JMIR Public Health and Surveillance

Impact Factor (2023): 3.5
Volume 8 (2022), Issue 4 ISSN 2369-2960 Editor in Chief: Travis Sanchez, PhD, MPH

Contents

Original Papers

Reviews

Original Paper

Interactive Versus Static Decision Support Tools for COVID-19: Randomized Controlled Trial

Alice Röbbelen[1*], MSc; Malte L Schmieding[2*], MD; Marvin Kopka[2,3], MSc; Felix Balzer[2*], MSc, MD, PhD; Markus A Feufel[1*], Dipl-lng (FH), MSc, PhD

[1]Division of Ergonomics, Department of Psychology and Ergonomics (IPA), Technische Universität Berlin, Berlin, Germany

[2]Institute of Medical Informatics, Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

[3]Cognitive Psychology and Ergonomics, Department of Psychology and Ergonomics (IPA), Technische Universität Berlin, Berlin, Germany

[*]these authors contributed equally

Corresponding Author:
Malte L Schmieding, MD
Institute of Medical Informatics
Charité – Universitätsmedizin Berlin
Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin
Charitéplatz 1
Berlin, 10117
Germany
Phone: 49 30 450 570 425
Email: malte.schmieding@charite.de

Abstract

Background: During the COVID-19 pandemic, medical laypersons with symptoms indicative of a COVID-19 infection commonly sought guidance on whether and where to find medical care. Numerous web-based decision support tools (DSTs) have been developed, both by public and commercial stakeholders, to assist their decision making. Though most of the DSTs' underlying algorithms are similar and simple decision trees, their mode of presentation differs: some DSTs present a static flowchart, while others are designed as a conversational agent, guiding the user through the decision tree's nodes step-by-step in an interactive manner.

Objective: This study aims to investigate whether interactive DSTs provide greater decision support than noninteractive (ie, static) flowcharts.

Methods: We developed mock interfaces for 2 DSTs (1 static, 1 interactive), mimicking patient-facing, freely available DSTs for COVID-19-related self-assessment. Their underlying algorithm was identical and based on the Centers for Disease Control and Prevention's guidelines. We recruited adult US residents online in November 2020. Participants appraised the appropriate social and care-seeking behavior for 7 fictitious descriptions of patients (case vignettes). Participants in the experimental groups received either the static or the interactive mock DST as support, while the control group appraised the case vignettes unsupported. We determined participants' accuracy, decision certainty (after deciding), and mental effort to measure the quality of decision support. Participants' ratings of the DSTs' usefulness, ease of use, trust, and future intention to use the tools served as measures to analyze differences in participants' perception of the tools. We used ANOVAs and *t* tests to assess statistical significance.

Results: Our survey yielded 196 responses. The mean number of correct assessments was higher in the intervention groups (interactive DST group: mean 11.71, SD 2.37; static DST group: mean 11.45, SD 2.48) than in the control group (mean 10.17, SD 2.00). Decisional certainty was significantly higher in the experimental groups (interactive DST group: mean 80.7%, SD 14.1%; static DST group: mean 80.5%, SD 15.8%) compared to the control group (mean 65.8%, SD 20.8%). The differences in these measures proved statistically significant in *t* tests comparing each intervention group with the control group ($P<.001$ for all 4 *t* tests). ANOVA detected no significant differences regarding mental effort between the 3 study groups. Differences between the 2 intervention groups were of small effect sizes and nonsignificant for all 3 measures of the quality of decision support and most measures of participants' perception of the DSTs.

Conclusions: When the decision space is limited, as is the case in common COVID-19 self-assessment DSTs, static flowcharts might prove as beneficial in enhancing decision quality as interactive tools. Given that static flowcharts reveal the underlying

decision algorithm more transparently and require less effort to develop, they might prove more efficient in providing guidance to the public. Further research should validate our findings on different use cases, elaborate on the trade-off between transparency and convenience in DSTs, and investigate whether subgroups of users benefit more with 1 type of user interface than the other.

Trial Registration: Deutsches Register Klinischer Studien DRKS00028136; https://tinyurl.com/4bcfausx (retrospectively registered)

(JMIR Public Health Surveill 2022;8(4):e33733) doi:10.2196/33733

KEYWORDS

clinical decision support; usability; COVID-19; consumer health; medical informatic; symptom checker; decision support; symptom; support; decision making; algorithm; flowchart; agent

Introduction

In December 2019, the spread of the novel lung disease COVID-19 caused by SARS-CoV-2 with previously unknown etiology was detected, and it developed into a global pandemic within a few weeks [1,2]. The disease courses of COVID-19 are heterogeneous. On the one hand, it is dangerous and can be lethal, even among previously healthy individuals. On the other hand, COVID-19 can present itself with unspecific and mild symptoms [3]. Given the dynamic development of the disease and public health measures, laypersons often felt uncertain about appropriate behavior, especially when a COVID-19 infection was suspected but not proven. This situation resulted in considerable uncertainty both about what level of medical care was needed (care-seeking behavior) and whether isolation or quarantine was required (social behavior). As a result, capacity utilization in the health care system increased [4].

The increased demand for medical advice could not be met through traditional routes (eg, telephone consultation, visit to the general practitioner, visit to the emergency room, local health authorities). As a consequence, various online services with information and self-assessments (ie, patient-facing clinical decision support tools [DSTs]) were developed. These DSTs are comparable to commonly known symptom checkers (SCs), which "are tools developed to provide support to laypersons. Users can enter their complaints and, with some SCs, demographic or health-related information (eg, age, gender, past medical history) to obtain advice on the urgency of their complaints (triage advice) and the most likely diagnosis" [5]. The DSTs considered in the context of COVID-19 often require information about possible exposure to the virus, in addition to standard inputs related to gender, age, and major complaints. As output, the DSTs provide the probability of a COVID-19 infection rather than a diagnosis. Though most of the underlying algorithms are rule-based, simple, and similar decision trees, implementing official guidelines (eg, of the Centers for Disease Control and Prevention [CDC] [6]), their mode of presentation differs: Some DSTs present the rule-based decision tree algorithm as a static flowchart [7-12], while others are designed as conversational agents (ie, similar to chatbots) with a graphical user interface, which guides the user through the decision tree's nodes step-by-step in an interactive manner [13-21].

Our review of the literature suggests that there is no research, neither on SCs nor on DSTs for COVID-19, indicating whether and how the interactivity of a user interface influences the decision outcome and decision support experience. Currently available studies assessing DSTs for COVID-19 report no findings on the influence between user interaction and quality of decision support [4,22-26]. However, many publications on so-called patient decision aids (PtDAs) exist that have examined web-based and paper-based decision aids. In these studies, web-based PtDAs showed no difference from paper-based PtDAs in their effect on participants' decision making [27]. In terms of web-based PtDAs, interactive and static formats have not yet been compared to the best of our knowledge.

However, the comparison between web-based tools is particularly relevant during the COVID-19 pandemic, as laypersons' decision making should ideally be supported at home to minimize the risks of unnecessary COVID-19 infections. The aim of this study is to assess whether static and interactive DSTs increase laypersons' accuracy and confidence when making COVID-19-related decisions via the internet and whether 1 type of DST (static or interactive) is superior to the other.

Methods

Study Design

We chose a between-subjects design with 1 three-level independent variable: participants had to solve fictional case vignettes while receiving different types of decision support (no support in the control group, a DST in the form of a static flowchart in the first intervention group, and an interactive DST in the second intervention group). Participants were randomly assigned to 1 of the 3 groups. Decision support quality was evaluated by analyzing the tools' effects on decision making and the participants' judgement of the tools with multiple dependent variables: Regarding effects on decision making, the number of correct appraisals (using official CDC recommendations from October 14, 2020, as the standard [28]), perceived certainty regarding the appraisal, and mental effort in making the appraisals were examined. Participants' judgments were collected on tool usefulness, ease of use, trust in the tools' recommendations, and intention to use the tools in the future.

Ethics Approval and Consent to Participate

The Ethics Committee of the Department of Psychology and Ergonomics (IPA) at the Technische Universität Berlin approved this study (tracking number ROEB_01_20200715). Participants volunteered to take part in the survey and provided informed consent before participating.

Development of the Decision Support Tools

We designed 2 mock DSTs. The design was intended to mimic existing tools available on the internet. Features of various tools related to COVID-19 were analyzed. Relevant common features of these tools were the use of short, precise questions or statements to which the participants could respond in agreement (yes) or disagreement (no), and sometimes also with uncertainty (unsure). These questions guided participants through a decision tree's nodes step-by-step until a recommendation could be provided. Input requirements and output content were mostly based on recommendations from public health agencies, such as the CDC. For the purpose of our study, the John Hopkins University Coronavirus Self-Checker was identified as a most suitable template because it is a decision tree–based interactive tool that requires a small amount of information at each step as input [18]. This facilitated the construction of a second DST providing identical advice in the form of a static flowchart, outlining the entire decision tree at once. In both cases, the content was based on the CDC recommendations from October 14, 2020 [28].

Considering these aspects, we developed 2 mock DSTs with identical content that differed only with respect to the interactivity of their user interface: a static flowchart where participants could follow the arrows of the appropriate path using their gaze (Multimedia Appendix 1). This flowchart was developed using Microsoft PowerPoint [29]. The other tool was implemented as a simple, interactive conversational agent, where participants clicked buttons to respond to questions, guiding them through the nodes of the decision tree step-by-step (Multimedia Appendix 2). The individual screens of the interactive DST were also designed using Microsoft PowerPoint and then linked together using InVision [30], enabling dynamic interaction. Mimicking the output of existing tools, 5 different recommendations were possible: "emergency care is required," "self-care and physical distancing are sufficient," "self-care is sufficient and quarantine necessary," "self-care is sufficient and isolation is necessary," and "nonemergency care and isolation are required." We anticipated that laypersons would struggle comprehending the difference between "quarantine" and "isolation," yet we decided to differentiate between them, as some COVID-19 DSTs use these terms without explaining them [7,11]. During data analyses, however, we rated answers as correct when participants confused these 2 terms (see the Data Analysis section). We developed mock tools rather than choosing existing DSTs from the internet in order to ensure comparability of the algorithm and design of the tools, as well as stable integration into the survey throughout the surveying period.

Preparing the Case Vignettes

We created 7 short patient descriptions (case vignettes). In this process, we extracted decision criteria identified by the CDC and its recommendations concerning help-seeking and social behavior in the case of a potential COVID-19 infection [28]. Decision criteria included the presence of typical COVID-19 symptoms with the following expressions: 0, no symptoms; 1, primary symptoms (typical for a COVID-19 infection); and 2, secondary symptoms (can also occur in a COVID-19 infection but are less typical) (see Multimedia Appendix 3) [28]. Another decision criterion represented the potential contact with the pathogen SARS-CoV-2 and the following expressions: 1, critical contact with a confirmed infected person has occurred; 2, contact may have occurred; or 3, critical contact can be excluded. The last important criterion is the presence or absence of risk factors classified by the Robert Koch Institute (RKI), Germany's federal public health authority (eg, certain pre-existing conditions such as chronic lung disease; see Multimedia Appendix 4) [3]. Through a pretest with 15 participants, in which we evaluated the response behavior of the participants as well as their free-text feedback, we revised the wording of the vignettes, improving intelligibility and removing errors spotted by the pretest participants. The vignettes included more information than necessary for the appraisal to better simulate a real-life decisional context and increase ecological validity. We also ensured that all information asked for in the developed DSTs was included in the vignettes and that the fictitious patients represented diverse age and gender groups. In addition, each possible outcome of the mock tools was covered at least once as a correct solution. Multimedia Appendices 5 and 6 contain an outline of the 7 case vignettes.

Data Collection

Participants

All participants lived in the United States, were at least 18 years old, and had no professional medical background. Our investigation was limited to US residents, as recommendations of the developed DSTs were based on CDC recommendations and guidelines and thus applied to the US region.

Survey and Instruments

We created an online survey using UNIPARK (QuestBack GmbH, Germany) [31]. To ensure differences between the 2 groups were not due to confounding variables, we surveyed variables we suspected to have an influence on the main outcome measures, namely sociodemographic factors, affinity toward interacting with technology, prior knowledge of COVID-19, prior use of a COVID-19 DST, and perceived threat of a COVID-19 infection. The sociodemographic variables were age, gender, US residency, and highest level of completed formal education. Affinity toward interacting with technology was surveyed with the Affinity for Technology Interaction Short Scale (ATI-S) [32]. The perceived threat of COVID-19 was surveyed with an instrument developed by Kim and Park [33]. Participants' prior knowledge of COVID-19 necessary to appraise the adequate health-seeking and social behavior for a suspected COVID-19 patient was assessed with 5 self-developed multiple-choice questions (see Multimedia Appendix 7).

In the survey's main part, participants were presented the 7 case vignettes in random order. In both the experimental groups and the control group, participants provided their personal appraisal of the adequate help-seeking behavior for each case vignette and, in a further question, of the adequate social behavior. In the 2 experimental groups, participants were also prompted to state which recommendations of the tool (supposedly) provided concerned help-seeking behavior and social behavior (see Multimedia Appendices 8 and 9).

As we anticipated participants commonly erring in determining the appropriate social behavior when required to differentiate between isolation and quarantine, we deemed both answers correct for our main analysis and conducted a second analysis without this adjustment (Multimedia Appendix 9).

After each case vignette, participants were further asked to rate their mental effort required in making decisions concerning the respective fictitious case presentation using a 9-point category scale ranging from very, very low mental effort to very, very high mental effort (see Multimedia Appendix 10) [34]. This scale by Paas et al [35] can be considered a "subjective, indirect measure of cognitive load" in making decisions regarding case vignettes. Since mental effort increases with higher perceived demands of a stimulus or task [35], it was included here as an indicator of the quality of decision support. We assume that a good DST guides people to the right decision without requiring high cognitive load.

Following all 7 case vignettes, participants' decisional uncertainty was assessed once using the O'Connor [36] Decisional Conflict Subscale. On 3 items with a 5-point scale, participants rated how confident they were in their decisions (see Multimedia Appendix 11). The scale was developed to evaluate health care consumer decision aids. We used this scale as an indicator of the quality of decision support, since we assume that a good DST helps medical laypersons to feel confident in their decisions. This assumption is supported by O'Connor's [36] claim that "decision aids should reduce uncertainty and confusion in choosing a course of action."

In addition to the effect of the tools on decision making and participants' perceptions of their decision making, we also asked directly about participant's perceptions of the tools in the 2 intervention groups. They were asked about the perceived usefulness and ease of use of the tools after having worked on all 7 cases. Perceived usefulness is defined as the individual's perception of the extent to which using the tool improves decision-making performance. Perceived ease of use implies that using the tool does not require any effort [37]. Both constructs were measured using scales from the Davis Technology Acceptance Model [38]. In addition, the intentions to use the DSTs in the future and the trust in the recommendations of the DSTs were measured as a subjective rating using 1 item each (Multimedia Appendix 12 presents the exact phrasing of the items).

Procedure

We recruited participants via the crowdsourcing platform Prolific.co [39]. This platform is aimed at researchers and allows them to conveniently recruit participants for online surveys and experiments. Prolific.co provides a pool of participants that can be screened based on demographic data. Researchers can make their survey or experiment available on Prolific.co, and users registered on Prolific.co as prospective participants can choose whether to participate in studies that they are eligible for according to the set prescreening criteria. We chose Prolific.co as it is characterized by a diverse population that provides high-quality data [40]. Each participant received a remuneration of US $2.00 for completing the survey. In addition, participants could earn a bonus of US $0.20 for each correct decision. To ensure that all participants answered the questions attentively, attention checks were added to the survey. Participants failing more than 1 of 3 attention checks were excluded. Furthermore, at the end of the survey, participants were asked to self-report whether they had participated in the survey attentively, honestly, and without external assistance, as suggested by Rouse [41]. Participants were remunerated independently of their answer, but only data from participants confirming this question were included in the analysis. By selecting the weekend day and early afternoon PDTs, we attempted to recruit a population as diverse as possible, as suggested by Casey et al [42]. The survey was released for participation on Prolific.co on 4 different days (November 21, 2020, at 1 PM Pacific Daylight Time [PDT]; November 22, 2020, at 11 AM PDT; November 28, 2020, at 12 PM PDT; and November 29, 2020, at 1 PM PDT). On each day, 50 participants were recruited within a few hours after release.

Data Analysis

Data were cleaned and explored using R 4.0.2 [43] and the *tidyverse* packages [44]. After examining the central tendencies and distributions of the variables separately for the 3 groups and plotting them using the *ggplot2* package [45], we performed a one-way analysis of variance to compare the 3 conditions separately for the dependent measures accuracy, decisional certainty, and mental effort. If significant group differences were present, we followed up with Bonferroni-corrected post hoc pairwise *t* tests. For the dependent variables' usefulness, ease of use, trust in tool recommendations, and future use intention, we conducted Welch two-sample *t* tests to compare the 2 intervention groups because the sample sizes of the groups were unequal [46]. Effect sizes were quantified with Cohen *d* and calculated using the *rstatix* package [47]. When results were not statistically significant, we performed a post hoc power analysis using the *pwr* package [48] with corresponding group sizes, an α level of .05, and a power 1-β of .80.

Results

Participant Characteristics

In total, our survey was accessed 233 times during the 4 days it was available. Overall, 37 participations could not be used for data analysis because they dropped out of the questionnaire (n=12, 32%), exceeded the maximum survey completion time (n=10, 27%), failed attention checks (n=6, 16%), or did not meet eligibility criteria (n=9, 25%). The remaining participants all affirmed that they had paid close attention during the survey and answered honestly. This yielded a total of 196 (84.1%) participants, who assessed all 7 case vignettes and could be included in the analysis (see Table 1 for details). The mean time for completion of the survey was 22 minutes and 3 seconds. Across the 3 groups, the participant characteristics were similar; see Table 1.

Table 1. Participant characteristics (N=196) of an experimental study assessing the influence of DSTs[a] on laypersons' COVID-19-related appraisals. Participants were nonmedically trained US inhabitants sampled online in November 2020.

Characteristics	Total sample	Group 1: control group (no DST)	Group 2: static DST	Group 3: interactive DST
Sample size, N	196	66	62	68
Age (years), median (IQR)	30 (18)	30 (17.2)	26.5 (13.2)	33 (20.5)
Gender, n (%)				
Female	94 (48)	31 (47)	27 (44)	36 (53)
Male	100 (51)	33 (50)	35 (56)	32 (47)
Other	2 (1)	2 (3)	0	0
Education, n (%)				
Non-high-school graduate	4 (2)	1 (2)	1 (2)	2 (3)
High school graduate	34 (17)	9 (14)	18 (29)	7 (10)
Some college	66 (34)	22 (33)	20 (32)	24 (35)
Bachelor's degree	49 (25)	20 (30)	12 (19)	17 (25)
Graduate degree or higher	43 (22)	14 (21)	11 (18)	18 (26)
Recent[b] experience with COVID-19 decisions, n (%)				
Recently faced a triage decision	70 (35)	25 (38)	27 (44)	18 (26)
Recently faced a social behavior decision	101 (52)	35 (53)	32 (52)	34 (50)
Recently consulted a static DST to face the decision	56 (29)	15 (23)	21 (34)	20 (29)
Recently consulted an interactive DST to face the decision	53 (27)	21 (32)	15 (24)	17 (25)
Medical training, n (%)				
No training	163 (83)	55 (83)	51 (82)	57 (84)
Basic first-aid training	33 (17)	11 (17)	11 (18)	11 (16)
Affinity for technology interaction on a scale of 1-6, median (IQR)[c]	4 (1.2)	4 (1.7)	4 (1)	4 (1)
Perceived threat of COVID-19 on a scale of 1-7, median (IQR)[d]	5 (2)	5.25 (1.6)	5 (2.1)	5 (2)
Prior knowledge of COVID-19 on a scale of 0-5, median (IQR)[e]	3 (2)	3 (2)	3 (1)	3 (1)

[a]DST: decision support tool.

[b]"Recent" was defined as "in the past 6 months."

[c]Measured by the Wessel Affinity for Technology Interaction Short Scale (ATI-S).

[d] Measured by a subjective self-assessment on 2 items on a scale of 1-7 adapted from Kim and Park [33].

[e]Measured by the number of correctly answered multiple-choice questions with reference to COVID-19.

Effects on Decision Making

Omnibus ANOVAs detected group differences with respect to decision accuracy (overall and separately for decisions about help-seeking behavior and social behavior) and perceived certainty but not for mental effort in decision making; see Table 2.

Table 2. Omnibus ANOVAs measuring the effects of different DSTs[a] on laypersons' ability to correctly appraise fictitious descriptions of patients with symptoms indicative of COVID-19 in an experimental study. In total, 196 participants (all US residents and nonmedically trained) were recruited online in November 2020 and asked to judge how fictitious patients with symptoms indicative of COVID-19 should behave. Participants were randomly assigned to 1 of 3 groups in which they either received support by a static DST (flowchart) or an interactive DST (mimicking a conversational agent) or received no support.

Dependent variables	Group 1: without DST	Group 2: static DST	Group 3: interactive DST	Test statistics of group comparison		
				$F_{2,193}$	*P* value	η^2
Accuracy, mean (SD)						
Total correct decisions (min=0, max=14)[b]	10.17 (2.00)	11.45 (2.48)	11.71 (2.37)	8.59	<.001	0.08
Correct decisions on help-seeking behavior (min=0, max=7)	4.82 (0.96)	5.47 (1.39)	5.54 (1.40)	6.58	<.001	0.002
Correct decisions on social behavior (min=0, max=7)[b]	5.35 (1.41)	5.98 (1.26)	6.16 (1.18)	7.33	<.001	0.02
Decisional certainty: certainty score from 0 to 100[c], mean (SD)	65.78 (20.78)	80.51 (15.89)	80.72 (14.08)	15.67	<.001	0.14
Mental effort on a scale from 1 to 9, mean (SD)	5.62 (1.57)	5.40 (1.68)	5.09 (1.78)	1.73	.18	N/A[d]

[a]DST: decision support tool.

[b]The response options "quarantine" and "isolation" were not differentiated for this analysis, that is, they were both considered correct because layperson participants commonly confuse these terms. An analysis without this adjustment is provided in Multimedia Appendix 9 and shows the same trend.

[c]Responses were transformed into a certainty score between 0 and 100 (ie, 0=person feels extremely uncertain about the best choice and 100=person feels extremely certain about the best choice).

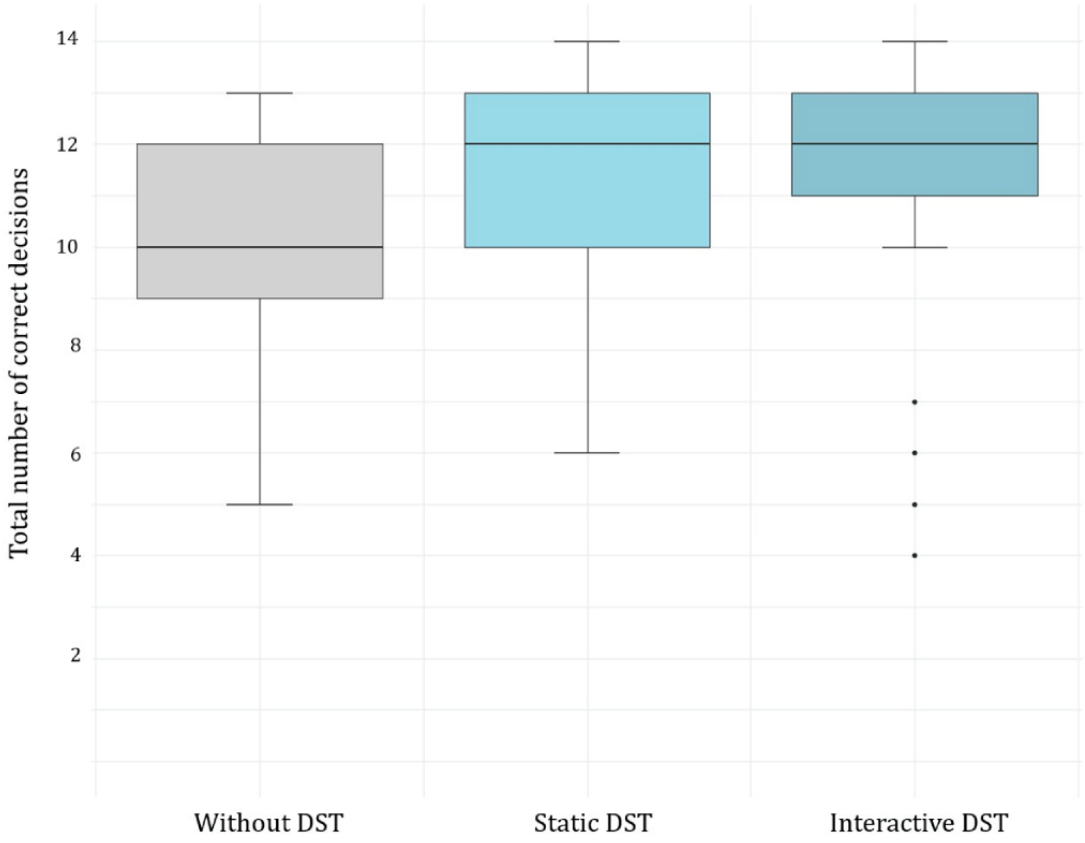
[d]N/A: not applicable.

Accuracy

Across all groups, participants' appraisals were commonly correct. The average participant decided correctly in more than 10 (70%) of the 14 decisions. Overall decision accuracy was higher in the experimental groups receiving support from a DST than in the unsupported control group; see Figure 1 and Table 2. Bonferroni-corrected pairwise Welch two-sample *t* tests indicated statistically significant differences with moderate effect sizes between control and experimental groups (control group vs static DST group: $t_{117.35}$=–3.21, *P*<.001, Cohen *d*=–0.57, 95% CI –1.1 to –0.14; control group vs interactive DST group: $t_{129.6}$=–4.06, *P*<.001, Cohen *d*=–0.70, 95% CI –1.31 to –0.21). The difference in decision accuracy between the 2 experimental groups showed a negligible effect size and was not significant ($t_{125.58}$=–0.59, *P*=.55, Cohen *d*=–0.10, 95% CI –0.54 to 0.34). Based on a post hoc power analysis, we estimated that on a population level, the real difference in accuracy was less than Cohen *d*=0.5 (equivalent to 1.22 correct responses), with a power of 0.8.

We found similar results when analyzing participants' accuracy for help-seeking behavior and social behavior separately; see Table 2. When the differences between the response options "isolation" and "quarantine" were considered, the pattern remained the same, but the gap in accuracy between experimental and control groups widened; see Multimedia Appendix 13.

Figure 1. Boxplot showing the distribution of the 196 participants' decision accuracy to appraise 7 fictitious descriptions of patients with symptoms indicative of COVID-19. Study participants (all US inhabitants, nonmedically trained, sampled online in November 2020) were tasked to answer 2 questions per patient description. We randomly assigned participants to 1 of 3 experimental groups; in 2 groups, they were supported by either a static DST (ie, a flowchart) or an interactive DST (ie, a conversational agent mimicking a chatbot). In the control group, they received no decision support. The boxplots' filled box represents the IQR, the horizontal line inside the box the median, the whiskers the maximum and minimum values within 1.5 IQR of the median, and the single dots the outliers of participants' total number of correct decisions. DST: decision support tool.

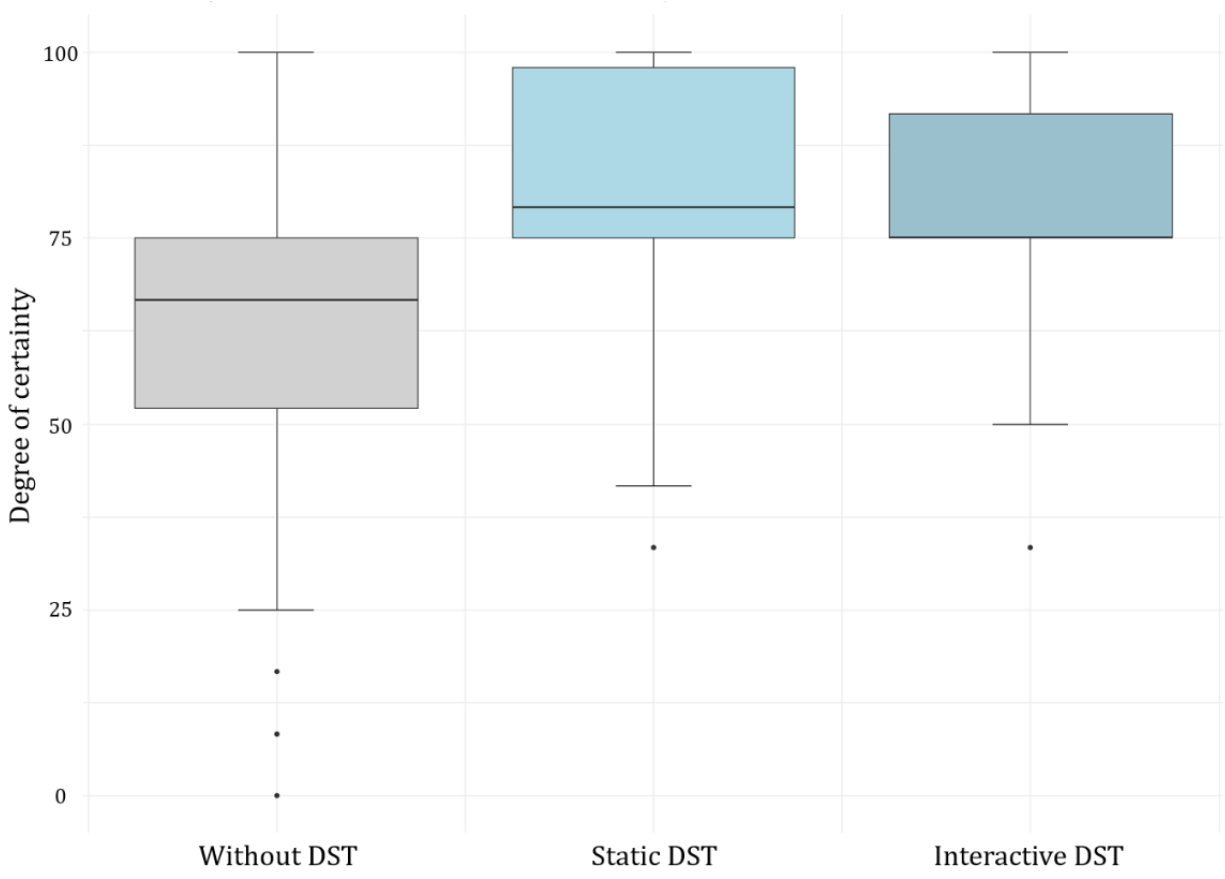


Decisional Certainty

Participants were commonly certain in their decision making. Less than 15 (8%) of the 196 participants indicated a certainty of less than 50 on a scale from 0 to 100, in contrast to 32 (16%) indicating maximum certainty (100/100). Participants' certainty in their decisions differed between the 3 groups: certainty in decision making was rated lower by participants without decision support than by those receiving decision support; see Figure 2 and Table 2. Bonferroni-corrected pairwise Welch two-sample *t* tests marked these differences as statistically significant with large effect sizes (control group vs static DST group: $t_{121.1}$=–4.51, *P*<.001, Cohen *d*=–0.79, 95% CI –1.22 to –0.37; control group vs interactive DST group: $t_{115.67}$=–4.67, *P*<.001, Cohen *d*=–0.81, 95% CI –1.27 to –0.44).

Decision certainty in the 2 experimental groups was nearly identical, with mean certainty scores of 80.51 (static DST) and 80.72 (interactive DST). The inferential analysis indicated this difference to be of negligible effect size and nonsignificant ($t_{123.68}$=–0.09, *P*=.92, Cohen *d*=0.01, 95% CI –0.47 to 0.43). Based on a post hoc power analysis, we estimated that on a population level, the real difference in decisional certainty between both experimental groups was less than Cohen *d*=0.5 (equivalent to 7.5 percentage points), with a power of 0.8.

Figure 2. Laypersons' perceived certainty in their own appraisals of COVID-19-related clinical decisions obtained in an experimental study in November 2020. The 196 study participants were US residents, nonmedically trained, and sampled online. Our study tasked them to assess 7 fictitious descriptions of patients with symptoms indicative of COVID-19. Participants were randomized to either receive support from a static DST (ie, a flowchart) or an interactive DST (ie, a conversational agent mimicking a chatbot). Following the 14 appraisals, we surveyed the participants' certainty in their answers using the Decisional Conflict Scale. A score of 0% indicated minimum certainty, while 100% indicated maximum certainty. The boxplots' filled box represents the IQR, the horizontal line inside the box the median, the whiskers the maximum and minimum values within 1.5 IQR of the median, and the single dots the outliers of participants' total number of correct decisions. DST: decision support tool.



Mental Effort

The perception of how demanding the task was varied among participants. Although a third (63/196, 32.1%) stated that the tasks required low mental effort (ie, they indicated an average mental effort score below 4.5 on a scale from 1 to 9), 4 (40%) of 10 considered the tasks to be at least somewhat demanding (indicating scores above 5.5).

Participants rated the mental effort of decision making lowest in the interactive DST group (mean 5.06) and highest in the control group unsupported by a DST (mean 5.62); see Table 2. However, the differences between group means were not significant in the omnibus ANOVA; see Table 2. Based on a post hoc power analysis, we estimated that on a population level, the effect size of the difference was not greater than η^2=0.047, with a power of 0.8.

Participants' Perceptions of the Tools

We investigated participants' perceptions of the mock tools with 4 metrics: perceived usefulness, perceived ease of use, trust, and intention to use the tools again in the future. On average, participants perceived the DSTs positively: their self-reported trust in the tools was generally high in both experimental groups as was the number of DST recommendations the participants followed (on average, in more than 12 [86%] of a total of 14 recommendations); see Table 3. Accordingly, almost all (126/130, 95%) participants judged the mock tools to be useful (ie, indicated a perceived usefulness score of above 3.0 on a scale from 1 to 5), with half of the participants scoring the usefulness at or above 4.75. Ease of use was rated similarly positive (see Table 3), and their intention to use the tools in the future was high, on average. Regarding differences between the 2 DSTs, the mock users' scores on perceived usefulness, trust, and future use intention were higher in the static than in the interactive DST group. Ease of use was perceived higher in the interactive DST group. However, all effect sizes of differences between the 2 DST groups were small (Cohen $d \leq 0.35$) and proved significant only for future use intention in a Welch two-sample t test; see Table 3.

Based on a post hoc power analysis, we estimated that on a population level, the real difference in these variables between both groups was less than Cohen d=0.5 (equivalent to a difference of 0.34 for perceived usefulness and perceived ease of use, 0.42 for self-reported trust, 0.96 for followed decisions, and 0.51 for future use intention), with a power of 0.8.

Table 3. Laypersons' perceptions of 2 mock DSTs[a] for COVID-19-related clinical decisions obtained in an experimental study in November 2020. The 196 study participants were US residents, nonmedically trained, and sampled online. Our study tasked them to assess 7 fictitious descriptions of patients with symptoms indicative of COVID-19. Participants were randomized to either receive support from a static DST (ie, a flowchart) or an interactive DST (ie, a conversational agent mimicking a chatbot) or receive no support (control group). Subsequently, participants in the intervention groups were asked to rate the given tools' usefulness and perceived ease of use, state their trust in the tools, and state their future intention to use the tools. We measured usefulness and perceived ease of use according to the Davis Technology Acceptance Model.

Dependent variables	Group 2: static DST	Group 3: interactive DST	Test statistics of group comparison		
			t_{df}	*P* value	Cohen *d*
Perceived usefulness on a scale from 1 to 5, mean (SD)	4.56 (0.65)	4.35 (0.71)	t_{128}=1.7	.09	0.30
Perceived ease of use on a scale from 1 to 5, mean (SD)	4.26 (0.71)	4.47 (0.58)	$t_{117.92}$=–1.90	.06	–0.34
Trust, mean (SD)					
Self-reported trust in the tools' recommendation on a scale from 1 to 7	6.08 (0.84)	5.83 (0.85)	$t_{127.15}$=1.74	.08	0.31
Decisions where the recommendation of the tools was followed on a scale from 0 to 14	12.73 (1.81)	12.5 (2.03)	$t_{127.94}$=0.67	.75	0.12
Future use intention on a scale from 1 to 7, mean (SD)	6.23 (0.88)	5.87 (1.16)	$t_{123.91}$=1.99	.05	0.35

[a]DST: decision support tool.

Discussion

Principal Findings

In our study, DSTs increased laypersons' accuracy and certainty in decision making. Potentially, DSTs can also reduce the mental effort in decision making, but this effect was statistically nonsignificant in our experiment. Thus, our experiment confirms the benefit of DSTs for COVID-19-related self-triage discussed in prior studies [4,23].

Regarding the question of whether 1 mode of presentation of DSTs (ie, static or interactive presentation) is superior to the other, our experiment produced evidence that differences in measures of the quality of decision support and participants' perception of the tools are small. Interactive DSTs are potentially more convenient to use. Users of the interactive DST rated the mental effort as lower and the perceived ease of use as higher than the users of the static DST. However, these differences were not significant and did not translate into a higher level of trust, greater perceived usefulness, or greater intention to use the tool in the future. On the contrary, users rated the static tool more favorably on these measures than the interactive DST. Overall, the effect sizes for differences in these measures were low and statistically nonsignificant.

To the best of our knowledge, our study is among the first to directly compare the effectiveness of different modes of presentation on the quality of support received from web-based, patient-facing clinical DSTs. Our results are in line with findings from similar research in a different field of application: In their 2018 meta-analysis on the use of PtDAs to support decisions concerning prostate cancer screening, Baptista et al [27] concluded that web-based decision aids reduce decisional conflict compared to no decision aid, but no more than static, printed-out decision aids. Comparability between Baptista et al [27] and our study is limited, as web-based decision aids and their paper-based printed-out versions assessed by Baptista et al [27] do not directly match the mock conversational agent and the static flowchart we presented digitally in our study. However, both results suggest that for a decision scenario with clear-cut options, the mode of presentation with a DST has little to no effect on how helpful it is for its users.

Practical Implications

First, our results underline the benefits of making DSTs available to laypersons for decisions with clear-cut options, such as those encountered in the COVID-19 pandemic, as both decisional accuracy and trust in the decision increased when laypersons were supported by both mock DSTs. Second, the quality of decision support provided by the static flowchart and the interactive tool did not differ significantly, and post hoc sensitivity analyses on effect sizes indicated that we can rule out large effects. Thus, factors such as development effort might potentially weigh more than the format of interaction when deciding on how to present decision support to laypersons.

When complexity is low, as in the case of the COVID-19 decisions, the static version may provide a better overview and thus make the decision more transparent to its users and be quicker and more cost-efficient to develop and publish than an interactive conversational agent. In contrast, for decisions with higher complexity, full transparency might compromise ease of use of a static DST. Thus, for more complex decisions, interactive tools that guide the user step-by-step may be more user friendly. However, our results suggest that interactivity is not an effective means in itself to increase the usefulness of a DST: First, the higher convenience of using the interactive DST did not translate into greater trust or perceived usefulness, nor did it yield a greater increase in decision accuracy or certainty than static flowcharts. As the latter entail less development efforts, while increasing transparency, they might be the preferable mode of DSTs for public health officials to implement on their websites to provide guidance to the public on decisions of low complexity and limited decision space.

Our study also raises 2 topics for further research. First, which factors increase users' trust in DSTs, and how are they weighed?

From our results, we can only speculate that users prefer transparency over convenience. Second, would subgroups of the population prefer 1 type of interface over the other? Although our results showed only minor differences between sample averages of users of the static and interactive DSTs, potentially subgroups of the population benefit more from 1 type of interface than the other. For example, users with low technological affinity might prefer a static flowchart over a conversational agent.

Limitations

The results of our study are mainly limited by concerns about external validity. That is, the interactive mock DST we developed does not fully represent the whole variety of interactive tools that are available on the internet. Although our DST mimics the interactive tool from Johns Hopkins University, other tools incorporate significantly more decision factors and possible outcomes. Second, the sampled study population's composition of educational background is not representative of the adult US population. Our study sample included a higher proportion of highly educated persons, with only 4 (2%) of 196 participants being nonhighschool graduates as compared to 9.8% in the US adult population [49]. In addition, the mock interactive tool we developed for this paper does not fully exhaust the potential of interactivity, as our participants only followed prompts to respond to binary questions by clicking either Yes or No buttons, while more interactive tools also require users to enter information manually (eg, age or current location). Finally, participants in this study did not use the tool for self-assessment but to appraise cases with fictitious patients. This means that they did not experience personal concerns, as would likely be the case with COVID-19 suspicion in a real-world use situation. To promote external validity, the recommendations of the DSTs we developed conform with all CDC guidelines and their interactive capacities are basic but mimic those of existing DSTs.

Conclusion

When the decision space is limited, a static flowchart potentially performs just as well as an interactive tool in enhancing the decision quality of laypersons with symptoms indicative of a COVID-19 infection. As static flowcharts reveal their underlying decision algorithm more transparently, they might prove to be more suitable in not only guiding laypersons through the health care system but also communicating the reasoning and thereby empowering patients. Further research should validate our findings on different use cases, elaborate on the trade-off between transparency and convenience in DSTs, and investigate whether subgroups of users benefit more from 1 type of user interface than the other by assessing interactions between outcome variables (eg, accuracy, mental effort, perceived usefulness) and participant characteristics (eg, age, eHealth literacy). We have made all data necessary to conduct these exploratory analyses and to reproduce our reported findings publicly available [50].

Acknowledgments

The authors express their gratitude to the participants and to Eike Richter for his advice on statistical methods. The project was funded by the home institutions of the last authors (MAF and FB). No external funding was required to conduct this study. The authors acknowledge support from the German Research Foundation (DFG) and the Open Access Publication Fund of Charité - Universitätsmedizin Berlin.

