Yemen: Mixed Methods Evaluation Study

Ghamdan Gamal Alkholidy<sup>1</sup>, MBBS; Labiba Saeed Anam<sup>1</sup>, MD; Ali Hamoud Almahaqri<sup>2</sup>, BSc; Yousef Khader<sup>3</sup>, ScD

<sup>1</sup>Yemen Field Epidemiology Training Program, Ministry of Public Health and Population, Sana'a, Yemen

<sup>2</sup>National Influenza Control Program, Ministry of Public Health and Population, Sana'a, Yemen

<sup>3</sup>Department of Community Medicine, Public Health and Family Medicine, Faculty of Medicine, Jordan University of Science & Technology, Irbidjordan, Jordan

**Corresponding Author:**

Ghamdan Gamal Alkholidy, MBBS  
Yemen Field Epidemiology Training Program  
Ministry of Public Health and Population  
Alsteen St, Bldg 23  
Sana'a  
Yemen  
Phone: 967 771530764  
Email: [ghamdangamal@gmail.com](mailto:ghamdangamal@gmail.com)

## Abstract

**Background:** The national severe acute respiratory illness (SARI) surveillance system in Yemen was established in 2010 to monitor SARI occurrence in humans and provide a foundation for detecting SARI outbreaks.

**Objective:** To ensure that the objectives of national surveillance are being met, this study aimed to examine the level of usefulness and the performance of the SARI surveillance system in Yemen.

**Methods:** The updated Centers for Disease Control and Prevention guidelines were used for the purposes of our evaluation. Related documents and reports were reviewed. Data were collected from 4 central-level managers and stakeholders and from 10 focal points at 4 sentinel sites by using a semistructured questionnaire. For each attribute, percent scores were calculated and ranked as follows: very poor ( $\leq 20\%$ ), poor (20%-40%), average (40%-60%), good (60%-80%), and excellent ( $>80\%$ ).

**Results:** As rated by the evaluators, the SARI surveillance system achieved its objectives. The system's flexibility (percent score: 86%) and acceptability (percent score: 82%) were rated as "excellent," and simplicity (percent score: 74%) and stability (percent score: 75%) were rated as "good." The percent score for timeliness was 23% in 2018, which indicated poor timeliness. The overall data quality percent score of the SARI system was 98.5%. Despite its many strengths, the SARI system has some weaknesses. For example, it depends on irregular external financial support.

**Conclusions:** The SARI surveillance system was useful in estimating morbidity and mortality, monitoring the trends of the disease, and promoting research for informing prevention and control measures. The overall performance of the SARI surveillance system was good. We recommend expanding the system by promoting private health facilities' (eg, private hospitals and private health centers) engagement in SARI surveillance, establishing an electronic database at central and peripheral sites, and providing the National Central Public Health Laboratory with the reagents needed for disease confirmation.

(JMIR Public Health Surveill 2021;7(7):e27621) doi:[10.2196/27621](https://doi.org/10.2196/27621)

**KEYWORDS**

evaluation; surveillance; Centers for Disease Control and Prevention guidelines; severe acute respiratory illness; Yemen

## Introduction

Worldwide, acute lower respiratory infection is the second commonest cause of morbidity and the third commonest cause of mortality in all age groups [1]. A significant proportion of

the global burden of acute lower respiratory infection, especially in children and older adults, is attributable to influenza and respiratory syncytial viruses.

The World Health Organization (WHO) estimated that worldwide annual influenza epidemics result in about 3 million

to 5 million cases of severe illness and about 250,000 to 500,000 deaths. In early 2019, the Global Burden of Disease study estimated that 99,000 to 200,000 annual deaths resulting from lower respiratory tract infections are directly attributable to influenza [1,2].

Estimates are rare in many countries, including countries in the Eastern Mediterranean Region. The influenza A (H1N1) pandemic highlighted the necessity of reliable estimates for the disease burden of severe acute respiratory illness (SARI) and influenza-associated SARI (F-SARI) in all countries and regions of the world [3].