Authors' Contributions

AR and MLS conceived the study. AR created the questionnaire, developed the mock decision support tool (DSTs), designed and conducted the analyses, and wrote the first draft of the paper. MLS oversaw the creation of the case vignette adaptations and aided in the design of the methods and manuscript development. MK assisted in statistical analyses and manuscript development. MAF and FB provided critical input and advised on the study and questionnaire design, analysis methods, and drafts of the paper. AR and MLS share the first authorship. MAF and FB share the last authorship. All authors accept full responsibility for the final version of the paper.

The lead authors affirm that this manuscript is an honest, accurate, and transparent account of the study being reported, that no important aspects of the study have been omitted, and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Conflicts of Interest

All authors have completed the International Committee of Medical Journal Editors uniform disclosure form and declare no support from any organization for the submitted work, no financial relationships with any organizations that might have an interest in the submitted work in the previous 3 years, and no other relationships or activities that could appear to have influenced the submitted work.

Editorial Notice

This randomized study was only retrospectively registered, given that the authors believed registration was unnecessary for this kind of trial with a theoretical case example and the particular outcomes which are measured. The editor granted an exception from ICMJE rules mandating prospective registration of randomized trials. However, readers are advised to carefully assess the

validity of any potential explicit or implicit claims related to primary outcomes or effectiveness, as retrospective registration does not prevent authors from changing their outcome measures retrospectively.

Multimedia Appendix 1
Static decision support flowchart for laypersons with suspected COVID-19 infection developed as part of the scientific study.
[PDF File (Adobe PDF File), 391 KB - publichealth_v8i4e33733_app1.pdf]

Multimedia Appendix 2
Interactive decision support tool for laypersons with suspected COVID-19 infection developed as part of the scientific study.
[PDF File (Adobe PDF File), 254 KB - publichealth_v8i4e33733_app2.pdf]

Multimedia Appendix 3
List of COVID-19 symptoms according to the Robert Koch Institute.
[DOCX File , 54 KB - publichealth_v8i4e33733_app3.docx]

Multimedia Appendix 4
List of risk factors for COVID-19 according to the Robert Koch Institute.
[DOCX File , 54 KB - publichealth_v8i4e33733_app4.docx]

Multimedia Appendix 5
Overview of the case vignettes developed as part of the scientific study.
[PDF File (Adobe PDF File), 262 KB - publichealth_v8i4e33733_app5.pdf]

Multimedia Appendix 6
Outline of all 7 case vignettes developed as part of the scientific study.
[DOCX File , 57 KB - publichealth_v8i4e33733_app6.docx]

Multimedia Appendix 7
Multiple-choice questions used to assess the participants' prior knowledge.
[PDF File (Adobe PDF File), 154 KB - publichealth_v8i4e33733_app7.pdf]

Multimedia Appendix 8
Decisions on each case vignette in the control group without decision support.
[PDF File (Adobe PDF File), 194 KB - publichealth_v8i4e33733_app8.pdf]

Multimedia Appendix 9
Decisions on each case vignette in the intervention groups with decision support, here illustrated by the flowchart example.
[PDF File (Adobe PDF File), 281 KB - publichealth_v8i4e33733_app9.pdf]

Multimedia Appendix 10
Single item for mental effort.
[PDF File (Adobe PDF File), 52 KB - publichealth_v8i4e33733_app10.pdf]

Multimedia Appendix 11
Decisional Conflict Subscale.
[PDF File (Adobe PDF File), 72 KB - publichealth_v8i4e33733_app11.pdf]

Multimedia Appendix 12
Items measuring participants’ perspectives on the 2 DSTs with items on ease of use, usefulness, trust, and future intention to use the DSTs. DST: decision support tool.
[PDF File (Adobe PDF File), 223 KB - publichealth_v8i4e33733_app12.pdf]

Multimedia Appendix 13
Boxplot for decision accuracy in regard to decision support type. Responses where participants confused "quarantine" and "isolation" were rated as wrong in this analysis.

[PDF File (Adobe PDF File), 90 KB - publichealth_v8i4e33733_app13.pdf]

Multimedia Appendix 14
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 9667 KB - publichealth_v8i4e33733_app14.pdf]

References

1. Sohrabi C, Alsafi Z, O'Neill N, Khan M, Kerwan A, Al-Jabir A, et al. World Health Organization declares global emergency: a review of the 2019 novel coronavirus (COVID-19). Int J Surg 2020 Apr 26;76:71-76 [FREE Full text] [doi: 10.1016/j.ijsu.2020.02.034] [Medline: 32112977]
2. World Health Organization. Pandemie der Coronavirus-Krankheit (COVID-19). URL: https://www.euro.who.int/de/health-topics/health-emergencies/coronavirus-covid-19 [accessed 2022-03-07]
3. Epidemiologischer Steckbrief zu SARS-CoV-2 und COVID-19. Demografische Faktoren, Symptome und Krankheitsverlauf. Robert-Koch-Institut. URL: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Steckbrief.html [accessed 2021-05-05]
4. Judson T, Odisho A, Neinstein A, Chao J, Williams A, Miller C, et al. Rapid design and implementation of an integrated patient self-triage and self-scheduling tool for COVID-19. J Am Med Inform Assoc 2020;27(6):866. [doi: 10.1093/jamia/ocaa051]
5. Schmieding ML, Mörgeli R, Schmieding MAL, Feufel MA, Balzer F. Benchmarking triage capability of symptom checkers against that of medical laypersons: survey study. J Med Internet Res 2021 Mar 10;23(3):e24475. [doi: 10.2196/24475]
6. What to Do if You Are Sick. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/if-you-are-sick/steps-when-sick.html [accessed 2021-05-05]
7. Coronavirus Disease 2019 (COVID-19) Risk Assessment and Public Health Management Decision Making. Centers for Disease Control and Prevention. URL: https://dhr.idaho.gov/wp-content/uploads/2020/COVID-19/CDCDecisionMakingModel.pdf [accessed 2020-10-14]
8. COVID-19 Flowchart. Jackson Free Press. URL: https://www.jacksonfreepress.com/photos/2020/mar/14/36329/ [accessed 2021-05-02]
9. North Dakota Department of Health. What To Do If You Are A Close Contact. URL: https://www.health.nd.gov/what-do-if-you-are-close-contact [accessed 2022-03-07]
10. COVID-19: Bin ich betroffen und was ist zu tun?: Orientierungshilfe für Bürgerinnen und Bürger. Robert-Koch-Institut. URL: https://edoc.rki.de/bitstream/handle/176904/6779.2/Orientierungshilfe_COVID_Buerger_20-05-06_doi_10.25646-6629-2_korr.pdf?sequence=1&isAllowed=y [accessed 2022-03-07]
11. Contact Tracing Flow Chart: For potential and actual exposures to COVID-19. State of Michigan. URL: https://www.michigan.gov/documents/coronavirus/Person_Exposure_Final_3-25-2020_684832_7.pdf [accessed 2021-05-02]
12. Coronavirus-Entscheidungshilfe. ZEIT ONLINE GmbH. URL: https://www.zeit.de/wissen/gesundheit/2020-03/Coronavirus-Entscheidungshilfe-2020-03-31.pdf [accessed 2021-05-02]
13. Ada Health GmBH. COVID-19 Screener. 2020. URL: https://ada.com/de/covid-19-screener/#assessment [accessed 2022-03-07]
14. Coronavirus Self-Checker. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/testing.html [accessed 2021-05-02]
15. CovApp: Haben Sie Symptome? CovApp zeigt Ihnen mögliche nächste Schritte. Charité – Universitätsmedizin Berlin. URL: https://covapp.charite.de/ [accessed 2022-03-07]
16. Emory University School of Medicine. Coronavirus Checker. URL: https://c19check.com/start [accessed 2022-03-07]
17. Together against Coronavirus: COVID-GUIDE. in4medicine AG. URL: https://covidguide.health/de/ [accessed 2021-05-02]
18. Coronavirus Self-Checker and COVID-19 Vaccine FAQ. The Johns Hopkins University, The Johns Hopkins Hospital, and The Johns Hopkins Health System Corporation. URL: https://www.hopkinsmedicine.org/coronavirus/covid-19-self-checker.html [accessed 2021-05-02]
19. Universitätsklinikum Giessen und Marburg GmbH. COVID-ONLINE. URL: https://covid-online.de/#/ [accessed 2022-03-07]
20. Covid Assessment. The USC Gehr Family Center for Health Systems Sciences and Innovation. URL: https://covidassessment.org/acknowledgement [accessed 2021-05-02]
21. COVID-19 Symptom Checker. WebMD LLC. URL: https://www.webmd.com/coronavirus/coronavirus-assessment/default.htm [accessed 2021-05-02]
22. Kahnbach L, Lehr D, Brandenburger J, Mallwitz T, Jent S, Hannibal S, et al. Quality and adoption of COVID-19 tracing apps and recommendations for development: systematic interdisciplinary review of European apps. J Med Internet Res 2021 Jun 02;23(6):e27989 [FREE Full text] [doi: 10.2196/27989] [Medline: 33890867]
23. Lunn PD, Timmons S, Julienne H, Belton CA, Barjaková M, Lavin C, et al. Using decision aids to support self-isolation during the COVID-19 pandemic. Psychol Health 2021 Feb 19;36(2):195-213. [doi: 10.1080/08870446.2020.1849701] [Medline: 33210950]

24. Munsch N, Martin A, Gruarin S, Nateqi J, Abdarahmane I, Weingartner-Ortner R, et al. Diagnostic accuracy of web-based COVID-19 symptom checkers: comparison study. J Med Internet Res 2020 Oct 06;22(10):e21299 [FREE Full text] [doi: 10.2196/21299] [Medline: 33001828]
25. Perlman A, Vodonos Zilberg A, Bak P, Dreyfuss M, Leventer-Roberts M, Vurembrand Y, et al. Characteristics and symptoms of app users seeking COVID-19-related digital health information and remote services: retrospective cohort study. J Med Internet Res 2020 Oct 20;22(10):e23197 [FREE Full text] [doi: 10.2196/23197] [Medline: 32961527]
26. Zens M, Brammertz A, Herpich J, Südkamp N, Hinterseer M. App-based tracking of self-reported COVID-19 symptoms: analysis of questionnaire data. J Med Internet Res 2020 Sep 9;22(9):e21956 [FREE Full text] [doi: 10.2196/21956] [Medline: 32791493]
27. Baptista S, Teles Sampaio E, Heleno B, Azevedo LF, Martins C. Web-based versus usual care and other formats of decision aids to support prostate cancer screening decisions: systematic review and meta-analysis. J Med Internet Res 2018 Jun 26;20(6):e228 [FREE Full text] [doi: 10.2196/jmir.9070] [Medline: 29945857]
28. If You Are Sick or Caring for Someone. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/if-you-are-sick/index.html [accessed 2021-05-02]
29. Microsoft Corporation. Microsoft PowerPoint (Version 2019). URL: https://www.microsoft.com/de-de/microsoft-365/powerpoint [accessed 2022-03-07]
30. InVision. InVisionApp Inc. URL: https://www.invisionapp.com/ [accessed 2020-08-15]
31. Über das Online-Befragungstool UNIPARK. Trivian XI GmbH. 2021. URL: https://www.unipark.com/ [accessed 2020-07-05]
32. Wessel D, Attig C, Franke T. ATI-S: an ultra-short scale for assessing Affinity for Technology Interaction in user studies. 2019 Presented at: Proceedings of Mensch und Computer 2019; 2019; Hamburg Germany. [doi: 10.1145/3340764.3340766]
33. Kim J, Park H. Development of a health information technology acceptance model using consumers' health behavior intention. J Med Internet Res 2012 Oct 01;14(5):e133 [FREE Full text] [doi: 10.2196/jmir.2143] [Medline: 23026508]
34. Paas F, Tuovinen JE, Tabbers H, Van Gerven PWM. Cognitive load measurement as a means to advance cognitive load theory. Educ Psychol 2003 Mar;38(1):63-71. [doi: 10.1207/S15326985EP3801_8]
35. Kirschner P, Kirschner F. Mental effort. In: Seel NM, editor. Encyclopedia of the Sciences of Learning. Boston, MA: Springer; 2012.
36. O'Connor AM. Validation of a Decisional Conflict Scale. Med Decis Making 2016 Jul 02;15(1):25-30. [doi: 10.1177/0272989x9501500105]
37. Ammenwerth E. Technology acceptance models in health informatics: TAMUTAUT. In: Scott P, de Keizer N, Georgiou A, editors. Applied Interdisciplinary Theory in Health Informatics. Amsterdam, the Netherlands: IOS Press BV; 2019:64.
38. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
39. Online Participant Recruitment For Surveys and Market Research. Prolific Academic. URL: https://www.prolific.co/ [accessed 2021-04-09]
40. Peer E, Brandimarte L, Samat S, Acquisti A. Beyond the Turk: alternative platforms for crowdsourcing behavioral research. J Exp Soc Psychol 2017 May;70:153-163. [doi: 10.1016/j.jesp.2017.01.006]
41. Rouse SV. A reliability analysis of Mechanical Turk data. Comput Hum Behav 2015 Feb;43:304-307. [doi: 10.1016/j.chb.2014.11.004]
42. Casey LS, Chandler J, Levine AS, Proctor A, Strolovitch DZ. Intertemporal differences among MTurk workers: time-based sample variations and implications for online data collection. SAGE Open 2017 Jun 14;7(2):215824401771277. [doi: 10.1177/2158244017712774]
43. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2021.
44. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, et al. Welcome to the Tidyverse. JOSS 2019 Nov;4(43):1686. [doi: 10.21105/joss.01686]
45. Wickham H. ggplot2: 3.3.5. Vienna, Austria: R Foundation for Statistical Computing; 2021.
46. Delacre M, Lakens D, Leys C. Why psychologists should by default use Welch's t-test instead of Student's t-test. Int Rev Soc Psychol 2017 Apr 05;30(1):92. [doi: 10.5334/irsp.82]
47. Kassambara A. rstatix: Pipe-Friendly Framework for Basic Statistical Tests. URL: https://CRAN.R-project.org/package=rstatix [accessed 2022-03-07]
48. Champely S. pwr: Basic Functions for Power Analysis. URL: https://cran.r-project.org/package=pwr [accessed 2022-03-07]
49. United States Census Bureau. Educational Attainment in the United States. URL: https://www.census.gov/data/tables/2020/demo/educational-attainment/cps-detailed-tables.html [accessed 2022-03-07]
50. Röbbelen A, Schmieding ML, Kopka M, Balzer F, Feufel MA. Data set supplementing "Interactive versus static decision support tools for COVID-19: An experimental comparison. Zenodo 2021 Nov 6. [doi: 10.5281/ZENODO.5651117]

Abbreviations

CDC: Centers for Disease Control and Prevention

DST: decision support tool
PtDA: patient decision aid
SC: symptom checker

Edited by T Sanchez; submitted 23.09.21; peer-reviewed by J Michel, A Hill; comments to author 01.10.21; revised version received 21.11.21; accepted 30.11.21; published 15.04.22.

Please cite as:
Röbbelen A, Schmieding ML, Kopka M, Balzer F, Feufel MA
Interactive Versus Static Decision Support Tools for COVID-19: Randomized Controlled Trial
JMIR Public Health Surveill 2022;8(4):e33733
URL: https://publichealth.jmir.org/2022/4/e33733
doi:10.2196/33733
PMID:34882571

Review

eHealth Interventions to Address HIV and Other Sexually Transmitted Infections, Sexual Risk Behavior, Substance Use, and Mental Ill-health in Men Who Have Sex With Men: Systematic Review and Meta-analysis

GJ Melendez-Torres[1], DPhil, MPH, RN; Rebecca Meiksin[2], MPH; T Charles Witzel[2], PhD; Peter Weatherburn[2], MSc; Jane Falconer[2], MA; Chris Bonell[2], PhD

[1]University of Exeter Medical School, Exeter, United Kingdom

[2]London School of Hygiene and Tropical Medicine, London, United Kingdom

Corresponding Author:
GJ Melendez-Torres, DPhil, MPH, RN
University of Exeter Medical School
St Luke's Campus
Heavitree Road
Exeter, EX1 2LU
United Kingdom
Phone: 44 07847360259
Email: g.j.melendez-torres@exeter.ac.uk

Abstract

Background: Men who have sex with men experience disproportionately high levels of HIV and other sexually transmitted infections (STIs), sexual risk behavior, substance use, and mental ill-health. These experiences are interrelated, and these interrelations are potentiated by structural conditions of discrimination, stigma, and unequal access to appropriate health services, and they magnify each other and have intersecting causal pathways, worsening both risk for each condition and risk for the negative sequelae of each condition. eHealth interventions could address these issues simultaneously and thus have wide-ranging and greater effects than would be for any 1 outcome alone.

Objective: We systematically reviewed the evidence for the effectiveness of eHealth interventions in addressing these outcomes separately or together.

Methods: We searched 19 databases for randomized trials of interactive or noninteractive eHealth interventions delivered via mobile phone apps, internet, or other electronic media to populations consisting entirely or principally of men who have sex with men to prevent HIV, STIs, sexual risk behavior, alcohol and drug use, or common mental illnesses. We extracted data and appraised each study, estimated meta-analyses where possible by using random effects and robust variance estimation, and assessed the certainty of our findings (closeness of the estimated effect to the true effect) by using GRADE (Grading of Recommendations, Assessment, Development and Evaluations).

Results: We included 14 trials, of which 13 included active versus control comparisons; none reported mental health outcomes, and all drew from 12 months or less of follow-up postintervention. Findings for STIs drew on low numbers of studies and did not suggest consistent short-term (<3 months postintervention; d=0.17, 95% CI –0.18 to 0.52; I^2=0%; 2 studies) or midterm (3-12 months postintervention, no meta-analysis, 1 study) evidence of effectiveness. Eight studies considering sexual risk behavior outcomes suggested a short-term, nonsignificant reduction (d=–0.14, 95% CI –0.30 to 0.03) with very low certainty, but 6 studies reporting midterm follow-ups suggested a significant impact on reducing sexual risk behavior (d=–0.12, 95% CI –0.19 to –0.05) with low certainty. Meta-analyses could not be undertaken for alcohol and drug use (2 heterogeneous studies) or for HIV infections (1 study for each of short-term or midterm follow-up), and alcohol outcomes alone were not captured in the included studies. Certainty was graded as low to very low for most outcomes, including all meta-analyses.

Conclusions: To create a comprehensive eHealth intervention that targets multiple outcomes, intervention evaluations should seek to generalize both mechanisms and components that are successfully used to achieve change in 1 outcome over multiple outcomes. However, additional evaluations of interventions seeking to address outcomes other than sexual risk behavior are needed before development and evaluation of a joined-up intervention.

(JMIR Public Health Surveill 2022;8(4):e27061) doi:10.2196/27061

KEYWORDS

men who have sex with men; HIV and sexually transmitted infections; mental health; substance use; mobile apps; HIV; eHealth; electronic media; mobile phone apps; sexual risk

Introduction

Men who have sex with men (MSM) experience HIV and other sexually transmitted infections (STIs), sexual risk behavior [1,2], substance use disorders [3-8], and mental ill-health at higher levels than the general population [9,10]. These experiences are interrelated, and these interrelations are potentiated by structural conditions of discrimination, stigma, and unequal access to appropriate health services, and they magnify each other and have intersecting causal pathways, worsening both risk for each condition and risk for the negative sequelae of each condition. This is known as a syndemic. MSM who report substance use are more likely to report condomless anal intercourse and HIV infection [11]. MSM reporting higher levels of anxiety and depression are more likely to have potential alcohol dependency [12], and MSM with depressive symptoms report more condomless anal intercourse [13]. Public health strategies to address these outcomes together therefore have the potential to achieve multiplicative effects.

eHealth interventions are those facilitated by electronic media and devices. Such interventions aim to promote healthy behaviors and mental health, for example, by challenging thought patterns that obstruct change, increasing or maintaining motivation, setting and reviewing goals, and providing feedback on behavior. eHealth interventions have the capacity to acceptably and cheaply address multiple risk behaviors and health states by assessing individual needs, identifying connections between these, and then tailoring support [14]. This has the potential for multiplicative effects, given the same individuals commonly experience multiple risk behaviors and health states where addressing on behavior or health state can predispose a change in others [15]. eHealth interventions targeting sexual health, substance use, and mental ill-health among MSM most commonly draw on the information-motivation-behavioral skills model and social cognitive theory, while also often drawing on other scientific theories of behavior and behavior change [16]. Systematic reviews focused on general or mixed populations report that eHealth interventions can reduce alcohol use [17] and address common causes of mental ill-health [18-24]. Emerging evidence also suggests that eHealth interventions might reduce drug use and sexual risk behavior. Given the clustered and interacting nature of these problems among MSM, evidence of the effectiveness of eHealth interventions for addressing each of these outcomes among MSM might suggest the value of an eHealth intervention that addresses these outcomes simultaneously and holistically. This intervention might also have multiplicative effects in ameliorating these syndemic conditions.

Our goal in undertaking this systematic review was to jointly consider the evidence for the effectiveness of eHealth interventions in addressing HIV and other STIs, sexual risk behavior, substance use, and mental ill-health separately or together in order to assess whether the evidence base supports the development of a single eHealth intervention for MSM for these syndemically related conditions.

Methods

This was part of a larger evidence synthesis project also examining theories of change and implementation studies for eHealth interventions for MSM. Our systematic review was registered in PROSPERO (International Prospective Register of Systematic Reviews, CRD42018110317).

Inclusion and Exclusion Criteria

We included randomized trials, without regard to comparator, of interactive or noninteractive eHealth interventions delivered via mobile phone apps, internet, or other electronic media (ie, electronic communication technology) and delivered to populations consisting entirely or principally of MSM. We also required interventions to address prevention of HIV, STIs, sexual risk behavior, alcohol and drug use, or common mental illnesses as outcomes representing, for example, incident cases of STIs and HIV, number of condomless sex partners, substance use frequency or prevalence, or symptomatology or diagnoses of anxiety or depression. We excluded (1) eHealth interventions merely facilitating one-off contact, (2) those addressing HIV self-testing, clinic attendance, or STI partner notification only, and (3) interventions delivered by human providers via electronic media. These interventions were excluded to better facilitate an understanding of the effectiveness of "fully eHealth" interventions delivered over a period of time on individual health behaviors rather than on the management of STIs—one-off health education interventions delivered via video or other opportunistic interventions.

Study Search and Selection

To locate studies, we searched 19 databases in October and November 2018, updating searches in April 2020. We also searched 3 trial registers to identify ongoing recently published or otherwise unindexed trials alongside searches for grey literature. The full details of the databases and a sample search string are shown in Multimedia Appendix 1. We searched included studies' reference lists and contacted subject experts. We deduplicated the retrieved references and uploaded these to EPPI-Reviewer (v 4.0, EPPI-Center). An inclusion criteria worksheet with guidance notes was prepared and piloted by 2 reviewers screening batches of the same 50 references. Where the 2 reviewers disagreed, they met to discuss this and if possible, reach a consensus, with recourse to a third reviewer if necessary. Once the inclusion criteria worksheet generated 95% agreement, each title and abstract was screened by 1 reviewer. Two reviewers assessed each full text that passed screening.

Data Extraction and Appraisal

Two reviewers independently extracted data from outcome evaluations by using existing tools, meeting to discuss in cases of disagreement. We extracted data on basic study details (target population, study location, timing and duration, research questions, or hypotheses); methods (design, sampling and sample size, data collection, and analysis); and intervention description (timing and duration, program development, theoretical framework/logic model, content and activities, providers, details of any intervention offered to the control group), as well as allocation; sequence generation and concealment; measures, follow-up, and blinding; retention; and data on outcomes. Trials were appraised using the Cochrane risk of bias tool [25]. Where there was a risk of missing data from published reports affecting our analysis, we contacted authors wherever possible to request additional information. The risk of bias domains were used to inform GRADE (Grading of Recommendations, Assessment, Development and Evaluations) tables, which reflects the certainty of evidence for each outcome and time point. GRADE is a tool for researchers to describe how close the effect estimated in a meta-analysis is to the "true" effect of the intervention.

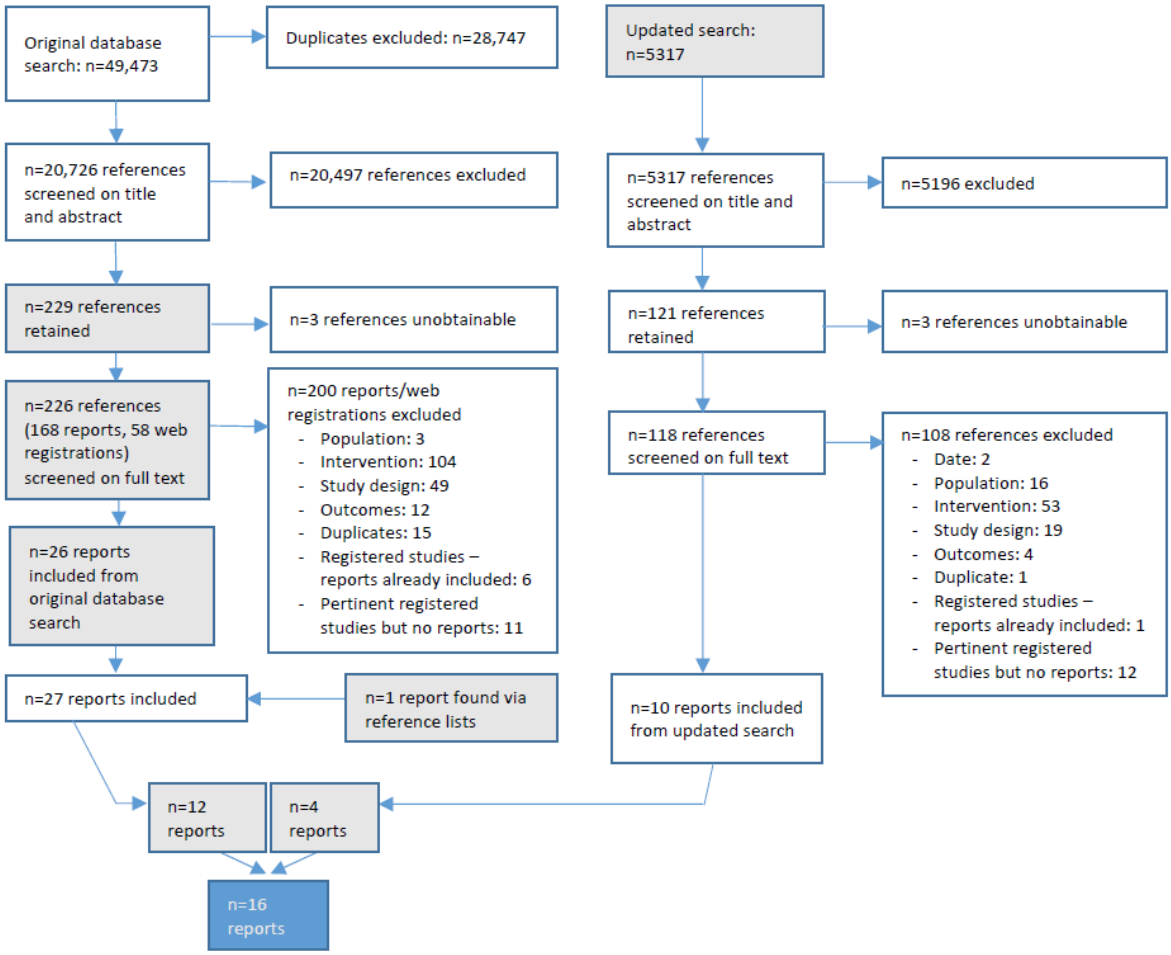
Meta-analysis Methods

We narratively synthesized outcome evaluations by type of comparison and then by outcome, thereby grouping effect estimates by a follow-up duration after the intervention (short-term involving less than 3 months and medium-term involving 3 months up to 1 year). We defined type of comparison as either active versus control, where interventions were compared against a business-as-usual or no-treatment control, and active versus active, where interventions were compared against other eligible eHealth interventions. Where necessary, we rebased follow-up times using the stated intervention length but report in our narrative synthesis follow-up times as described in original reports. After converting effect estimates to standardized mean differences, we used a robust variance estimation meta-analysis with random effects to synthesize outcomes by follow-up time and then across all follow-up times and quantified heterogeneity by using I^2. Robust variance estimation permits the inclusion of multiple effect sizes per study by accounting for nonindependence in standard errors. If an indication of substantial heterogeneity was determined with fewer than 3 studies (eg, study-level I^2 greater than 50%), we did not present a pooled estimate by follow-up time or across follow-up times. We were unable to assess publication bias owing to the number of effect sizes in any 1 meta-analysis. We intended to undertake a network meta-analysis but were unable to do so owing to the low number and variable quality of studies.

Results

Our original search yielded 49,473 references, with 20,726 remaining after deduplication (see Figure 1). We included a further 5317 references as a result of our updated search. From these, 344 references were not excluded on title and abstract and were screened on full texts over both waves of searching, yielding 16 records of 14 trials included in our syntheses.

Figure 1. Study flowchart: outcome evaluations of eHealth interventions to address HIV and other sexually transmitted infections, sexual risk behavior, substance use, and mental ill-health in men who have sex with men.



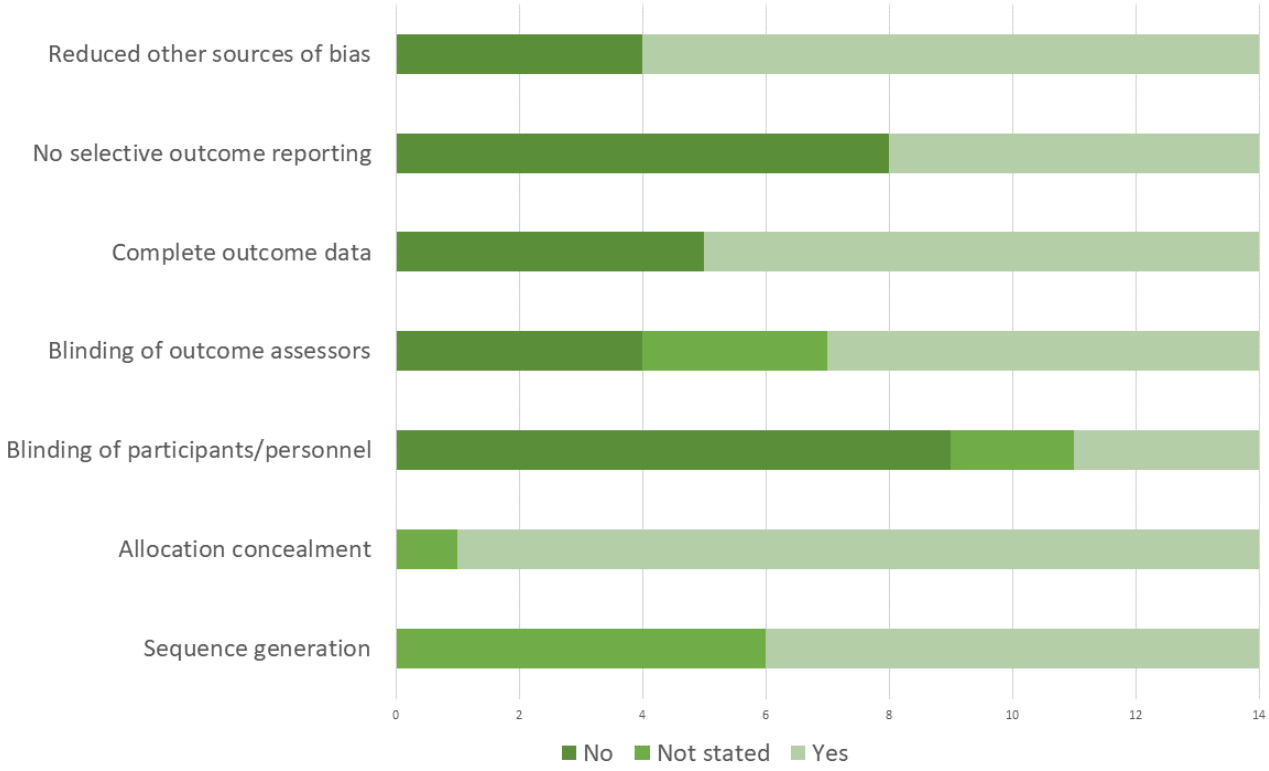
Included Trials and Interventions

The included trials [26-39] were published between 2006 and 2020 (see Multimedia Appendix 2). Of the included trials and included arms within those trials, 13 compared active intervention versus no treatment [26-38], while 2 [31,39] compared active interventions against each other. All but 4 trials were conducted in the United States: 1 was undertaken in the Netherlands [31], 1 in China [28], 1 in Taiwan [29], and 1 in Sweden [38]. We categorized interventions into time-limited or modular (guiding participants sequentially through intervention content from beginning to end; 11 trials) [26-28,30-32,34,35,37-39] and open-ended interventions (either user-responsive [36], static website content [33], or a multifeature app [29]; 3 trials). Interventions were generally delivered via website (including mobile website), with the exception of TXT-Auto [36], Safe Behavior and Screening (SBS) [29], and 1 downloaded videogame [30]. Within the 11 trials of time-limited or modular interventions, interventions had different foci. For example, myDEx [26] focused on comprehensive sexual education for young people, and China-Gates HIV Prevention Program [28], Hot and Safe M4M.org [27], Keep it Up! [34,35], Sexpulse [37], and WRAPP [38,39] presented general HIV prevention and sexual health content—all in interactive web-based modular formats—whereas SOLVE [30] used a videogame format. Davidovich et al [31] tested a tailored and a nontailored version of a web-based modular noninteractive intervention termed as cognitive vaccine, while Hirshfield et al [32] presented a modular video series, SexPositive. The 3 trials using open-ended interventions were delivered using text messages (TXT-Auto) [36], a mobile multifeature app (SBS) [29], or a static website [33].

Quality of the Included Trials

We present in Figure 2 a summary risk of the bias graph, presenting review-level findings by domain. Trials generally had acceptable sequence generation and allocation concealment; however, blinding of outcome assessors was not always reported. Most trials presented complete outcome data, but 5 trials had high attrition, with imbalanced attrition between arms. Eight of the included trials appeared to be at high risk of bias owing to selective outcome reporting, including differences between reported outcomes and outcomes listed in the protocol. Study-level judgments and GRADE tables for each analysis are available in Multimedia Appendix 3.

Figure 2. Summary risk of bias graph: outcome evaluations of eHealth interventions to address HIV and other sexually transmitted infections, sexual risk behavior, substance use, and mental ill-health in men who have sex with men.



Mental Health

No included studies presented relevant mental health outcomes.

Alcohol and Drug Use

Two studies [29,36] presented the estimates for drug use, both focused on open-ended interventions, but none included alcohol use specifically. Although both studies presented short-term results, only Reback et al [36] presented midterm results. Effect estimates in none of the follow-up categories suggested consistent evidence of effectiveness (see Table 1).

Table 1. Meta-analysis findings by outcome and follow-up time: eHealth interventions in men who have sex with men.

Outcome	Follow-up[a]	Results	Certainty
Alcohol and drug use	Short-term	No meta-analysis; 2 heterogeneous studies [29,36]	Very low
Alcohol and drug use	Midterm	No meta-analysis; 1 study [36]	Very low
HIV infections	Short-term	No meta-analysis; 1 study [29]	Low
HIV infections	Midterm	No meta-analysis; 1 study [35]	Very low
Sexually transmitted infections	Short-term	d=0.17, 95% CI –0.18 to 0.52, I^2=0%, 2 studies [29,33]	Very low
Sexually transmitted infections	Midterm	No meta-analysis; 1 study included [35]	Moderate
Sexually transmitted infections	Overall	d=0.07, 95% CI –0.79 to 0.94, I^2=16%, 3 studies [29,33,35]	N/A[b]
Sexual risk	Short-term	d=–0.14, 95% CI –0.30 to 0.03, I^2=61%, 8 studies [26,27,29,32,34-37]	Very low
Sexual risk	Midterm	d=–0.12, 95% CI –0.19 to –0.05, I^2=27%, 6 studies [28,31,32,35-37]	Low
Sexual risk	Overall	d=–0.15, 95% CI –0.26 to –0.05, I^2=56%, 10 studies [26-29,31,32,34-37]	

[a]Short-term: up to 3 months postintervention; midterm: 3 months to a year after the intervention.

[b]N/A: Not applicable.

Short-term Results

After 6 months of app use (ie, at 6 months postrandomization), Chiou et al [29] found that the SBS open-ended intervention with general content reduced drug use as measured on a 5-point Likert scale (β=–1.19, SE 0.204, P<.001). In Reback et al [36], neither at the postintervention follow-up at 8 weeks postrandomization (d=0.15, 95% CI –0.16 to 0.45) nor at 3

months postrandomization (*d*=–0.03, 95% CI –0.34 to 0.28) was there evidence of a difference between TXT-Auto and the assessment-only condition on days of methamphetamine use. Statistical heterogeneity precluded meta-analysis of the results, and certainty of the assessment of evidence was graded as very low owing to risk of bias (details of randomization), inconsistency of studies, and imprecision of effect estimates.

Midterm Results

At 6 months postrandomization, there was no evidence of a difference between the TXT-Auto intervention tested in Reback et al [36] against the assessment-only condition on days of methamphetamine use (*d*=0.23, 95% CI –0.07 to 0.54). A similar estimate was produced at 9 months postrandomization (*d*=0.28, 95% CI –0.02 to 0.59). Although both estimates suggested a possible intervention effect of increased days of methamphetamine use in the intervention arm, authors noted that this could have been due to baseline imbalance.

HIV Infections

Two studies presented findings relating to incident HIV infections, the former reporting on an open-ended intervention and the latter on a time-limited modular intervention [29,35]. In the short-term (at postintervention follow-up), Chiou et al [29] found an incidence rate ratio of 1.56 (95% CI 0.26-9.56) comparing SBS against a control group (see Table 1). In the midterm (at 6 months postintervention), Mustanski et al [35] found that incident HIV diagnoses were not different in the Keep It Up! intervention arm (9 diagnoses over 384 person-years) than in the control arm (8 diagnoses over 410 person-years). Included interventions did not have an overall impact on HIV infections, with an increase in incidence of HIV infections equivalent to 0.12 standard deviations but an imprecisely estimated confidence interval that included the point of no effect (95% CI –0.34 to 0.59). Certainty in the assessment of the evidence ranged from very low to low owing to risk of bias (selective outcome reporting) and imprecision of effect estimates.

STIs

Three studies presented estimates for STIs: Chiou et al [29], Milam et al [33], and Mustanski et al [35]. Of these, Chiou et al [29] and Milam et al [33] presented short-term results on open-ended interventions, whereas Mustanski et al [35] presented midterm results on a time-limited/modular intervention. There was no evidence of short-term impacts on incident STIs, while there was some evidence, albeit from 1 study, for midterm impacts on incident STIs.

Short-term Results

At postintervention, Chiou et al [29] reported an incidence rate ratio of 1.39 (95% CI 0.31-6.37) for incident syphilis infections in the SBS group as compared to control. Milam et al [33] reported rates of any incident bacterial STIs (syphilis, gonorrhea, or chlamydia) over 12 months, which was the intervention period, finding that the proportions in each arm (30% intervention vs 25% control) were not statistically different (*P*=.50) and that the distribution of visits with new STIs per subject did not differ between arms (*P*=.57).

Midterm Results

Mustanski et al [35] reported results for several STIs, both individually and as a composite outcome, at 6 months postintervention. Findings were principally reported as risk ratio (RR) and suggested a statistically significant 40% difference in risk of any STI diagnosis (RR 0.60, 95% CI 0.38-0.95). Findings for individual STIs were not significant—specifically, urethral chlamydia (RR 0.60, 95% CI 0.13-2.34), urethral gonorrhea (RR 0.35, 95% CI 0.01-4.33), rectal chlamydia (RR 0.61, 95% CI 0.34-1.06), or rectal gonorrhea (RR 0.91, 95% CI 0.40-2.05). Another analysis of the outcome drew on a matched-pair analysis and estimated a within-subject reduction in risk of 68% for any STI (95% CI 0.40-0.83).

Meta-analysis

A meta-analysis of effect sizes with follow-up of less than 3 months (see Table 1) included 2 effect sizes from 2 studies reporting on open-ended interventions and suggested a nonsignificant increase in STIs in the intervention group as compared to that in the control group (*d*=0.17, 95% CI –0.18 to 0.52), with heterogeneity not meaningfully present in this meta-analysis (I^2=0%) [29,33]. The certainty of evidence was very low due to risk of bias (details of randomization, selective outcome reporting) and imprecision of effect estimates. The overall analysis across short- and medium-term follow-up and intervention types drew on 3 studies contributing 7 effect sizes and suggested a small and nonsignificant increase in STIs in the intervention group as compared to the control group (*d*=0.07, 95% CI –0.79 to 0.94) and low heterogeneity (I^2=16%) [29,33,35]. Plots are presented in Multimedia Appendix 4 [26-39].

Sexual Risk Outcomes

Eleven studies presented estimates for sexual risk outcomes [26-32,34-36]. One further study [38] intended to present short-term results relating to sexual risk outcomes but did not estimate an effect owing to an unexpectedly low sample size. Sexual risk outcomes are organized into condomless anal intercourse (at the time of included studies' publication usually referred to as unprotected anal intercourse), condom use, serononconcordant sex acts, and sex acts under the influence of drugs and alcohol.

Short-term Results

Nine studies presented short-term results [26,27,29,30,32,34-37]. Of these, Chiou et al [29] and Reback et al [36] report on open-ended interventions while the rest report on time-limited/modular interventions. Results are presented by type of outcome: condomless sex, condom use during sex, HIV serononconcordant sex, and sex under the influence of drugs.

Six studies presenting short-term results for condomless sex yielded inconsistent evidence as to the effectiveness of interventions on this outcome [26,27,34-37]. All [26,27,34,35,37] but Reback et al [36] reported on time-limited/modular interventions. First, in their evaluation of the myDEx intervention, Bauermeister et al [26] found that at 90 days after the randomization (ie, at postintervention), intervention recipients had significantly lower odds than

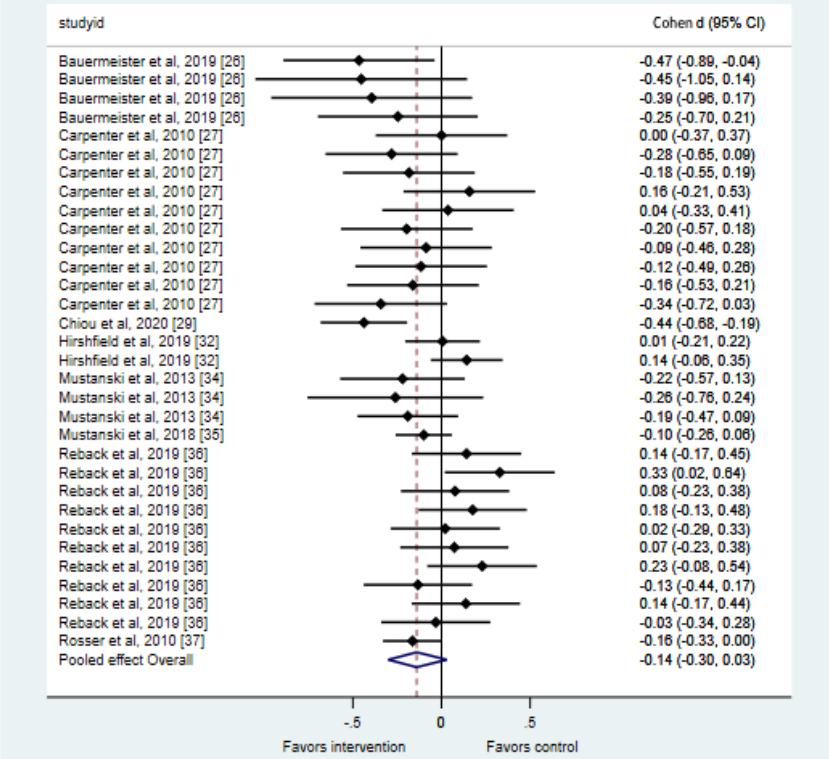
attention control recipients of any condomless receptive anal intercourse during the prior 3 months (odds ratio [OR] 0.43, 95% CI 0.20-0.94). The effect was lower in magnitude and nonsignificant for any condomless insertive anal intercourse (OR 0.64, 95% CI 0.28-1.44). However, at 3 months postbaseline, Carpenter et al [27] did not find that the Hot and Safe M4M intervention generated significant differences between groups on any condomless anal intercourse, condomless insertive anal intercourse, condomless receptive anal intercourse, condomless insertive oral intercourse, or condomless receptive oral intercourse. Specific significance tests per outcome were not provided, though calculated standardized mean differences (see Figure 3 for specific estimates) did not suggest a significant impact of the intervention. In the first evaluation of the Keep it Up! intervention, Mustanski et al [34] found that at 12 weeks postintervention, those who received the program had a lower rate ratio of condomless anal intercourse acts (rate ratio=0.56, *P*=.04; n=63). However, in the second evaluation of Keep it Up! [35], there was no significant difference between groups on any condomless anal intercourse, numbers of male condomless anal intercourse partners overall, or number of casual condomless anal intercourse partners at 3 months postrandomization, although specific significance tests were not reported for these outcomes. Calculated standardized mean differences between groups on reports of any condomless anal intercourse did not suggest a significant difference (*d*=–0.10, 95% CI –0.26 to 0.06). Rosser et al [37] estimated that the Sexpulse intervention reduced the number of male condomless anal intercourse partners by 16.8% at 3 months postbaseline, though this effect was only marginally significant (95% CI 0.691-1.000). Finally, reporting on an open-ended intervention, Reback et al [36] did not undertake endpoint-specific tests for condomless anal intercourse outcomes; however, we calculated that the intervention did not reduce episodes of condomless anal intercourse with main partners, anonymous partners, partners for transactional sex, or casual partners at 8 weeks or 3 months postbaseline (see Figure 3 for specific estimates). There was some suggestion of a harmful effect in terms of the intervention group having a higher number of condomless anal intercourse episodes with casual partners at 8 weeks postbaseline, but this may have been due to substantial baseline imbalance.

A study by Christensen et al [30] examined the mediating impact of shame on number of condomless anal intercourse events in the prior 3 months in the SOLVE time-limited/modular intervention at 3 months postbaseline. Estimates of the intervention's total impact on condomless anal intercourse were not presented, but the significantly reported indirect effect on condomless anal intercourse through shame suggests a significant total effect of the intervention on condomless anal intercourse. However, these estimates are not directly comparable to the other tests of intervention impact presented here.

Two studies presented short-term results for condom use, and together, they reported mixed evidence of effectiveness on this outcome [29,34]. At 6 months postbaseline, Chiou et al [29] found that the SBS open-ended intervention increased the proportion of anal intercourse encounters where condoms were used by 20.7% (SE 0.058, *P*=.001) [29]. Similarly, in the first evaluation of the Keep it Up! time-limited/modular intervention, Mustanski et al [34] showed a reduced number of condom use errors (*d*=–0.19, *P*=.56) and condom failures (*d*=–0.22, *P*=.30), but not significantly.

Three studies presented short-term results for HIV serononconcordant sex acts and yielded mixed evidence on the effectiveness of time-limited/modular interventions for this outcome [26,27,32]. In their evaluation of the myDEx intervention, Bauermeister and colleagues [26] found that at 90 days postrandomization (ie, at postintervention), intervention recipients had lower odds than attention control recipients of any condomless receptive anal intercourse with serodiscordant or serounknown partners not known to be on pre-exposure prophylaxis or virally suppressed during the prior 3 months but not significantly so (OR 0.44, 95% CI 0.15-1.31); a similar pattern was found for insertive anal intercourse (OR 0.49, 95% CI 0.17-1.33). However, at 3 months postbaseline, Carpenter et al [27] found that the Hot and Safe M4M intervention generated greater reductions in all condomless anal intercourse events with partners of positive or unknown serostatus (group by time F_{1101}=7.59, *P*=.007), including condomless insertive anal intercourse (group by time F_{1101}=7.24, *P*=.008) but not condomless receptive anal intercourse (group by time F_{1101}=1.35, *P*=.25). The intervention group also reported reduced condomless insertive oral intercourse events (group by time F_{1101}=7.45, *P*=.007) and reduced condomless receptive oral intercourse events with partners of positive or unknown serostatus (group by time F_{1101}=8.45, *P*=.004). Hirshfield et al [32] found that at 3 months postbaseline, the SexPositive intervention did not reduce either the number of condomless anal intercourse partners known to be serodiscordant (adjusted standardized β=.003, 95% CI –.168 to .178) or the number of condomless anal intercourse partners not known to be seroconcordant (adjusted standardized β=–.073, 95% CI –.332 to .051). Reback et al [36] was the only study to present short-term results for this category of sexual risk outcomes, reporting on an open-ended intervention. At neither 8 weeks postbaseline (*d*=0.23, 95% CI –0.08 to 0.54) nor 3 months postbaseline (*d*=0.08, 95% CI –0.23 to 0.38) was there a significant difference between groups on episodes of sex while on methamphetamine.

Figure 3. Short-term estimates of effects of eHealth interventions on sexual risk behaviors in men who have sex with men [26,27,29,32,34-37].



Midterm Results

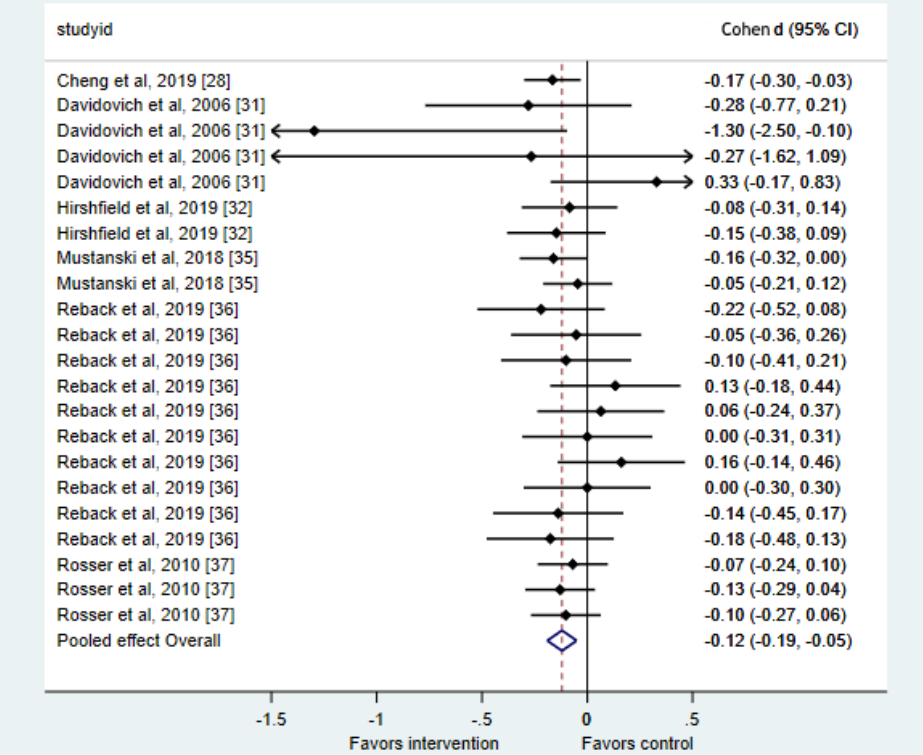
Six studies presented midterm results—all but the last focused on time-limited/modular interventions [28,31,32,35-37]. Results are presented by type of outcome: condomless sex, condom use, serononconcordant sex acts, and sex acts under the influence of drugs.

Four studies presenting midterm results for condomless sex yielded inconsistent evidence as to the effectiveness of interventions on this outcome [28,35-37]. Let us first consider time-limited/modular interventions. First, in the evaluation of the China-Gates HIV Prevention Program, intervention participants were less likely than control participants at 6 months postbaseline to report condomless anal intercourse in the last 3 months with a risk difference of 9.3% (95% CI 1.1-17.5); estimates using multiple imputation to include the entire sample generated a similar estimate (8.9%, 95% CI 1.2-16.6) [28]. In the second evaluation of Keep it Up!, there was no significant difference between groups on numbers of male casual condomless anal intercourse acts and number of condomless anal intercourse partners at 6 or 12 months postrandomization, though specific significance tests were not reported for these outcomes [35]. However, at 12 months after the randomization, intervention participants were 17% less likely to report any condomless anal intercourse in the prior 3 months (95% CI 0.70-0.99). Rosser et al [37] estimated that the Sexpulse intervention did not reduce the number of male partners with condomless anal intercourse at 12 months postbaseline (incidence rate ratio 0.998, 95% CI 0.952-1.046). We also calculated that there was no evidence of a significant effect at 6 months (d=–0.13, 95% CI –0.29 to 0.04) or 9 months (d=–0.10, 95% CI –0.27 to 0.06) postbaseline. Regarding open-ended interventions, Reback et al [36] did not undertake endpoint-specific tests for condomless anal intercourse outcomes; however, we calculated that the intervention did not reduce episodes of condomless anal intercourse with main partners, anonymous partners, partners for transactional sex, or casual partners 6 months or 9 months postbaseline (see Figure 4 for specific estimates).

Studies of this outcome focused on time-limited/modular interventions only. Davidovich et al [31] presented midterm results for condom use with no evidence of effectiveness, whereas this study and that by Hirshfield et al [32] presented inconsistent evidence of intervention effectiveness on midterm results for serononconcordant sex acts. In Davidovich et al [31], intervention participants receiving the tailored intervention were significantly more likely than control participants to practice negotiated safety (seroconcordant condomless anal intercourse or no condom use only in the context of monogamous relationships) as compared to condomless anal intercourse with partners of unknown HIV concordance if they received the tailored intervention (OR 10.50, 95% CI 1.19-92.72); however, intervention participants receiving the nontailored intervention did not show a significant difference on this outcome as compared to control recipients (OR 1.62, 95% CI 0.14-19.07) [31]. In this same study, intervention participants were not significantly different in the odds of condom use as compared to the control participants at 6 months postbaseline. This was the case comparing either the tailored version of the intervention (OR 1.66, 95% CI 0.68-4.02) or the nontailored version of the intervention (OR 0.55, 95% CI 0.22-1.37) against control, with OR values greater than 1, suggesting increased condom use. However, it should be noted that in this analysis, negotiated safety, condom use, and other condomless anal intercourse were mutually exclusive categories estimated as part of a multinomial regression model. In addition, Hirshfield et al [32] found that

at 12 months postbaseline, the SexPositive intervention did not significantly reduce either the number of condomless anal intercourse partners known to be serodiscordant (adjusted standardized β=–.073, 95% CI –.332 to .051) or the number not specifically known to be seroconcordant (adjusted standardized β=–.084, 95% CI –.399 to .045). Reback et al [36] was the only study to present medium-term results for this category of sexual risk outcomes, reporting on an open-ended intervention. We calculated the differences by using endpoint means. At neither 6 months postbaseline (*d*=–0.10, 95% CI –0.41 to 0.21) nor 9 months postbaseline (*d*=–0.18, 95% CI –0.48 to 0.13) was there a significant difference between groups on episodes of sex while using methamphetamine.

Figure 4. Midterm estimates of effects of eHealth interventions on sexual risk behaviors in men who have sex with men [28,31,32,35-37].



Meta-analysis

The effect sizes for sexual risk outcomes are presented in Table 1 as well as in Figure 3 for outcomes at follow-ups of less than 3 months postintervention (short-term) and in Figure 4 for outcomes between 3 months and 1 year after the intervention (midterm). In both plots, negative effect sizes represent benefits. A meta-analysis drawing on 32 effect sizes from 8 studies with less than 3 months follow-up [26,27,29,32,34-37] found a suggestion of effectiveness, albeit not statistically significant (*d*=–0.14, 95% CI –0.30 to 0.03) with substantial heterogeneity (I^2=61%). The certainty of evidence for this meta-analysis was graded as very low owing to risk of bias (details of randomization, selective outcome reporting), inconsistency of studies, and publication bias arising from the noninclusion of 2 studies [30,38]. A meta-analysis drawing on 22 effect sizes from 6 studies with 3 months to 1 year of follow-up suggested a significant impact on reducing sexual risk (*d*=–0.12, 95% CI –0.19 to –0.05) but with low heterogeneity (I^2=27%). The certainty of evidence for this meta-analysis was graded as low owing to risk of bias (details of randomization, selective outcome reporting). We then pooled estimates regardless of follow-up time. Based on 54 effect sizes from 10 studies [26-29,31,32,34-37], interventions significantly reduced the sexual risk as compared to control groups (*d*=–0.15, 95% CI –0.26 to 0.05). This meta-analysis had substantial heterogeneity (I^2=56%). To explore this heterogeneity, we compared interactive interventions [26-29,34-37] with noninteractive interventions [31,32]. A random-effects meta-regression did not suggest that this characteristic accounted for heterogeneity, with noninteractive interventions not meaningfully worse than interactive interventions (β=.12, 95% CI –.66 to .89).

Active Versus Active Comparisons

Two included studies included comparisons between active interventions, which were time-limited/modular in approach [31,39]. Both studies reported sexual risk outcomes only, and only Davidovich et al [31] presented extractable outcome data. Thus, meta-analysis was not undertaken. In Davidovich et al [31], we calculated that participants receiving the tailored intervention were more likely than participants receiving the nontailored intervention to report, at 6 months postbaseline, the practice of negotiated safety (defined above) as compared to condomless anal intercourse with other partners (OR 6.50, 95% CI 2.49-16.90) and the practice of condom use as compared to condomless anal intercourse (OR 2.98, 95% CI 1.74-5.12). In their evaluation of time-limited interactive web-based modular interventions, Bowen et al [39] tested the impact of ordering modules with content about HIV knowledge, content about risk in casual or new partnerships, and content about contexts of risk on the proportion of anal intercourse partners with whom a condom was used every time. At postintervention, there was

no statistical difference between modules on sexual risk, but specific group differences were not presented.

Discussion

Summary of Findings

Our systematic review of intervention effectiveness included 14 trials, of which 13 included active versus control comparisons. Trials included substance use, HIV, and STIs, and sexual risk behavior outcomes but not mental health outcomes. Substance use outcomes did not include alcohol use. Furthermore, all outcome estimates drew from 12 months or less of follow-up postintervention. A further 2 trials presented active versus active comparisons [31,39]. Neither trial found a difference between tested interventions on sexual risk outcomes and thus are not discussed further. Moreover, equity-relevant characteristics, for example, moderation of intervention effectiveness by income, ethnicity, and other social variables were not meaningfully addressed by this body of evidence.

In active versus control comparisons, findings for drug use drew on 2 studies in short-term follow-up [29,36], which could not be meta-analyzed owing to extreme heterogeneity, and 1 study in midterm follow-up [36]. Together, these studies did not present consistent evidence of effectiveness, and the GRADE profile for both analyses suggested the certainty of evidence was very low. Analysis for HIV infection also drew on 2 studies—1 in short-term follow-up [29] and 1 in midterm follow-up [35]. Neither study suggested that an eHealth intervention was effective at reducing infections, though short follow-up times and low event rates precluded meaningful comparison. Again, the GRADE profile suggested that the certainty of these findings was low or very low. Analyses for STIs were similarly scant, drawing on 2 trials in the short-term [29,33] and 1 trial in the midterm follow-up [35]. Although a pooled analysis of short-term follow-ups suggested no impact of the interventions on incident STIs with very low certainty, the 1 trial informing the midterm follow-up did suggest a meaningful and statistically significant reduction in incident STIs, with corresponding moderate certainty [35].

The largest analyses assessed sexual risk behavior outcomes. Although the GRADE profile suggested that the certainty of the conclusions was very low or low owing primarily to the risk of bias in the included trials and possible publication bias, pooled estimates from midterm follow-ups drawing on 6 trials suggested a small and statistically significant impact of eHealth interventions in reducing the sexual risk behavior (d=–0.12). Pooled estimates from short-term follow-ups drawing on 8 trials did not reach statistical significance but suggested a trend toward reductions in sexual risk behavior of similar magnitude (d=–0.14). We tested whether the interactivity of the interventions was related to intervention effectiveness on sexual risk behaviors; however, a meta-regression did not suggest significant differences between interventions on the basis of this characteristic.

Implications

The quality and quantity of evidence supporting the effectiveness on eHealth interventions was low or, in the case of mental health outcomes, nonexistent for most of the outcomes we set out to analyze. Even where meta-analyses drew on multiple studies, issues with included trials precluded certainty in the evidence presented. Moreover, despite substantial heterogeneity in meta-analyses for sexual risk behavior outcomes, we were unable to explain this heterogeneity. Another key gap in this systematic review related to the inclusion of outcomes that accurately reflect current knowledge about minimizing sexual risk. For example, a focus on condom use does not reflect that risk for HIV can be managed through effective biomedical means such as adherence to HIV treatment for people living with HIV or pre-exposure prophylaxis. It is likely that interventions designed in the current context would more explicitly acknowledge biomedical approaches to managing risk.

One of the questions we set out to address in this systematic review was if the existing evidence overall suggests that our scoped outcomes could coherently, feasibly, and effectively be addressed by a single health intervention. It is clear based on the meta-analyses presented that the evidence does not, as yet, suggest that this is the case, despite the consistent acceptability of these interventions we documented in a linked systematic review of process evaluations [40]. This is largely because interventions, with few exceptions, were focused on individual outcomes as well as the patchy effects for outcomes that were assessed. Only 2 interventions captured substance use outcomes alongside sexual risk behavior outcomes [29,36], where trials otherwise reported outcomes over multiple categories; this was between sexual risk behavior and either HIV or other STIs. No study reported outcomes for depression or anxiety, despite poor mental health being a key syndemic condition nor did any study report outcomes relating to alcohol use, despite this being the most commonly misused substance in high-income country settings.

In order to remedy this, future trials of eHealth interventions should include several considerations. First, given the complete lack of evidence in this area, trials should consider how to address poor mental health in MSM, with a focus on how determinants of poor mental health in MSM both relate to outcomes considered here (sexual ill-health, substance use) and other antecedents (eg, stigma). Second, trials should address a range of substance use behaviors that are syndemically linked to other relevant outcomes (mental well-being and sexual health), including alcohol use. Third, trials should incorporate follow-ups long enough and sample sizes large enough to detect a meaningful impact on HIV and other STIs, given the time needed to detect meaningful differences in HIV infection incidence. Alternatively, given the challenge associated with attaining sample sizes large enough to demonstrate impacts on STIs, studies could include a wider range of behavioral indicators closely linked with reduced STIs to improve confidence that a distal improvement in incident infections is likely. Most included studies reported only 1 or 2 relevant outcomes relating to sexual risk outcomes. Fourth, trials should ensure that interventions are not inequality-generating and that this is examined empirically. This is important because interventions may unintentionally exacerbate inequalities within groups due to, for example, differential access to mental or

sexual health services or substance use services. Finally, to generate a joined-up comprehensive eHealth intervention that targets multiple outcomes, intervention evaluations should seek to generalize both mechanisms and components that are successfully used to achieve change in 1 outcome over multiple outcomes and should draw on current technologies and modalities. For example, a videogame download for a desktop computer is unlikely to be as acceptable as a specific app-based on text message–based intervention. The majority of the interventions used websites, but it was frequently unclear based on intervention descriptions how relevant these interventions would be in the current mobile health environment. It remains important to note that this is a rapidly developing field of research, and it is likely that a range of "next-generation" eHealth interventions will be evaluated in the near future. This will be valuable to understand how more recent technologies, including greater use of apps as opposed to websites, can support reducing health inequalities in MSM.

Limitations

Although we undertook an exhaustive search, the possibility remains that we were unable to include all relevant outcome evaluations. In addition, our meta-analysis drew on evidence of variable quality, with limited scope for meta-analyses. We were unable to account for heterogeneity between studies where this was present owing to either scarcity of evidence or limitations of our evidence base. In particular, we were unable to undertake network meta-analysis, and most meta-analyses had too few studies to make meta-regression, for example, comparing intervention type on outcomes, meaningful. In our analysis of sexual risk outcomes, which was the one model where we were able to undertake meta-regression, we were unable to explain heterogeneity. Meta-regressions by outcome type to determine differential effectiveness on outcomes within sexual risk would have been uninterpretable owing to the statistical methods used for meta-analysis and owing to multiple studies reporting outcomes across several domains. Probable publication and selective reporting bias across studies meant that several estimates of intervention effectiveness from the included studies could not be included in our meta-analyses; in at least one case, outcomes stated in a trial protocol were not published in the main trial report. Finally, we were unable to locate evidence for some scoped outcomes.

Conclusion

Our systematic review found that while some evidence exists for the effectiveness of eHealth interventions in addressing sexual risk in MSM, the quality of evidence was poor, as was the quality of the evidence for the range of outcomes considered. eHealth interventions present a potentially powerful avenue for addressing a range of interrelated health challenges; however, more needs to be understood about how interventions work for individual outcomes before progressing a comprehensive intervention.

Acknowledgments

CB conceptualized and led the design of the study. CB, TCW, GJMT, JF, PW, and RM contributed to the development of the study's methods. GJMT and RM quality assessed and extracted the studies. GJMT led the statistical analysis and drafting of the manuscript, with significant input from JF, CB, TCW, RM, and PW. This work was funded by the National Institute for Health Research Public Health Research Program (NIHR PHR 17/44/48). This report presents independent research commissioned by the NIHR. The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the National Health Service, the NIHR, Medical Research Council, Central Commissioning Facility, NIHR Evaluation, Trials and Studies Coordinating Centre, NIHR PHR, or the Department of Health. The funder had no role in study design, in the collection, analysis or interpretation of data, or in writing this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional search details.
[DOCX File , 18 KB - publichealth_v8i4e27061_app1.docx]

Multimedia Appendix 2
Table of included studies.
[DOCX File , 40 KB - publichealth_v8i4e27061_app2.docx]

Multimedia Appendix 3
Study-level risk of bias judgments.
[DOCX File , 23 KB - publichealth_v8i4e27061_app3.docx]

Multimedia Appendix 4
Forest plots for sexually transmitted infection outcomes.
[DOCX File , 301 KB - publichealth_v8i4e27061_app4.docx]

References

1. Aghaizu A, Brown A, Nardone A, Delpech V. HIV in the United Kingdom: 2013 report. Public Health England. URL: https://www.unaids.org/sites/default/files/country/documents/GBR_narrative_report_2014.pdf [accessed 2022-03-22]
2. Beyrer C, Baral SD, van Griensven F, Goodreau SM, Chariyalertsak S, Wirtz AL, et al. Global epidemiology of HIV infection in men who have sex with men. The Lancet 2012 Jul;380(9839):367-377. [doi: 10.1016/s0140-6736(12)60821-6]
3. Buffin J, Roy A, Williams H, Yorston C. Part of the picture: lesbian, gay and bisexual people's alcohol and drug use in England (2009-2011). The National LGB Drug & Alcohol Database. 2012. URL: http://fileserver.idpc.net/library/Part_of-The-Picture_4th_Year_ReportV4-F%20(1).pdf [accessed 2022-03-22]
4. Guasp A. The gay and bisexual men's health survey. Stonewall UK. 2012. URL: https://www.stonewall.org.uk/resources/gay-and-bisexual-men%E2%80%99s-health-survey-2013 [accessed 2022-03-22]
5. Hickson F, Bonell C, Weatherburn P, Reid D. Illicit drug use among men who have sex with men in England and Wales. Addiction Research & Theory 2010 Jan 08;18(1):14-22. [doi: 10.3109/16066350902770433]
6. Lee J, Griffin GK, Melvin C. Tobacco use among sexual minorities in the USA, 1987 to May 2007: a systematic review. Tobacco Control 2009 Aug;18(4):275-282. [doi: 10.1136/tc.2008.028241] [Medline: 19208668]
7. Schmidt AJ, Bourne A, Weatherburn P, Reid D, Marcus U, Hickson F, EMIS Network. Illicit drug use among gay and bisexual men in 44 cities: Findings from the European MSM Internet Survey (EMIS). Int J Drug Policy 2016 Dec;38:4-12. [doi: 10.1016/j.drugpo.2016.09.007] [Medline: 27788450]
8. Vosburgh HW, Mansergh G, Sullivan PS, Purcell DW. A review of the literature on event-level substance use and sexual risk behavior among men who have sex with men. AIDS Behav 2012 Aug;16(6):1394-1410. [doi: 10.1007/s10461-011-0131-8] [Medline: 22323004]
9. King M, Semlyen J, Tai SS, Killaspy H, Osborn D, Popelyuk D, et al. A systematic review of mental disorder, suicide, and deliberate self harm in lesbian, gay and bisexual people. BMC Psychiatry 2008 Aug 18;8(70):1-17. [doi: 10.1186/1471-244x-8-70]
10. McFall S. Understanding Society. URL: https://www.understandingsociety.ac.uk/ [accessed 2022-03-22]
11. Santos G, Do T, Beck J, Makofane K, Arreola S, Pyun T, et al. Syndemic conditions associated with increased HIV risk in a global sample of men who have sex with men. Sex Transm Infect 2014 May;90(3):250-253. [doi: 10.1136/sextrans-2013-051318] [Medline: 24431183]
12. Brogan N, Paquette D, Lachowsky N, Blais M, Brennan D, Hart T, et al. Canadian results from the European Men-who-have-sex-with-men Internet survey (EMIS-2017). Can Commun Dis Rep 2019 Nov 07;45(11):271-282 [FREE Full text] [doi: 10.14745/ccdr.v45i11a01] [Medline: 31755878]
13. Semple SJ, Pitpitan EV, Goodman-Meza D, Strathdee SA, Chavarin CV, Rangel G, et al. Correlates of condomless anal sex among men who have sex with men (MSM) in Tijuana, Mexico: The role of public sex venues. PLoS One 2017;12(10):e0186814 [FREE Full text] [doi: 10.1371/journal.pone.0186814] [Medline: 29065132]
14. Champion KE, Parmenter B, McGowan C, Spring B, Wafford QE, Gardner LA, et al. Effectiveness of school-based eHealth interventions to prevent multiple lifestyle risk behaviours among adolescents: a systematic review and meta-analysis. The Lancet Digital Health 2019 Sep;1(5):e206-e221. [doi: 10.1016/s2589-7500(19)30088-3]
15. Prochaska JJ, Spring B, Nigg CR. Multiple health behavior change research: an introduction and overview. Prev Med 2008 Mar;46(3):181-188 [FREE Full text] [doi: 10.1016/j.ypmed.2008.02.001] [Medline: 18319098]
16. Meiksin R, Melendez-Torres GJ, Falconer J, Witzel TC, Weatherburn P, Bonell C. Theories of change for e-health interventions targeting HIV/STIs and sexual risk, substance use and mental ill health amongst men who have sex with men: systematic review and synthesis. Syst Rev 2021 Jan 11;10(1):21 [FREE Full text] [doi: 10.1186/s13643-020-01523-2] [Medline: 33423693]
17. Riper H, Blankers M, Hadiwijaya H, Cunningham J, Clarke S, Wiers R, et al. Effectiveness of guided and unguided low-intensity internet interventions for adult alcohol misuse: a meta-analysis. PLoS One 2014;9(6):e99912 [FREE Full text] [doi: 10.1371/journal.pone.0099912] [Medline: 24937483]
18. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 2009;38(4):196-205. [doi: 10.1080/16506070903318960] [Medline: 20183695]
19. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. PLoS One 2010 Oct 13;5(10):e13196 [FREE Full text] [doi: 10.1371/journal.pone.0013196] [Medline: 20967242]
20. Arnberg FK, Linton SJ, Hultcrantz M, Heintz E, Jonsson U. Internet-delivered psychological treatments for mood and anxiety disorders: a systematic review of their efficacy, safety, and cost-effectiveness. PLoS One 2014;9(5):e98118 [FREE Full text] [doi: 10.1371/journal.pone.0098118] [Medline: 24844847]
21. Christensen H, Batterham P, Calear A. Online interventions for anxiety disorders. Current Opinion in Psychiatry 2014;27:7-13. [doi: 10.1097/yco.0000000000000019]
22. Kaltenthaler E, Parry G, Beverley C, Ferriter M. Computerised cognitive-behavioural therapy for depression: systematic review. Br J Psychiatry 2008 Sep;193(3):181-184. [doi: 10.1192/bjp.bp.106.025981] [Medline: 18757972]

23. Păsărelu CR, Andersson G, Bergman Nordgren L, Dobrean A. Internet-delivered transdiagnostic and tailored cognitive behavioral therapy for anxiety and depression: a systematic review and meta-analysis of randomized controlled trials. Cogn Behav Ther 2017 Jan 07;46(1):1-28. [doi: 10.1080/16506073.2016.1231219] [Medline: 27712544]
24. Spijkerman M, Pots W, Bohlmeijer E. Effectiveness of online mindfulness-based interventions in improving mental health: A review and meta-analysis of randomised controlled trials. Clin Psychol Rev 2016 Apr;45:102-114 [FREE Full text] [doi: 10.1016/j.cpr.2016.03.009] [Medline: 27111302]
25. Higgins I, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 updated March. Oxford: Wiley-Blackwell; 2011.
26. Bauermeister JA, Tingler RC, Demers M, Connochie D, Gillard G, Shaver J, et al. Acceptability and Preliminary Efficacy of an Online HIV Prevention Intervention for Single Young Men Who Have Sex with Men Seeking Partners Online: The myDEx Project. AIDS Behav 2019 Nov;23(11):3064-3077 [FREE Full text] [doi: 10.1007/s10461-019-02426-7] [Medline: 30762190]
27. Carpenter KM, Stoner SA, Mikko AN, Dhanak LP, Parsons JT. Efficacy of a web-based intervention to reduce sexual risk in men who have sex with men. AIDS Behav 2010 Jun;14(3):549-557 [FREE Full text] [doi: 10.1007/s10461-009-9578-2] [Medline: 19499321]
28. Cheng W, Xu H, Tang W, Zhong F, Meng G, Han Z, et al. Online HIV prevention intervention on condomless sex among men who have sex with men: a web-based randomized controlled trial. BMC Infect Dis 2019 Jul 19;19(1):644 [FREE Full text] [doi: 10.1186/s12879-019-4251-5] [Medline: 31324233]
29. Chiou P, Liao P, Liu C, Hsu Y. Effects of mobile health on HIV risk reduction for men who have sex with men. AIDS Care 2020 Mar;32(3):316-324. [doi: 10.1080/09540121.2019.1668531] [Medline: 31558040]
30. Christensen JL, Miller LC, Appleby PR, Corsbie-Massay C, Godoy CG, Marsella SC, et al. Reducing shame in a game that predicts HIV risk reduction for young adult men who have sex with men: a randomized trial delivered nationally over the web. Journal of the International AIDS Society 2013 Nov 13;16:18716. [doi: 10.7448/ias.16.3.18716]
31. Davidovich U, de Wit JBF, Stroebe W. Using the internet to reduce risk of HIV-infection in steady relationships: A randomized controlled trial of a tailored intervention for gay men. Amsterdam: Roel & Uigeefprojecten; 2006. URL: https://www.scienceopen.com/document?vid=9ffb6cc6-eb1e-492c-a92d-092ce99148e5 [accessed 2022-03-22]
32. Hirshfield S, Downing MJ, Chiasson MA, Yoon IS, Houang ST, Teran RA, et al. Evaluation of Sex Positive! A Video eHealth Intervention for Men Living with HIV. AIDS Behav 2019 Nov;23(11):3103-3118 [FREE Full text] [doi: 10.1007/s10461-019-02498-5] [Medline: 31011912]
33. Milam J, Morris S, Jain S, Sun X, Dubé MP, Daar ES, CCTG 592 Team. Randomized Controlled Trial of an Internet Application to Reduce HIV Transmission Behavior Among HIV Infected Men Who have Sex with Men. AIDS Behav 2016 Jun;20(6):1173-1181 [FREE Full text] [doi: 10.1007/s10461-015-1215-7] [Medline: 26487300]
34. Mustanski B, Garofalo R, Monahan C, Gratzer B, Andrews R. Feasibility, acceptability, and preliminary efficacy of an online HIV prevention program for diverse young men who have sex with men: the keep it up! intervention. AIDS Behav 2013 Nov;17(9):2999-3012 [FREE Full text] [doi: 10.1007/s10461-013-0507-z] [Medline: 23673793]
35. Mustanski B, Parsons JT, Sullivan PS, Madkins K, Rosenberg E, Swann G. Biomedical and Behavioral Outcomes of Keep It Up!: An eHealth HIV Prevention Program RCT. Am J Prev Med 2018 Aug;55(2):151-158 [FREE Full text] [doi: 10.1016/j.amepre.2018.04.026] [Medline: 29937115]
36. Reback CJ, Fletcher JB, Swendeman DA, Metzner M. Theory-Based Text-Messaging to Reduce Methamphetamine Use and HIV Sexual Risk Behaviors Among Men Who Have Sex with Men: Automated Unidirectional Delivery Outperforms Bidirectional Peer Interactive Delivery. AIDS Behav 2019 Jan;23(1):37-47 [FREE Full text] [doi: 10.1007/s10461-018-2225-z] [Medline: 30006792]
37. Rosser BRS, Oakes JM, Konstan J, Hooper S, Horvath KJ, Danilenko GP, et al. Reducing HIV risk behavior of men who have sex with men through persuasive computing: results of the Men's INTernet Study-II. AIDS 2010 Aug 24;24(13):2099-2107 [FREE Full text] [doi: 10.1097/QAD.0b013e32833c4ac7] [Medline: 20601853]
38. Schonnesson LN, Bowen AM, Williams ML. Project SMART: Preliminary Results From a Test of the Efficacy of a Swedish Internet-Based HIV Risk-Reduction Intervention for Men Who Have Sex With Men. Arch Sex Behav 2016 Aug;45(6):1501-1511. [doi: 10.1007/s10508-015-0608-z] [Medline: 26545912]
39. Bowen AM, Williams ML, Daniel CM, Clayton S. Internet based HIV prevention research targeting rural MSM: feasibility, acceptability, and preliminary efficacy. J Behav Med 2008 Dec;31(6):463-477 [FREE Full text] [doi: 10.1007/s10865-008-9171-6] [Medline: 18770021]
40. Meiksin R, Melendez-Torres GJ, Falconer J, Witzel TC, Weatherburn P, Bonell C. eHealth Interventions to Address Sexual Health, Substance Use, and Mental Health Among Men Who Have Sex With Men: Systematic Review and Synthesis of Process Evaluations. J Med Internet Res 2021 Apr 23;23(4):e22477 [FREE Full text] [doi: 10.2196/22477] [Medline: 33890855]

Abbreviations

GRADE: Grading of Recommendations, Assessment, Development and Evaluations

MSM: men who have sex with men
OR: odds ratio
PROSPERO: Prospective Register of Systematic Reviews
RR: risk ratio
SBS: Safe Behavior and Screening
STI: sexually transmitted infection

Edited by T Sanchez; submitted 09.01.21; peer-reviewed by J Jain, J Mitchell; comments to author 09.03.21; revised version received 20.03.21; accepted 30.11.21; published 06.04.22.

Please cite as:
Melendez-Torres GJ, Meiksin R, Witzel TC, Weatherburn P, Falconer J, Bonell C
eHealth Interventions to Address HIV and Other Sexually Transmitted Infections, Sexual Risk Behavior, Substance Use, and Mental Ill-health in Men Who Have Sex With Men: Systematic Review and Meta-analysis
JMIR Public Health Surveill 2022;8(4):e27061
URL: https://publichealth.jmir.org/2022/4/e27061
doi:10.2196/27061
PMID:35384845

Original Paper

Long-term Changes in the Premature Death Rate in Lung Cancer in a Developed Region of China: Population-based Study

Wenjing Ye[1*], MD; Weiwei Lu[2*], MA; Xiaopan Li[3,4*], PhD; Yichen Chen[4,5], PhD; Lin Wang[6,7], BSc; Guangwang Zeng[6,7], BSc; Cheng Xu[6,7], MA; Chen Ji[8], PhD; Yuyang Cai[9], PhD; Ling Yang[6,7], MD, PhD; Zheng Luo[10], MD

[1]Department of Respiratory Medicine, Xinhua Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[2]Department of General Practice, Xinhua Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[3]Department of Health Management Center, Zhongshan Hospital, Shanghai Medical College of Fudan University, Shanghai, China

[4]Office of Scientific Research and Information Management, Pudong Institute of Preventive Medicine, Shanghai, China

[5]Center for Disease Control and Prevention, Pudong New Area, Shanghai, China

[6]Department of Geriatrics, Xinhua Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[7]Institute of Hospital Development Strategy, China Hospital Development Institute, Shanghai Jiao Tong University, Shanghai, China

[8]Warwick Clinical Trials Unit, Warwick Medical School, Coventry, United Kingdom

[9]School of Public Health, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[10]Shanghai University of Medicine & Health Sciences Affiliated Zhoupu Hospital, Pudong New Area, Shanghai, China

[*]these authors contributed equally

Corresponding Author:
Zheng Luo, MD
Shanghai University of Medicine & Health Sciences Affiliated Zhoupu Hospital
Pudong New Area
No.1500 Zhouyuan Rd
Shanghai, 201318
China
Phone: 86 18008497518
Email: zhengluo86@foxmail.com

Abstract

Background: Lung cancer is a leading cause of death worldwide, and its incidence shows an upward trend. A study of the long-term changes in the premature death rate in lung cancer in a developed region of China has great exploratory significance to further clarify the effectiveness of intervention measures.

Objective: This study examined long-term changes in premature lung cancer death rates in order to understand the changes in mortality and to design future prevention plans in Pudong New Area (PNA), Shanghai, China.

Methods: Cancer death data were collected from the Mortality Registration System of PNA. We analyzed the crude mortality rate (CMR), age-standardized mortality rate by Segi's world standard population (ASMRW), and years of life lost (YLL) of patients with lung cancer from 1973 to 2019. Temporal trends in the CMR, ASMRW, and YLL rate were calculated by joinpoint regression expressed as an average annual percentage change (AAPC) with the corresponding 95% CI.

Results: All registered permanent residents in PNA (80,543,137 person-years) from 1973 to 2019 were enrolled in this study. There were 42,229 deaths from lung cancer. The CMR and ASMRW were 52.43/10^5 and 27.79/10^5 person-years, respectively. The YLL due to premature death from lung cancer was 481779.14 years, and the YLL rate was 598.16/10^5 person-years. The CMR and YLL rate showed significantly increasing trends in men, women, and the total population (P<.001). The CMR of the total population increased by 2.86% (95% CI 2.66-3.07, P<.001) per year during the study period. The YLL rate increased with an AAPC of 2.21% (95% CI 1.92-2.51, P<.001) per year. The contribution rates of increased CMR values caused by demographic factors were more evident than those caused by nondemographic factors.

Conclusions: Lung cancer deaths showed an increasing trend in PNA from 1973 to 2019. Demographic factors, such as the aging population, contributed more to an increased CMR. Our research can help us understand the changes in lung cancer mortality and can be used for similar cities in designing future prevention plans.

(JMIR Public Health Surveill 2022;8(4):e33633) doi:10.2196/33633

KEYWORDS

lung cancer; mortality; years of life lost; trend analysis; decomposition method

Introduction

Lung cancer is the leading cause of cancer-related deaths worldwide, with 2,206,771 new lung cancer cases and 1,796,144 deaths in 2020 [1]. Many studies have made tremendous efforts in the discovery of potential biomarkers for the detection, classification, and progression monitoring of lung cancer. Lung cancer treatment has made great progress over the past decade. Study of the long-term changes in the disease burden of lung cancer has great exploratory significance to further clarify the epidemiological characteristics of lung cancer and improve patient survival times [2].