Many countries have established sentinel sites for influenza epidemiological surveillance. The data captured from sentinel sites have been used by WHO member states to estimate disease burden at the national level and to compare data between countries.

Due to the fact that many emerging and reemerging diseases classified under the International Health Regulations are of an acute respiratory nature (eg, SARS [severe acute respiratory syndrome], MERS-CoV [Middle East respiratory syndrome coronavirus], and novel influenza pathogens such as H5N1 and H7N7), it is necessary to strengthen surveillance systems for acute respiratory infections and influenza in all WHO member states. This will enable countries to produce more accurate estimations of SARI and F-SARI burden [1,2].

Based on the Worldometer elaboration of the latest United Nations data in 2020, the current population size of Yemen is 29,771,764 [4]. Yemen has 4 seasons, but it is likely that the influenza virus is being circulated throughout the year. Thus, there is a great probability that different patterns of influenza virus circulation occur throughout the year [5]. Yemen is one of the countries that experience a high number of deaths resulting from acute and chronic respiratory infections. In 2008, the Ministry of Public Health and Population initiated the National Influenza Control Program to enhance the country's capacity in monitoring influenza diseases among community, guide the country in reducing morbidity and mortality from influenza diseases through the early detection of emerging novel influenza subtypes, provide a timely response for influenza prevention and control, and provide recommendations for improving influenza surveillance.

In 2013, more than 250 cases of influenza and 10 deaths were reported. A total of 1811 patients with SARI were admitted from 2011 to 2016. Of these patients, 1413 (78%) were aged <15 years, 89 (4.9%) patients had influenza viruses, and 655 (36.2%) had noninfluenza viruses. Further, the case fatality rate was 8% [6].

The Surveillance Department of the Ministry of Public Health proposed the inclusion of influenza in the Joint Program Review and Planning Mission [6]. The Central Public Health Laboratory was recognized as a national influenza center where facilities for carrying out polymerase chain reaction and serology tests are available and functional [6,7].

The WHO recommends that countries should perform surveillance for SARI and F-SARI and that surveillance systems should undergo periodic comprehensive evaluations. To ensure

that the objectives of national surveillance are being met, this study aimed to examine the level of usefulness and the performance of the SARI surveillance system.

## Methods

### Study Design

A descriptive study was conducted to evaluate the SARI surveillance system from October to December 2018 based on the updated Centers for Disease Control and Prevention guidelines for the evaluation of a public health surveillance system [8]. Mixed methods with quantitative and qualitative components were used for the evaluation. SARI sentinel sites at four public hospitals—Al Joumhour, Al Wahda, Al Swaide, and Al Thawra—in four governorates (Sana'a city, Aden, Taiz, and Al Hodeida) were studied. All possible stakeholders, including National Influenza Control Program managers, data entry staff, Ministry of Public Health and Population staff, and members of focal points in sentinel sites, were enrolled in this study.

### Data Collection

A desk review of all documents, guidelines, strategies, and pertinent scientific literature on influenza programs was conducted. Data were collected via in-depth interviews and semistructured questionnaires (Multimedia Appendix 1) with stakeholders and members of focal points at the sentinel sites, respectively. In addition, a review of the SARI system database was conducted.

### System Attributes

A total of 9 surveillance system attributes that can affect usefulness were assessed. Quantitative analysis was used to assess data quality, timeliness, sensitivity, and positive predictive values. Qualitative analysis was used to assess representativeness, simplicity, flexibility, acceptability, and stability.