In recent decades, China has witnessed rapid economic development, and tremendous changes have occurred in the population and epidemics [3]. Shanghai is an economic, science and technology, industrial, financial, and exhibition center and is 1 of the earliest cities in China to enter an aging society. Shanghai is a representative city for modernization development in China. In the next 20 years, other cities in China and other low- and middle-income countries (LMICs) are likely to follow the development characteristics of Shanghai [4]. The Shanghai Pudong New Area (PNA) officially began to develop on April 18, 1990. Since then, Pudong has become 1 of the areas with the fastest urbanization and economic growth in China. Economic growth and total scale have highlighted the speed and height of Pudong's development and opening up. Over the past 30 years, PNA has adhered to reform, expansion, and strengthened innovation. It has developed from a field into a modern urban area with concentrated elements and advanced facilities. It has become the epitome of Shanghai's modernization and the symbol of China's reform and expansion [5,6].

Cancer is a disease that seriously endangers human health. In the past decade, cancer has been the main cause of death in China [7]. The treatment of cancer has become the most important research direction in the world to improve life expectancy [1]. Years of life lost (YLL) refers to the loss of life caused by early death. It can more accurately reflect the burden on society [7]. The aging population, progress of treatment, smoking, and environmental pollution may be factors that affect long-term changes in the premature death rate in lung cancer. An epidemiological study of lung cancer mortality trends over time may help quantify its impacts on public health and society, promote the assessment of current protocols for lung cancer, and define high-risk populations that will benefit from early detection programs for lung cancer [8]. Our study examined long-term changes in the premature death rate in lung cancer from 1973 to 2019 in PNA, Shanghai, China, to improve public health.

Methods

Data Source

Cancer death data were collected from the Mortality Registration System of PNA, including age, gender, date, and cause of death. The complete population data were provided by the Statistics Bureau and the Public Security Bureau of PNA [9]. Periodic evaluations, data cleaning, and compilation were performed to ensure completeness of the registration system. The per capita gross domestic product (GDP) of Shanghai and PNA were collected from the Shanghai Municipal Bureau of Statistics [10] and the Shanghai PNA Bureau of Statistics [11].

Deaths from malignant neoplasm of the trachea (C33) and bronchus and lung (C34) were classified by the underlying cause of lung cancer deaths according to the *International Classification of Diseases 10th Revision* (ICD-10) [12]. Since the data covered a long time span of 47 years, data before 1975, coded based on ICD-8, and data for 1975-2001, coded based on ICD-9, were converted to ICD-10 codes.

Causes of death were coded by rigorously trained clinicians, and each record was further verified by the Center for Disease Control and Prevention (CDC).

Ethics Approval and Consent to Participate

Our study did not involve any intervention in human participants. The surveillance protocol was approved by the ethical committee of the Shanghai PNA Center CDC. Strict confidentiality of individual data was practiced throughout the study.

Statistical Analyses

The crude mortality rate (CMR) and the age-standardized mortality rate by Segi's world standard population (ASMRW) of neurological disorders were calculated per 100,000 ($/10^5$). The CMR and ASMRW between genders were compared using the Poisson approximation method and the Mantel-Haenszel test, respectively.

The YLL was calculated according to the original method described by Murray and Lopez [13]. The equation used to calculate the YLL is as follows [14]:

$$YLL = KCe^{ra}/(r+\beta)^2 \{ e^{-(r+\beta)(L+a)}[-(r+\beta)(L+a) - 1] - e^{-(r+\beta)a}[-(r+\beta)a - 1]\} + [(1-k)/r] \times (1 - e^{-rL}),$$

where a is the age at death, β is the age weighting parameter (β=.04), C is the age weighting fit with constant (C=0.1658), r is the discount rate (r=3%), L is the standard life expectancy at the age of death according to the standard reference life table for the Global Burden of Disease (GBD) study [15], and e is the Napier constant.

The calculation of the YLL was performed using the World Health Organization (WHO) template [14].

Temporal trends in the CMR, ASMRW, and YLL rate were calculated using joinpoint Regression 4.3.1.0 (National Cancer Institute, Bethesda, MD, USA) and expressed as an average annual percentage change (AAPC) with a corresponding 95% CI. The Z test was performed to assess whether the AAPC was statistically different from 0. The terms "increase" and "decrease" were used to describe a statistically significant (P<.05) AAPC, while "stable" was used for not statistically significant trends.

Age was classified into 8 groups: 0-4, 5-14, 15-29, 30-44, 45-59, 60-69, 70-79, and 80+ years. Age-specific CMRs were calculated for each age group. Changes in the mortality rates of each period in 5 years from 1973 to 2019 were compared with the period before it or the data from 1973 to 1979, and causes from demographic and nondemographic factors were estimated by the decomposition method, in which mortality rates were calculated and compared for each 5-year age group, from 0-4 to 85+ years [16]. All statistical analyses were conducted using SPSS Statistics version 21.0 (SPSS, Inc, Chicago, IL, USA) and R version 3.4.3 (R Core Team). Statistical significance was set at P<.05.

Results

Baseline Characteristics of Underlying Death from Lung Cancer

From 1973 to 2019, all registered permanent residents in PNA, with a total of 80,543,137 person-years, were enrolled in this study. There were 42,229 deaths from lung cancer. Of these, 30,638 (72.55%) patients were men. The median age at death from lung cancer was 72.10 years, and the average age at death was 70.96±11.21 years. The CMR and ASMRW of lung cancer were 52.43/10^5 and 27.79/10^5 person-years, respectively. In addition, the CMR and ASMRW were 77.04/10^5 and 44.27/10^5 person-years, respectively, in men, while the corresponding rates were 28.43/10^5 and 13.77/10^5 person-years, respectively, in women (Table 1).

Table 1. Baseline characteristics of deaths (1973-2019).

Characteristic	Deaths, n (%)	Age (years), mean (SD)	Age (years), median	CMR[a] (/10^5 person-years)	ASMRW[b] (/10^5 person-years)	YLL[c] (years)	YLL rate (/10^5 person-years)
Gender							
Male	30,638 (72.55)	70.54 (10.74)	71.56	77.04	44.27	343,728.73	864.30
Female	11,591 (27.45)	72.08 (12.30)	73.73	28.43	13.77	138,050.40	338.58
Period							
1973-1979	929 (2.20)	64.88 (10.31)	66.38	21.57	23.40	12,915.91	299.88
1980-1984	969 (2.29)	65.75 (11.14)	67.08	29.29	28.79	13,006.25	393.17
1985-1989	969 (2.29)	66.47 (10.98)	67.48	31.51	28.45	12,711.93	413.33
1990-1994	1890 (4.48)	67.24 (10.42)	68.18	40.02	31.10	24,095.98	510.23
1995-1999	4797 (11.36)	68.32 (10.70)	69.77	43.16	30.70	59,014.12	530.92
2000-2004	6363 (15.07)	69.73 (11.10)	71.43	52.98	31.23	74,513.88	620.42
2005-2009	7595 (17.99)	71.22 (11.33)	73.46	57.88	27.68	84,179.23	641.55
2010-2014	8832 (20.91)	71.66 (11.47)	73.34	63.10	26.28	96,872.05	692.14
2015-2019	9885 (23.41)	72.65 (10.85)	72.62	66.40	24.02	104,535.93	702.19
Total	42,229 (100.00)	70.96 (11.21)	72.10	52.43	27.79	481,779.14	598.16

[a]CMR: crude mortality rate.

[b]ASMRW: age-standardized mortality rate by Segi's world standard.

[c]YLL: years of life lost.

Age-specific Mortality in Lung Cancer

The CMRs in the age groups of 0-4, 5-14, 15-29, 30-44, 45-59, 60-69, 70-79, and ≥80 years were 0.09/10^5, 0.02/10^5, 0.36/10^5, 4.26/10^5, 35.97/10^5, 132.70/10^5, 302.76/10^5, and 372.98/10^5 person-years, respectively (Table 2).

Table 2. Number of deaths, CMR[a], YLL[b], and YLL rates (1973-2019).

Group by age (years)	Deaths, n (%)	CMR (/10^5 person-years)	YLL (years)	YLL rate (/10^5 person-years)
Age (years)				
0-4	3 (0.01)	0.09	90.93	2.61
5-14	1 (0.002)	0.02	58.54	0.73
15-29	57 (0.13)	0.36	1543.65	9.62
30-44	826 (1.96)	4.26	19,559.72	100.95
45-59	6385 (15.12)	35.97	119,690.09	674.25
60-69	11,521 (27.28)	132.70	159,523.69	1837.34
70-79	14,642 (34.67)	302.76	134,731.32	2785.91
≥80	8793 (20.82)	372.98	46,581.20	1975.87
Total	42,229 (100.00)	52.43	481,779.14	598.16

[a]CMR: crude mortality rate.

[b]YLL: years of life lost.

Burden of Premature Death from Lung Cancer

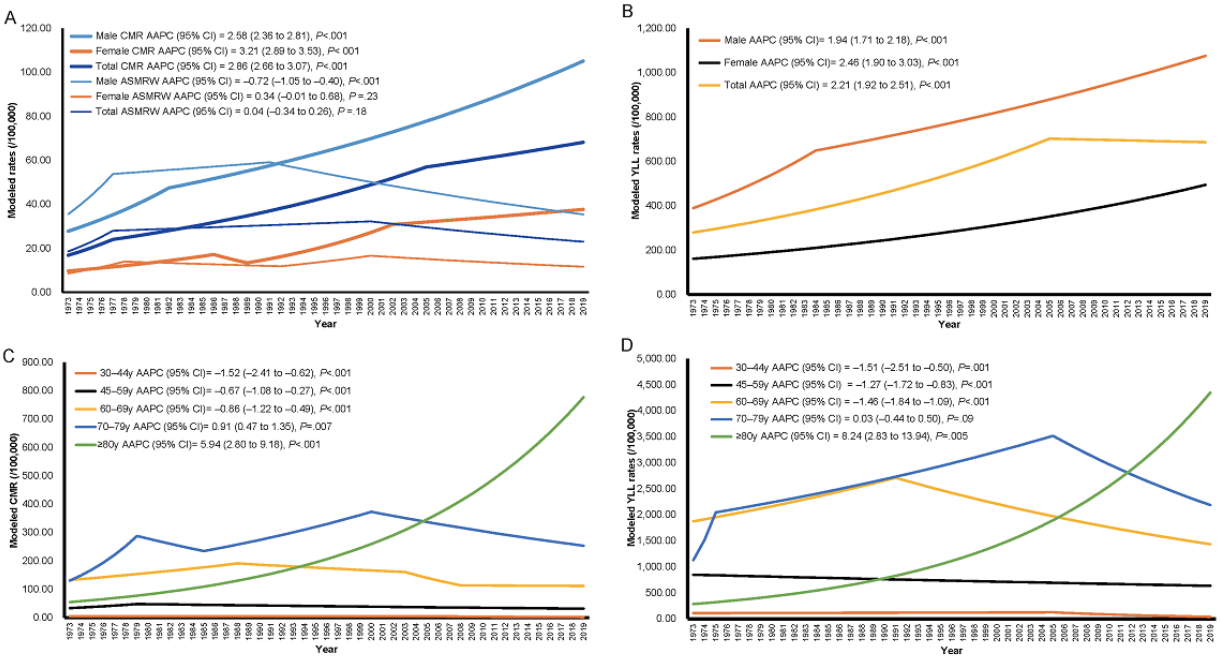
From 1973 to 2019, the YLL due to premature death from lung cancer was 481779.14 years and the YLL rate was 598.16/10^5 person-years. The YLL and rate of YLL in men (343,728.73 years and 864.30/10^5 person-years, respectively) were higher than those in women (138,050.40 years and 338.58/10^5 person-years, respectively); see Table 1. In terms of age, the top 3 YLL were in the age groups of 60-69, 70-79, and 45-59 years, which were 159,523.69, 134,731.32, and 119,690.09 years, respectively. The top 3 YLL rates were in the age groups of 70-79, 80+, and 60-69 years, which were 2785.91/10^5, 1975.87/10^5, and 1837.34/10^5, respectively (Table 2).

Trends in Mortality and YLL in Lung Cancer

The temporal trends in the CMR, ASMRW, and YLL rate were expressed based on the modeled CMR, ASMRW, and YLL rate, as shown in Figure 1. The CMR and YLL rate for deaths from lung cancer showed significantly increasing trends in men and women, and the total population during 1973-2019 (all P<.001). The ASMRW decreased in men by 0.72% (95% CI –1.05 to –0.40, P<.001) per year, while the ASMRW in women and the total population during 1973-2019 was not statistically significant (P=.23 and .18, respectively). The CMR in lung cancer in the total population increased by 2.86% (95% CI 2.66-3.07, P<.001) per year during the study period. The YLL rate increased with an AAPC of 2.21% (95% CI 1.92-2.51, P<.001) per year from 1973 to 2019 (Figures 1A and 1B).

Regarding age-specific mortality, the YLL, CMR, and ASMRW of the total population were observed from 1973 to 2019 (Figures 1C and 1D). The increasing trends in the CMR were also observed in the age groups of 70-79 years (P=.01) and 80+ years (P<.001). The 30-44-, 45-59-, and 60-69-year age groups had statistically decreasing trends in the CMR (P<.001). The YLL rate increased by 8.24% (95% CI 2.83-13.94, P=.01) per year in the age group of 80+ years and 0.03% (95% CI –0.44 to 0.50, P=.09) per year in the age group of 70-79 years. However, the YLL rate decreased by 1.51% (95% CI –2.51 to 0.05, P=.001) per year in the age group of 30-44 years, 1.27% (95% CI –1.72 to 0.83, P<.001) per year in the age group of 45-59 years, and 1.46% (95% CI –1.84 to 1.09, P<.001) per year in the age group of 60-69 years (Figures 1C and 1D).

Figure 1. Trends in the CMR (/10^5 person-years), ASMRW (/10^5 person-years), and YLL (/10^5 person-years) rate of persons dying from lung cancer by pathology type and age group in Shanghai PNA from 1973 to 2019. (A) CMR and ASMRW in lung cancer, (B) YLL rate in lung cancer, (C) CMR of age groups, and (D) YLL rate of age groups. AAPC: average annual percentage change; ASMRW: age-standardized mortality rate by Segi's world standard population; CMR: crude mortality rate; PNA: Pudong New Area; YLL: years of life lost.



Quantitatively Impacts of Demographic and Nondemographic Factors on Increased CMRs

The increasing CMRs caused by nondemographic and demographic factors are shown in Figure 2. Based on the CMR in lung cancer in 1973-1979, no statistically significant trend was found caused by nondemographic factors in the total population, with an AAPC of 0.17% (95% CI –11.34 to 13.16, *P*=.97) from 1980 to 2019, but a significant upward trend was observed in the increased CMR caused by demographic factors (AAPC [95% CI]=51.70% [35.48-69.88], *P*<.001). In men, the increased CMR caused by nondemographic factors decreased by 32.96% (95% CI –51.68 to –6.99, *P*=.02) during 1980-2019, and the CMR caused by demographic factors increased by 46.42% (95% CI 32.23-62.03, *P*<.001). In women, the increased CMR caused by nondemographic factors showed an upward trend with an AAPC of 24.24% (95% CI 2.60-50.44, *P*=.03), and the CMR caused by demographic factors also increased (AAPC [95% CI]=55.63% [38.54-74.83], *P*<.001); see Table 3. Figure 2B-D shows the proportion of increased CMR values caused by nondemographic and demographic factors. From 1985 to 2019, demographic factors played a decisive role in the contribution of the CMR compared to 1973-1979.

Figure 2. Increased CMRs caused by demographic and nondemographic factors and their proportion during the period from 1973 to 2019 compared with the CMR in lung cancer during 1973-1979 in Shanghai PNA. (A) Increased CMRs, (B) contribution of increased CMRs in the total population, (C) contribution of increased CMRs in men, and (D) contribution of increased CMRs in women. CMR: crude mortality rate; PNA: Pudong New Area.

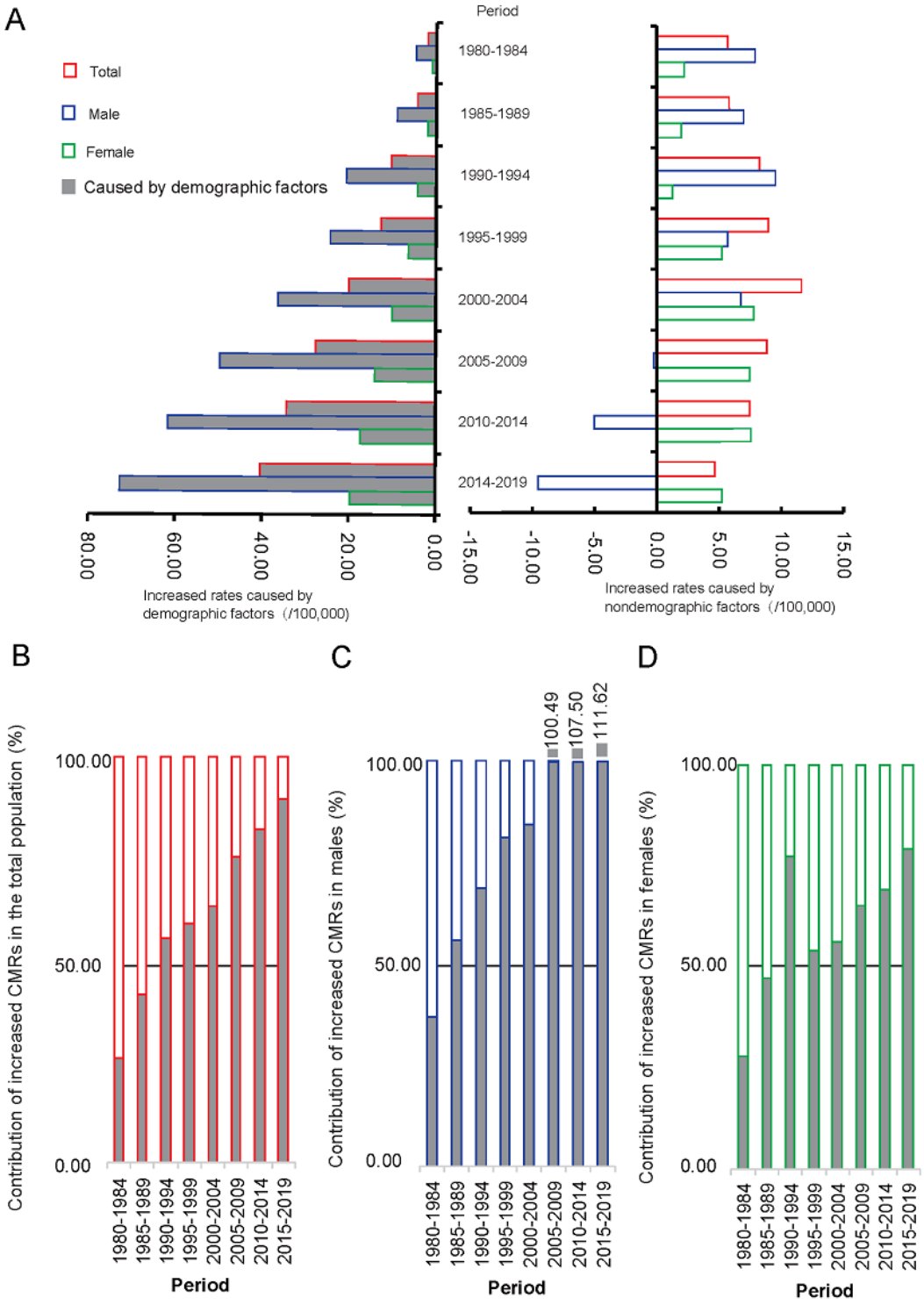


Table 3. Increased CMRs[a] caused by demographic and nondemographic factors and their contribution during the period from 1973 to 2019 compared with the CMR in lung cancer during 1973-1979 or the period before it in Shanghai PNA[b].

Comparison periods	CMR of the based period (/10^5 person-years)	CMR of the other period (/10^5 person-years)	*D* value of mortality (/10^5)	Impact of demographic factors		Impact of nondemographic factors	
				Added value (/10^5)	Contribution rate (%)	Added value (/10^5)	Contribution rate (%)
Panel A (based on the first period)							
1980-1984 vs 1973-1979	21.57	29.29	7.72	2.00	25.68	5.73	74.14
1985-1989 vs 1973-1979	21.57	31.51	9.94	4.11	41.38	5.83	58.62
1990-1994 vs 1973-1979	21.57	40.02	18.45	10.20	55.28	8.25	44.72
1995-1999 vs 1973-1979	21.57	43.16	21.59	12.70	58.83	8.89	41.17
2000-2004 vs 1973-1979	21.57	52.98	31.41	19.86	63.21	11.55	36.79
2005-2009 vs 1973-1979	21.57	57.88	36.31	27.45	75.58	8.87	24.42
2010-2014 vs 1973-1979	21.57	63.10	41.53	34.10	82.11	7.43	17.89
2015-2019 vs 1973-1979	21.57	66.40	44.83	40.24	89.76	4.59	10.24
Panel B (based on the last period)							
1980-1984 vs 1973-1979	21.57	29.29	7.72	2.00	25.68	5.73	74.14
1985-1989 vs 1980-1984	29.29	31.51	2.22	2.66	114.33	–0.44	-14.33
1990-1994 vs 1985-1989	31.51	40.02	8.51	7.61	89.34	0.91	10.66
1995-1999 vs 1990-1994	40.02	43.16	3.14	2.72	86.91	0.41	13.09
2000-2004 vs 1995-1999	43.16	52.98	9.82	8.13	82.79	1.69	17.21
2005-2009 vs 2000-2004	52.98	57.88	4.90	10.12	134.01	–5.21	34.01
2010-2014 vs 2005-2009	57.88	63.10	5.22	7.86	125.16	–2.64	25.16
2015-2019 vs 2010-2014	63.10	66.40	3.30	7.80	136.61	–4.50	36.61

[a]CMR: crude mortality rate.

[b]PNA: Pudong New Area.

Discussion

Principal Findings

It is crucial to understand the long-term changes in the rate of premature death from lung cancer for medical treatment research to formulate future preventive measures. The goal of "Healthy China 2030" has been to reduce the premature death rate of noncommunicable diseases by 30% [17]. Since 2000, many cities in China have gradually entered an aging society, and Shanghai is the first to do so. In recent years, the aging in Shanghai has not been alleviated but has gradually increased. Since 2018, PNA as a miniature Shanghai has already entered a superaging society, with a proportion of over 20% (Multimedia Appendix 1). We concluded that the increasing trends in the CMR were seen in the age groups of 70-79 years (*P*=.01) and 80+ years (*P*<.001) in terms of age-specific mortality and burden. The YLL rate increased by 8.24% (95% CI 2.83-13.94, *P*=.01) per year in the age group of 80+ years and 0.03% (95% CI –0.44 to 0.50, *P*=.09) per year in the age group of 70-79 years. The proportion of individuals aged 70-79 years (Multimedia Appendix 2) was almost the largest since 1995. The age groups under 70 years had statistically decreasing trends in lung cancer CMR (*P*<.001). It is evident that aging has contributed to an increase in lung cancer mortality.

Economic development, such as an increased GDP, improves public health. However, some factors related to the developed economy, such as lifestyle factors and environmental and medical levels, may also influence the mortality rate. Globally, smoking-attributable deaths have increased by 20.1% (15.3-25.2) since 1990, with most deaths occurring in China [18]. In China, smoking has either peaked or continued to increase [19]. According to our statistics over the past 50 years, the YLL rate increased with an AAPC of 2.21% per year from 1973 to 2019, and this may be associated with the increased smoking rate. The YLL rate of men is higher than that of women (Figure 1), which is consistent with the difference in the smoking proportion between men and women.

Outdoor air pollution exposure is a clear carcinogen to humans [20]. For decades, after rapid industrialization and urbanization, air pollution in China has worsened [21]. Air pollution has significantly affected the health of Chinese people as 1 of the top 10 risk factors for death [18]. Several large cohorts confirmed that the particulate matter ($PM_{2.5}$) concentration in the environment is associated with the risk of lung

adenocarcinoma in nonsmokers and lung cancer mortality in lifelong nonsmokers [20,22]. When the development of PNA started in 1990, the secondary industry accounted for more than 75% and the tertiary industry accounted for only 20% of the GDP of PNA. In 2018, the status of the secondary and tertiary industries was reversed, and the proportion of the tertiary industry exceeded 75% of the PNA GDP. However, the development of secondary industries cannot prevent air and environmental pollution. A long-period comparative analysis of air pollution in Shanghai analyzed the continuous Morlet wavelet transform on the time series of a 5274-day air pollution index from 2000. The monthly variation in air pollution in Shanghai was not significant, and air pollution in Shanghai showed an increasing trend, but the situation has reversed since 2015 [23]. It is evident that in the past 50 years, especially before 2015, air pollution in PNA has led to an increasing trend in the YLL due to premature death in people with lung cancer in Shanghai.

The initial treatment of lung cancer is relatively simple, including surgery, chemotherapy, and radiotherapy. In the early 2000s, the key genes for lung cancer were identified. In this field, molecular detection is the basic method for guiding and individualizing treatment selection tools. Epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors (TKIs) are small molecule–targeted drugs widely used in advanced non-small-cell lung cancer (NSCLC). Its effectiveness can be shown in symptom improvement, lesion control, and prolongation of progression-free survival (PFS). Women and nonsmoking patients were the dominant groups in TKI treatment [24]. The introduction of immune checkpoint inhibitors in 2015 was an important milestone in the treatment of lung cancer. Immunotherapy has been proven to have a long-lasting positive effect on patients with NSCLC and has been rapidly upgraded to first-line treatment after the success of second-line and backline treatments [25,26]. Since 2003, lung cancer–targeting drugs have entered the scope of Shanghai medical insurance and are widely used in lung cancer treatment. Furthermore, immunotherapy and antiangiogenic drugs have also been gradually introduced in Shanghai. The diversity of drug selection and the individualization and accuracy of treatment schemes directly affect the survival of patients with lung cancer. This may explain why the YLL rate showed a downward trend with an AAPC of the total population after 2005 compared with that before 2005. We found that, except for patients over 80 years old, the CMRs and YLL rate decreased after 2000 in almost all age groups. The changes in these data are related to the rapid development of tumor-targeted therapy and immunotherapy over the past 20 years.

Lung cancer is the leading cause of cancer-related deaths worldwide, with 2,206,771 new lung cancer cases and 1,796,144 deaths in 2020. Lung cancer has a high incidence rate and contributes to 30% of all cancer-related deaths in China. The mortality trends in lung cancer in the United States have gradually decreased [27]. Mortality rates increased from 1990 by 3.91% per year and decreased from 2004 by 1.95% in Montenegro [28]. Some studies have observed a sharp increase in the lung cancer mortality rate since 2000 in China [27]. The lung cancer mortality rate increased from 30.18% in 2004 to 36.10% in 2010 [29]. In 2018, the mortality rate of Chinese men was 68% higher than their US counterparts, while that of Chinese women was similar to that of US women. The difference in men may be related to smoking [27]. Lung cancer mortality in China may increase by 40% between 2015 and 2030 [30]. Our study also found that mortality increased significantly in both genders in PNA. In addition to reducing the proportion of smokers, the widespread application of chest computerized tomography (CT) screening will impact lung cancer mortality in China [27].

Strengths and Limitations

This study has major strengths, including a large population size (over 8 million) and a relatively long time span (47 years). However, there are several limitations to the study. First, all our data were from a single district, although this district is the largest in Shanghai. Second, there were no data on lifestyle, histological typing, and disease history in our study, so it was impossible to determine the role of each risk factor that may lead to changes in lung cancer mortality [31]. Nonetheless, our study was based on complete and accurate public data over 4 decades from the government surveillance system, and high data quality was ensured.

Conclusion

This population-based study found increasing trends in the mortality and burden of lung cancer in men and women, and the total population in a developed region of China from 1973 to 2019. Demographic factors, particularly aging, contributed to an increase in mortality. Our study can contribute to a better understanding of lung cancer and can be used in similar cities to design future prevention plans.

Acknowledgments

The authors thank all the staff in the vital statistics system of Pudong New Area (PNA) from 1973 to 2019 for their great work in data collection and ensuring high data quality.

This work was supported by the General Program of Health Bureau of Shanghai PNA (PW2021A-68), the Shanghai Specialized Research Fund for Integrated Chinese and Western Medicine in General Hospitals (ZHYY-ZXYJHZX-201914), the General Program of the National Natural Science Foundation of China (81673250), key disciplines of public health in Shanghai (GWV-10.1-XK17), the National Science and Technology Major Project of the Ministry of Science and Technology of China (2017ZX10201201-006-001), the Shanghai Public Health System Construction Three-year Action Plan Outstanding Youth Talent Training Program (GWV-10.2-YQ43), and the Reserve Academic Leaders Training Program of the Pudong New Area Center for Disease Control and Prevention (PDCDC-HBXD2020-05).

Data Availability

The data that support the findings of this study are available from the Center for Disease Control and Prevention (CDC) of Pudong New Area (PNA), Shanghai, but restrictions apply to the availability of these data, which were used under license for this study; therefore, they are not publicly available. However, data are available from the authors upon reasonable request and with permission from the CDC of PNA.

Authors' Contributions

WY, WL, and XL drafted the manuscript. XL, WY, ZL, and YC participated in the collection, analysis, and interpretation of data. YC, ZL, WY, GZ, LW, CX, and XL contributed to the data collection and suggestions for analysis. CJ, LY, YC, and ZL conceived the study, participated in its design, and coordination and critically revised the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Trends in the proportion of the ≥65-year age group in the total and female population in Shanghai PNA and the capital per GDP in Shanghai and Shanghai PNA from 1995 to 2018. GDP: gross domestic product; PNA: Pudong New Area.
[PNG File , 24 KB - publichealth_v8i4e33633_app1.png]

Multimedia Appendix 2
Age composition of the population in Shanghai PNA from 1973 to 2019. PNA: Pudong New Area.
[PNG File , 20 KB - publichealth_v8i4e33633_app2.png]

References

1. World Health Organization International Agency for Research on Cancer. GLOBOCAN: Lung Cancer Fact Sheet 2020. URL: https://gco.iarc.fr/today/data/factsheets/cancers/15-Lung-fact-sheet.pdf [accessed 2020-12-01]
2. Osmani L, Askin F, Gabrielson E, Li QK. Current WHO guidelines and the critical role of immunohistochemical markers in the subclassification of non-small cell lung carcinoma (NSCLC): moving from targeted therapy to immunotherapy. Semin Cancer Biol 2018 Oct;52(Pt 1):103-109 [FREE Full text] [doi: 10.1016/j.semcancer.2017.11.019] [Medline: 29183778]
3. Wang J, Gao S, Wang Y, Zhou J, Lin J, Wang L, et al. Cancer incidence and mortality patterns in Luwan district of Shanghai during 2002-2011. Drug Discov Ther 2018 May 13;12(2):77-87 [FREE Full text] [doi: 10.5582/ddt.2018.01009] [Medline: 29681580]
4. Li X, Deng Y, Tang W, Sun Q, Chen Y, Yang C, et al. Urban-rural disparity in cancer incidence, mortality, and survivals in Shanghai, China, during 2002 and 2015. Front Oncol 2018 Dec 3;8:579-579 [FREE Full text] [doi: 10.3389/fonc.2018.00579] [Medline: 30560091]
5. Chen H, Hao L, Yang C, Yan B, Sun Q, Sun L, et al. Understanding the rapid increase in life expectancy in Shanghai, China: a population-based retrospective analysis. BMC Public Health 2018 Feb 14;18(1):256-256 [FREE Full text] [doi: 10.1186/s12889-018-5112-7] [Medline: 29444657]
6. Li X, Qian M, Zhao G, Yang C, Bao P, Chen Y, et al. The performance of a community-based colorectal cancer screening program: evidence from Shanghai Pudong New Area, China. Prev Med 2019 Jan;118:243-250. [doi: 10.1016/j.ypmed.2018.11.002] [Medline: 30412744]
7. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018 Nov 12;68(6):394-424 [FREE Full text] [doi: 10.3322/caac.21492] [Medline: 30207593]
8. GBD 2016 Causes of Death Collaborators. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017 Sep;390(10100):1151-1210. [doi: 10.1016/S0140-6736(17)32152-9]
9. People's Government of Pudong New Area. Shanghai Statistical Yearbook. URL: http://www.pudong.gov.cn/shpd/InfoOpen/newTongJiList.aspx?CategoryNum=014004002002
10. Shanghai Government. Annual Report of Finance Bureau. URL: https://www.shanghai.gov.cn/2021ndbg/index.html
11. People's Government of Pudong New Area. Information disclosure of Finance Bureau. URL: https://www.pudong.gov.cn/zwgkzfxxgknbzfbm_czj/index.html
12. Paoin W, Yuenyongsuwan M, Yokobori Y, Endo H, Kim S. Development of the ICD-10 simplified version and field test. Health Inf Manag 2018 May 06;47(2):77-84. [doi: 10.1177/1833358317701277] [Medline: 28537209]

13. Murray CJ, Lopez AD, World Health Organization , World B, Harvard School of Public Health. The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020. Summary. Geneva, Switzerland: World Health Organization; 1996:1-41.
14. World Health Organization. Health Statistics and Information Systems, Disease Burden and Mortality Estimates. URL: https://www.who.int/healthinfo/global_burden_disease/estimates/en/index1.html [accessed 2020-07-31]
15. World Health Organization. Global Health Estimates 2016: Disease Burden by Cause, Age, Sex, by Country and by Region, 2000-2016. Geneva: World Health Organization; 2018.
16. Cheng X, Yang Y, Schwebel DC, Liu Z, Li L, Cheng P, et al. Population ageing and mortality during 1990-2017: a global decomposition analysis. PLoS Med 2020 Jun 8;17(6):e1003138-e1003138 [FREE Full text] [doi: 10.1371/journal.pmed.1003138] [Medline: 32511229]
17. Tan X, Zhang Y, Shao H. Healthy China 2030, a breakthrough for improving health. Glob Health Promot 2019 Dec 03;26(4):96-99. [doi: 10.1177/1757975917743533] [Medline: 29297762]
18. GBD 2016 Risk Factors Collaborators. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017 Sep;390(10100):1345-1422. [doi: 10.1016/S0140-6736(17)32366-8]
19. Jha P. Avoidable global cancer deaths and total deaths from smoking. Nat Rev Cancer 2009 Sep 20;9(9):655-664. [doi: 10.1038/nrc2703] [Medline: 19693096]
20. Gharibvand L, Lawrence Beeson W, Shavlik D, Knutsen R, Ghamsary M, Soret S, et al. The association between ambient fine particulate matter and incident adenocarcinoma subtype of lung cancer. Environ Health 2017 Jun 24;16(1):71-71 [FREE Full text] [doi: 10.1186/s12940-017-0268-7] [Medline: 28646928]
21. Su T, Chen S, Chan C. Progress of ambient air pollution and cardiovascular disease research in Asia. Prog Cardiovasc Dis 2011 Mar;53(5):369-378. [doi: 10.1016/j.pcad.2010.12.007] [Medline: 21414472]
22. Turner MC, Krewski D, Pope CA, Chen Y, Gapstur SM, Thun MJ. Long-term ambient fine particulate matter air pollution and lung cancer in a large cohort of never-smokers. Am J Respir Crit Care Med 2011 Dec 15;184(12):1374-1381. [doi: 10.1164/rccm.201106-1011oc]
23. Su Y, Sha Y, Zhai G, Zong S, Jia J. Comparison of air pollution in Shanghai and Lanzhou based on wavelet transform. Environ Sci Pollut Res Int 2019 Jun 21;26(17):16825-16834 [FREE Full text] [doi: 10.1007/s11356-017-8959-8] [Medline: 28432621]
24. Duma N, Santana-Davila R, Molina JR. Non-small cell lung cancer: epidemiology, screening, diagnosis, and treatment. Mayo Clin Proc 2019 Aug;94(8):1623-1640. [doi: 10.1016/j.mayocp.2019.01.013] [Medline: 31378236]
25. Reck M, Rodríguez-Abreu D, Robinson AG, Hui R, Csőszi T, Fülöp A, et al. Pembrolizumab versus chemotherapy for PD-L1–positive non–small-cell lung cancer. N Engl J Med 2016 Nov 10;375(19):1823-1833. [doi: 10.1056/nejmoa1606774]
26. Borghaei H, Paz-Ares L, Horn L, Spigel DR, Steins M, Ready NE, et al. Nivolumab versus docetaxel in advanced nonsquamous non–small-cell lung cancer. N Engl J Med 2015 Oct 22;373(17):1627-1639. [doi: 10.1056/nejmoa1507643]
27. Yang D, Liu Y, Bai C, Wang X, Powell CA. Epidemiology of lung cancer and lung cancer screening programs in China and the United States. Cancer Lett 2020 Jan 01;468:82-87. [doi: 10.1016/j.canlet.2019.10.009] [Medline: 31600530]
28. Nedović-Vuković M, Laušević D, Ljaljević A, Golubović M, Trajković G. Lung cancer mortality in Montenegro, 1990 to 2015. Croat Med J 2019 Feb;60(1):26-32. [doi: 10.3325/cmj.2019.60.26]
29. Chen W, Zhang S, Zou X, Zhao P. An analysis of lung cancer mortality in China, 2004 - 2005. Zhonghua Yu Fang Yi Xue Za Zhi 2010 May;44(5):378-382. [Medline: 20654224]
30. Martín-Sánchez JC, Lunet N, González-Marrón A, Lidón-Moyano C, Matilla-Santander N, Clèries R, et al. Projections in breast and lung cancer mortality among women: a Bayesian analysis of 52 countries worldwide. Cancer Res 2018 Jul 31;78(15):4436-4442. [doi: 10.1158/0008-5472.can-18-0187]
31. Global Burden of Disease Cancer Collaboration. Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 29 cancer groups, 1990 to 2017: a systematic analysis for the Global Burden of Disease Study. JAMA Oncol 2019 Dec 01;5(12):1749-1768 [FREE Full text] [doi: 10.1001/jamaoncol.2019.2996] [Medline: 31560378]

Abbreviations

AAPC: average annual percentage change
ASMRW: age-standardized mortality rate by Segi's world standard population
CMR: crude mortality rate
GDP: gross domestic product
ICD-10: International Classification of Diseases 10th Revision
NSCLC: non-small-cell lung cancer
PNA: Pudong New Area
TKI: tyrosine kinase inhibitor
YLL: years of life lost

Edited by H Bradley; submitted 18.09.21; peer-reviewed by Y Deng, A Reyna Sevilla; comments to author 08.11.21; revised version received 20.11.21; accepted 24.02.22; published 20.04.22.

Please cite as:
Ye W, Lu W, Li X, Chen Y, Wang L, Zeng G, Xu C, Ji C, Cai Y, Yang L, Luo Z
Long-term Changes in the Premature Death Rate in Lung Cancer in a Developed Region of China: Population-based Study
JMIR Public Health Surveill 2022;8(4):e33633
URL: https://publichealth.jmir.org/2022/4/e33633
doi:10.2196/33633
PMID:35442209

Original Paper

Sexual Partner Referral for HIV Testing Through Social Networking Platforms: Cross-sectional Study

Piao-Yi Chiou[1,2], PhD; Chien-Ching Hung[3,4], MD, PhD; Chien-Yu Chen[5], BSc

[1]School of Nursing, National Taiwan University College of Medicine, Taipei, Taiwan

[2]Department of Nursing, National Taiwan University Hospital, Taipei, Taiwan

[3]Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan

[4]Department of Tropical Medicine and Parasitology, National Taiwan University College of Medicine, Taipei, Taiwan

[5]Department of Nursing, National Taipei University of Nursing and Health Sciences, Taipei, Taiwan

Corresponding Author:
Piao-Yi Chiou, PhD
School of Nursing
National Taiwan University College of Medicine
No. 2-1, Xuzhou Rd.
Taipei, 100
Taiwan
Phone: 886 2 23123456 ext 88427
Fax: 886 2 82584729
Email: purechiou2@gmail.com

Abstract

Background: Men who have sex with men (MSM) who undergo voluntary HIV counseling and testing (VCT) often report condomless anal sexual intercourse, having many sexual partners, and being exposed to risky sexual networks. Limited research has discussed the application of motivational interviewing and convenience referral platforms to facilitate the referral of sexual partners for HIV testing among MSM.

Objective: This study aimed to evaluate the effects of VCT referral by sexual partners through social networking platforms and the test results after elicited interviews with MSM; compare the characteristics and risk behaviors among MSM tested without referral, index subjects, and referred sexual partners; and explore unknown sexual affiliations through visualizing and quantifying the social network graph.

Methods: This was a cross-sectional study. Purposeful sampling was used to recruit index subjects from a community HIV screening station frequented by MSM in Taipei City on Friday and Saturday nights. Respondent-driven sampling was used to recruit sexual partners. Partner-elicited interviews were conducted by trained staff before VCT to motivate MSM to become index subjects and refer sexual partners via the Line app, or to disclose the accounts and profiles of sexual partners on relevant social networking platforms. Referred sexual partners received rapid HIV testing, and the recruitment process was repeated until leads were exhausted.

Results: After the interviews, 28.2% (75/266) of MSM were successfully persuaded to become index subjects in the first wave, referring 127 sexual partners via the Line app for rapid HIV testing and disclosing 40 sexual partners. The index subjects and tested sexual partners had more sexual partners (F_2=3.83, P=.02), more frequent anal intercourse (F_2=10.10, P<.001), and higher percentages of those who had not previously received HIV testing (χ^2_1=6.1, P=.047) compared with MSM tested without referrals. The new HIV-seropositivity rate among tested sexual partners was 2.4%, which was higher than the rate in the other 2 groups. The social network analysis revealed the following 4 types of sexual affiliation: chain, Y, star, and complicated. Among the HIV-negative sexual partners, 26.9% (43/160) had sexual affiliations with HIV-positive nodes, and 40% (10/25) were untested sexual partners with a direct sexual affiliation with an HIV-positive node. Four transmission bridges were found in the network graph.

Conclusions: Partner-elicited interviews can effectively promote referral for HIV testing and case identification via Line, and can clarify unknown sexual affiliations of MSM to facilitate the development of a tailored prevention program. Social network analysis is needed for an insightful understanding of the different network structures.

(JMIR Public Health Surveill 2022;8(4):e32156) doi:10.2196/32156

KEYWORDS

HIV testing; men who have sex with men; mobile health; motivational interviewing; referral and consultation; risk behavior; sexual partners; social networking

Introduction

Background

Voluntary counseling and testing (VCT) for HIV, a process aimed at enabling an individual to make an informed choice regarding HIV testing, and condom use are key strategies for achieving the World Health Organization's goal of 95% of all persons with HIV being aware of their status and being successfully treated by 2025, which is ultimately aimed at eliminating HIV by 2030 [1,2]. In 2019, it was estimated that 81% of people with HIV worldwide and 88% in Taiwan knew their infection status [3,4]. In addition to the current VCT model, a proactive approach should be applied to reach and test those who have not yet been diagnosed.

HIV could have been primarily transmitted through condomless sex in a social network characterized by similar and related risk behaviors [5,6]. Men who have sex with men (MSM) and who undergo VCT often report condomless anal sexual intercourse and having multiple sexual partners [7,8]. They comprise the main at-risk group for HIV transmission globally [9]. The active provision of VCT to the sexual partners of MSM who have been screened for HIV could promote the identification of HIV cases and the obstruction of transmission chains [10]. The referral of sexual partners is a delicate and private matter, and is difficult to compare against social network referral relating to nonsexual relationships [11,12]. There is a need for guiding frameworks for motivational interviewing on sensitive matters, including for sexual partners [13], and limited research has been conducted in this regard.

Motivational interviewing is based on the theory of helpfulness and disclosure, which emphasizes that self-disclosure and helping others could be conducive to interviewees being willing to share intimate topics like sexual relationships [14]. According to this theory, concepts, such as verbal persuasion, an emphasis on benefits, and the motivation to be a helper, contribute to the generation of interview content that encourages MSM to refer their sexual partners for VCT [15,16]. A convenient platform for initiating referrals is also necessary [17].

Social networking platforms that include mobile instant messaging apps (eg, Line and WeChat), mobile geosocial network apps (eg, Grindr, Jack'd, and Hornet), and web-based communities (eg, Facebook, Twitter, and Instagram), have become popular platforms for MSM to seek and meet sexual partners and receive health information [18-21]. Previous studies confirmed that posting screening advertisements, sending messages, or engaging in web-based discussions can effectively improve self-reported HIV testing behaviors among people engaging in condomless sex [22-25]. Such platforms provide a convenient private approach for delivering testing messages, and as the testing process is coordinated directly among MSM, their sexual partners, and testing staff, they could be used as referral platforms. Cross-referral between MSM and their sexual partners via social networking platforms presents unique opportunities to access connected communities and intervene in their risk-taking behavior.

Social network analysis is usually applied to reveal previously unknown facets about risky behaviors, specifically regarding affiliations between an individual (node) and the rest of the network [5,6]. The analysis of each node, which is defined as an individual engaging in certain types of actions in a network, provides valuable information on individual risk features and behaviors in a social network [26]. A social network graph can be developed for more diversified analyses, such as presenting the network type, and quantifying links between HIV-negative and HIV-positive nodes to understand the possible distances of HIV transmission [27].

Objective

The purpose of this study was to evaluate the effects of sexual partners' referrals for HIV testing through social networking platforms, along with the generated test results, after conducting motivational interviewing with MSM. Specifically, we compared differences in the characteristics and risk behaviors of the MSM tested without referral, index subjects, and referred sexual partners in this network. Finally, we aimed to reveal the unknown sexual affiliations of each HIV-negative node to HIV-positive nodes through visualizing and quantifying the social network graph.

Methods

Study Design

The study followed a cross-sectional design. Purposeful sampling was used to recruit participants, while respondent-driven sampling was used to recruit sexual partners, who were referred via social networking platforms. Partner-elicited interviews were conducted until no new sexual partners were referred or no social networking accounts and profiles of sexual partners on platforms where they first met were disclosed by the index subjects to the research team.

Participants

Participants included those who had undergone VCT by trained staff, had self-identified as MSM, had engaged in condomless anal sexual intercourse, were over 20 years old, were literate, owned a mobile phone, lived in Taipei, and had agreed to participate in this research. Men self-identifying as transgender were excluded. Participants who received VCT during the first wave and referred or disclosed their sexual partners via social networking platforms were defined as index subjects, while sexual partners were defined as those with whom the referrer had engaged in sexual intercourse during the previous 3 months. Tested sexual partners were defined as those who had been referred through the referrer and who had completed the rapid

HIV test by the trained staff. This process continued until all leads were exhausted. Participants who were unwilling to refer or disclose their sexual partners during the first wave were defined as MSM tested without referrals.

Procedure and Data Collection

This research was conducted from August 2017 to January 2018. MSM who underwent VCT and met the inclusion criteria were recruited at a community screening station in a gay village, mainly a business community, frequented by MSM in Taipei City on Friday and Saturday nights from 6 PM to 10 PM. The reason for choosing this location is that the station provides HIV screening services for the target group from New Taipei City and Taipei City, which have the highest HIV prevalence rates in Taiwan. After explaining the research purpose and procedure, written informed consent was obtained from the participants. Individuals who agreed to participate in the research then completed a questionnaire, and a pretest counseling session was conducted by trained staff. The 20-30–min partner-elicited interviews were performed and information on the sexual partners was collected by the same staff. A free and anonymous rapid HIV test and posttest counseling session were subsequently provided to the participants. The reasons for refusal to be recruited, or not agreeing to refer or disclose any information on sexual partners, were also discussed and recorded in the first wave. For those who refused to participate in the research, the HIV testing procedure was completed without the partner-elicited interview.

The Line app is one of the most popular mobile instant messaging apps in Taiwan. It can be downloaded free of charge and was therefore used as a referral platform in this study. Successful referrers were added as friends to our official Line account, to open a chat room and send messages. Then, a QR code for our official Line account was provided to index subjects. They could post a message to share the QR code or our official Line account directly to their sexual partners via the in-app function. After reading our recruitment message, the sexual partners could then add our official account as a friend and make an appointment for HIV testing with our staff through the chat room. The referral process is depicted in Figures 1 and 2.

Figure 1. Flowchart of the referral process using the Line app.

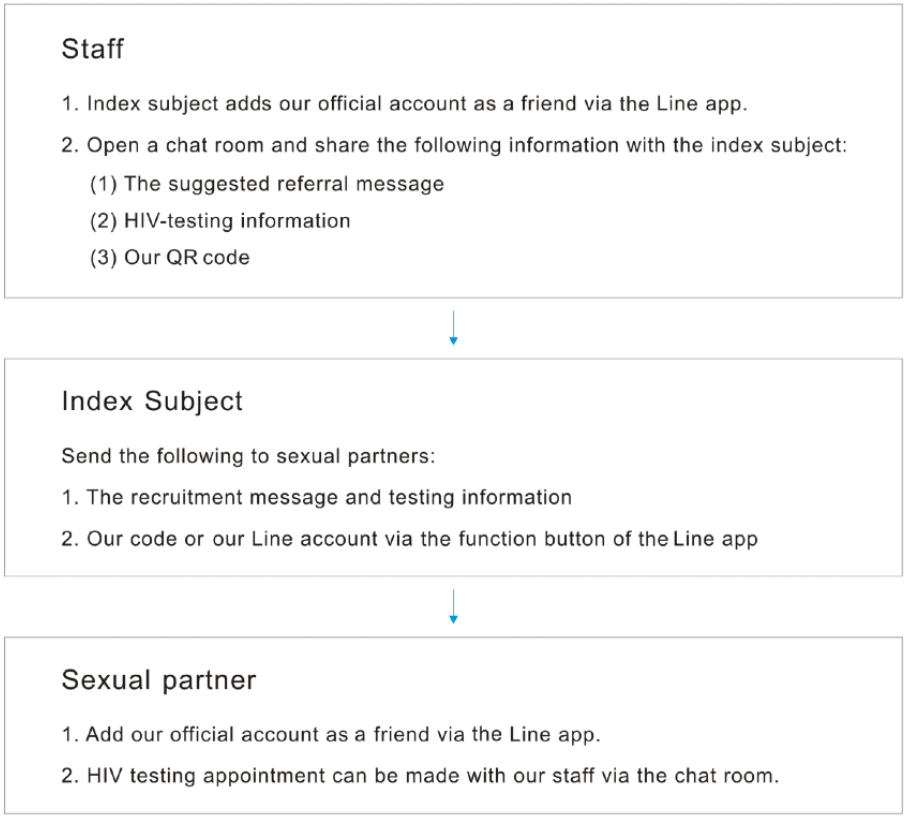
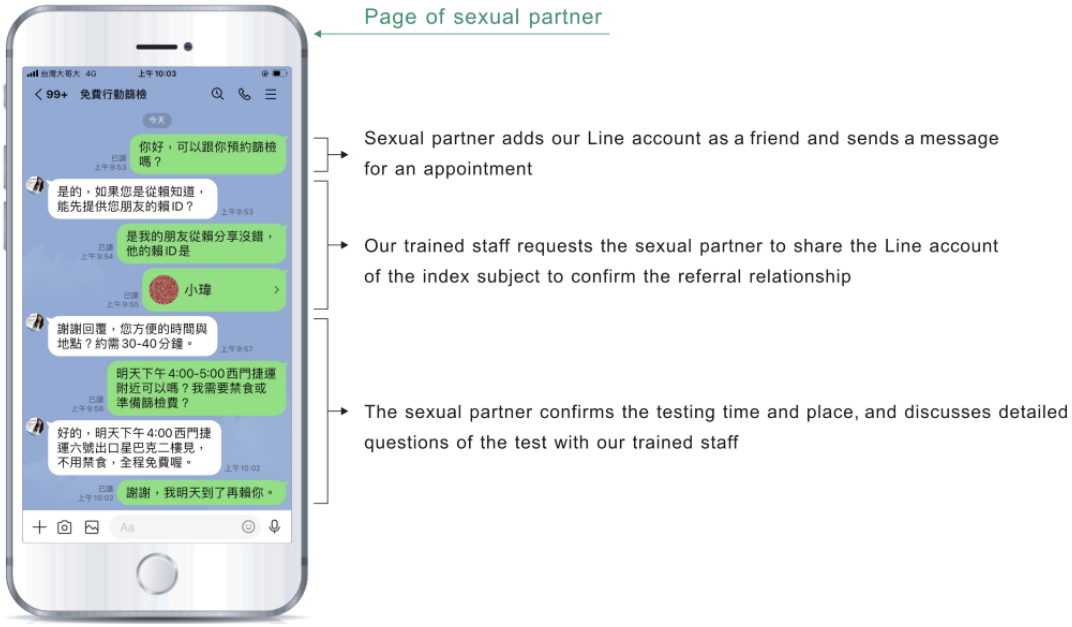


Figure 2. A screenshot of confirmation of referral and the testing appointment.



Partner-elicited interviews were conducted with the tested sexual partners to encourage further referrals. Sexual partners' test appointments and results were checked weekly. If the referred sexual partner did not contact our official Line account, the referrer was requested to remind the partner to contact us.

The trained staff accompanied and referred all participants with HIV-positive results to the HIV case manager at the hospital for HIV confirmation, diagnosis, assessment, and treatment.

Partner-Elicited Interviews

The partner-elicited interviews were designed around the theoretical concept of helpfulness and disclosure [14-16] (Multimedia Appendix 1). The interview process proceeded smoothly as it commenced with an empathic and inspiring statement, followed by an in-depth discussion guided by open-ended questions [28]. Participants were motivated to become referrers via verbal persuasion, which promoted self-awareness about their HIV status and facilitated discussions on safe sexual behaviors and HIV prevention measures in partner relationships. The benefits for their referred sexual partners were then explained, including 2 easy steps via the Line app to link their sexual partners with our staff to make an appointment and receive a free and anonymous rapid HIV test in a convenient and personalized manner. If the participants were unwilling to directly refer their sexual partners, the staff elicited them as helpers by asking them to disclose information about their sexual partners, such as their accounts and profiles on specific social networking platforms, and their current HIV status, thus helping to discover unknown HIV-related sexual affiliations within the MSM community.

The content validity of partner-elicited interviews was examined through theoretical consistency and expert validity. The trained staff who conducted the interviews had completed a 5-hour training course related to HIV testing, counseling, and practice, presented by the Taiwan AIDS Nurse Association, and an additional 8 hours of internship for consistency training before participating in the formal research.

The Demographics, Risk Behaviors, and HIV-Testing Experience Questionnaire

The demographics portion of the questionnaire asked participants about their age, education, employment, religion, marital status, sexual orientation, and sexual preferences, while the risk behavior section asked about their experiences with finding sexual partners through social media, age of sexual debut, number of sexual partners in the past 3 months, frequency of anal intercourse and condom use in the past 3 months, and experiences with recreational drug use. The HIV-testing experience questionnaire asked about experiences with HIV testing and frequency of HIV testing.

Sexual Partner Record

A data collection form was used to record information about referred/disclosed sexual partners of index subjects. Contents included nicknames and/or contact information (eg, user account and profile information on social media platforms or the last 4 digits of their phone number), or any other information about sexual partners, including the selection of HIV testing appointments (descriptions of the time/location/date and any special requirements for delivering an HIV test) and the current HIV status.

Rapid HIV Testing Equipment

The third generation of the Alere Determine HIV-1/2 was used to test for HIV-1/HIV-2 antibodies in human serum or plasma. It was licensed by the Taiwan Food and Drug Administration, with a sensitivity of 99.75% and a specificity of 99.87% [29]. Tests took approximately 15 minutes to complete, and the test results were managed and recorded by trained staff.

Analysis Procedure

The primary outcome was HIV-positive results. Secondary outcomes were the referral rate; the comparison of demographics, risk behaviors, and HIV-testing experiences; and quantified findings of the social network graph. Demographic data were analyzed by calculating the means of

continuous variables and the percentages of categorical variables. A one-way analysis of variance, Fisher least significant difference test, and chi-square test were used to investigate the differences and associations in participants' characteristics, risk behaviors, HIV-testing experiences, and HIV test results between different groups.

The UCINET software package [30] was used to analyze the social network at the ego level and the output of the social network graph. A node was defined as an individual in the network. The numbers of the intermediate nodes linking each HIV-negative node of tested and untested sexual partners to the closest HIV-positive node were calculated (the calculation was not repeated), to present the influence of each node in the possible distance of HIV transmission among the social networking platform users. Unique codes were used within the network graph to maintain anonymity.

Ethical Statement

This study was approved by the institutional review board of MacKay Memorial Hospital (17MMHISO68e). Written informed consent was obtained from all participants, who were given information about the study's purpose, procedures, potential benefits and harms, and confidentiality. All information was collected and processed anonymously and replaced with unique codes. At the end of the research, all the accounts and records of the research participants and untested sexual partners were deleted to protect their privacy. Each participant received a compensatory voucher worth US $6.66 after completing their interviews and rapid HIV tests. Participants who successfully referred their sexual partners received an additional voucher of the same value.

Results

During the study period, 279 MSM visited our screening stations for rapid HIV tests, all of whom met the inclusion criteria. Of these, 4.7% (13/279) refused to participate owing to not having enough time to attend the interviews and 95.3% (266/279) agreed to participate in the partner-elicited interviews, with 28.2% (75/266) agreeing to refer/disclose information about their sexual partners and become index subjects. The remaining 71.8% (191/266) of subjects without referral provided the following reasons for declining: 32.5% (62/191) reported that their sexual partners had recently undergone HIV tests, 30.4% (58/191) preferred to respect the privacy of their sexual partners, 25.7% (49/191) did not have any information about their sexual partners, and 11.5% (22/191) simply declined to answer.

At the end of the recruitment period, a total of 167 sexual partners were referred/disclosed, and of these, 127 had been referred via the Line app and 40 were disclosed by index subjects. All 127 who were referred via the app made appointments and completed the rapid HIV test conducted by our trained staff, at a time and place designated for their convenience. As for the 40 partners who had not been referred for HIV testing, their social networking accounts and profiles on platforms where the index subjects first met them were disclosed by the index subjects, including geosocial network apps (Hornet: 14/40, 35.0%; Grindr: 7/40, 17.5%; and Jack'd: 4/40, 10.0%), online communities (Facebook: 6/40, 15.0% and Twitter: 5/40, 12.5%), and the Line app (4/40, 10.0%). Figure 3 presents the detailed recruitment process.

Figure 3. Diagram of the recruitment process. MSM: men who have sex with men.

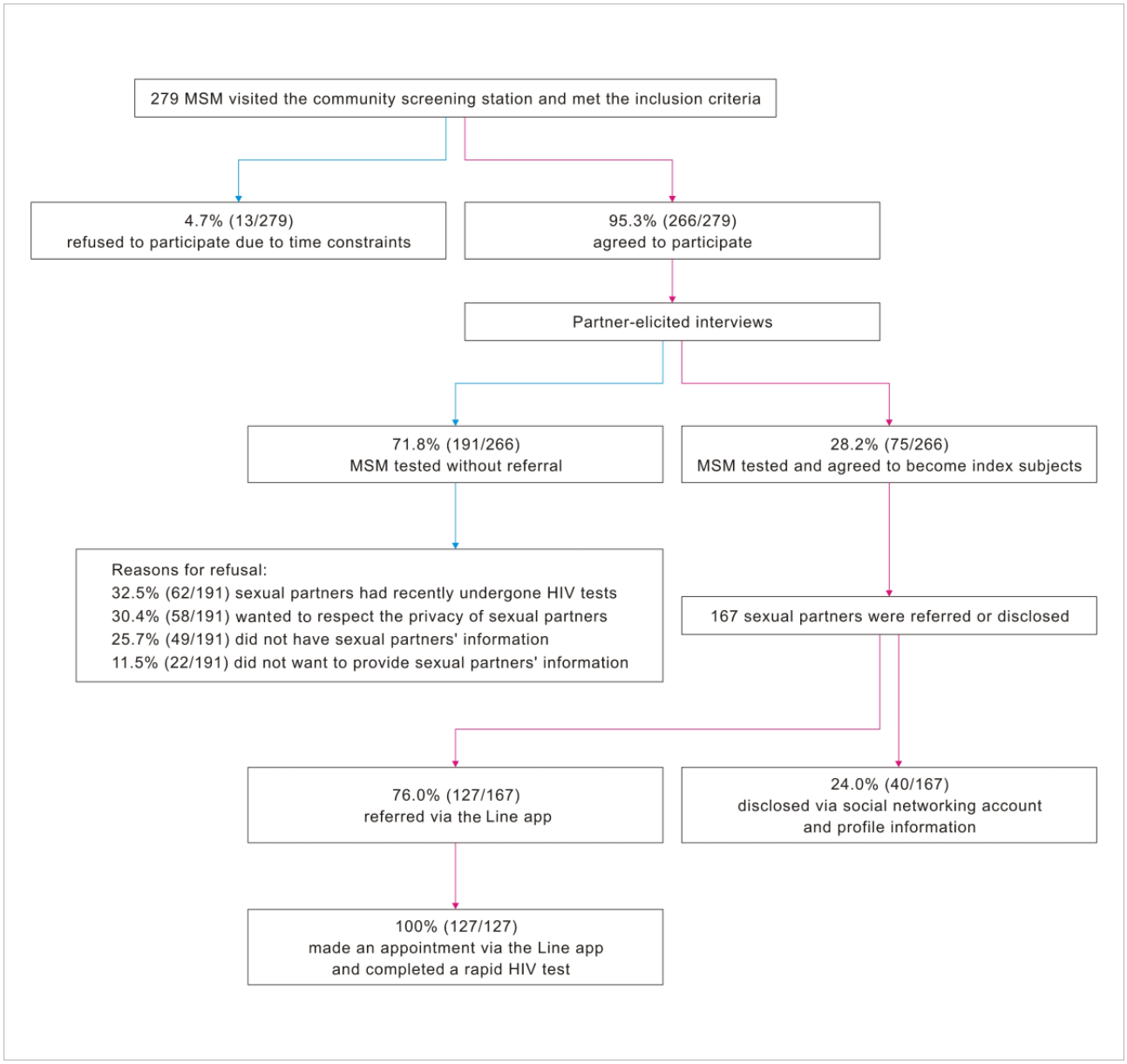


Table 1 presents comparisons among the MSM tested without referrals, index subjects, and tested sexual partners. As indicated, the mean age (F_2=6.9, P=.001) and age of sexual debut (F_2=5.5, P=.005) of the index subjects and tested sexual partners were significantly lower than those of the MSM tested without referrals.

Table 1. Comparisons of participants' demographics, risk behaviors, HIV-testing experiences, and testing results (N=393).

Variable	MSM[a] tested without referral[b] (n=191)	Index subject[c] (n=75)	Tested sexual partner[d] (n=127)	LSD test[e]	*F* value (*df*)	χ^2 (*df*)	*P* value
Age (years), mean (SD)	32.7 (8.5)	29.5 (6.3)	29.9 (7.5)	b > c, d	6.9 (2)	N/A[f]	.001
Education, n (%)				N/A	N/A	4.6 (2)	.33
Above university	37 (19.4)	16 (21.3)	20 (15.7)				
College or university	133 (69.6)	48 (64.0)	83 (65.4)				
High school or lower	21 (11.0)	11 (14.7)	24 (18.9)				
Employment, n (%)				N/A	N/A	3.7 (2)	.44
Employed	140 (73.3)	55 (73.3)	83 (65.4)				
Unemployed	26 (13.6)	11 (14.7)	27 (21.3)				
Student	25 (13.1)	9 (12.0)	17 (13.4)				
Marital status, n (%)				N/A	N/A	1.9 (2)	.75
Single	179 (93.7)	73 (97.3)	121 (95.3)				
Married	9 (4.7)	1 (1.3)	4 (3.1)				
Divorced	3 (1.6)	1 (1.3)	2 (1.6)				
Sexual orientation, n (%)				N/A	N/A	0.2 (1)	.89
Homosexual	166 (86.9)	65 (86.7)	108 (85.0)				
Bisexual	25 (13.1)	10 (13.3)	19 (15.0)				
Preferred position, n (%)				N/A	N/A	8.9 (2)	.06
Versatile	100 (52.4)	32 (42.7)	62 (48.8)				
Top	51 (26.7)	31 (41.3)	31 (24.4)				
Bottom	40 (20.9)	12 (16.0)	34 (26.8)				
Finding sexual partners through social media[g,h], n (%)				N/A	N/A	0.7 (1)	.72
Yes	131 (68.6)	48 (64.0)	83 (65.4)				
No	60 (31.4)	27 (36.0)	44 (34.6)				
Age of sexual debut (years), mean (SD)	21.5 (5.2)	19.7 (3.7)	20.0 (4.6)	b > c, d	5.5 (2)	N/A	.005
Number of sexual partners[g], mean (SD)	2.1 (1.9)	3.1 (5.5)	2.9 (2.5)	b < c, d	3.8 (2)	N/A	.02
Number of times engaging in anal intercourse[h], mean (SD)	4.0 (3.2)	6.8 (7.0)	6.1 (6.1)	b < c, d	10.1 (2)	N/A	<.001
Frequency of condom use during anal intercourse[h,i], n (%)				N/A	N/A	7.97 (3)	.24
Used every time	69 (41.1)	25 (39.1)	46 (42.2)				
Frequently used	55 (32.7)	17 (26.6)	25 (22.9)				
Rarely used	40 (23.8)	17 (26.6)	29 (26.6)				
Never used	4 (2.4)	5 (7.8)	9 (8.3)				
Experiences with recreational drug use, n (%)				N/A	N/A	1.1 (1)	.55
No	168 (88.0)	63 (84.0)	107 (84.3)				
Yes	23 (12.0)	12 (16.0)	20 (15.7)				
Experiences of HIV testing, n (%)				N/A	N/A	6.1 (1)	.047
Yes	176 (92.1)	64 (85.3)	106 (83.5)				
No	15 (7.9)	11 (14.7)	21 (16.5)				

Variable	MSM[a] tested without referral[b] (n=191)	Index subject[c] (n=75)	Tested sexual partner[d] (n=127)	LSD test[e]	*F* value (*df*)	χ^2 (*df*)	*P* value
Interval of regular HIV testing (n=176), n (%)				N/A	N/A	15.4 (3)	.02
Three months	41 (23.3)	18 (28.6)	14 (13.1)				
Half year	62 (35.2)	21 (33.3)	41 (38.3)				
One year	35 (19.9)	6 (9.5)	13 (12.1)				
Not regularly tested	38 (21.6)	18 (28.6)	39 (36.4)				
HIV-testing results, n (%)				N/A	N/A	0.9 (1)	.64
Negative	189 (99.0)	74 (98.7)	124 (97.6)				
Newly positive	2 (1.0)	1 (1.3)	3 (2.4)				

[a]MSM: men who have sex with men.

[b]Those who did not refer or disclose any sexual partner.

[c]Those who agreed to refer or disclose their sexual partners during the first wave.

[d]Those who were referred through the Line app and completed the rapid HIV test by the trained staff.

[e]Fisher least significant difference test.

[f]N/A: not applicable.

[g]Including mobile instant messaging or geosocial apps and online communities.

[h]In the past 3 months.

[i]N=341; due to missing data.

Further, the mean number of sexual partners (F_2=3.8, *P*=.02) and mean frequency of anal intercourse (F_2=10.1, *P*<.001) during the preceding 3 months were significantly higher for both the index subjects and tested sexual partners than for the MSM tested without referrals. The percentages of those who had not previously undergone HIV testing (χ^2_1=6.1, *P*=.047) and did not regularly undergo testing (χ^2_2=15.4, *P*=.02) were also significantly higher among the index subjects and tested sexual partners than the MSM tested without referrals. The HIV-seropositivity rate of the tested sexual partners was 2.4%, which was higher than the rates for both the index subjects (1.3%) and the MSM tested without referrals (1.0%); these did not show statistically significant differences.

Figure 4 shows the graph after partner-elicited interviews of 75 index subjects, to which were added 127 tested and 40 untested sexual partners. Three newly diagnosed HIV nodes and 4 previously known HIV-positive nodes were discovered, and this resulted in an HIV-positive rate of 4.2% (7/167) for all sexual partners. The sociometric structure in this social network graph included 59 pairs and groups of chain types (n=134). There were 3 Y-type groups (n=16), 4 star-type groups (n=32), and 2 groups of the complicated type (n=60). In the largest complicated type (n=48), the number of HIV-positive nodes was also the largest (n=5), and 4 sexual partners created a bridge with direct sexual affiliations between an HIV-positive node and the other nodal clusters, thus playing an important role as HIV transmission intermediaries. Of these, 2 tested bridges had been referred via the Line app and 2 untested bridges had been disclosed, namely the Hornet and Grindr apps, where the index subject first met the sexual partners.

Figure 4. The social network graph.

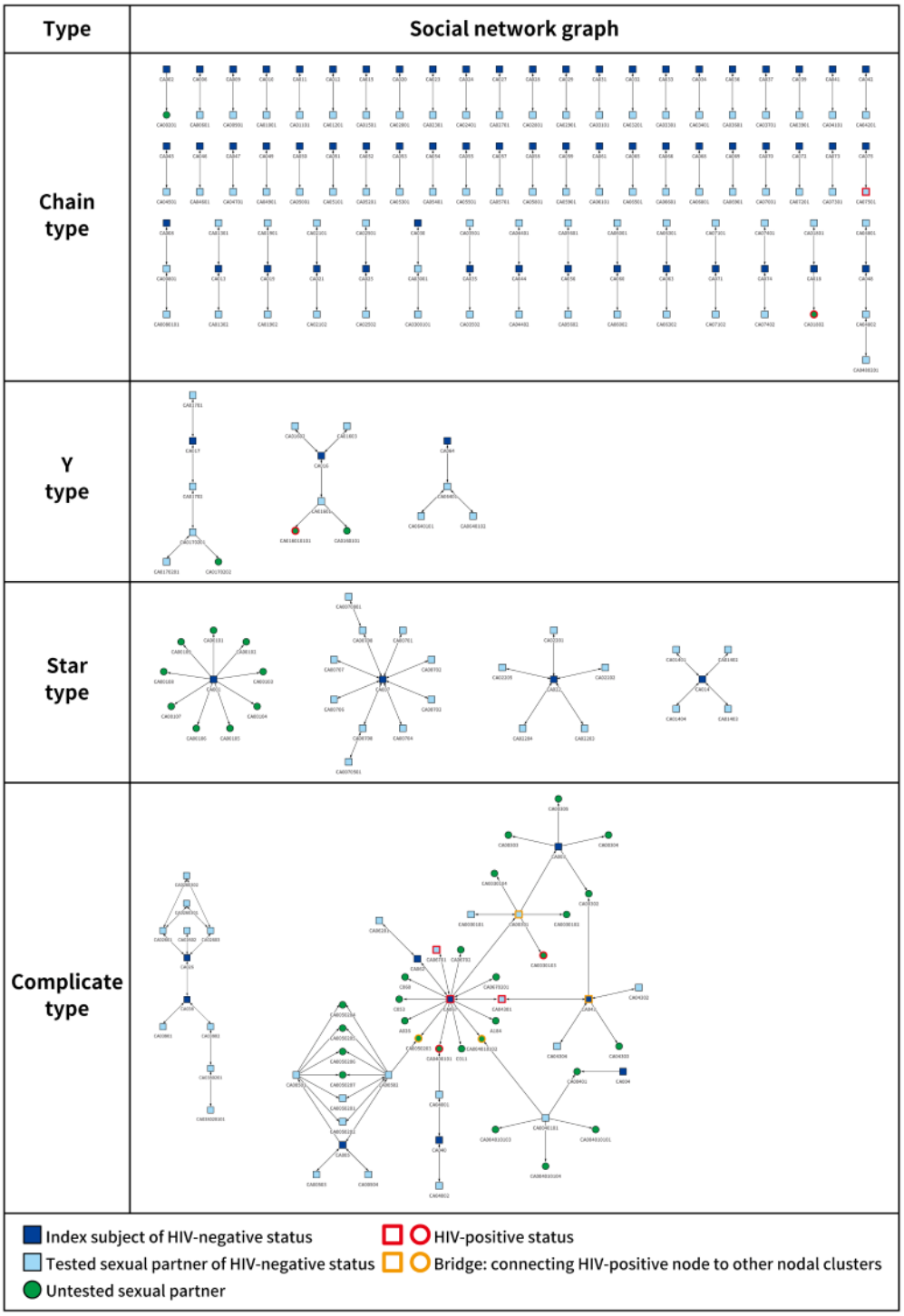


Table 2 presents the network graph analysis of the number of intermediate nodes, which link 160 HIV-negative tested (n=124) and untested (n=36) sexual partners to the closest HIV-positive nodes (n=8). Of them, 73.1% (117/160) had no sexual affiliation with the HIV-positive node and 26.9% (43/160) had sexual affiliations with the HIV-positive node directly or through 1 to 3 intermediate nodes. Untested sexual partners reflected a higher proportion (10/25, 40%) of direct sexual affiliations with HIV-positive nodes than tested sexual partners (3/18, 17%).

Table 2. Data on intermediate nodes linking HIV-negative nodes of sexual partners to the closest HIV-positive nodes.

Variable[a]	Total HIV-negative sexual partners (n=160), n (%)	Tested sexual partners[b] (n=124), n (%)	Untested sexual partners[c] (n=36), n (%)
Not linking to any HIV-positive node	117 (73.1)	106 (90.6)	11 (9.4)
Linking to the closest HIV-positive node	43 (26.9)	18 (41.9)	25 (58.1)
Directly	13 (30.2)	3 (16.7)	10 (40.0)
Through 1 intermediate node	11 (25.6)	7 (38.9)	4 (16.0)
Through 2 intermediate nodes	16 (37.2)	5 (27.8)	11 (44.0)
Through 3 intermediate nodes	3 (7.0)	3 (16.7)	0 (0)

[a]Each node was calculated at the closest distance to an HIV-positive node; the calculation was not repeated.

[b]Those who were referred through the Line app and completed the rapid HIV test by our trained staff.

[c]Those whose accounts and profiles on social media platforms where they first met the referrer were disclosed by the referral to our staff.

Discussion

Principal Findings

This study's main finding was that the partner-elicited interview successfully persuaded MSM to refer their sexual partners through the Line app to promote HIV testing. Eliciting referral via tested MSM could target sexual partners with behaviors carrying higher HIV risk. The data collected through this referral mode may contribute to understanding hitherto unknown social network affiliations.

The concept of helping and disclosing proved successful for developing and guiding the partner-elicited interviews in this research. Feedback from the referrers who had agreed to refer or disclose after the partner-elicited interviews indicates that they felt altruistic, had noncritical attitudes, and felt empowered to share the messages from the trained staff during the interview process. Moreover, the application of easy referral steps via the familiar and private interface [31] of the Line app was a key factor that motivated the referrers to be willing to refer their sexual partners for HIV testing. The interviews were conducted prior to HIV testing and subsequent test results. Referrers did not have to bear the negative consequences of notifying their sexual partners of an HIV diagnosis [32], which may have been a contributing factor to securing referrers' cooperation.

All the sexual partners (127/127, 100.0%) referred via the Line app completed the rapid HIV tests by our trained staff during the research period, which signifies that the referral strategy was highly feasible and that the recruitment exercise had been successful. One possible reason for the high testing rate of the sexual partners could be the thorough preparation of the recruiters and the trust relationship between the referrers and peers [33]. We provided a suggested referral message to share with sexual partners to explain the reason for the referral and endorse the testing experience. The relationship tends to be more intimate than a causal relationship when the Line app is used to deliver private messages related to sexual behavior and HIV testing [34]. In addition, the homepage of our official Line account had a public and credible introduction [35], and both these methods increased the sexual partners' trust to add the recommended Line account. Our model emphasized the web discussion of the HIV testing appointment individually in the private chat room of the Line app and a free, anonymous, rapid HIV test at a designated time and place, which contributed to the willingness of MSM to undergo the HIV test [17,36].

The partner-elicited interviews could be better used to target at-risk MSM and promote the HIV case findings. These results are similar to those of previous studies that used contact tracing and partner notification of HIV-positive index subjects in China [37,38]. Although the new HIV-positive rate of 10.5% to 11.1% among sexual partners in the previously mentioned study is higher than the rate of 2.4% in our study, our results still revealed that HIV case findings can be mobilized earlier among MSM via social networking platforms prior to an HIV-positive diagnosis. During the study period, Taiwan began to promote pre-exposure prophylaxis for the HIV risk group, accompanied by a large number of mobile screening programs in the community, which resulted in a sharp drop in the HIV-positive rate [39,40]. This may also be one of the reasons for the low positive rate in this study.

The accounts and profiles from social networking platforms used to refer/disclose sexual partners can be captured and exported to the network graph to present unknown and complex sexual affiliations [41]. This study uncovered a diverse network structure. The application of social network analysis (eg, multilevel modeling) is needed in the future to obtain insights into the different network structures. A high percentage of the tested MSM did not refer/disclose their sexual partners in this study, and we therefore cannot exclude the possibility that most HIV risk-taking information within the social network remains unknown. Moreover, it is worth noting that more untested sexual partners were at high risk of HIV infection via direct sexual affiliations with HIV-positive nodes. A survey also found that a high percentage (82.3%) of people newly diagnosed with HIV/AIDS had sex without a condom with an average of 2 sexual partners [42]. In the current network, 2 untested bridges need urgent testing and education on safe sex, thus helping to break the current HIV transmission risks affecting the nodal cluster [43]. However, due to the law of personal data protection in Taiwan, the trained staff are not allowed to contact the 2 untested bridges. Hence, the index subject forms an important channel through which health care providers may be able to provide an HIV self-test tool or link these sexual partners, who cannot be contacted otherwise, to testing resources. In addition

to sexual partners' referral, providing education to change risky behaviors, such as condomless anal sex, during pretest and posttest counseling may improve treatment and adherence to highly active antiretroviral therapy in the future [44].

Most participants in this study met new sexual partners through social media, which reflects the findings of previous studies [45]. In addition to continuously eliciting network information from MSM during interviews, it is important to develop innovative measures for health care providers to directly reach out and provide testing resources through these same applications and communities.

Limitations

This study has several limitations. First, we directly recruited MSM who were willing to refer or disclose their sexual partners, which was not a random selection. This excluded many individuals who were unwilling to participate, thus affecting the study's generalizability. Second, this study was conducted at a community screening station in Taipei City, which limits the inferences that can be made. Third, the social network was analyzed at the end of the data collection process, which resulted in static presentation. For this reason, causal relationships could not be confirmed. Fourth, the research period lasted only 6 months, which may have prevented a thorough exploration of the social networks. Future related studies should thus be conducted at multiple screening stations over a longer research period, thereby increasing the amount of data available to complete and analyze the social networks. Finally, there are currently no specific criteria to determine the validity of referrals. A comparison with traditional approaches is needed.

Conclusion

Partner-elicited interviews could promote the referral/disclosure of sexual partners via social networking platforms among MSM. The motivational interviewing and referral process via the Line app used in this research can be directly applied in the education and practical training of HIV-screening consultants. Further, this may enhance the uptake of HIV testing and help target HIV-at-risk sexual networks. This study's social network analysis revealed the types and unknown details of the direct and indirect sexual affiliations of each HIV-negative and HIV-positive node, thus promoting a better understanding of the possible distance of HIV transmission within the network. Emerging screening methods (eg, tests delivered through peer referrals or the direct recruitment of MSM through social media) are needed to actively target untested MSM engaging in condomless anal sex. These results may be useful for those attempting to improve current HIV screening programs and contact tracing for newly diagnosed case findings and surveillance.