### Analysis Methods

To determine the level of usefulness, the system was considered useful if it met at least one of its objectives and one of its planned uses. With regard to qualitative attributes, stakeholders were asked to rate the degree to which they agreed with attributes' specific indicators by using a 5-point Likert scale (1=strongly disagree; 2=disagree; 3=neutral; 4=agree; 5=strongly agree). Higher scores indicated better performance in terms of the studied attribute. The scores of all indicators for each attribute were summed and divided by the maximum scores to produce a percent score. The percent score was used to rank each attribute. The final rank of each attribute was classified as follows: excellent (attribute score: >80%), good (attribute score: 60%-80%), average (attribute score: 40%-60%), poor (attribute score: 20%-40%), and very poor (attribute score: ≤20%).

With regard to quantitative attributes, data quality was assessed by measuring the completeness of patient interview forms, form transmission data, and respiratory specimen collection and testing data and by checking signs and symptoms records and primary diagnoses to determine whether surveillance case definitions had been adhered to properly. The data collected by

the system were compared against the minimum data collection standards for SARI surveillance. Timeliness was assessed by calculating the percentage of specimens that were collected and sent to the laboratory within 72 hours.

## Results

### Findings From the Desk Review for Describing the SARI System

#### The Main Purpose and Objectives of the SARI System

The National Influenza Sentinel Surveillance system was established in 2010. The main purpose and objectives of the SARI system are to monitor influenza occurrence in humans and to provide a foundation for detecting outbreaks of novel strains of influenza.

#### SARI Case Definition

A case of SARI was defined as a person meeting the case definition of influenza-like illness (ie, the sudden onset of a fever of >38 °C and at least 1 of the following respiratory symptoms: dry cough, sore throat in the absence of another diagnosis, and shortness of breath or difficulty in breathing requiring hospital admission).

#### Sources of Data for the SARI System

Surveillance for influenza-like illness and SARI was carried out in 4 sentinel sites. Aggregated data were collected from and reported by each sentinel site. The data included the following:

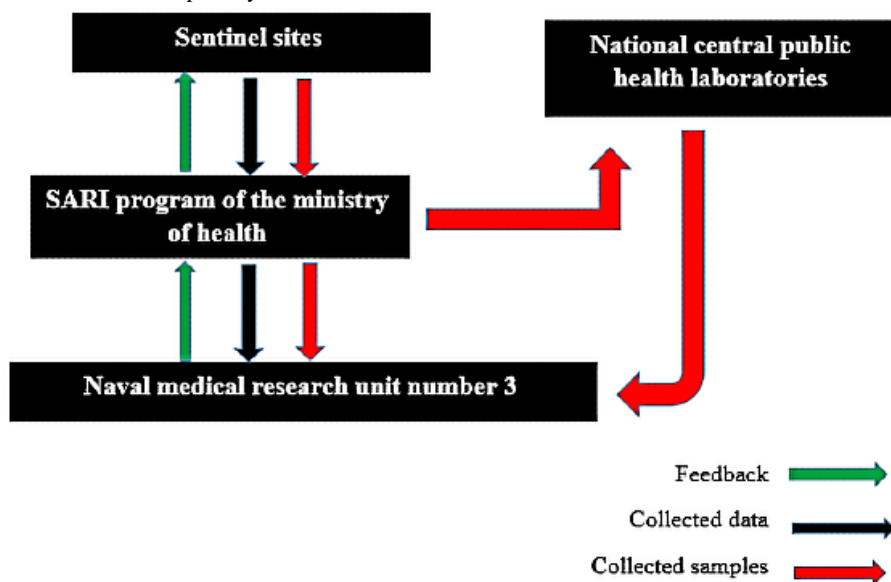
- the number of new SARI cases during the reported week
- the number of new SARI cases in which specimens were collected during the reported week
- the total number of new hospital admissions to wards in which SARI surveillance is being conducted
- specimen and epidemiological data

Specimens and epidemiological data are collected from the sentinel sites and transported to national public health laboratories. At the laboratory, specimens are tested for the influenza A and B viruses and are further subtyped if they test positive for the influenza A virus. Epidemiological and virological data collected from the sentinel sites should be collected and reported regularly to the national health authorities on a weekly basis throughout the year.

#### Data Flow and Feedback in the SARI Surveillance System

The system was designed so that each sentinel site could send its reports to central sites within 1 week. Further, samples sent to National Central Public Health laboratories within 72 hours are sent to US Naval Medical Research Unit Number 3 for confirmation, as shown in Figure 1. At the central sites, data are reviewed, organized, and analyzed, as required. Feedback is then sent to sentinel sites.

**Figure 1.** This figure illustrates the data flow and feedback in the influenza-like illness and SARI surveillance program of the Ministry of Public Health and Population in Yemen. SARI: severe acute respiratory illness.



### Demographic Characteristics of the Participants

A total of 3 persons at the central sites and 10 persons at the peripheral sites (3 pediatricians, 3 nurses, 2 laboratorian specialists, 3 health system directors, and 2 medical specialists) evaluated the system. The mean age of the participants was 44.5 years.

### Findings From In-depth Interviews at Central and Peripheral Sites

#### Usefulness

Of the 4 SARI stakeholders at the central sites, 3 (75%) agreed that the SARI surveillance system met its objectives. The usefulness percent score was 75%, indicating that the usefulness of the SARI system was good.

### Flexibility

All 13 respondents agreed that the system could easily adapt to changes in the SARI case definition, include other diseases, and accommodate any changes in data with less effort and minimal

costs. The overall percent score for flexibility was 86%, which indicated excellent performance (Table 1). At sentinel sites, the percent score, as determined by 10 stakeholders, was 71%. This indicated that the system had good performance at the peripheral sites.

**Table 1.** Flexibility of the severe acute respiratory illness (SARI) surveillance system (N=13).<sup>a</sup>

Indicator	Total score	Percent score	Rank
The system could adapt easily to changes in the SARI case definition	34	83	Excellent
The system can adapt to the integration of other surveillance systems	35	84	Excellent
The system adapted to accommodate new, additional information (eg, variation in resources)	41	90	Excellent

<sup>a</sup>The total score, percent score, and rank of the system were 120, 86, and excellent, respectively.

### Stability

All 4 stakeholders at the central sites agreed that the system could adapt to changes in resources, but half of the stakeholders

(2/4, 50%) stated that the system mainly depends on external funds. With regard to the scoring system, the overall stability percent score was 75% (Table 2). This indicated that the stability of the SARI system was good.

**Table 2.** Stability of the severe acute respiratory illness system (n=4).<sup>a</sup>

Indicator	Total score	Percent score	Rank
The system can adapt to changes in resources	4	100	Excellent
The system can adapt to funding withdrawal	2	50	Average

<sup>a</sup>The total score, percent score, and rank of the system were 6, 75, and good, respectively.

### Findings From the Self-Administered Semistructured Questionnaire for Sentinel Site Focal Points

#### Simplicity

In total, 2 indicators of simplicity were ranked as excellent, 3 were ranked as good, and 1 was ranked as average. The

simplicity percent score of the system was 74% (Table 3), indicating that the SARI system's simplicity attribute was good.

**Table 3.** Simplicity of the severe acute respiratory illness (SARI) system based on total scores, percent scores, and rank.<sup>a</sup>

Indicator	Total score	Percent score	Rank
There is the existence of a SARI case definition	30	60	Average
Using the SARI case definition is easy	37	74	Good
The SARI system uses an easy and understandable format	37	74	Good
Writing a SARI report does not take much time	40	80	Excellent
The trainees had training	42	80	Excellent

<sup>a</sup>The total score, percent score, and rank of the system were 186, 74, and good, respectively.

#### Acceptability

The two indicators of acceptability (ie, the willingness to participate among people within the system and satisfaction

with the SARI surveillance system) were ranked as excellent. The acceptability percent score of the SARI system was 82% (ie, the SARI system had excellent performance; Table 4).