Acknowledgments

The authors thank the Centers for Disease Control, Taiwan, ROC (MOHW106-CDC-C-114-000115) for providing research funding. The authors sincerely thank all study participants for providing important information about their sexual partners.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The contents of partner-elicited interviews.
[DOCX File , 19 KB - publichealth_v8i4e32156_app1.docx]

References

1. 2025 AIDS Targets. United Nations Programme on HIV/AIDS. URL: https://aidstargets2025.unaids.org/ [accessed 2021-07-21]
2. Nguyen PT, Gilmour S, Le PM, Onishi K, Kato K, Nguyen HV. Progress toward HIV elimination goals: trends in and projections of annual HIV testing and condom use in Africa. AIDS 2021 Jul 01;35(8):1253-1262. [doi: 10.1097/QAD.0000000000002870] [Medline: 33730746]
3. The 2020 HIV Self-Screening Program is launched! Regular screening for peace of mind. Taiwan Centers for Disease Control. URL: https://www.cdc.gov.tw/Bulletin/Detail/P7tQd645gMwWVa8jCDkJ_Q?typeid=9 [accessed 2020-05-01]
4. 90-90-90: Treatment for all. United Nations Programme on HIV/AIDS. URL: https://www.unaids.org/en/resources/presscentre/featurestories/2020/september/20200921_90-90-90 [accessed 2020-09-21]
5. Friedman SR, Aral S. Social networks, risk-potential networks, health, and disease. J Urban Health 2001 Sep;78(3):411-418 [FREE Full text] [doi: 10.1093/jurban/78.3.411] [Medline: 11564845]
6. Lee SS, Tam DKP, Ho RLM, Wong KH. Social network methodology for studying HIV epidemiology in men having sex with men. J Infect Public Health 2009;2(4):177-183 [FREE Full text] [doi: 10.1016/j.jiph.2009.09.002] [Medline: 20701880]
7. Pham MD, Aung PP, Paing AK, Pasricha N, Agius PA, Tun W, et al. Factors associated with HIV testing among young men who have sex with men in Myanmar: a cross-sectional study. J Int AIDS Soc 2017 Nov;20(3):e25026 [FREE Full text] [doi: 10.1002/jia2.25026] [Medline: 29105323]

8. Wimonsate W, Pattanasin S, Ungsedhapand C, Pancharoen K, Luechai P, Satumay K, et al. Repeat HIV testing among HIV-uninfected men who have sex with men attending Silom Community Clinic, Bangkok, 2011 - 2014. Int J STD AIDS 2018 Dec;29(14):1417-1423 [FREE Full text] [doi: 10.1177/0956462418788724] [Medline: 30115000]
9. United Nations Programme on HIV/AIDS. URL: http://www.unaids.org/en/resources/fact-sheet [accessed 2021-07-21]
10. Niekamp A, Mercken LA, Hoebe CJ, Dukers-Muijrers NH. A sexual affiliation network of swingers, heterosexuals practicing risk behaviours that potentiate the spread of sexually transmitted infections: A two-mode approach. Social Networks 2013 May;35(2):223-236. [doi: 10.1016/j.socnet.2013.02.006]
11. Centers for Disease Control and Prevention (CDC). Use of social networks to identify persons with undiagnosed HIV infection--seven U.S. cities, October 2003-September 2004. MMWR Morb Mortal Wkly Rep 2005 Jun 24;54(24):601-605 [FREE Full text] [Medline: 15973240]
12. Girault P, Green K, Clement NF, Rahman YAA, Adams B, Wambugu S. Piloting a Social Networks Strategy to Increase HIV Testing and Counseling Among Men Who Have Sex with Men in Greater Accra and Ashanti Region, Ghana. AIDS Behav 2015 Nov;19(11):1990-2000. [doi: 10.1007/s10461-015-1069-z] [Medline: 25903507]
13. de Munnik S, den Daas C, Ammerlaan HSM, Kok G, Raethke MS, Vervoort SCJM. Let's talk about sex: A qualitative study exploring the experiences of HIV nurses when discussing sexual risk behaviours with HIV-positive men who have sex with men. Int J Nurs Stud 2017 Nov;76:55-61. [doi: 10.1016/j.ijnurstu.2017.09.002] [Medline: 28938103]
14. Neequaye DA, Ask K, Granhag PA, Vrij A. Facilitating disclosure in intelligence interviews: The joint influence of helpfulness priming and interpersonal approach. J Investig Psychol Offender Profil 2018 Sep 20;15(3):319-334. [doi: 10.1002/jip.1515]
15. Arieli S, Grant AM, Sagiv L. Convincing Yourself to Care About Others: An Intervention for Enhancing Benevolence Values. J Pers 2014 Feb 08;82(1):15-24. [doi: 10.1111/jopy.12029] [Medline: 23301825]
16. Capraro V, Smyth C, Mylona K, Niblo GA. Benevolent characteristics promote cooperative behaviour among humans. PLoS One 2014 Aug 20;9(8):e102881 [FREE Full text] [doi: 10.1371/journal.pone.0102881] [Medline: 25140707]
17. Chiou PY, Ko NY, Chien CY. Mobile HIV Testing Through Social Networking Platforms: Comparative Study. J Med Internet Res 2021 Mar 26;23(3):e25031 [FREE Full text] [doi: 10.2196/25031] [Medline: 33769298]
18. Beymer MR, Rossi AD, Shu SB. Assessing Self-Control and Geosocial Networking App Behavior Among an Online Sample of Men Who Have Sex with Men. J Urban Health 2016 Aug;93(4):698-708 [FREE Full text] [doi: 10.1007/s11524-016-0056-7] [Medline: 27184573]
19. Chan PA, Towey C, Poceta J, Rose J, Bertrand T, Kantor R, et al. Online Hookup Sites for Meeting Sexual Partners Among Men Who Have Sex with Men in Rhode Island, 2013: A Call for Public Health Action. Public Health Rep 2016 Mar;131(2):264-271 [FREE Full text] [doi: 10.1177/003335491613100210] [Medline: 26957661]
20. Holloway IW, Rice E, Gibbs J, Winetrobe H, Dunlap S, Rhoades H. Acceptability of smartphone application-based HIV prevention among young men who have sex with men. AIDS Behav 2014 Feb;18(2):285-296 [FREE Full text] [doi: 10.1007/s10461-013-0671-1] [Medline: 24292281]
21. Ko NY, Tseng PC, Huang YC, Chen YC, Hsu ST. Seeking sex partners through the internet and mobile phone applications among men who have sex with men in Taiwan. AIDS Care 2016 Jul;28(7):927-931. [doi: 10.1080/09540121.2015.1131969] [Medline: 26754350]
22. Young SD, Cumberland WG, Nianogo R, Menacho LA, Galea JT, Coates T. The HOPE social media intervention for global HIV prevention in Peru: a cluster randomised controlled trial. The Lancet HIV 2015 Jan;2(1):e27-e32 [FREE Full text] [doi: 10.1016/S2352-3018(14)00006-X] [Medline: 26236767]
23. Ko NY, Hsieh CH, Wang MC, Lee C, Chen CL, Chung AC, et al. Effects of Internet popular opinion leaders (iPOL) among Internet-using men who have sex with men. J Med Internet Res 2013 Feb 25;15(2):e40 [FREE Full text] [doi: 10.2196/jmir.2264] [Medline: 23439583]
24. Rhodes SD, McCoy TP, Tanner AE, Stowers J, Bachmann LH, Nguyen AL, et al. Using Social Media to Increase HIV Testing Among Gay and Bisexual Men, Other Men Who Have Sex With Men, and Transgender Persons: Outcomes From a Randomized Community Trial. Clin Infect Dis 2016 Jun 01;62(11):1450-1453 [FREE Full text] [doi: 10.1093/cid/ciw127] [Medline: 26980878]
25. Zou H, Wu Z, Yu J, Li M, Ablimit M, Li F, et al. Internet-facilitated, voluntary counseling and testing (VCT) clinic-based HIV testing among men who have sex with men in China. PLoS One 2013;8(2):e51919 [FREE Full text] [doi: 10.1371/journal.pone.0051919] [Medline: 23418417]
26. Conserve DF, Bay C, Kilonzo MN, Makyao NE, Kajula L, Maman S. Sexual and Social Network Correlates of Willingness to Self-Test for HIV Among Ever-Tested and Never-Tested Men: Implications for the Tanzania STEP Project. AIDS Care 2019 Feb;31(2):169-176 [FREE Full text] [doi: 10.1080/09540121.2018.1537466] [Medline: 30362377]
27. Zweig KA. Graph Theory, Social Network Analysis, and Network Science. In: Zweig KA, editor. Network Analysis Literacy. Lecture Notes in Social Networks. Vienna: Springer; 2016:23-55.
28. Beach MC, Laws MB, Rose G, Roter D, Lee Y, Chander G, et al. Effects of Minimal Versus Intensive Intervention to Enhance Motivational Interviewing in HIV Care. AIDS Behav 2018 Jan;22(1):276-286 [FREE Full text] [doi: 10.1007/s10461-017-1794-6] [Medline: 28578544]

29. "Abbott" DETERMINE HIV-1/2. Taiwan Food and Drug Administration. URL: https://www.fda.gov.tw/MLMS/H0001D3.aspx?LicId=06010855 [accessed 2021-07-29]
30. Borgatti SP, Everett MG, Freeman LC. Ucinet for Windows: Software for Social Network Analysis. UCINET Software. URL: https://sites.google.com/site/ucinetsoftware/home [accessed 2021-07-29]
31. Iribarren SJ, Ghazzawi A, Sheinfil AZ, Frasca T, Brown W, Lopez-Rios J, et al. Mixed-Method Evaluation of Social Media-Based Tools and Traditional Strategies to Recruit High-Risk and Hard-to-Reach Populations into an HIV Prevention Intervention Study. AIDS Behav 2018 Jan;22(1):347-357 [FREE Full text] [doi: 10.1007/s10461-017-1956-6] [Medline: 29124420]
32. Zhang K, Zhao J, Li X, Chen X, Wang H, Williams AB, et al. Perceived Facilitators and Barriers regarding Partner Notification in People Living With HIV in Hunan, China: A Qualitative Study From the Patient Perspective. J Assoc Nurses AIDS Care 2019 Nov;30(6):658-667. [doi: 10.1097/JNC.0000000000000093] [Medline: 31574528]
33. Stojanovski K, Naja-Riese G, King EJ, Fuchs JD. A Systematic Review of the Social Network Strategy to Optimize HIV Testing in Key Populations to End the Epidemic in the United States. AIDS Behav 2021 Sep 19;25(9):2680-2698 [FREE Full text] [doi: 10.1007/s10461-021-03259-z] [Medline: 33871730]
34. Macapagal K, Coventry R, Puckett JA, Phillips G, Mustanski B. Geosocial Networking App Use Among Men Who Have Sex With Men in Serious Romantic Relationships. Arch Sex Behav 2016 Aug;45(6):1513-1524 [FREE Full text] [doi: 10.1007/s10508-016-0698-2] [Medline: 26969319]
35. Martinez O, Wu E, Shultz AZ, Capote J, López Rios J, Sandfort T, et al. Still a hard-to-reach population? Using social media to recruit Latino gay couples for an HIV intervention adaptation study. J Med Internet Res 2014 Apr 24;16(4):e113 [FREE Full text] [doi: 10.2196/jmir.3311] [Medline: 24763130]
36. Han L, Wei C, Muessig KE, Bien CH, Meng G, Emch ME, et al. HIV test uptake among MSM in China: Implications for enhanced HIV test promotion campaigns among key populations. Glob Public Health 2017 Jan;12(1):31-44 [FREE Full text] [doi: 10.1080/17441692.2015.1134612] [Medline: 26785328]
37. Fu X, Qi J, Hu Y, Pan X, Li Y, Liu H, et al. Partner notification in cooperation with community-based organizations among HIV-positive men who have sex with men in two Chinese cities. Int J STD AIDS 2016 Sep;27(10):821-831 [FREE Full text] [doi: 10.1177/0956462416648827] [Medline: 27140946]
38. Lian Y, Zhao Y, Wang J, Shi C, Wang K, Wang X, et al. A health communication intervention to integrate partner testing with antiretroviral therapy service among men who have sex with men in China: an observational cohort study. BMC Public Health 2018 Nov 06;18(1):1235 [FREE Full text] [doi: 10.1186/s12889-018-6147-5] [Medline: 30400787]
39. New strategies to prevent and treat AIDS, the CDC has launched a Pre-Exposure Prophylaxis administration program, which enables high-risk group to reduce the risk of infection. Taiwan Centers for Disease Control. URL: https://www.mohw.gov.tw/fp-2633-10279-1.html [accessed 2020-05-20]
40. Overview of the AIDS Epidemic Situation in Taiwan in 2018. Taiwan Centers for Disease Control. URL: https://www.mohw.gov.tw/dl-56145-bcfce12b-2271-4b43-b2f4-0210123721b9.html [accessed 2020-06-22]
41. Pennise M, Inscho R, Herpin K, Owens J, Bedard BA, Weimer AC, et al. Using smartphone apps in STD interviews to find sexual partners. Public Health Rep 2015 May;130(3):245-252 [FREE Full text] [doi: 10.1177/003335491513000311] [Medline: 25931628]
42. Nguyen PT, Gilmour S, Le PM, Nguyen TT, Tanuma J, Nguyen HV. Factors associated with high-risk behaviors of people newly diagnosed with HIV/AIDS: results from a cross-sectional study in Vietnam. AIDS Care 2021 May;33(5):607-615. [doi: 10.1080/09540121.2020.1761519] [Medline: 32397740]
43. Fujimoto K, Valente TW. Social network influences on adolescent substance use: disentangling structural equivalence from cohesion. Soc Sci Med 2012 Jun;74(12):1952-1960 [FREE Full text] [doi: 10.1016/j.socscimed.2012.02.009] [Medline: 22475405]
44. Le PM, Nguyen PT, Nguyen HV, Bui DH, Vo SH, Nguyen NV, et al. Adherence to highly active antiretroviral therapy among people living with HIV and associated high-risk behaviours and clinical characteristics: A cross-sectional survey in Vietnam. Int J STD AIDS 2021 Sep;32(10):911-918. [doi: 10.1177/09564624211002405] [Medline: 33861666]
45. Chow EP, Cornelisse VJ, Read TR, Hocking JS, Walker S, Chen MY, et al. Risk Practices in the Era of Smartphone Apps for Meeting Partners: A Cross-Sectional Study Among Men Who Have Sex with Men in Melbourne, Australia. AIDS Patient Care STDS 2016 Apr;30(4):151-154. [doi: 10.1089/apc.2015.0344] [Medline: 27028181]

Abbreviations

MSM: men who have sex with men
VCT: voluntary counseling and testing

Edited by Y Khader; submitted 30.07.21; peer-reviewed by P Nguyen, Y Hu, G Phillips; comments to author 31.10.21; revised version received 29.11.21; accepted 18.12.21; published 05.04.22.

Please cite as:
Chiou PY, Hung CC, Chen CY
Sexual Partner Referral for HIV Testing Through Social Networking Platforms: Cross-sectional Study
JMIR Public Health Surveill 2022;8(4):e32156
URL: https://publichealth.jmir.org/2022/4/e32156
doi:10.2196/32156
PMID:35380540

Original Paper

Support for Texting-Based Condom Negotiation Among Forcibly Displaced Adolescents in the Slums of Kampala, Uganda: Cross-sectional Validation of the Condom Use Negotiated Experiences Through Technology Scale

Moses Okumu[1,2], PhD; Carmen H Logie[3,4], PhD; David Ansong[5], PhD; Simon Mwima[1,6,7], MPH; Robert Hakiza[8], BSc; Peter A Newman[3], PhD

[1]School of Social Work, University of Illinois, Urbana-Champaign, Urbana, IL, United States
[2]School of Social Sciences, Uganda Christian University, Mukono, Uganda
[3]Factor Inwentash Faculty of Social Work, University of Toronto, Toronto, ON, Canada
[4]United Nations University Institute for Water, Environment, and Health, Hamilton, ON, Canada
[5]School of Social Work, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States
[6]AIDS Control Program, Ministry of Health, Kampala, Uganda
[7]Bukedi Prevention Institute, Kampala, Uganda
[8]Young African Refugees for Integral Development, Kampala, Uganda

Corresponding Author:
Moses Okumu, PhD
School of Social Work, University of Illinois, Urbana-Champaign
1010 W Nevada St
Urbana, IL, 61801
United States
Phone: 1 217 333 2213
Email: okumu@illinois.edu

Abstract

Background: Promoting sexual health among forcibly displaced adolescents is a global public health priority. Digital sexual communication strategies (eg, sexting) may increase adolescents' confidence in discussing sexual health issues and negotiating condom use. However, limited evidence exists describing validated measures for text-based condom negotiation in the literature.

Objective: This study helps fill this gap by adapting and examining the psychometric properties of a condom use experience through technology (condom use negotiated experiences through technology [CuNET]) scale.

Methods: Using peer network sampling, 242 forcibly displaced adolescents (aged 16-19 years) living in Kampala's slums were recruited for participation between January and March 2018. A subscale (embarrassment to negotiate condom use) of the Multidimensional Condom Attitudes Scale was adapted to incorporate sexting, yielding CuNET. Participants were randomly assigned to calibration and validation subsamples to conduct exploratory and confirmatory factor analyses to establish and validate the scale. CuNET measured participants' support levels for texting-based condom negotiation via sexting based on gender, and multivariable logistic regression was used to explore its associations with sexual health outcomes (recent consistent condom use, access to sexual and reproductive health services, and lifetime sexually transmitted infection testing).

Results: The one-factor CuNET with the validation sample was valid (χ^2_4=5.3; *P*=.26; root mean square error of approximation=0.05, 90% CI 0.00-0.16; comparative fit index=0.99; Tucker-Lewis index=0.99; standardized root mean square residual=0.006), and reliability (Cronbach α=.98). Adolescent girls showed significantly lower levels of support for using sexting to negotiate condom use (mean 13.60, SE 0.70 vs mean 21.48, SE 1.23; *P*=.001). In multivariable analyses, a 1-point increase in the CuNET score was associated with increased odds of recent consistent condom use (adjusted odds ratio [aOR] 1.73, 95% CI 1.24-2.41) but not with access to sexual and reproductive health services (aOR 1.51, 95% CI 0.99-2.30) or lifetime sexually transmitted infection testing (aOR 0.90, 95% CI 0.64-1.26).

Conclusions: The unidimensional CuNET scale is valid and reliable for forcibly displaced adolescents living in slums in Kampala, gender-sensitive, and relevant for predicting consistent condom use among urban displaced and refugee adolescents. Further development of this scale will enable a better understanding of how adolescents use digital tools for condom negotiation.

(JMIR Public Health Surveill 2022;8(4):e27792) doi:10.2196/27792

KEYWORDS

condom negotiation; sexting; refugee and displaced adolescents; digital sexual communication; HIV prevention; gender

Introduction

Background

Promoting sexual health among forcibly displaced adolescents is a global public health priority. A 2019 systematic review of sexual health interventions in humanitarian settings highlighted the urgent need for sexual health interventions targeting adolescents [1]. Uganda hosts over 1 million forcibly displaced persons, over 72,000 of whom live in Kampala [2], with an HIV prevalence rate of more than twice (13.9%) [3] the national average (6.2%) [4]. In a 2019 cross-sectional study of youth (n=1134) living in Kampala's slums, 9.9% of participants reported HIV or sexually transmitted infection (STI) co-infection [5]. Despite the heightened risk of HIV among adolescents, condom use remains sporadic [6]. In a sample of slum-dwelling youth in Kampala attending vocational training, Swahn et al [7] found that 1 in 3 participants did not use condoms, whereas 44.9% had sex with multiple partners and 23.2% engaged in transactional sex. To improve the sexual health of vulnerable adolescents, the World Health Organization [8] is calling on public health advocates to leverage digital technologies, as digital media use is widespread among adolescents.

At the same time, there is scant evidence of valid and reliable measures for evaluating the effectiveness of digital sexual health interventions. Despite preliminary studies investigating the effect of digital sexual communication on condom self-efficacy and use, there is limited evidence of validated measures to study this effect [9,10]. This knowledge gap impedes researchers' understanding of why and how today's adolescents are deploying digital sexual communication to negotiate for safer sex.

Digital Sexual Communication and Safer Sex

Digital technologies are (1) integral socialization tools for adolescents, many of whom regularly use texting apps and phone-based social media apps [11], and (2) becoming increasingly recognized as useful sexual health delivery tools because of their low cost and the privacy they provide, especially for adolescents [12,13]. For instance, using baseline data from a longitudinal study of 284 US adolescents, Widman et al [14] found that many adolescents used technology to negotiate for safer sex practices such as condom use, HIV or STI testing, and limiting sexual partners. Moreover, sexually active adolescents are more likely to engage in *sexting* (ie, a digital sexual communication strategy that involves sending or receiving sexually explicit materials through their mobile technologies) [10] to promote condom use [15]. A recent cross-sectional study of forcibly displaced adolescents in Kampala found that adolescents who engaged in sexting reported condom use compared with those who did not sext [16]. Adolescents' digital communication practices require validated measures for evaluation.

Measuring Digital Sexual Communication: Condom Use Negotiated Experiences Through Technology

This study combined items from one subscale (ie, embarrassment about using and negotiating condom use) of Helweg-Larsen and Collins [17] Multidimensional Condom Attitudes Scale (MCAS) and sexting data to yield the condom use negotiated experiences through technology (CuNET) scale, and then tested the psychometric properties and utility of the CuNET scale. The MCAS is the most commonly used scale to measure individual beliefs about and attitudes toward condoms [17]. The MCAS comprises five subscales measuring respondents' opinions regarding (1) the reliability and effectiveness of condoms, (2) sexual pleasure associated with condom use, (3) stigma attached to condom use, (4) embarrassment about the purchase of condoms, and (5) embarrassment about condom use and condom negotiation [17,18]. Studies have found that the fifth subscale, embarrassment surrounding condom use and negotiation, is the most important factor in predicting condom use [17,19]. Given this subscale's predictive power, measures of condom negotiation should incorporate contextual (ie, cultural socialization) and individual (ie, knowledge of and confidence with condoms) attributes that could affect condom use.

Preventive interventions informed by Social Cognitive Theory (SCT) [20] have demonstrated effectiveness in increasing condom use confidence and reducing STI incidence by equipping individuals with negotiation skills and increasing their confidence and ability to use condoms effectively. Yet measures of condom use likelihood have not been updated considering the significant changes in young people's (often technology-based) negotiations of sexual activity. For instance, digital technologies may have shifted young people's perceptions of sexual risk practices and attitudes toward condoms [10,21,22]. Digital technologies may alter adolescents' attitudes and embarrassment about negotiating condom use by providing a digital environment for these discussions. As sexual health scholars and preventionists evaluate how best to leverage these new digital environments for sexual health interventions [11], they will need valid and reliable measures that recognize these emergent digital environments.

Gender Sensitivity

Valid and reliable measures or scales for future digital sexual health interventions must also acknowledge how gender affects attitudes toward sexual activity, condom use, and negotiation. Gendered socialization in many societies emphasizes the control

and dominance of men, with different gender-based expectations (eg, concerning sex) assigned to men and women [23,24]. Even though adolescent girls may be using new digital spaces to challenge traditional cultural sexual scripts that cast them as passive actors and sexual gatekeepers for men, evidence shows that sexting scripts have many similarities to traditional scripts [25]. In 2017, a review of qualitative literature, including 8 studies mostly from the United States (age range 12-25 years), showed that adolescent girls were asked far more often than adolescent boys to send nude pictures [26]. Similarly, a cross-sectional study conducted among 1653 adolescents in Sweden found that 26.2% of girls aged between 15 and 16 years in romantic relationships sent sexts to their partners, compared with 19.6% of boys [27]. Although the study authors did not provide a rationale for these different sexting rates, they are indicative of power imbalances and cultural expectations that compel girls to be responsive to the sexual needs of boys rather than asserting their own sexual needs. Collectively, these findings emphasize that any scale developed to assess sexting practices and inform digital sexual health interventions must be gender sensitive.

This Study

This study contributes to the current sexting, digital sexual communication, and sexual health literature by modifying and validating a gender-sensitive measure that evaluates adolescents' attitudes about using sexting to negotiate for condom use. Drawing from sexting literature [9,10,28], we hypothesized that adolescents who engage in sexting might demonstrate higher rates of condom use, as indicated by their low levels of embarrassment in using sexting to negotiate condom use. Using our adapted CuNET scale (ie, low embarrassment to negotiate and discuss condom use over sexting), we compared support for condom negotiation via sexting across gender among a sample of forcibly displaced adolescents in Kampala, Uganda, and tested the relationship between the CuNET scale and sexual health outcomes.

Methods

Participant Recruitment

From January to March 2018, in collaboration with 3 refugee-serving organizations and 2 government agencies, 242 forcibly displaced adolescents were recruited to participate in the study. Eligible participants ranged from the ages 16 to 19 years and (1) self-identified as a refugee or displaced person or having refugee or displaced parents, (2) reported that they lived in 1 of 5 slums settlements (Kabalagala, Rubaga, Kansanga, Katwe, and Nsambya), and (3) were able to provide informed consent. A total of 12 peer research assistants (PRAs), who self-identified as refugees or displaced persons aged 18 to 24 years, were also recruited and trained in research methods and confidentiality to help with participant recruitment and survey administration.

Ethics Approval

The University of Toronto (#35405) and the Uganda Ministry of Health (ADM: 105/261/01) granted ethical approval for the research. Before completing the surveys, all participants provided written informed consent. Following guidance and waivers from the ethics boards, participants aged between 16 and 17 years were deemed capable of providing informed consent.

Sampling and Data Collection

A peer-networking technique [29], an effective strategy to recruit and include marginalized populations, such as refugees and displaced youth, was used to recruit adolescents. Participants received recruitment vouchers from PRAs to help recruit 2 to 5 other forcibly displaced adolescents in their social network until the target number of participants was reached. Participants chose a private space for the PRA to administer the 35- to 45-minute structured survey on tablets in English or Swahili. Sexual health toolkits, counseling services, and a transport refund of USh 12,500 (approximately US $ 3.74) were provided to all participants who completed the survey.

Measures

The MCAS's embarrassment about negotiation and condom use subscale [17]—a 5-item scale that uses a Likert-type scale ranging from *strongly disagree*=1 to *strongly agree*=7; Cronbach α=.83—has been identified as the strongest predictor of condom use among all MCAS subscales. Therefore, this subscale was selected for adaptation to include sexting. As part of the adaptation process, feedback was solicited from practitioners, PRAs, and experts in Uganda before pilot testing the updated instrument among PRAs (n=12) and forcibly displaced adolescents (n=4) in Nsambya. Participants read the instructions and each item aloud, commented in their own words on what they were being instructed to do and what each item brought to their mind, picked a response option, and explained why they chose the option. After pilot testing the instrument on a number of people, the researchers conducted a debriefing session with participants to look for patterns in the feedback. Specifically, the researchers wanted to know whether participants had similar hesitations, requests for clarification, and whether they had any suggestions for different wording. Feedback during the pilot study suggested that statements should be reverse worded to avoid the response set. For instance, "When I suggest using a condom, I am almost always embarrassed" became "While sexting, I am not embarrassed to suggest using condoms to my partner" (Textbox 1). This recommendation is supported by Boateng et al [30], who argue that negatively worded or reverse-scored items have the potential to negatively impact a scale's psychometric properties.

Participants in the pilot test also recommended that detailed response anchors be provided for clarity. The instrument was revised to include detailed responses with corresponding numbers: *strongly disagree*=1, *disagree*=2, *somewhat disagree*=3, *neither agree nor disagree*=4, *somewhat agree*=5, *agree*=6, and *strongly agree*=7. After the participants approved the final revised measure, it was named CuNET. The CuNET scale response items include, "While sexting, it is really easy to bring up issues of using condoms to my partners," and "While sexting, I know what to say to my partner when I want to talk about condoms or other protections." Textbox 1 shows the final version of the CuNET scale used in this study.

Textbox 1. Condom use negotiated experiences through technology observed indicators adapted for a cross-sectional sample of forcibly displaced adolescents in the slums of Kampala, Uganda.

Original items from Multidimensional Condom Attitudes Scale

- When I suggest using a condom, I am almost always embarrassed
- It is really hard to bring up the issue of using condoms to my partner
- It is easy to suggest to my partner that we use a condom
- I'm comfortable talking about condoms with my partner
- I never know what to say when my partner and I need to talk about condoms or other protection

Items used in this study

- While sexting, I am not embarrassed to suggest using condoms to my partner
- While sexting, it is really easy to bring up issues of using condoms to my partner
- While sexting, it is easy to suggest to my partner that we use a condom
- While sexting, I am comfortable talking about condoms with my partner
- While sexting, I know what to say to my partner when I want to talk about condoms or other protections

To assess validity, associations between CuNET scores and sexual health outcomes (ie, access to sexual and reproductive health (SRH) services, lifetime STI testing, and recent consistent condom use) were examined among participants who were sexually active. Access to SRH was measured using a single self-reported item asking whether participants had ever accessed SRH services in the previous 3 months. Lifetime STI testing was assessed using a single item asking participants whether they had ever received an STI test. Recent consistent condom use was assessed using a single item asking participants whether they consistently used condoms in the previous 3 months. For all 3 items, the responses included yes=1 and no=0.

Statistical Analyses

Descriptive, bivariate analyses, multivariable regression and exploratory factor analysis (EFA) were conducted using Stata 14 (StataCorp), whereas confirmatory factor analysis (CFA) was conducted using Mplus 7.4 [31]. Missing data were less than 5% that any procedure used for handling missing data would have resulted in similar results. First, descriptive analyses of all the variables were conducted. The steps were followed as outlined in Bowen and Guo [32] for testing measurement models by first creating two subsamples using a random process: calibration and validation. Specifically, Bowen and Guo recommend four steps for establishing a measure's validity and reliability: (step 1) using EFA and calibration sampling to test the factor structure of the data, (step 2) conducting CFA and calibration sampling to test for construct validity of the scale, (step 3) using the full sample, chi-square independence test and 2-tailed independent *t* tests to conduct a gender sensitivity analysis of individual items and the entire scale, respectively, and (step 4) using the full sample and multivariable logistic regression to examine the predictive validity of the scale. The following steps allow a comprehensive understanding of the validity and reliability of the scale.

Factor Structure

In step 1, the calibration subsample (N=121) was used to conduct an EFA to classify individual items into CuNET factors. Factor loadings that were ≥0.30 were retained [33]. The number of factors was determined by identifying eigenvalues >1 (Kaiser criterion) and via visual examination of the scree plot [34]. To test whether the sample was sufficient for conducting factor analysis, the Kaiser-Meyer-Olkin test (score above 0.60 recommended) and Bartlett test of sphericity (acceptable if statistically significant) were used [33].

Construct Validity

Following the EFA, CFA was conducted in step 2 using the validation subsample (N=121) to verify the one-factor structure from the EFA results. As the CuNET response set was ordinal (ie, 7 categories), weighted least square mean and variance adjusted estimators were used for the CFA because weighted least square mean and variance adjusted estimators provide a more accurate parameter estimate and a model fit that is more robust to ordinal data [31]. Model fit was assessed using (1) comparative fit index (CFI), with adequate fit represented as CFI≥0.90; (2) standardized root mean square residual (SRMR), acceptable if ≤0.08; and (c) the root mean square error of approximation (RMSEA), acceptable if ≤0.08 [35,36]. A good model fit is indicated when CFI≥0.95, RMSEA≤0.05, and SRMR≤0.06 [36].

Reliability

Calibration and validation data were combined, and items were summed up before assessing the reliability of the CuNET scale. Reliability was evaluated using Cronbach α, with values higher than .70 deemed acceptable [37]. Given that all items in the CuNET scale start with the words "While sexting," Cronbach α may be inflated because of similarities in some of the wordings of the scale. Therefore, reliability was further evaluated using the composite reliability index, with a value of or above 0.70 deemed acceptable [33]. In addition, convergent validity was assessed using the average variance extracted (AVE), with a value above 0.50 deemed acceptable [33].

Gender Sensitivity

In step 3, after confirming the one-factor CuNET scale, the full sample (combining the calibration and validation data) was used to examine gender differences. Here, a higher score indicated positive support for texting-based condom negotiation. Independent *t* tests were then used to compare the mean scores of the CuNET scale by gender.

The next aim was to analyze gender differences among participants who reported positive versus negative support for texting-based condom negotiation. After validating the CuNET scale, a dichotomous variable was created to categorize participants as *positive supporters* or *negative supporters*. Positive supporters (coded as 1) included participants who selected the response anchors *somewhat agree*=5, *agree*=6, and *strongly agree*=7. Negative supporters (coded as 0) included participants who selected the response anchors *strongly disagree*=1, *disagree*=2, *somewhat disagree*=3, and *neither agree or disagree*=4. A chi-square independence test was then conducted.

Predictive Validity

In step 4, using the full sample, we conducted multivariable logistic regression to examine associations between CuNET and sexual health outcomes of recent consistent condom use, access to SRH services, and lifetime STI testing, adjusting for sociodemographic factors (eg, age, gender, education, and mobile phone use) factors. For this analysis, we tested 3 models among a sample of only adolescents who indicated to have engaged in sexual intercourse. We also calculated adjusted odds ratios (aORs), highlighting those significant at the $P<.05$ level. Missing responses were excluded from the analyses; the number of complete responses was reported for each variable.

Results

Sample Characteristics

As illustrated in Table 1, the average age of the participants was 17.56 (SD 1.10) years. More than 3 in 4 (196/242, 81%) were identified as adolescent girls, 77.8% (179/230) had less than secondary/ secondary education, 61.2% (148/242) were originally from the Democratic Republic of the Congo, and 63.2% (153/242) had been in a dating relationship in the last 12 months. About 61.6% (149/242) owned and used mobile phones, whereas over half (122/242, 50.5%) used more than one type of mobile app. Approximately 15.3% (37/242) engaged in digital sexting and 44.2% (107/242) self-reported having had sexual intercourse in their lifetime.

Table 1. Characteristics of forcibly displaced adolescents in the slums of Kampala, Uganda: cross-sectional findings (N=242).

Variables	Values
Sociodemographic factors	
Age (years), mean (SD; range)	17.56 (1.10; 16-19)
Gender, n (%)	
Girls	196 (81)
Boys	46 (19)
Education (n=230), n (%)	
Secondary or less than secondary school	179 (77.8)
Postsecondary education	51 (22.2)
Place of birth (n=242), n (%)	
South Sudan	22 (9.1)
Burundi	50 (20.7)
Democratic Republic of the Congo	148 (61.2)
Rwanda	9 (3.7)
Others	13 (5.4)
Time in Uganda (years; n=242), n (%)	
<1	16 (6.6)
1-5	137 (56.8)
>5	88 (36.5)
Immigration status (n=242), n (%)	
Refugees	216 (90)
Seeking asylum or undocumented	24 (10)
Dating relationship, n (%)	
Not dating	88 (36.5)
Dating	153 (63.2)
Technology use	
Mobile phone ownership and use (n=242), n (%)	
No	93 (38.4)
Yes	149 (61.6)
Average text messages sent per day	2.90 (1.75)
Mobile App use (n=241), n (%)	
No app	69 (28.5)
1 type of app	51 (21.1)
2-3 types of apps	89 (36.8)
4 or more types of apps	33 (13.6)
Digital sexual communication, n (%)	
Sexting patterns (n=242)	
Nonsexters	205 (84.7)
Sexters	37 (15.3)
Sexual health, n (%)	
Ever had sex (n=241)	
No	134 (55.6)
Yes	107 (44.4)

Variables	Values
Recent consistent condom use (n=100)	
No	74 (74)
Yes	26 (26)
Access to SRH[a] services (n=100)	
No	71 (71)
Yes	29 (29)
Lifetime STI[b] testing (n=100)	
No	93 (93)
Yes	7 (7)

[a]SRH: sexual and reproductive health.

[b]STI: sexually transmitted infection.

EFA Results

An EFA with 5 items and oblique rotation revealed a single-factor solution of the CuNET construct using the calibration sample (Table 2). The Kaiser-Meyer-Olkin measure of sampling adequacy was 0.89, exceeding the commonly recommended value of 0.60, and Bartlett test of sphericity was significant: χ^2_{10}=3418.9 (P<.001), eigenvalue=4.48, and communalities above 0.30.

Table 2. Summary of exploratory factor analysis results for CuNET[a] scale using a calibration sample of forcibly displaced adolescents living in the slums of Kampala, Uganda: cross-sectional findings (N=121).

Item	Factor loading	KMO[b]
CuNET[b]1: While sexting, I am not embarrassed to suggest using condoms to my partner	0.91	0.93
CuNET2: While sexting, it is really easy to bring up issues of using condoms to my partner	0.96	0.93
CuNET3: While sexting, it is easy to suggest to my partner that we use a condom	0.97	0.85
CuNET4: While sexting, I am comfortable talking about condoms with my partner	0.96	0.83
CuNET5: While sexting, I know what to say to my partner when I want to talk about condoms or other protections	0.94	0.92
KMO	N/A[c]	0.89

[a]CuNET: condom use negotiated experiences through technology.

[b]KMO: Kaiser-Meyer-Olkin measure of sample adequacy.

[c]N/A: not applicable.

CFA Results

The one-factor CuNET with the validation sample showed good fit: χ^2_4=5.3 (P=.26), RMSEA=0.05 (90% CI 0.00-0.16), CFI=0.99, Tucker–Lewis index=0.99, and SRMR=0.006 (Tables 3 and 4).

Factor loadings ranging from 0.91 to 0.98 exceeded the recommended 0.30 cutoff for the modified instruments. Evidence from the CFA using the validation subsample provides statistical support to the unidimensional model of the CuNET construct in the sample.

Table 3. Covariance matrix of the confirmatory factor analysis for the CuNET[a] scale using a validation sample of forcibly displaced adolescents living in the slums of Kampala, Uganda: cross-sectional findings (n=121).

	CuNET 1	CuNET 2	CuNET 3	CuNET 4	CuNET 5
CuNET 1	4.07	—[b]	—	—	—
CuNET 2	3.86	4.58	—	—	—
CuNET 3	3.6	4.17	4.28	—	—
CuNET 4	3.67	4.05	4.02	4.36	—
CuNET 5	3.69	4.17	3.92	3.91	4.45

[a]CuNET: condom use negotiated experiences through technology.

[b]Not applicable.

Table 4. Confirmatory factor analysis for the CuNET[a] scale using a cross-sectional validation sample of forcibly displaced adolescents living in the slums of Kampala, Uganda (n=121).

CuNET scale items	Factor loadings
CuNet 1: While sexting, I am not embarrassed to suggest using condoms to my partner	0.91
CuNet 2: While sexting, it is really easy to bring up issues of using condoms to my partner	0.98
CuNet 3: While sexting, it is easy to suggest to my partner that we use a condom	0.96
CuNet 4: While sexting, I am comfortable talking about condoms with my partner	0.96
CuNet 5: While sexting, I know what to say to my partner when I want to talk about condoms or other protections	0.93
Fit indices from CFA[b]; one latent factor	
Chi-square	5.29[c]
RMSEA[d]	0.05
CFI[e]	0.99
TLI[f]	0.99
Internal consistency of the final model	
Number of items	5
Cronbach α	0.98
Average variance extracted	0.90
Composite reliability index	0.98

[a]CuNET: condom use negotiated experiences through the technology scale.

[b]CFA: confirmatory factor analysis.

[c]*P*=.26.

[d]RMSEA: root mean square error approximation.

[e]CFI: comparative fit index.

[f]TLI: Tucker–Lewis index.

Internal Reliability Results

Using the full sample, the values for Cronbach α (.98) and composite reliability index (0.98) exceeded the 0.70 thresholds, indicating that the CuNET scale had high internal consistency and reliability (Table 4). The results also provided evidence of convergent validity (AVE=0.90), as the CuNET scale exceeded the AVE threshold of 0.50.

Support for Texting-Based Condom Negotiation

Table 5 presents levels of support for texting-based condom negotiation. Overall, approximately 1 in 4 respondents reported support for using sexting to negotiate for condom use.

Using independent *t* tests, the gender-based analysis showed that the combined CuNET scale was gender-sensitive: adolescent girls showed significantly lower levels of support for using sexting to negotiate condom use compared with adolescent boys (mean 13.60; SE 0.70 vs 21.48; SE 1.23; *P*<.001). Using CuNET scale item analysis, the chi-square test of independence revealed that adolescent boys showed significantly higher levels of support for using sexting to negotiate condom use in all 5 items. For instance, 61% (28/46) of adolescent boys agreed that “While sexting, I know what to say to my partner when I want to talk about condoms or other protections,” compared with 21% (41/196) of adolescent girls. Adolescent girls reported the lowest

percentage (34/196, 17.4%) of support for the item asking if sexting could reduce their discomfort with suggesting condom use to their partners, compared with adolescent boys (21/46, 46%).

Table 5. Levels of CuNET[a] among a cross-sectional sample of forcibly displaced adolescents living in the slums of Kampala, stratified by gender (5 items; N=242)[b].

Items	Statements	Girls (n=196), n (% agreed)	Boys (n=46), n (% agreed)	Total, n (% agreed)	*P* value
CuNet 1	While sexting, I am not embarrassed to suggest using condoms to my partner	34 (17.4)	21 (45.7)	55 (22.8)	<.001
CuNet 2	While sexting, it is really easy to bring up issues of using condoms to my partner	38 (19.5)	24 (52.2)	62 (25.7)	<.001
CuNet 3	While sexting, it is easy to suggest to my partner that we use a condom	39 (20)	25 (54.3)	64 (26.6)	<.001
CuNet 4	While sexting, I am comfortable talking about condoms with my partner	43 (22.1)	26 (56.5)	69 (28.6)	<.001
CuNet 5	While sexting, I know what to say to my partner when I want to talk about condoms or other protections	41 (21)	28 (60.9)	69 (28.6)	<.001

[a]CuNET: condom use negotiated experiences through the technology scale.

[b]A chi-square independence test was conducted to examine how CuNET items differed by gender; agreed percentages were calculated by creating a categorical variable using a cutoff of 5 and above, which indicated positive support. The total is the percentage of participants who provided positive attitudes toward using sexing for condom negotiation.

Associations Between Support for Texting-Based Condom Negotiation and Sexual Health Outcomes

In adjusted analyses, a 1-point increase in the CuNET score was associated with higher odds of recent consistent condom use versus no recent consistent condom use (aOR 1.73, 95% CI 1.24-2.41). Holding other factors constant, we found no association between CuNET and access to SRH services versus no access to SRH services (aOR 1.51, 95% CI 0.99-2.30) or recent lifetime STI testing versus never lifetime STI testing (aOR 0.90, 95% CI 0.64-1.26; Table 6).

Table 6. Independent association between CuNET[a] and sexual health factors among a cross-sectional sample of sexually active forcibly displaced adolescents living in the slums of Kampala (N=100)[b].

	Values	Recent consistent condom use		Values	Access to SRH[c] services		Values	STI[d] testing (Ever)	
		aOR[e] (95% CI)	*P* value		aOR (95% CI)	*P* value		aOR (95% CI)	*P* value
CuNet, mean (SD)	24.31 (8.46)	1.73 (1.24-2.41)	*.001* [f]	21.13 (8.60)	1.51 (0.99-2.30)	.05	14.31 (11.14)	0.90 (0.64-1.26)	.54
Age (years), mean (SD)	17.96 (0.96)	0.92 (0.33-4.05)	.84	17.60 (1.07)	0.26 (0.09-0.78)	*.02*	18.34 (0.97)	2.49 (1.19-5.22)	*.02*
Gender (ref boys), n (%)									
Girls	21 (80.8)	0.83 (0.19-3.67)	.81	19 (73.1)	0.79 (0.08-8.14)	.85	48 (96)	0.09 (0.01-0.89)	*.04*
Boys	5 (19.2)	N/A[g]	N/A	7 (26.9)	N/A	N/A	2 (4)	N/A	N/A
Mobile phone ownership (ref no mobile phone), n (%)									
Yes	24 (92.3)	0.23 (0.04-1.51)	.13	12 (80)	0.37 (0.04-3.51)	.39	27 (84.4)	0.51 (0.51-12.36)	.26
No	2 (8.7)	N/A	N/A	3 (20)	N/A	N/A	5 (15.6)	N/A	N/A
Dating relationship (ref no dating relation), n (%)									
Dating	26 (100)	1.15 (0.33-2.41)	.83	N/A	N/A	N/A	30 (93.8)	0.28 (0.07-1.16)	.08
Not dating	0 (0)	N/A	N/A	N/A	N/A	N/A	2 (6.3)	N/A	N/A
Education (ref postsecondary education), n (%)									
Secondary or less than secondary school	12 (48)	0.74 (0.16-3.42)	.70	N/A	N/A	N/A	23 (71.9)	0.64 (0.12-3.43)	.61
Postsecondary education	13 (52)	N/A	N/A	N/A	N/A	N/A	9 (28.1)	N/A	N/A

[a]CuNET: condom use negotiated experiences through technology.

[b]CuNET scale summarized scores were calculated for the 5 items; higher scores indicated higher support for using sexing to negotiate condom use. Due to distribution in the access to sexual and reproductive health services variable, only adjusted for 3 variables.

[c]SRH: sexual and reproductive health.

[d]STI: sexually transmitted infection.

[e]aOR: adjusted odds ratio.

[f]Statistically significant.

[g]N/A: not applicable.

Discussion

Principal Findings

Accurate measurement of forcibly displaced adolescents' support for condom negotiation via digital tools may be important for the development and evaluation of digital sexual health interventions that aim to promote condom use. The findings from this study show that the one-factor CuNET scale is (1) valid and reliable for forcibly displaced adolescents living in informal settlements in Kampala, Uganda, (2) gender sensitivity, and (3) a portable tool for identifying factors associated with recent condom use among urban forcibly displaced adolescents.

Comparison With Prior Work

The findings of this study depart from sexting literature that focuses solely on the harmful effects of technology on sexuality [15,21,22] by highlighting the positive applications that technology can have among forcibly displaced adolescents. Among adolescents who may be socialized to view conversations surrounding sex as taboo [23], sexting may provide them with convenient ways to communicate with each other privately. Indeed, a recent meta-analysis of 23 studies concluded that sexting was part of adolescents' development and sexual exploration [10]. This study extends the measurement of condom negotiation strategies from face-to-face interactions to include interactions in digital environments and updates the MCAS—embarrassment to negotiate the condom use subscale [17] by recognizing sexting as a condom negotiation strategy. That is, sexting can be an important method by which adolescents influence their partners to use condoms, and thus a potentially high-impact target of sexual health interventions. As this study only focused on the utility of participants' embarrassment to negotiate condom use via sexting, there is a need for future studies that develop a more comprehensive condom negotiation scale that encompasses other digital sexual communication strategies.

A closer look at the CuNET scale items indicates that confidence, knowledge, and persuasion are the key traits measured by the scale. These traits align with an important tenet of SCT—that for an individual to negotiate for condom use, they must have enough knowledge about the subject and confidence in their ability to engage their partner in conversations about sex-related topics. Importantly, this study shows that the venue of communication (ie, digital or in-person) may influence confidence levels surrounding these conversations. Digital technologies may afford adolescents with a continuous, digital, and real time venues to negotiate condom use with less embarrassment. Embarrassment may be addressed through the development and evaluation of interventions that encourage the exploration of sexual health subjects through technological interfaces.

Another primary goal of this study was to examine the gender sensitivity of the CuNET scale. The findings highlight gender differences in the endorsement of the CuNET scale items. Adolescent boys viewed the use of sexting for condom negotiation more favorably than adolescent girls. This finding is consistent with the sexual script theory [23], which highlights the double standards in sexual expectations for girls and boys. Adolescent girls may be socialized to avoid initiating any discussions of sex-related topics, regardless of venue. It is plausible that girls may support negotiating condom use via digital tools; however, they may not be willing to disclose this to peer interviewers because of the negative judgments attached to them doing so [38].

Findings from the psychometrics testing of the CuNET scale highlight how the changing landscape of digital sexual communications may provide an important opportunity to revisit condom negotiation strategies and interventions. Consistent with Widman et al [14], we found that favorable support for texting-based condom negotiation was associated with condom use. Thus, although texting-based condom negotiation may provide creative approaches that sexually active adolescents can use to negotiate condom use with their partners, interventions that increase condom education and access are critical. Therefore, interventionists need to consider gender differences and sociocultural norms that influence condom use and condom negotiation among adolescents.

Limitations

The findings of this study should be interpreted within the context of its limitations. This study relies on data from a nonprobable sample of forcibly displaced adolescents living in slums of Kampala, Uganda. This sample may not necessarily be representative of forcibly displaced communities in Kampala, Uganda, and does not capture the diversity of the adolescent refugee population globally. Our study had more adolescent girls than adolescent boys, which limits our ability to compare the 2 genders. Originally, the study was designed to include only adolescent girls because of the high HIV rates, which was 3-fold than that of adolescent boys in Uganda. However, our local collaborators asked us to also include adolescent boys. As a result, we oversampled the proportion of girls to reflect the local HIV epidemic characteristics. The use of a cross-sectional design means that the causality of the predictive validity of the CuNET scale cannot be inferred. Future longitudinal studies are needed to explore whether CuNET scale scores predict later condom use. The inability to test group invariance because of the small sample size calls for future studies that use rigorous methods to validate the CuNET scale. Furthermore, although the scale was pilot-tested and modified according to participant feedback on its face and content validity, there is a need for future studies that draw together adolescent participants and adolescent health experts to develop a culturally responsive CuNET scale that can capture multiple possible ways in which adolescents engage in condom negotiation. In addition, the high reliability and internal consistency of the CuNET scale indicate that some items may have a close meaning. Future studies should further refine the CuNET scale to reduce redundancy in scale items.

Conclusions

Despite these limitations, the strengths of this study include the validity and reliability of a new gender-sensitive scale (CuNET) useful for measuring support for texting-based condom negotiation among marginalized adolescents. The CuNET scale can be used to collect more targeted data for prevention studies addressing adolescents' support for negotiating condom use via digital tools. The fact that CuNET is a digital sexual communication instrument responds to recent calls to leverage digital tools for the sexual health of adolescents [11,12]. On the basis of this study's findings, more research is needed to develop further and validate a gender-responsive CuNET scale for adolescents. The current iteration of the CuNET scale appears to be a valid measure for assessing support for texting-based condom negotiation among forcibly displaced adolescents living in HIV hyperendemic settings in Kampala, Uganda, with strong potential for future adaptation to a host of settings and populations.

Acknowledgments

This study was supported by the Canadian Institutes of Health Research. CHL's role was also supported by funding from the Canada Foundation for Innovation, Canada Research Chairs, and the Ontario Ministry of Research and Innovation. The authors acknowledge all the peer navigators and participants, as well as collaborating agencies: Ugandan Ministry of Health, Office of the Prime Minister, Young African Refugees for Integral Development, Tomorrow Vijana, Most At Risk Populations Initiative, and InterAid Uganda.

Conflicts of Interest

None declared.

References

1. Jennings L, George AS, Jacobs T, Blanchet K, Singh NS. A forgotten group during humanitarian crises: a systematic review of sexual and reproductive health interventions for young people including adolescents in humanitarian settings. Confl Health 2019;13:57 [FREE Full text] [doi: 10.1186/s13031-019-0240-y] [Medline: 31788022]
2. Uganda country refugee response plan. United Nations High Commission for Refugees. 2019. URL: https://data2.unhcr.org/en/documents/download/67314 [accessed 2022-02-12]
3. Swahn MH, Culbreth R, Salazar LF, Kasirye R, Seeley J. Prevalence of HIV and associated risks of sex work among youth in the slums of Kampala. AIDS Res Treat 2016;2016:5360180-5360188 [FREE Full text] [doi: 10.1155/2016/5360180] [Medline: 27239340]
4. Uganda population-based HIV impact assessment : UPHIA 2016-2017 : summary sheet : preliminary findings. Uganda Ministry of Health. 2017. URL: https://stacks.cdc.gov/view/cdc/50004 [accessed 2022-02-12]
5. Culbreth R, Swahn MH, Salazar LF, Ametewee LA, Kasirye R. Risk factors associated with HIV, sexually transmitted infections (STI), and HIV/STI co-infection among youth living in the slums of Kampala, Uganda. AIDS Behav 2020 Apr;24(4):1023-1031. [doi: 10.1007/s10461-019-02444-5] [Medline: 30825036]
6. Uganda Violence Against Children Survey. Ministry of Gender, Labour and Social Development. 2018. URL: https://www.unicef.org/uganda/reports/uganda-violence-against-children-survey [accessed 2022-02-12]
7. Swahn MH, Buchongo P, Kasirye R. Risky behaviors of youth living in the slums of Kampala: a closer examination of youth participating in vocational training programs. Vulnerable Child Youth Stud 2018;13(3):276-290 [FREE Full text] [doi: 10.1080/17450128.2018.1489168] [Medline: 31452668]
8. Classification of digital health interventions v1.0: a shared language to describe the uses of digital technology for health. World Health Organization. 2018. URL: https://apps.who.int/iris/handle/10665/260480 [accessed 2022-02-12]
9. Drouin M. Sexual communication in the digital age. In: The Wiley Handbook of Psychology, Technology, and Society. Chichester, UK: John Wiley & Sons, Ltd; 2015.
10. Mori C, Temple JR, Browne D, Madigan S. Association of sexting with sexual behaviors and mental health among adolescents: a systematic review and meta-analysis. JAMA Pediatr 2019 Aug 01;173(8):770-779 [FREE Full text] [doi: 10.1001/jamapediatrics.2019.1658] [Medline: 31206151]
11. Wadham E, Green C, Debattista J, Somerset S, Sav A. New digital media interventions for sexual health promotion among young people: a systematic review. Sex Health 2019 Apr;16(2):101-123. [doi: 10.1071/SH18127] [Medline: 30819326]
12. Newby K, Crutzen R, Brown K, Bailey J, Saunders J, Szczepura A, et al. An intervention to increase condom use among users of chlamydia self-sampling websites (wrapped): intervention mapping and think-aloud study. JMIR Form Res 2019 May 01;3(2):e11242 [FREE Full text] [doi: 10.2196/11242] [Medline: 31042156]
13. Guse K, Levine D, Martins S, Lira A, Gaarde J, Westmorland W, et al. Interventions using new digital media to improve adolescent sexual health: a systematic review. J Adolesc Health 2012 Dec;51(6):535-543. [doi: 10.1016/j.jadohealth.2012.03.014] [Medline: 23174462]
14. Widman L, Nesi J, Choukas-Bradley S, Prinstein MJ. Safe sext: adolescents' use of technology to communicate about sexual health with dating partners. J Adolesc Health 2014 May;54(5):612-614 [FREE Full text] [doi: 10.1016/j.jadohealth.2013.12.009] [Medline: 24512716]
15. Marume A, Maradzika J, January J. Adolescent sexting and risky sexual behaviours in Zimbabwe: a cross-sectional study. Sexuality Culture 2018 Feb 27;22(3):931-941. [doi: 10.1007/s12119-018-9508-4]
16. Okumu M. The Role of Digital Sexual Communication in the Sexual Health of Forcibly Displaced Adolescents Living in Informal Urban Settlements of Kampala, Uganda. Toronto, Canada: University of Toronto; 2020.
17. Helweg-Larsen M, Collins BE. The UCLA Multidimensional Condom Attitudes Scale: documenting the complex determinants of condom use in college students. Health Psychol 1994;13(3):224-237. [doi: 10.1037/0278-6133.13.3.224]
18. French SE, Holland KJ. Condom negotiation strategies as a mediator of the relationship between self-efficacy and condom use. J Sex Res 2013;50(1):48-59. [doi: 10.1080/00224499.2011.626907] [Medline: 22126308]
19. Conley TD, Collins BE. Differences between condom users and condom nonusers in their multidimensional condom attitudes. J Appl Social Pyschol 2005 Mar;35(3):603-620. [doi: 10.1111/j.1559-1816.2005.tb02137.x]
20. Bandura A. Social cognitive theory in cultural context. Applied Psychol 2002 Apr;51(2):269-290. [doi: 10.1111/1464-0597.00092]
21. Cooper K, Quayle E, Jonsson L, Svedin CG. Adolescents and self-taken sexual images: a review of the literature. Comput Human Behav 2016 Feb;55:706-716. [doi: 10.1016/j.chb.2015.10.003]
22. Kosenko K, Luurs G, Binder AR. Sexting and sexual behavior, 2011-2015: a critical review and meta-analysis of a growing literature. J Comput Mediat Commun 2017 May 15;22(3):141-160. [doi: 10.1111/jcc4.12187]
23. Simon W, Gagnon JH. Sexual scripts. Society 1984 Nov;22(1):53-60. [doi: 10.1007/bf02701260]
24. Connell RW, Messerschmidt JW. Hegemonic masculinity. Gender Soc 2016 Jun 29;19(6):829-859. [doi: 10.1177/0891243205278639]
25. Symons K, Ponnet K, Walrave M, Heirman W. Sexting scripts in adolescent relationships: is sexting becoming the norm? New Media Soc 2018 Mar 26;20(10):3836-3857. [doi: 10.1177/1461444818761869]

26. Anastassiou A. Sexting and young people: a review of the qualitative literature. Qual Report 2017 Aug;22(8):8-16 [FREE Full text] [doi: 10.46743/2160-3715/2017.2951]
27. Burén J, Lunde C. Sexting among adolescents: a nuanced and gendered online challenge for young people. Comput Human Behav 2018 Aug;85:210-217. [doi: 10.1016/j.chb.2018.02.003]
28. Barrense-Dias Y, Berchtold A, Surís JC, Akre C. Sexting and the definition issue. J Adolesc Health 2017 Nov;61(5):544-554. [doi: 10.1016/j.jadohealth.2017.05.009] [Medline: 28734631]
29. Magnani R, Sabin K, Saidel T, Heckathorn D. Review of sampling hard-to-reach and hidden populations for HIV surveillance. AIDS 2005 May;19 Suppl 2:S67-S72. [doi: 10.1097/01.aids.0000172879.20628.e1] [Medline: 15930843]
30. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best practices for developing and validating scales for health, social, and behavioral research: a primer. Front Public Health 2018;6:149 [FREE Full text] [doi: 10.3389/fpubh.2018.00149] [Medline: 29942800]
31. Muthén L, Muthén B. Mplus User's Guide. 8th edition. Los Angeles, CA: Muthén & Muthén; 2017.
32. Bowen N, Guo S. Structural Equation Modeling. Oxfordshire, England, UK: Oxford University Press; 2011.
33. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. J Market Res 2018 Nov 28;18(1):39-50. [doi: 10.1177/002224378101800104]
34. Mengo C, Okumu M, Ombayo B, Nahar S, Small E. Marital rape and HIV risk in Uganda: the impact of women's empowerment factors. Violence Against Women 2019 Dec;25(15):1783-1805. [doi: 10.1177/1077801218821444] [Medline: 30672397]
35. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Structural Equation Model Multidisciplinary J 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
36. Kline R. Principles and Practice of Structural Equation Modeling, 4th edition. New York: The Guilford press; 2015.
37. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika 1951 Sep;16(3):297-334. [doi: 10.1007/BF02310555]
38. Ringrose J, Harvey L, Gill R, Livingstone S. Teen girls, sexual double standards and 'sexting': gendered value in digital image exchange. Feminist Theory 2013 Nov 14;14(3):305-323. [doi: 10.1177/1464700113499853]

Abbreviations

aOR: adjusted odds ratio
AVE: average variance extracted
CFA: confirmatory factor analysis
CFI: comparative fit index
CuNET: condom use negotiated experiences through technology
EFA: exploratory factor analysis
MCAS: Multidimensional Condom Attitudes Scale
PRA: peer research assistant
RMSEA: root mean square error of approximation
SCT: Social Cognitive Theory
SRH: sexual and reproductive health
SRMR: standardized root mean square residual
STI: sexually transmitted infection

Edited by T Sanchez; submitted 06.02.21; peer-reviewed by W Byansi, B Tonui; comments to author 12.05.21; revised version received 16.10.21; accepted 27.12.21; published 06.04.22.

Please cite as:
Okumu M, Logie CH, Ansong D, Mwima S, Hakiza R, Newman PA
Support for Texting-Based Condom Negotiation Among Forcibly Displaced Adolescents in the Slums of Kampala, Uganda: Cross-sectional Validation of the Condom Use Negotiated Experiences Through Technology Scale
JMIR Public Health Surveill 2022;8(4):e27792
URL: https://publichealth.jmir.org/2022/4/e27792
doi:10.2196/27792
PMID:35384852

Original Paper

Toward Using Twitter for PrEP-Related Interventions: An Automated Natural Language Processing Pipeline for Identifying Gay or Bisexual Men in the United States

Ari Z Klein[1], PhD; Steven Meanley[2], PhD, MPH; Karen O'Connor[1], MS; José A Bauermeister[2], PhD, MPH; Graciela Gonzalez-Hernandez[1], PhD

[1]Department of Biostatistics, Epidemiology, and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

[2]Department of Family and Community Health, School of Nursing, University of Pennsylvania, Philadelphia, PA, United States

Corresponding Author:
Ari Z Klein, PhD
Department of Biostatistics, Epidemiology, and Informatics
Perelman School of Medicine
University of Pennsylvania
Blockley Hall, 4th Floor
423 Guardian Drive
Philadelphia, PA, 19104
United States
Phone: 1 215 746 1101
Email: ariklein@pennmedicine.upenn.edu

Abstract

Background: Pre-exposure prophylaxis (PrEP) is highly effective at preventing the acquisition of HIV. There is a substantial gap, however, between the number of people in the United States who have indications for PrEP and the number of them who are prescribed PrEP. Although Twitter content has been analyzed as a source of PrEP-related data (eg, barriers), methods have not been developed to enable the use of Twitter as a platform for implementing PrEP-related interventions.

Objective: Men who have sex with men (MSM) are the population most affected by HIV in the United States. Therefore, the objectives of this study were to (1) develop an automated natural language processing (NLP) pipeline for identifying men in the United States who have reported on Twitter that they are gay, bisexual, or MSM and (2) assess the extent to which they demographically represent MSM in the United States with new HIV diagnoses.

Methods: Between September 2020 and January 2021, we used the Twitter Streaming Application Programming Interface (API) to collect more than 3 million tweets containing keywords that men may include in posts reporting that they are gay, bisexual, or MSM. We deployed handwritten, high-precision regular expressions—designed to filter out noise and identify actual self-reports—on the tweets and their user profile metadata. We identified 10,043 unique users geolocated in the United States and drew upon a validated NLP tool to automatically identify their ages.

Results: By manually distinguishing true- and false-positive self-reports in the tweets or profiles of 1000 (10%) of the 10,043 users identified by our automated pipeline, we established that our pipeline has a precision of 0.85. Among the 8756 users for which a US state–level geolocation was detected, 5096 (58.2%) were in the 10 states with the highest numbers of new HIV diagnoses. Among the 6240 users for which a county-level geolocation was detected, 4252 (68.1%) were in counties or states considered priority jurisdictions by the *Ending the HIV Epidemic* initiative. Furthermore, the age distribution of the users reflected that of MSM in the United States with new HIV diagnoses.

Conclusions: Our automated NLP pipeline can be used to identify MSM in the United States who may be at risk of acquiring HIV, laying the groundwork for using Twitter on a large scale to directly target PrEP-related interventions at this population.

(JMIR Public Health Surveill 2022;8(4):e32405) doi:10.2196/32405

KEYWORDS

natural language processing; social media; data mining; PrEP; pre-exposure prophylaxis; HIV; AIDS

Introduction

Pre-exposure prophylaxis (PrEP) with antiretroviral drugs is highly effective at preventing the acquisition of HIV in men who have sex with men (MSM) [1]. There is a substantial gap, however, between the number of people in the United States who have indications for PrEP, including 25% of MSM [2], and the number of them who are prescribed PrEP [3]; approximately one-third of primary care physicians (PCPs) in the United States who are aware of PrEP have prescribed PrEP or referred a patient for PrEP [4]. Although efforts should be made to increase PCPs' adoption of PrEP recommendations into routine clinical practice, PCP-based interventions are limited because some MSM, especially younger men, face challenges when disclosing their same-sex sexual behaviors to their PCPs [5]. Based on the findings of a recent study by Reuter et al [6] that examined Twitter users' attitudes toward being monitored for health-related research, some MSM may be more open to PrEP-related interventions on social media, such as targeted messages or advertisements.

Hannaford et al [7] found that social media can help identify factors for implementing PrEP-related interventions that are not captured by traditional research methods, and they suggested that social media may present novel opportunities to implement PrEP-related interventions. Although Twitter content has been analyzed as a source of PrEP-related data (eg, barriers) [8,9], to our knowledge, methods have not been developed to enable the use of Twitter as a platform for PrEP-related interventions. The foremost requirement for implementing PrEP-related interventions on Twitter is to identify users in the populations that have indications for PrEP. Given that MSM are the population most affected by HIV in the United States [10], the objectives of this study were to (1) develop an automated natural language processing (NLP) pipeline for identifying men in the United States who have reported on Twitter that they are gay, bisexual, or MSM and (2) assess the extent to which they demographically represent MSM in the United States with new HIV diagnoses. This study seeks to lay the groundwork for using Twitter on a large scale to directly target PrEP-related interventions at MSM who may be at risk of acquiring HIV.

Methods

Ethical Considerations

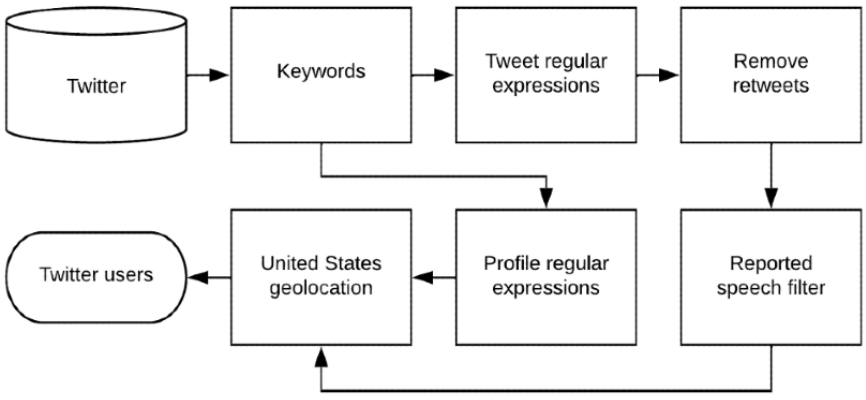
The Institutional Review Board of the University of Pennsylvania reviewed this study and deemed it exempt human subjects research under Category (4) of Paragraph (b) of the US Code of Federal Regulations Title 45 Section 46.101 for publicly available data sources (45 CFR §46.101(b)(4)).

Data Collection

Between September 2020 and January 2021, we used the Twitter Streaming Application Programming Interface (API) to collect more than 3 million tweets containing keywords that men may include in posts reporting that they are gay, bisexual, or MSM. As a preliminary approach, we deployed handwritten, high-precision regular expressions—search patterns designed to automatically match text strings—on the 3 million tweets to filter out noise and identify actual self-reports (Multimedia Appendix 1). After automatically removing retweets and "reported speech" (eg, quotations, news headlines) [11], the regular expressions matched 8603 tweets that were posted by 6358 users geolocated in the United States [12].

In addition to tweet-based regular expressions, we also deployed handwritten regular expressions on the user profile metadata of the 3 million tweets collected from the Twitter Streaming API (Multimedia Appendix 1). The regular expressions matched the profile metadata of 4127 users geolocated in the United States [12]. After removing duplicate users from our tweet- and profile-based searches, we identified a total of 10,043 unique users. Figure 1 illustrates our automated pipeline for identifying men in the United States who have reported on Twitter that they are gay, bisexual, or MSM. To assess the extent to which they demographically represent MSM in the United States with new HIV diagnoses, we analyzed the state- and county-level geolocations [12] of these 10,043 users and drew upon a validated NLP tool [13] to automatically identify their ages.

Figure 1. Automated natural language processing pipeline for identifying men in the United States who have reported on Twitter that they are gay, bisexual, or men who have sex with men.



Results

Pipeline Evaluation

True positives and false positives were manually distinguished by 2 annotators in a random sample of 1000 (10%) of the 10,043 users that were identified by our automated pipeline, consisting of 500 matching tweets and 500 matching profiles. *True positives* were defined as tweets or profiles in which the users reported that they are gay, bisexual, or MSM. Overall interannotator agreement (Cohen κ) based on independent, dual annotations for all 1000 users was 0.81, which is deemed to be "almost perfect agreement" [14]. More specifically, interannotator agreement was 0.83 for the 500 tweets and 0.79 for the 500 profiles. Upon resolving the disagreements, 417 (83.4%) tweets and 430 (86%) profiles were annotated as true positives and 83 (16.6%) tweets and 70 (14%) profiles were annotated as false positives. Based on this evaluation, our automated pipeline has an overall precision of 0.85, where *precision = true positives / (true positives + false positives)*. Table 1 provides examples of tweets and profiles that were manually annotated as true or false positives. The majority of the profiles that were annotated as false positives were users that mentioned being transgender or nonbinary—populations that are beyond the scope of this study.

Table 1. Sample manual annotations of tweets and profiles.

Type	Text	Label
Tweet	End the FDA's discriminatory and unscientific policy against gay men like me donating blood.	True positive
Tweet	As a bi guy we get so little representation, and almost all of its negative. It's frustrating.	True positive
Tweet	Today, we remember Matthew Shepard who's life was cut short as a result of a hate crime due to his identity as a gay male.	False positive
Profile	A proud black gay guy.	True positive
Profile	50+ gay trans man, writer, film and food lover. He/him OR they/them.	False positive

Demographics

To assess the utility of our automated pipeline for identifying MSM in the United States who may be particularly at risk of acquiring HIV, we analyzed their state- and county-level geolocations and ages. We detected a US state–level geolocation for 8756 (87.6%) of the 10,043 users identified by our automated pipeline, including users from all 50 states and the District of Columbia. As Figure 2 illustrates, the largest numbers of users were detected in California, New York, Texas, Florida, Illinois, Pennsylvania, Ohio, and Georgia. We detected a county-level geolocation for 6240 (71.2%) of these 8756 users. Table 2 presents the 15 counties for which we detected at least 100 users. We detected an age of ≥13 years [10] for 4782 (47.6%) of the 10,043 users, with a mean age of 31.9 (SD 13.1) years and a median age of 29 years. Table 3 presents the age distribution, based on each user's most recent tweet containing a self-report of age.

Figure 2. Number of Twitter users, by state, identified by our automated pipeline between September 2020 and January 2021.

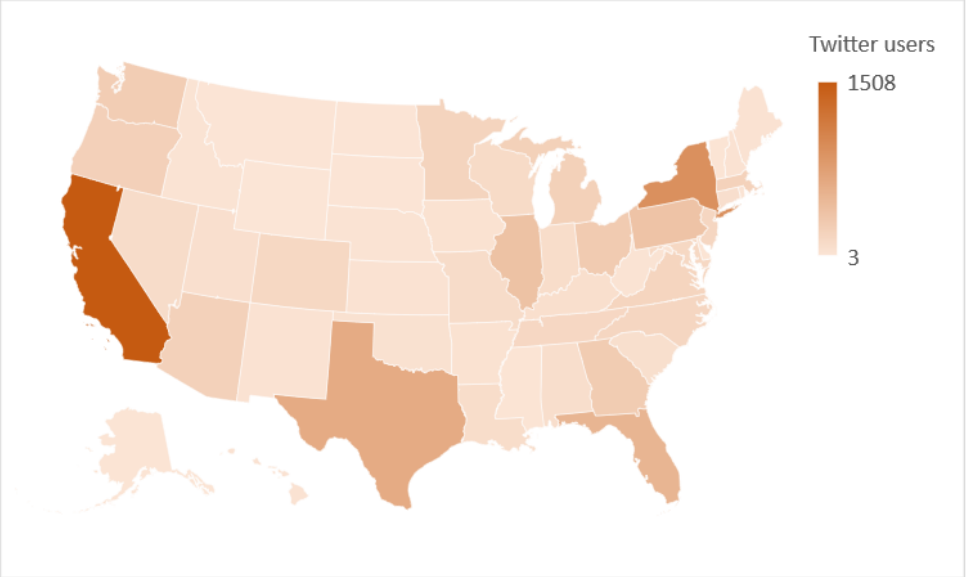


Table 2. Counties with at least 100 Twitter users identified by our automated pipeline between September 2020 and January 2021.

US county	Users (N=6240), n (%)
Los Angeles County, CA	535 (8.6)
New York County, NY	417 (6.7)
Cook County, IL	318 (5.1)
District of Columbia, DC	237 (3.8)
King County, WA	192 (3.1)
Fulton County, GA	155 (2.5)
San Mateo County, CA	151 (2.4)
Multnomah County, OR	128 (2.1)
Kings County, NY	127 (2)
Dallas County, TX	123 (2)
Philadelphia County, PA	121 (1.9)
Harris County, TX	116 (1.9)
Maricopa County, AZ	111 (1.8)
Suffolk County, MA	110 (1.8)
Travis County, TX	109 (1.7)

Table 3. Age distribution of Twitter users identified by our automated pipeline between September 2020 and January 2021.

Age group (years)	Users (N=4782), n (%)
13-24	1630 (34.1)
25-34	1644 (34.4)
35-44	704 (14.7)
45-54	449 (9.4)
≥55	355 (7.4)

Discussion

Principal Findings

Our study demonstrates that gay men, bisexual men, or MSM in the United States publicly report their sexual orientation on Twitter and that these users can be accurately identified on a large scale. Moreover, among the 8756 users for which our automated pipeline detected a US state–level geolocation, 5096 (58.2%) were in the 10 states with the highest numbers of new HIV diagnoses [10]. Among the 6240 users for which a county-level geolocation was detected, 4252 (68.1%) were in counties or states considered priority jurisdictions by the *Ending the HIV Epidemic* initiative [15]. Furthermore, the age distribution of the users reflected the ranking of the most frequent age groups with new HIV diagnoses among MSM in the United States [10], with the 25-34 years age group first and the 13-24 years age group second. More specifically, these 2 age groups represent both the majority of the users in this study and the majority of MSM with new HIV diagnoses [10]. The mean (31.9 years) and median (29 years) ages of the users are within the age group (25-34 years) with the largest number of new HIV diagnoses, which is also the only age group in which HIV infections have increased since 2014 [10]. Therefore, our automated pipeline can be used as the basis for PrEP-related interventions targeted directly at MSM who are largely in the regions and age groups most affected by HIV in the United States, including younger men who may face challenges when discussing their same-sex sexual behaviors with their PCPs [5].

Conclusions

This paper presented an automated NLP pipeline that can be used to identify MSM in the United States who may be at risk of acquiring HIV, laying the groundwork for using Twitter on a large scale to directly target PrEP-related interventions at this population.

Acknowledgments

The authors would like to thank Ivan Flores for contributing to software development and Alexis Upshur for contributing to annotating the Twitter data for validation of the pipeline. This research was supported by a grant from the Penn Center for AIDS Research, a National Institutes of Health–funded program (P30 AI 045008).

Authors' Contributions

AZK contributed to designing the pipeline, developing the sets of regular expressions, preparing the data set for validation, resolving the annotators' disagreements, analyzing the demographics, and writing the manuscript. SM contributed to guiding data collection from Twitter and data validation and editing the manuscript. KO contributed to annotating the Twitter data for validation, calculating interannotator agreement, and editing the manuscript. JB contributed to guiding the overall study design and data collection from Twitter and editing the manuscript. GGH contributed to conceptualizing the research study, guiding the overall study design and data collection from Twitter, and editing the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Regular expressions.
[DOCX File , 12 KB - publichealth_v8i4e32405_app1.docx]

References

1. Grant RM, Lama JR, Anderson PL, McMahan V, Liu AY, Vargas L, iPrEx Study Team. Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. N Engl J Med 2010 Dec 30;363(27):2587-2599 [FREE Full text] [doi: 10.1056/NEJMoa1011205] [Medline: 21091279]
2. Smith DK, Van Handel M, Wolitski RJ, Stryker JE, Hall HI, Prejean J, et al. Vital signs: estimated percentages and numbers of adults with indications for preexposure prophylaxis to prevent HIV acquisition--United States, 2015. MMWR Morb Mortal Wkly Rep 2015 Nov 27;64(46):1291-1295 [FREE Full text] [doi: 10.15585/mmwr.mm6446a4] [Medline: 26606148]
3. Huang YA, Zhu W, Smith DK, Harris N, Hoover KW. HIV preexposure prophylaxis, by race and ethnicity - United States, 2014-2016. MMWR Morb Mortal Wkly Rep 2018 Oct 19;67(41):1147-1150 [FREE Full text] [doi: 10.15585/mmwr.mm6741a3] [Medline: 30335734]
4. Blackstock OJ, Moore BA, Berkenblit GV, Calabrese SK, Cunningham CO, Fiellin DA, et al. A cross-sectional online survey of HIV pre-exposure prophylaxis adoption among primary care physicians. J Gen Intern Med 2017 Jan;32(1):62-70 [FREE Full text] [doi: 10.1007/s11606-016-3903-z] [Medline: 27778215]
5. Petroll AE, Mitchell JW. Health insurance and disclosure of same-sex sexual behaviors among gay and bisexual men in same-sex relationships. LGBT Health 2015 Mar;2(1):48-54 [FREE Full text] [doi: 10.1089/lgbt.2013.0050] [Medline: 26790018]
6. Reuter K, Zhu Y, Angyan P, Le N, Merchant AA, Zimmer M. Public concern About monitoring Twitter users and their conversations to recruit for clinical trials: survey study. J Med Internet Res 2019 Oct 30;21(10):e15455 [FREE Full text] [doi: 10.2196/15455] [Medline: 31670698]
7. Hannaford A, Lipshie-Williams M, Starrels JL, Arnsten JH, Rizzuto J, Cohen P, et al. The use of online posts to identify barriers to and facilitators of HIV pre-exposure prophylaxis (PrEP) among men who have sex with men: a comparison to a systematic review of the peer-reviewed literature. AIDS Behav 2018 Apr;22(4):1080-1095 [FREE Full text] [doi: 10.1007/s10461-017-2011-3] [Medline: 29285638]
8. McLaughlin ML, Hou J, Meng J, Hu CW, 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] [Medline: 26756069]
9. 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] [Medline: 27295507]
10. Centers for Disease Control and Prevention. Diagnoses of HIV infection in the United States and dependent areas, 2018 (Updated). HIV Surveillance Report, 2018 (Updated). 2020 May. URL: https://www.cdc.gov/hiv/library/reports/hiv-surveillance/vol-31/index.html [accessed 2021-03-23]
11. Klein AZ, Cai H, Weissenbacher D, Levine LD, Gonzalez-Hernandez G. A natural language processing pipeline to advance the use of Twitter data for digital epidemiology of adverse pregnancy outcomes. J Biomed Inform 2020;112S:100076 [FREE Full text] [doi: 10.1016/j.yjbinx.2020.100076] [Medline: 34417007]
12. Dredze M, Paul MJ, Bergsma S, Tran H. Carmen: a Twitter geo-location system with applications to public health. 2013 Presented at: The Association for the Advancement of Artificial Intelligence (AIII) Workshop Expanding the Boundaries of Health Informatics Using AI; Jul 14-15, 2013; Bellevue, WA.
13. Klein AZ, Magge A, Gonzalez-Hernandez G. ReportAGE: Automatically extracting the exact age of Twitter users based on self-reports in tweets. PLoS One 2022;17(1):e0262087 [FREE Full text] [doi: 10.1371/journal.pone.0262087] [Medline: 35077484]
14. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. Fam Med 2005 May;37(5):360-363 [FREE Full text] [Medline: 15883903]

15. Centers for Disease Control and Prevention. Ending the HIV epidemic in the U.S. - Jurisdictions. URL: https://www.cdc.gov/endhiv/jurisdictions.html [accessed 2021-05-10]

Abbreviations

API: application programming interface
MSM: men who have sex with men
NLP: natural language processing
PCP: primary care physician
PrEP: pre-exposure prophylaxis

Edited by H Bradley; submitted 26.07.21; peer-reviewed by J Banda Orozco, P Serrano; comments to author 08.11.21; revised version received 19.11.21; accepted 24.02.22; published 25.04.22.

Please cite as:
Klein AZ, Meanley S, O'Connor K, Bauermeister JA, Gonzalez-Hernandez G
Toward Using Twitter for PrEP-Related Interventions: An Automated Natural Language Processing Pipeline for Identifying Gay or Bisexual Men in the United States
JMIR Public Health Surveill 2022;8(4):e32405
URL: https://publichealth.jmir.org/2022/4/e32405
doi:10.2196/32405
PMID:35468092

Review

Factors Affecting Human Papillomavirus Vaccination in Men: Systematic Review

Hyunjeong Shin[1], RN, PhD; Songi Jeon[2], RN, CNM, PhD; Inhae Cho[1], RN, PhD; HyunJi Park[3], RN, MSN

[1]College of Nursing, Korea University, Seoul, Republic of Korea

[2]Department of Nursing, Catholic Kwandong University, Gangneung, Republic of Korea

[3]C&R Research, Seoul, Republic of Korea

Corresponding Author:
Inhae Cho, RN, PhD
College of Nursing
Korea University
145 Anam-ro, Sungbuk-gu
Seoul, 02841
Republic of Korea
Phone: 82 2 3290 4912
Email: inhae05@gmail.com

Abstract

Background: Despite the high risks associated with human papillomavirus (HPV), the HPV vaccination rate of men is far lower than women. Most previous review studies have focused on female vaccination and related affecting factors. However, previous studies have reported that the factors affecting HPV vaccination differ by gender.

Objective: The aim of this review was to identify the factors affecting HPV vaccine initiation in men through a systematic review approach.

Methods: A literature review was conducted across 3 central electronic databases for relevant articles. A total of 30 articles published between 2013 and 2019 met the inclusion criteria and were reviewed in this study.

Results: In total, 50 factors affecting HPV vaccination in men were identified, including 13 sociodemographic factors and social structure factors, 12 belief-related variables, 4 family factors, 4 community factors, 14 variables related to needs, and 3 environmental factors.

Conclusions: To increase HPV vaccination rates in men, strategies targeting young males and their families should consider frequent visits to or contact with health care providers so that health care professionals can provide recommendations for HPV vaccination.

(JMIR Public Health Surveill 2022;8(4):e34070) doi:10.2196/34070

KEYWORDS

health service use; men; papillomavirus; papillomavirus vaccines; systematic review; vaccination; vaccine; HPV; review; gender

Introduction

Background

Human papillomavirus (HPV) infection is one of the most common sexually transmitted infections (STIs) in all genders [1,2]. The available data show that approximately 13 million people are diagnosed with HPV infections annually in the United States [3], and these infections range from genital warts to cancers of the cervix, penis, anus, and head and neck. Some studies reported that HPV infection is associated with 5.2% of newly diagnosed cancers worldwide [4]. However, despite the high incidence of and risks associated with HPV-related diseases for both sexes [3-5], globally, the control of HPV infection is commonly considered for women only.

A crucial strategy for protecting people against HPV is the prophylactic HPV vaccination, which can prevent initial HPV infection [6-8]. Population uptake of HPV vaccination has demonstrated significant effectiveness in preventing HPV-related diseases, such as cervical intraepithelial neoplasia [9]. The World Health Organization has recommended HPV vaccination as a routine immunization for girls and female adolescents [6], and most countries have implemented national HPV vaccination programs for women [10]. More recently, HPV vaccination programs inclusive of all genders have been

discussed as being more effective for successfully acquiring herd immunity against HPV [10-12]. However, few countries have included young males in their national HPV vaccination programs [2,10,13]. According to a review that reported the HPV vaccination rates in adolescents, only the United States and Canada have reported rates in male adolescents, and these HPV vaccination rates (1.1%-31.7%) in males are confirmed to be very low compared to females (2.4%-94.4%) [14].

To increase the uptake of HPV vaccination among males, it is essential to understand the factors affecting HPV vaccine initiation. Previous review studies have reported that race and ethnicity, age, health insurance status, previous vaccination history, personal knowledge and awareness of HPV, and parental knowledge and education levels are significant predictors for HPV vaccination uptake [15,16]. Nevertheless, most of the published research to date has focused on female vaccination and the factors affecting it. Although some studies have reported that the factors affecting HPV vaccination differ by gender [17,18], they have rarely identified specific evidence regarding the factors affecting male HPV vaccination. In addition, those results that have reported factors affecting male HPV vaccination have been inconsistent. For example, regarding age, Thomas and colleagues reported that younger males had a higher rate of vaccination [19]. Conversely, other research has shown that it was older participants who had a higher rate of vaccination [20].

Increasing HPV vaccination rates in men requires identifying the unique factors affecting men's HPV vaccine initiation and establishing public health policies accounting for those factors. The Behavioral Model of Health Service Use (BMHSU) offers a multidimensional explanation about a person's use of health services. This study uses the BMHSU to structure the factors influencing HPV vaccination in men and provide a comprehensive understanding of these factors. Therefore, this study aimed to (1) confirm the factors affecting HPV vaccination in men through a systematic review of previous studies and (2) identify which components of the BMHSU were explored and influenced HPV vaccination in men.

BMHSU Framework

This study uses the BMHSU developed by Andersen [21] as a framework to explain the factors associated with HPV vaccination in men. The BMHSU has been widely used in studies on a person's access to health care and is a reliable model that could help understand determinants of health service usage. Since Andersen's behavioral model was introduced, it has been used to reveal barriers or determinants to accessing medical care in various medical situations, from mental health to cancer screening [22-25]. Hence, it is possible to explain differences in access to health care using model components. Moreover, this model provides a comprehensive and multifaceted understanding of access to health services in several review studies [26,27].

The BMHSU explains that health behaviors can be influenced by individual characteristics and the surrounding environment. In this model, the domain of population characteristics includes predisposing characteristics, enabling resources, and needs. Predisposing characteristics are described as personal propensities reflecting sociodemographic factors (eg, age, gender), social structure (eg, education, race, family size), and beliefs (eg, attitudes, knowledge, values). The enabling resources refer to the means that individuals have available to them, such as insurance, income, and community facilities. The needs refer to personally perceived and evaluated needs. Additionally, the domain of environment includes the health care system and external environment.

Methods

This study is a systematic review conducted to identify the factors influencing HPV vaccination in men. This study was performed according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [28]; however, this protocol was not registered.

Search Strategy

The data search was performed comprehensively across three electronic databases: PubMed, Embase, and CINAHL complete. Articles from the inception of the databases until July 2020 were searched. We applied search terms according to our research question: “What factors affect men's HPV vaccination uptake?” In accordance with the search terms, the MeSH (Medical Subject Headings), Emtree, and free terms were collected from the relevant literature and electronic databases before the search. A keyword search strategy was developed with the following terms: “Papillomavirus Infections,” “Papillomavirus,” “Human Papillomavirus,” “Wart virus,” “HPV,” “Vaccination,” “Vaccine,” “Immunization,” “Male,” “Men,” and “Boy.” The search strategies are given in detail in Multimedia Appendix 1.