**Table 4.** Acceptability of the severe acute respiratory illness (SARI) surveillance system among sentinel sites (n=10).<sup>a</sup>

Indicator	Total score	Percent score	Rank
Are you willing to participate within the system?	40	80	Excellent
Are you satisfied with the SARI surveillance system?	42	84	Excellent

<sup>a</sup>The total score, percent score, and rank of the system were 82, 82, and excellent, respectively.

## Findings From the Review of the SARI System Database

### Data Quality

Data quality was evaluated by assessing the percentage of complete forms and the missing variable data in the forms. All of the patients included in the central sites' database (N=245) had complete forms (245 forms; 100% completeness). With regard to missing data, 22 case report forms from 2018 were selected randomly and reviewed. No missing variables were found, and the variables in these forms were in line with those of the database (completeness: 100%; accuracy: 97%). The overall data quality percent score was 98.5% (excellent).

### Timeliness

The percentage of collected specimens at the health facilities that sent samples to the laboratory within 72 hours was used as an indicator of timeliness. Of the 182 collected samples, 42 (23.1%) samples were sent to the laboratory within 72 hours. The percent score for timeliness was 23% in 2018, indicating that the system's timeliness was poor.

### Overall Performance of the SARI Surveillance System

The overall performance of the SARI surveillance system had a percent score of 79% (ie, the SARI system had good performance; Table 5).

**Table 5.** The overall performance of the severe acute respiratory illness system in Yemen.

Indicators	Score	Percent score	Rank
Usefulness	15	75	Good
Flexibility	120	86	Excellent
Stability	6	75	Good
Simplicity	186	74	Good
Acceptability	82	82	Excellent
Timeliness	23	23	Poor
Data quality	99	99	Excellent
Overall performance	394	79	Good

## Strengths and Weaknesses of the SARI Surveillance System

Despite its many strengths, the SARI system has some weaknesses. For example, it depends on irregular external financial support. To act as a platform for the surveillance of other respiratory illnesses, the SARI surveillance system integrates an influenza-like illness surveillance system with the electronic Disease Early Warning System. This has several benefits. First, it allows for efficient laboratory data collection and transportation. Second, the SARI system uses resources more efficiently than other systems. These widespread benefits enhance the usefulness of the system and allow the system to meet its own surveillance objectives and address broader national priorities.

## Discussion

### Principal Findings

The surveillance systems at sentinel sites are tools for the early detection of disease, the monitoring of trends in the burden of diseases, and the generation of recommendations for the prevention and control of diseases. The evaluation of surveillance systems helps decision makers to set priorities for future planning, resource allocation, and future interventions for preventing the spread of diseases.

Overall, this study showed that the performance of the SARI surveillance system was good. Studies from Zambia [9] and the Democratic Republic of Congo [10] have reported similar findings. The SARI system was found to be useful in detecting trends and signal changes in the occurrence of SARI, estimating

the magnitude of morbidity and mortality related to SARI, and promoting research for informing prevention and control measures for SARI. Similarly, a previous evaluation in Zambia demonstrated the usefulness of the system [9].

The SARI surveillance system was shown to be simple. The case definition and case report forms were available and easy to use. Similarly, the simplicity of the SARI surveillance system was documented in previous evaluations conducted in the Democratic Republic of Congo [10] and Zambia [10].

The SARI surveillance system's flexibility was excellent. The system appeared to be able to adapt easily to changes in the SARI case definition and accommodate changes in data with less effort and minimal costs. This finding is consistent with the findings of a study from South Africa [11]. However, it is not consistent with the findings of studies from Zambia [9] and the Democratic Republic of Congo [10], which reported that the flexibility of the evaluated systems ranged from moderate to good. The acceptability of the SARI system was excellent, as reflected by the willingness of stakeholders to participate in the system and their satisfaction with the SARI surveillance system.

The stability of the SARI surveillance system was good. It was found that the system was stable and could adapt to changes in resources (eg, donors withdrawing their support). This finding is in line with the findings of previous studies from South Africa [11], the Democratic Republic of Congo [10], and Zambia [9]. However, it is not in line with the findings of a study from Pakistan [12], which reported average stability.



Our findings showed that the timeliness of the SARI system was very poor. This might be due to the lack of laboratory components that are essential for sampling. Previous evaluations in South Africa [11], China [13], Zambia [9], and the Democratic Republic of Congo [10] reported moderate to good timeliness. The quality and completeness of SARI surveillance system data were excellent.

### Limitations

We could not calculate positive predictive values and assess sensitivity because the samples have not been tested since 2016 due to a lack of reagents.

### Conclusion

Overall, the SARI surveillance system was useful in estimating morbidity and mortality, monitoring the trends of the disease, and promoting research for informing prevention and control measures. The overall performance of the SARI surveillance system was good. We recommended expanding the system by promoting private health facilities' (eg, private hospitals and private health centers) engagement in SARI surveillance, establishing an electronic database at central and peripheral sites, and providing the National Central Public Health Laboratory with the reagents needed for disease confirmation.

### Acknowledgments

The authors would like to acknowledge the Global Health Development/Eastern Mediterranean Public Health Network for their technical support.

### Conflicts of Interest

None declared.

### Multimedia Appendix 1

Questionnaire.

[DOCX File, 107 KB - [publichealth\\_v7i7e27621\\_app1.docx](#)]

### References

1. GBD 2016 Lower Respiratory Infections Collaborators. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory infections in 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Infect Dis* 2018 Nov;18(11):1191-1210 [FREE Full text] [doi: [10.1016/S1473-3099\(18\)30310-4](#)] [Medline: [30243584](#)]
2. Paget J, Spreeuwenberg P, Charu V, Taylor RJ, Iuliano AD, Bresee J, Global Seasonal Influenza-associated Mortality Collaborator NetworkGLaMOR Collaborating Teams\*. Global mortality associated with seasonal influenza epidemics: New burden estimates and predictors from the GLaMOR Project. *J Glob Health* 2019 Dec;9(2):020421 [FREE Full text] [doi: [10.7189/jogh.09.020421](#)] [Medline: [31673337](#)]
3. Gouya M, Rezaei F, Haghdoost A, Nabavi M, Farahi KS, Mostafavi E, et al. Estimation of influenza and severe acute respiratory illness incidence (burden) in three provinces of the Islamic Republic of Iran, 2012 and 2013. *East Mediterr Health J* 2016 Oct 02;22(7):432-439 [FREE Full text] [doi: [10.26719/2016.22.7.432](#)] [Medline: [27714736](#)]
4. Yemen population. Worldometer. URL: <https://www.worldometers.info/world-population/yemen-population/> [accessed 2020-06-05]
5. Surveillance, forecasting and response. World Health Organization. URL: <http://www.emro.who.int/surveillance-forecasting-response/pandemic-influenza/pip-yemen.html> [accessed 2018-10-13]
6. Al Amad MA, Al Mahaqri AA, Al Serouri AA, Khader YS. Severe acute respiratory infections with influenza and noninfluenza respiratory viruses: Yemen, 2011-2016. *Inquiry* 2019;56:46958019850731 [FREE Full text] [doi: [10.1177/0046958019850731](#)] [Medline: [31137990](#)]
7. Craig A, Kool J, Nilles E. The Pacific experience: supporting small island countries and territories to meet their 2012 International Health Regulations (2005) commitments. *Western Pac Surveill Response J* 2013 Jul 03;4(3):14-18 [FREE Full text] [doi: [10.5365/WPSAR.2012.3.4.007](#)] [Medline: [24319608](#)]
8. German RR, Lee LM, Horan JM, Milstein RL, Pertowski CA, Waller MN, Guidelines Working Group Centers for Disease Control and Prevention (CDC). Updated guidelines for evaluating public health surveillance systems: recommendations from the Guidelines Working Group. *MMWR Recomm Rep* 2001 Jul 27;50(RR-13):1-35; quiz CE1-7. [Medline: [18634202](#)]
9. Simusika P, Tempia S, Chentulo E, Polansky L, Mazaba ML, Ndumba I, et al. An evaluation of the Zambia influenza sentinel surveillance system, 2011-2017. *BMC Public Health Serv Res* 2020 Jan 13;20(1):35 [FREE Full text] [doi: [10.1186/s12913-019-4884-5](#)] [Medline: [31931793](#)]
10. Babakazo P, Kabamba-Tshilobo J, Wemakoy EO, Lubula L, Many LK, Ilunga BK, et al. Evaluation of the influenza sentinel surveillance system in the Democratic Republic of Congo, 2012-2015. *BMC Public Health* 2019 Dec 10;19(1):1652 [FREE Full text] [doi: [10.1186/s12889-019-8008-2](#)] [Medline: [31823763](#)]

11. Budgell E, Cohen AL, McAnerney J, Walaza S, Madhi SA, Blumberg L, et al. Evaluation of two influenza surveillance systems in South Africa. *PLoS One* 2015 Mar 30;10(3):e0120226. [doi: [10.1371/journal.pone.0120226](https://doi.org/10.1371/journal.pone.0120226)] [Medline: [25822719](https://pubmed.ncbi.nlm.nih.gov/25822719/)]
12. Noreen N, Khan AW, Badar N, Khan FK, Khudaidad F, Khan NU, et al. Evaluation of lab-based influenza surveillance system in Pakistan, 2017. *Global Biosecurity* 2019 Jul 03;1(2):1-7 [FREE Full text] [doi: [10.31646/gbio.26](https://doi.org/10.31646/gbio.26)]
13. Li D, Zhang LJ, Qin Y, Jian H, Li ZJ, Feng LZ, et al. Evaluation on data quality of severe acute respiratory infection sentinel surveillance system in China, 2015-2016. *Disease Surveillance* 2017;32(12):914-916. [doi: [10.3784/j.issn.1003-9961.2017.12.005](https://doi.org/10.3784/j.issn.1003-9961.2017.12.005)]

## Abbreviations

**F-SARI:** influenza-associated severe acute respiratory illness  
**MERS-CoV:** Middle East respiratory syndrome coronavirus  
**SARI:** severe acute respiratory illness  
**SARS:** severe acute respiratory syndrome  
**WHO:** World Health Organization

*Edited by M Algunaïd; submitted 31.01.21; peer-reviewed by F Lami, H Taha, R Bright; comments to author 06.03.21; revised version received 25.04.21; accepted 15.05.21; published 09.07.21.*

*Please cite as:*

*Alkholidy GG, Anam LS, Almahqri AH, Khader Y*

*Performance of the Severe Acute Respiratory Illness Sentinel Surveillance System in Yemen: Mixed Methods Evaluation Study*

*JMIR Public Health Surveill* 2021;7(7):e27621

URL: <https://publichealth.jmir.org/2021/7/e27621>

doi: [10.2196/27621](https://doi.org/10.2196/27621)

PMID: [34255695](https://pubmed.ncbi.nlm.nih.gov/34255695/)

©Ghamdan Gamal Alkholidy, Labiba Saeed Anam, Ali Hamoud Almahqri, Yousef Khader. Originally published in *JMIR Public Health and Surveillance* (<https://publichealth.jmir.org>), 09.07.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Public Health and Surveillance*, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.

---

Publisher:  
JMIR Publications  
130 Queens Quay East.  
Toronto, ON, M5A 3Y5  
Phone: (+1) 416-583-2040  
Email: [support@jmir.org](mailto:support@jmir.org)

---

<https://www.jmirpublications.com/>