Eligibility Criteria

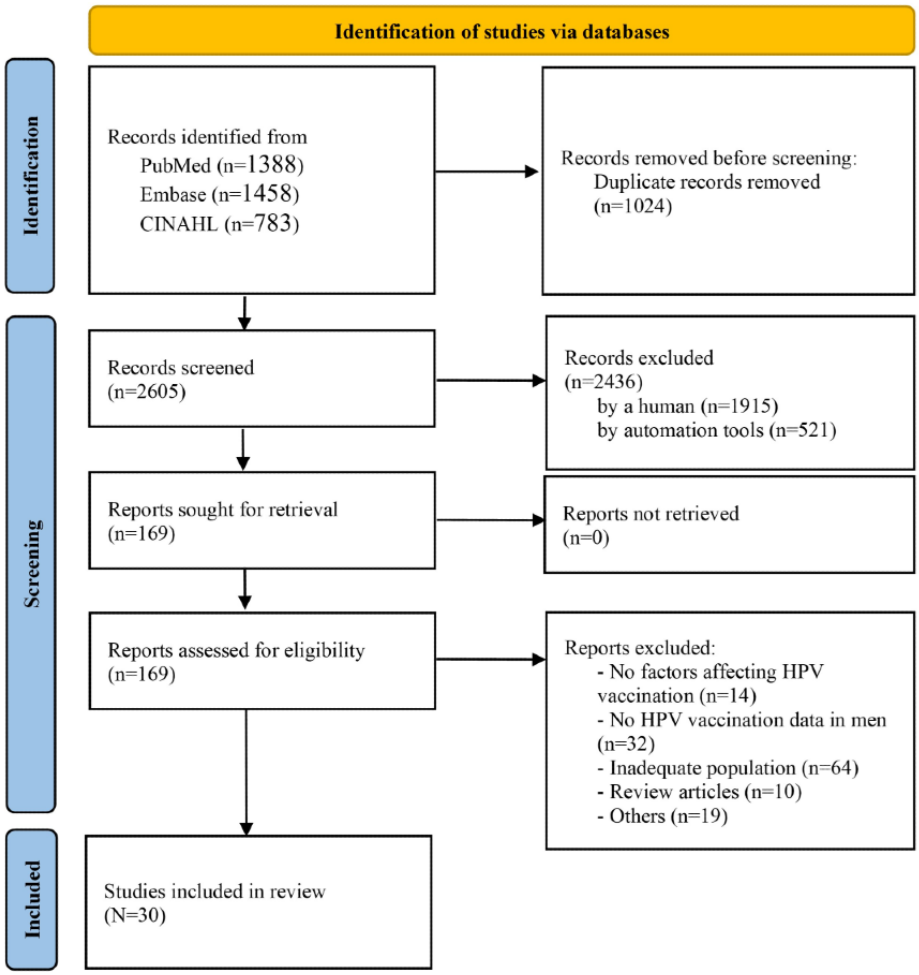
We set the inclusion criteria for identifying publications in accordance with the purpose of the review, as follows: (1) studies involving heterosexual male participants, (2) availability of HPV vaccine initiation data of the participants, (3) studies reporting factors or predictors associated with the HPV vaccination rate of men, and (4) peer-reviewed articles. The exclusion criteria were as follows: (1) studies presenting data excluding gender classification, (2) gray literature such as conference abstracts, (3) articles not published in English, and (4) articles describing experimental studies. All the articles were cross-checked by the research team.

Data Extraction and Synthesis

The data were extracted independently by 4 authors (HS, SJ, IC, and HJP). Duplicates were removed first using bibliography software (Endnote Version X9.1, Clarivate Analytics). After removing duplicates, the titles and abstracts were screened using the preset criteria, and irrelevant articles were excluded. A total of 30 studies were finally selected after assessing full-text articles. Figure 1 shows the flow diagram of article selection. In case of a discrepancy over selection, we first tried to reach an agreement through discussion. If the discrepancies remained unresolved, the principal investigator (HS) made the final decision.

For synthesis of the extracted data, the included studies were respectively coded using a predesigned template comprising the study authors, year of publication, country of data collection, study type, data source, age of the participants, sample size, and study results. The factors affecting HPV vaccination in men were classified as the key results. Factors were recorded for every association between factors and any corresponding *P* values used to identify which influencing factors had been studied and where the evidence was statistically significant were recorded. This information was then integrated to review the effects of different factors across the literature. Then, all the factors identified were grouped reflecting the components of the BMHSU.

Figure 1. Article selection flow diagram for the systematic review. HPV: Human papillomavirus.



Study Quality Assessment

We used the Mixed Methods Appraisal Tool (MMAT) version 2018 for quality assessment of the eligible articles [29]. The MMAT was designed to assess the quality of a variety of studies including qualitative, quantitative, and mixed-method research designs. Each quality criterion consisted of the following five items: (1) whether the sampling strategy was relevant to address the research question; (2) whether the sample was representative of the target population; (3) whether the measurements were appropriate; (4) whether there were risks of bias or confounders; (5) whether the statistical analysis was appropriate to answer the research question. Each item had "Yes," "No," or "Cannot tell" as the responses. To evaluate the quality of the studies, we determined the number of "Yes" responses to the items. The quality of the articles was assessed independently by 3 authors (HS, SJ, and HJP). The authors continuously discussed differences in evaluation until an agreement was reached.

Results

Study Characteristics

A total of 30 studies were included in the review [17,18,30-57]. The characteristics of the included studies are presented in Table 1. The studies were published between 2013 and 2019. All of them were conducted in the United States except for 1 study that was conducted in Denmark. Of the 30 studies, 26 were cross-sectional studies; the rest were cohort studies (n=3) and longitudinal studies (n=1). The participants of the studies were boys and male adults aged 9 to 34 years. The sample size of each study ranged from 168 to 809,656. The study quality scores obtained using the MMAT ranged from 2 to 5; the quality scores of most studies (n=24) were evaluated as very high (score=5) or high (score=4), 5 studies were rated moderate (score=3), and only 1 was rated low (score=2). The quality scores and the funding sources of the reviewed studies are presented in Table 1 and Multimedia Appendix 2.

Table 1. Characteristics of the reviewed studies (N=30)[a].

Study	Study type	Data source	Age (years)	Sample size (N)	Quality appraisal score[b]
Adjei Boakye et al (2018) [30]	Cross-sectional	2014-2015 National Health Interview Survey	18-26	7588	5
Adjei Boakye et al (2019) [31]	Cross-sectional	2014-2017 National Health Interview Survey	18-34	14,056	5
Agawu et al (2015) [32]	Cohort	2009-2013 large primary care network	11-18	58,757	5
Agénor et al (2015) [33]	Cross-sectional	2013-2014 National Health Interview Survey	18-31	6812	5
Bernat et al (2013) [34]	Cross-sectional	2010 web-based survey	18-24	1682	3
Bollerup et al (2017) [35]	Cohort	2006-2014 Civil Registration System	9-26	809,656	5
Burdette et al (2017) [36]	Cross-sectional	2008-2013 National Immunization Survey-Teen	13-17	56,632	5
Charlton et al (2017) [37]	Cohort	1996-2014 Growing Up Today Study	14-27	3342	3
Choi et al (2016) [17]	Cross-sectional	2012-2013 National Immunization Survey-Teen	13-17	20,355	5
Clarke et al (2016) [38]	Cross-sectional	2012-2013 Electronic medical records from Johns Hopkins Community Physicians clinics	11-26	14,688	5
Daniel-Ulloa et al (2016) [39]	Cross-sectional	2013 National Health Interview Survey	18-30	3003	5
Dela Cruz et al (2018) [40]	Cross-sectional	2014 population-based telephone survey in Hawaii	11-18	Parents (n=799) and their sons (n=467)	3
Fuller and Hinyard (2017) [41]	Cross-sectional	2013 Behavioral Risk Factor Surveillance System	18-26	1624	5
Hechter et al (2013) [42]	Cross-sectional	2009-2010 electronic health record from Kaiser Permanente Southern California	9-17	254,489	5
Johnson et al (2017) [18]	Cross-sectional	2013 National Immunization Survey-Teen	13-17	9554	5
Kepka et al (2016) [43]	Cross-sectional	2012 National Immunization Survey-Teen	13-17	10,141	5
Landis et al (2018) [44]	Cross-sectional	2014 National Immunization Survey-Teen	13-17	10,743	5
Lu et al (2013) [45]	Cross-sectional	2010 National Health Interview Survey	18-26	1741	5
Lu et al (2019) [46]	Cross-sectional	2011-2016 National Immunization Survey-Teen	13-17	9712	5
Morrow (2019) [47]	Cross-sectional	2013-2017 hospital-based teen health center, health department sexually transmitted disease clinic, and the general community	13-16	747	3
Pérez et al (2018) [48]	Cross-sectional	2011-2015 National Health Interview Survey	18-32	15,967	5
Ragan et al (2018) [49]	Cross-sectional	2014 self-administered HPV[c] Vaccine and Decision-Making Behaviors Survey at 2 universities	18-26	168	4
Rahman et al (2015) [50]	Cross-sectional	2011 National Immunization Survey-Teen	13-17	12,328	5
Ratanasiripong (2015) [51]	Cross-sectional	2012 web-based survey in California	18-26	189	2
Reiter et al (2013) [52]	Longitudinal	2010-2011 internet-based survey from a national sample of parents with sons	11-17	Parents (n=327) and their sons (n=228)	3
Reiter et al (2014) [53]	Cross-sectional	2010-2012 National Immunization Survey-Teen	13-17	4238	5
Sanders Thompson et al (2017) [54]	Cross-sectional	2012 National College Health Assessment IIb Survey	18-26	5013	5
Thompson et al (2016) [55]	Cross-sectional	2009-2013 National College Health Assessment II	18-26	31,130 (28.9% of 107,716)	5
Thompson et al (2019) [56]	Cross-sectional	2016 National Health Interview Survey	18-26	1714	4
Vu et al (2019) [57]	Cross-sectional	2016 data from Project DECOY	18-25	845	4

[a]Except for the study by Bollerup et al [35] that was carried out in Denmark, all the others were conducted in the United States.

[b] Scores were determined using the MMAT. All scores are quantitative.

[c]HPV: human papillomavirus.

Factors Affecting HPV Vaccination in Men

We adopted Anderson's BMHSU to structure the results of the review and reveal the relationships among the factors identified. We grouped the identified 50 factors into two components of the BMHSU: 47 population characteristics and 3 environmental factors. The factors identified in this review are presented in Textbox 1 and Table 2.

Textbox 1. Summary of the factors identified in this review. HPV: human papillomavirus; STI: sexually transmitted infection; Pap smear: Papanicolaou smear.

Population characteristics (47 factors)

- **Predisposing (25 factors)**
  - **Sociodemographic and social structure factors (13 factors)**
    - Race and ethnicity
    - Age
    - Parental education level
    - Sexual orientation
    - Nativity status
    - Relationship and marital status
    - Parental age
    - Parental marital status
    - Education level
    - Employment status
    - Number of children in household
    - Language proficiency of caregivers
    - Parental employment status
  - **Beliefs (12 factors)**
    - Parental awareness of HPV vaccine
    - Attitude toward HPV vaccination
    - Awareness of HPV or HPV vaccine
    - Perceived behavioral control
    - Receipt of STI information
    - Searching for health information
    - Subjective norms about getting HPV vaccine
    - Parents' perceived effectiveness of HPV vaccine
    - Parents' perceived risks of HPV
    - Parental talks with sons about HPV vaccine
    - Parental willingness to get sons free HPV vaccine
    - Parents' perceived severity of HPV-related cancers
- **Enabling (8 factors)**
  - **Family (4 factors)**
    - Household income
    - Insurance type
    - Having clinics for usual health care
    - Current state of insurance
  - **Community (4 factors)**
    - Region
    - Size and type of educational institution
    - Medical accessibility
    - Availability of cost-free vaccination programs

- **Need (14 factors)**
  - **Perceived (13 factors)**
    - Vaccinations other than HPV
    - Number of visits to clinics
    - Sexual behaviors
    - Time from the last check-up visit
    - Perceived personal health status
    - HIV testing history
    - Check-up at 11-12 years (eg, well-child visit)
    - History of urinary tract infections
    - Cigarette smoking
    - Reasons for clinic visits (preventive or acute)
    - Maternal vaccination for influenza
    - Maternal STI history
    - Maternal Pap smear screening
  - **Evaluated (1 factor)**
    - Mother’s abnormal Pap smear result

Environment (3 factors)

- **Health care system (3 factors)**
  - Health care provider’s recommendation for HPV vaccine
  - Year that vaccinations were performed
  - Facility type

Table 2. Studies researching factors affecting human papillomavirus vaccination in men.

Factors	References	Number of studies
Population characteristics		
Predisposing		
Sociodemographic and social structure factors		
Race and ethnicity	[17[a],18[a],30,34[a],36[a],38[a],39,40[a],41[a],43[a],44[a],45,46[a],48,49,52[a],53,54[a],55,57[a]]	20
Age	[17[a],18,30[a],34,36[a],38[a],40[a],43,46[a],47[a],48,49,51[a],54,55[a],57]	16
Parental education level	[17[a],18,35[a],36[a],40,43[a],44[a],46[a],53,57]	10
Sexual orientation	[33[a],34,37[a],39,54[a],55,57]	7
Nativity status	[30,31[a],35[a],46,48[a],53,57]	7
Relationship and marital status	[18,36[a],40,43,44,46[a]]	6
Parental age	[30,34,45,48[a],54,55[a]]	6
Parental marital status	[18[a],35[a],36[a],43,44[a],46]	6
Education level	[18,30[a],41,45,48[a]]	5
Employment status	[30,45,48[a],54[a],56]	5
Number of children in household	[36,53]	2
Language proficiency of caregivers	[18[a],53]	2
Parental employment status	[35][a]	1
Beliefs		
Parental awareness of HPV[b] vaccine	[40,50][a]	2
Attitude toward HPV vaccination	[51][a]	1
Awareness of HPV or HPV vaccine	[45][a]	1
Perceived behavioral control	[51][a]	1
Receipt of STI[c] information	[54][a]	1
Searching for health information	[56][a]	1
Subjective norms about getting HPV vaccine	[51]	1
Parents' perceived effectiveness of HPV vaccine	[40][a]	1
Parents' perceived risks of HPV	[40][a]	1
Parental talks with sons about HPV vaccine	[52]	1
Parental willingness to get sons free HPV vaccine	[52]	1
Parents' perceived severity of HPV-related cancers	[40]	1
Enabling		
Family		
Household income	[18,35[a],36[a],41,43,44[a],45[a],46,49,53[a],56]	11
Insurance type	[17[a],38[a],43[a],44[a],45,46[a],47[a],53,54[a]]	9
Having clinics for usual health care	[30,31,41[a],45,48,56]	6
Current state of insurance	[18,30,31,41,48[a],54[a]]	6
Community		
Region	[18,30,36,41,44,48,54][a]	7

Factors	References	Number of studies
Size and type of educational institution	[49,54[a],57]	3
Medical accessibility	[38,46][a]	2
Availability of cost-free vaccination programs	[18][a]	1
Need		
Perceived		
Vaccinations other than HPV	[32[a],43[a],44[a],49,54[a]]	5
Number of visits to clinics	[30,31,32,38,46][a]	5
Sexual behaviors	[34[a],37[a],45,49]	4
Time from the last check-up visit	[18[a],53[a],56]	3
Perceived personal health status	[32[a],54]	2
HIV testing history	[34,41][a]	2
Check-up at 11-12 years (eg, well-child visit)	[44,46[a]]	2
History of urinary tract infections	[54][a]	1
Cigarette smoking	[47][a]	1
Reasons for clinic visits (preventive or acute)	[32][a]	1
Maternal vaccination for influenza	[42][a]	1
Maternal STI history	[42]	1
Maternal Pap smear[d] screening	[42][a]	1
Evaluated		
Mother's abnormal Pap smear result	[42]	1
Environment		
Health care system		
Health care provider's recommendation for HPV vaccine	[17,18,36,44,46,50,53,57][a]	8
Year when vaccinations were performed	[17,32,36,53][a]	4
Facility type	[18,43,46]	3

[a]Studies that obtained statistically significant findings (P<.05).

[b]HPV: human papillomavirus.

[c]STI: sexually transmitted infection.

[d]Pap smear: Papanicolaou smear

Population Characteristics

According to the BMHSU, the domain of population characteristics includes predisposing factors, enabling factors, and need factors.

Predisposing Factors

Among the predisposing factors, social structure reflected the social position of the individuals in their society. We combined sociodemographic factors and social structure components of the BMHSU because it may be difficult to distinguish them in different cultural contexts. We then found 13 factors that were categorized as sociodemographic factors and social structure factors, including age, sexual orientation, relationship or marital status, parental age, race or ethnicity, education level, employment status, nativity status, number of children in the household, parental education level, parental marital status, language proficiency of caregivers, and parental employment status. Approximately half of the studies (n=16) explored age as a factor related to the HPV vaccination initiation [17,18,30,34,36,38,40,43,46-49,51,54,55,57]. Among them, 9 studies showed significant associations between age and vaccination [17,30,36,38,40,46,47,51,55]. However, the 5 studies with teenagers reported more vaccinations at older ages [17,36,40,46,47], whereas the other 4 studies covering populations aged up to 26 years reported the opposite relationship [30,38,51,55]. The most frequently mentioned factor was race or ethnicity (n=20) [17,18,30,34,36,38-41,43-46,48,49,52-55,57]; of the 20 studies, 13 showed significant relationships between race or ethnicity

and vaccination [17,18,34,36,38,40,41,43,44,46,52,54,57]. However, some of the results are inconsistent, with 7 studies reporting that Hispanics were more vaccinated than non-Hispanics [17,18,36,43,44,46,54]; other studies reported conflicting results that African Americans [34,38] were more vaccinated or Caucasians [40,52] were less vaccinated than other ethnic groups. As for parental education level, 6 out of 10 studies revealed significant associations [17,35,36,43,44,46]. The association between parents' marital status and HPV vaccination was explored in 6 studies [18,35,36,43,44,46]. A significant association was reported in 4 of them [18,35,36,44], and married parents were less likely to have their sons vaccinated according to 3 studies [18,36,44], but 1 study in Denmark reported the opposite result [35].

We identified 12 belief factors, namely attitude toward HPV vaccination, awareness of HPV or HPV vaccine, perceived behavioral control, receipt of STI information, searching for health information, subjective norms about getting HPV vaccine, parental awareness of HPV vaccine, parents' perceived effectiveness of HPV vaccine, parents' perceived risks of HPV, parental talks with sons about HPV vaccine, parental willingness to get sons vaccinated, and parents' perceived severity of HPV-related cancers. Beliefs refer to individual values about health, attitudes toward health services, or knowledge about disease. Significantly, sons of parents with high awareness of the HPV vaccines were more likely to be vaccinated [40,50]. However, other significant factors belonging to the belief domain were reported in only 1 study.

Enabling Factors

We found four enabling factors that were ultimately categorized as family factors: household income, health insurance types, having clinics for usual health care, and current state of insurance. As for household income, those with low economic status had a higher likelihood of being vaccinated according to 4 studies [36,44,45,53]; however, 1 study conducted in Denmark reported conflicting results [35]. Regarding the relationships between insurance type and HPV vaccination, 7 out of 9 studies reported that men with Medicaid or public insurance were significantly more likely to get vaccinated than men with private insurance [17,38,43,44,46,47,54]. In addition, uninsured people were less likely to take HPV vaccinations [48,54].

The remaining four factors identified were categorized as community factors: region (eg, specific city or state), size and type of educational institution, accessibility to medical facilities, and availability of cost-free vaccination programs. The rate of HPV vaccination was significantly higher in urban areas [38,46] where accessibility to medical facilities was better.

Need Factors

The need factors of populations could be classified into perceived and evaluated. In this study, 13 factors were identified as perceived needs that affect HPV vaccination in men, including vaccinations other than HPV, the number of visits to clinics, sexual behaviors, time from the last check-up visit to a clinic, perceived personal health status, history of urinary tract infections, participants' HIV testing history, check-up at 11 to 12 years (eg, well-child visit), cigarette smoking, reasons for clinic visits, maternal vaccination for influenza, maternal STI history, and maternal Papanicolaou (Pap) smear screening. The reviewed studies reported that men who received other vaccinations (eg, influenza; meningitis; hepatitis B; and tetanus, diphtheria, and pertussis vaccines) were more likely to take HPV vaccines [32,43,44,54]. All 5 studies that explored the number of visits to clinics as the factors affecting HPV vaccination in men [30-32,38,46] consistently reported that more frequent clinic visits had a significant relationship with more HPV vaccinations. It was also found that more HPV vaccinations were given to those who had undergone HIV tests [34,41] and those whose mothers had undergone Pap smear tests [42].

In this study, one factor was classified as an evaluated need: the mother's abnormal Pap smear result. It was found that initiation of HPV vaccination was associated with abnormal maternal Pap smear results [42].

Environment

The BMHSU suggests the health care system and external environment as the environmental factors that affect individual health behaviors [21]. This review did not reveal any significant external environmental factors. The following were identified as health care system factors: health care provider's recommendations for HPV vaccination, the year that the vaccinations were performed, and the types of health care facilities that men usually use. The most frequently reported factor was the health care provider's recommendations for HPV vaccination (n=8). Those who received recommendations for HPV vaccination showed more HPV vaccination uptake [17,18,36,44,46,50,53,57]. The studies also revealed that HPV vaccination rates have gradually increased since the year when the government recommended HPV vaccination for males in the United States [17,32,36,53].

Discussion

Principal Findings

This systematic review identified 30 peer-reviewed research articles published from 2013 to 2019 that contained studies analyzing the factors affecting male HPV vaccination. We identified a total of 50 modifiable and nonmodifiable factors across a wide range of domains assessed by the authors of those studies as being associated with male HPV vaccination. For this study, we used Andersen's BMHSU to structure the identified factors. The highest rates of vaccination tended to be associated with the following identified factors: men aged 10 to 20 years, Hispanic race, adolescents with single parents, higher parental knowledge and awareness of HPV vaccination, low economic status, individuals with public insurance, males living in urban areas, individuals receiving other vaccinations, frequent visits to clinics, and receiving health care providers' recommendations. Although many factors have been identified in previous studies and could be potential affecting factors (eg, attitude toward HPV vaccination, parental perceptions of the risks of HPV), they were often only assessed by single studies featuring small sample sizes, thus limiting the generalizability of these factors to the larger male population.

We found that health care providers' recommendations are major facilitators of HPV vaccinations in men. These findings coincided with those of prior review studies noting that the health care provider's recommendation was the most significant factor leading to the initiation of HPV vaccination [58-60]. One of the studies reviewed in this work reported that receiving the health care provider's recommendation was the most influential factor in parental decisions on vaccinating children for HPV [59]. Another review study reported that people who were advised to get HPV vaccinations by health care providers were more likely to be vaccinated than people with no recommendations, with an odds ratio of 10.1 [60]. However, despite positive evidence of health care providers' recommendations, the delivery of recommendations associated with HPV vaccine in men is a challenge that must be overcome. Some studies explain a gap associated with recommendations according to sex that originated from the provider's lack of knowledge or opinions [61,62]. Perkins and colleagues also reported that a provider-focused intervention targeting health care providers improved the initiation of the HPV vaccine prominently for boys [63]. Thus, a first step to enhance the rates of male HPV vaccination could be to induce consistent recommendations for immunization from the health care providers through training for health professionals.

This review suggests that frequent contact with health care providers is a promising factor for promoting HPV vaccination rates in men. We found that the number of visits to clinics and accessibility to health care facilities have significant relationships with the initiation of HPV vaccinations in men [30-32,38,46,54]. Furthermore, studies revealed that those males who had taken vaccinations other than those for HPV (eg, tetanus, diphtheria, and pertussis; meningitis; influenza; and hepatitis B) had higher HPV vaccination rates. This might have shown the potential of a program developed to link the initiation of HPV vaccination to other vaccinations conducted at the same age [13]. Thus, we should seek frequent contact with health care professionals and explore feasible programs at the same time for men based on these results.

This review also suggests that parental awareness of HPV vaccination is an essential facilitator of HPV vaccinations in children. Several previous studies on men's HPV vaccinations have noted that a lack of knowledge regarding the HPV vaccine is one of the causes of low HPV vaccination rates in men [64-66]. Radisic and colleagues also reported that parental awareness of HPV was strongly associated with vaccinating children [67]. Because it is advisable to initiate vaccination for early adolescents aged 9 to 14 years [13], it is important to increase parental awareness of vaccination after the national introduction of HPV vaccination in men. Thus, these findings suggest the necessity of parental education to promote HPV vaccination in men effectively. To promote rates of HPV vaccine uptake in men, national policies should include parents as decision-makers and aim to increase their knowledge and awareness.

In this study, some factors (eg, household income, marital status of parents) were significant but inconsistent in their direction of association with male HPV vaccination. In a study performed in the United States, household income had a negative relationship with HPV vaccination, whereas in a study performed in Denmark, a positive one was found. Regarding ethnic groups, the reviewed studies suggest various ethnic groups to be the most vaccinated, for example, Hispanics or African Americans. One of the reasons is that most of the reviewed studies were conducted in the United States. Since 2011, the United States has included men in the national HPV vaccination program. Accordingly, the HPV vaccination rate in men has changed over time, and it can be expected that the findings are diverse according to the time of the studies included in this review. For more nuanced and conclusive results on these aspects, future research should be performed in a wider range of countries and target families with different socioeconomic statuses and ethnic backgrounds.

Limitations

This review has some limitations. First, most of the included studies, except for 1, were conducted in the United States. Despite the comprehensive findings on the factors affecting male HPV vaccination, it might be difficult to generalize these results. Second, most studies that were reviewed in this work performed secondary data analyses and extracted the variables from available large-scale data sets that had already been constructed for different purposes. Thus, a wide range of psychosocial variables or factors were not reflected in the reviewed studies. Third, we only included heterosexual participants. Men who have sex with men (MSM) are at high risk for HPV infection [68,69]. HPV vaccination was emphasized for the MSM population more than for heterosexual males, as seen in the differences with respect to the recommended age [70]. For this reason, we assumed that the factors affecting HPV vaccine initiation in men might be different in the case of MSM. However, further research should include various sexual orientations and gender groups to understand and unify disparities in HPV vaccine initiation in men.

Conclusions

This review analyzed 30 research articles using a systematic approach and identified a total of 50 factors affecting HPV vaccination in men. Based on our results, to increase the rates of male HPV vaccination, strategies targeting health care providers should be considered so that health care professionals can provide consistent and evidence-based recommendations for HPV vaccination, including frequent visits or contact with health care providers. In addition, national policies should include parents as decision-makers and increase their knowledge and awareness of HPV vaccination and its effects on preventing a variety of cancers.

Acknowledgments

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (grant NRF-2019R1I1A1A01063595).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy used in electronic databases.
[DOCX File , 16 KB - publichealth_v8i4e34070_app1.docx]

Multimedia Appendix 2
Funding sources of the reviewed studies (N=30).
[DOCX File , 18 KB - publichealth_v8i4e34070_app2.docx]

References

1. Dunne EF, Markowitz LE, Saraiya M, Stokley S, Middleman A, Unger ER, Centers for Disease Control and Prevention (CDC). CDC grand rounds: reducing the burden of HPV-associated cancer and disease. MMWR Morb Mortal Wkly Rep 2014 Jan;63(4):69-72 [FREE Full text] [Medline: 24476977]
2. Prue G, Grimes D, Baker P, Lawler M. Access to HPV vaccination for boys in the United Kingdom. Medicine Access @ Point of Care 2018 Sep;2:1-9. [doi: 10.1177/2399202618799691]
3. Kreisel KM, Spicknall IH, Gargano JW, Lewis FM, Lewis RM, Markowitz LE, et al. Sexually transmitted infections among US women and men: prevalence and incidence estimates, 2018. Sexual Trans Dis 2021 Jan;48(4):208-214. [doi: 10.1097/olq.0000000000001355]
4. Chaturvedi AK. Beyond cervical cancer: burden of other HPV-related cancers among men and women. J Adolesc Health 2010 Apr;46(4):S20-S26. [doi: 10.1016/j.jadohealth.2010.01.016]
5. Taylor S, Bunge E, Bakker M, Castellsagué X. The incidence, clearance and persistence of non-cervical human papillomavirus infections: a systematic review of the literature. BMC Infect Dis 2016 Jun;16:293. [doi: 10.1186/s12879-016-1633-9]
6. World Health Organization. Global strategy to accelerate the elimination of cervical cancer as a public health problem. 2020. URL: https://apps.who.int/iris/bitstream/handle/10665/336583/9789240014107-eng.pdf [accessed 2021-03-04]
7. Donovan B, Franklin N, Guy R, Grulich AE, Regan DG, Ali H, et al. Quadrivalent human papillomavirus vaccination and trends in genital warts in Australia: analysis of national sentinel surveillance data. Lancet Infect Dis 2011 Jan;11(1):39-44. [doi: 10.1016/S1473-3099(10)70225-5]
8. Giuliano AR, Palefsky JM, Goldstone S, Moreira ED, Penny ME, Aranda C, et al. Efficacy of quadrivalent HPV vaccine against HPV Infection and disease in males. N Engl J Med 2011 Feb;364(5):401-411 [FREE Full text] [doi: 10.1056/NEJMoa0909537] [Medline: 21288094]
9. Arbyn M, Bryant A, Beutels P, Martin-Hirsch PP, Paraskevaidis E, Van Hoof E, et al. Prophylactic vaccination against human papillomaviruses to prevent cervical cancer and its precursors. Cochrane Database Syst Rev 2018 May;4:CD009069 [FREE Full text] [doi: 10.1002/14651858.CD009069] [Medline: 25267916]
10. Stanley M. Perspective: vaccinate boys too. Nature 2012 Aug;488(7413):S10. [doi: 10.1038/488S10a] [Medline: 22932433]
11. Elfström KM, Lazzarato F, Franceschi S, Dillner J, Baussano I. Human papillomavirus vaccination of boys and extended catch-up vaccination: effects on the resilience of programs. J Infect Dis 2015 Jul;213(2):199-205. [doi: 10.1093/infdis/jiv368]
12. Bogaards JA, Wallinga J, Brakenhoff RH, Meijer CJLM, Berkhof J. Direct benefit of vaccinating boys along with girls against oncogenic human papillomavirus: bayesian evidence synthesis. BMJ 2015 May;350:h2016. [doi: 10.1136/bmj.h2016]
13. World Health Organization. Human papillomavirus vaccines: WHO position paper, May 2017–recommendations. Vaccine 2017 Oct;35(43):5753-5755. [doi: 10.1016/j.vaccine.2017.05.069]
14. Loke AY, Kwan ML, Wong Y, Wong AKY. The uptake of human papillomavirus vaccination and its associated factors among adolescents: a systematic review. J Prim Care Community Health 2017 Nov;8(4):349-362. [doi: 10.1177/2150131917742299]
15. Kessels SJM, Marshall HS, Watson M, Braunack-Mayer AJ, Reuzel R, Tooher RL. Factors associated with HPV vaccine uptake in teenage girls: a systematic review. Vaccine 2012 May;30(24):3546-3556. [doi: 10.1016/j.vaccine.2012.03.063]
16. Rodriguez SA, Mullen PD, Lopez DM, Savas LS, Fernández ME. Factors associated with adolescent HPV vaccination in the U.S.: a systematic review of reviews and multilevel framework to inform intervention development. Prev Med 2020 Feb;131:105968 [FREE Full text] [doi: 10.1016/j.ypmed.2019.105968] [Medline: 31881235]
17. Choi Y, Eworuke E, Segal R. What explains the different rates of human papillomavirus vaccination among adolescent males and females in the United States? Papillomavirus Res 2016 Dec;2:46-51 [FREE Full text] [doi: 10.1016/j.pvr.2016.02.001] [Medline: 29074185]

18. Johnson KL, Lin M, Cabral H, Kazis LE, Katz IT. Variation in human papillomavirus vaccine uptake and acceptability between female and male adolescents and their caregivers. J Community Health 2017 Jun;42(3):522-532. [doi: 10.1007/s10900-016-0284-5]
19. Thomas R, Higgins L, Ding L, Widdice LE, Chandler E, Kahn JA. Factors associated with HPV vaccine initiation, vaccine completion, and accuracy of self-reported vaccination status among 13- to 26-year-old men. Am J Men's Health 2018 Jul;12(4):819-827 [FREE Full text] [doi: 10.1177/1557988316645155] [Medline: 27106515]
20. Krakow M, Beavis A, Cosides O, Rositch AF. Characteristics of adolescents lacking provider-recommended human papillomavirus vaccination. J Adolesc Health 2017 May;60(5):619-622. [doi: 10.1016/j.jadohealth.2016.11.028] [Medline: 28073618]
21. Andersen RM. Revisiting the behavioral model and access to medical care: Does it matter? J Health Soc Behav 1995 Mar;36:1-10. [doi: 10.2307/2137284]
22. Babitsch B, Gohl D, von Lengerke T. Re-revisiting Andersen's behavioral model of health services use: a systematic review of studies from 1998–2011. Psychosoc Med 2012 Oct;9:Doc11-Doc15 [FREE Full text] [doi: 10.3205/psm000089] [Medline: 23133505]
23. Fortin M, Bamvita JM, Fleury MJ. Patient satisfaction with mental health services based on Andersen's Behavioral Model. Can J Psychiatry 2017 Oct;63(2):103-114. [doi: 10.1177/0706743717737030]
24. Ogunsanya ME, Jiang S, Thach AV, Bamgbade BA, Brown CM. Predictors of prostate cancer screening using Andersen's Behavioral Model of Health Services Use. Urol Oncol 2016 Dec;34(12):529.e9-529.e14. [doi: 10.1016/j.urolonc.2016.06.016]
25. Hong YR, Samuels SK, Huo JH, Lee N, Mansoor H, Duncan RP. Patient-centered care factors and access to care: a path analysis using the Andersen behavior model. Public Health 2019 Jun;171:41-49. [doi: 10.1016/j.puhe.2019.03.020]
26. Bina R. Predictors of postpartum depression service use: a theory-informed, integrative systematic review. Women Birth 2020 Feb;33(1):e24-e32. [doi: 10.1016/j.wombi.2019.01.006] [Medline: 30718105]
27. Lederle M, Tempes J, Bitzer EM. Application of Andersen's behavioural model of health services use: a scoping review with a focus on qualitative health services research. BMJ Open 2021 May;11(5):e045018. [doi: 10.1136/bmjopen-2020-045018]
28. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
29. Hong QN, Fàbregues S, Bartlett G, Boardman F, Cargo M, Dagenais P, et al. The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. EFI 2018 Dec;34(4):285-291. [doi: 10.3233/EFI-180221]
30. Adjei Boakye E, Lew D, Muthukrishnan M, Tobo BB, Rohde RL, Varvares MA, et al. Correlates of human papillomavirus (HPV) vaccination initiation and completion among 18-26 year olds in the United States. Hum Vaccin Immunother 2018 Jun;14(8):2016-2024 [FREE Full text] [doi: 10.1080/21645515.2018.1467203] [Medline: 29708826]
31. Adjei Boakye E, Zeng W, Governor S, Nagendra S, Tobo BB, Simpson MC, et al. Differences in human papillomavirus (HPV) vaccine uptake by nativity status among men aged 18-34 years. Prev Med Rep 2019 Dec;16:101010 [FREE Full text] [doi: 10.1016/j.pmedr.2019.101010] [Medline: 31799106]
32. Agawu A, Buttenheim AM, Taylor L, Song L, Fiks AG, Feemster KA. Sociodemographic differences in human papillomavirus vaccine initiation by adolescent males. J Adolesc Health 2015 Nov;57(5):506-514. [doi: 10.1016/j.jadohealth.2015.07.002] [Medline: 26381919]
33. Agénor M, Peitzmeier SM, Gordon AR, Charlton BM, Haneuse S, Potter J, et al. Sexual orientation identity disparities in human papillomavirus vaccination initiation and completion among young adult US women and men. Cancer Causes Control 2016 Aug;27(10):1187-1196. [doi: 10.1007/s10552-016-0796-4]
34. Bernat DH, Gerend MA, Chevallier K, Zimmerman MA, Bauermeister JA. Characteristics associated with initiation of the human papillomavirus vaccine among a national sample of male and female young adults. J Adolesc Health 2013 Nov;53(5):630-636 [FREE Full text] [doi: 10.1016/j.jadohealth.2013.07.035] [Medline: 24138764]
35. Bollerup S, Baldur-Felskov B, Dehlendorff C, Kjaer SK. Socioeconomic predictors of human papillomavirus vaccination in Danish men - a nationwide study. Papillomavirus Res 2017 Jun;3:18-23 [FREE Full text] [doi: 10.1016/j.pvr.2016.11.004] [Medline: 28720452]
36. Burdette AM, Webb NS, Hill TD, Jokinen-Gordon H. Race-specific trends in HPV vaccinations and provider recommendations: persistent disparities or social progress? Public Health 2017 Jan;142:167-176. [doi: 10.1016/j.puhe.2016.07.009] [Medline: 27592005]
37. Charlton BM, Reisner SL, Agénor M, Gordon AR, Sarda V, Austin SB. Sexual orientation disparities in human papillomavirus vaccination in a longitudinal cohort of U.S. males and females. LGBT Health 2017 Jun;4(3):202-209 [FREE Full text] [doi: 10.1089/lgbt.2016.0103] [Medline: 28467238]
38. Clarke MA, Coutinho F, Phelan-Emrick DF, Wilbur M, Chou B, Joshu CE. Predictors of human papillomavirus vaccination in a large clinical population of males aged 11 to 26 years in Maryland, 2012–2013. Cancer Epidemiol Biomarkers Prev 2015 Dec;25(2):351-358. [doi: 10.1158/1055-9965.epi-15-0983]
39. Daniel-Ulloa J, Gilbert PA, Parker EA. Human papillomavirus vaccination in the United States: uneven uptake by gender, race/ethnicity, and sexual orientation. Am J Public Health 2016 Apr;106(4):746-747. [doi: 10.2105/ajph.2015.303039]

40. Dela Cruz MRI, Braun KL, Tsark JAU, Albright CL, Chen JJ. HPV vaccination prevalence, parental barriers and motivators to vaccinating children in Hawai'i. Ethn Health 2020 Oct;25(7):982-994 [FREE Full text] [doi: 10.1080/13557858.2018.1473556] [Medline: 29745749]
41. Fuller KM, Hinyard L. Factors associated with HPV vaccination in young males. J Community Health 2017 Dec;42(6):1127-1132. [doi: 10.1007/s10900-017-0361-4] [Medline: 28432549]
42. Hechter RC, Chao C, Sy LS, Ackerson BK, Slezak JM, Sidell MA, et al. Quadrivalent human papillomavirus vaccine uptake in adolescent boys and maternal utilization of preventive care and history of sexually transmitted infections. Am J Public Health 2013 Sep;103(9):e63-e68. [doi: 10.2105/ajph.2013.301495]
43. Kepka D, Ding Q, Hawkins AJ, Warner EL, Boucher KM. Factors associated with early adoption of the HPV vaccine in US male adolescents include Hispanic ethnicity and receipt of other vaccines. Prev Med Rep 2016 Dec;4:98-102 [FREE Full text] [doi: 10.1016/j.pmedr.2016.05.014] [Medline: 27413668]
44. Landis K, Bednarczyk RA, Gaydos LM. Correlates of HPV vaccine initiation and provider recommendation among male adolescents, 2014 NIS-Teen. Vaccine 2018 Jun;36(24):3498-3504. [doi: 10.1016/j.vaccine.2018.04.075]
45. Lu P, Williams WW, Li J, Dorell C, Yankey D, Kepka D, et al. Human papillomavirus vaccine initiation and awareness: U.S. young men in the 2010 National Health Interview Survey. Am J Prev Med 2013 Apr;44(4):330-338 [FREE Full text] [doi: 10.1016/j.amepre.2012.11.027] [Medline: 23498097]
46. Lu P, Yankey D, Fredua B, O'Halloran AC, Williams C, Markowitz LE, et al. Association of provider recommendation and human papillomavirus vaccination initiation among male adolescents aged 13-17 years-United States. J Pediatr 2019 Mar;206:33-41.e1 [FREE Full text] [doi: 10.1016/j.jpeds.2018.10.034] [Medline: 30448270]
47. Morrow C, Ding L, Gorbach P, Chandler E, Crumpton M, Kahn JA. Rates and predictors of human papillomavirus vaccination among young men recruited from urban clinical and community settings. J Men's Health 2019 Mar;15(1):30-44. [doi: 10.22374/jomh.v15i1.114]
48. Pérez AE, Agénor M, Gamarel KE, Operario D. Nativity disparities in human papillomavirus vaccination among U.S. adults. Am J Prev Med 2018 Feb;54(2):248-258 [FREE Full text] [doi: 10.1016/j.amepre.2017.10.019] [Medline: 29241719]
49. Ragan KR, Bednarczyk RA, Butler SM, Omer SB. Missed opportunities for catch-up human papillomavirus vaccination among university undergraduates: identifying health decision-making behaviors and uptake barriers. Vaccine 2018 Jan;36(2):331-341. [doi: 10.1016/j.vaccine.2017.07.041]
50. Rahman M, Laz TH, McGrath CJ, Berenson AB. Provider recommendation mediates the relationship between parental human papillomavirus (HPV) vaccine awareness and HPV vaccine initiation and completion among 13- to 17-year-old US adolescent children. Clin Pediatr (Phila) 2014 Sep;54(4):371-375. [doi: 10.1177/0009922814551135]
51. Ratanasiripong NT. Factors related to human papillomavirus (HPV) vaccination in college men. Public Health Nurs 2015 Apr;32(6):645-653. [doi: 10.1111/phn.12198]
52. Reiter PL, McRee A, Pepper JK, Gilkey MB, Galbraith KV, Brewer NT. Longitudinal predictors of human papillomavirus vaccination among a national sample of adolescent males. Am J Public Health 2013 Aug;103(8):1419-1427 [FREE Full text] [doi: 10.2105/AJPH.2012.301189] [Medline: 23763402]
53. Reiter PL, Brewer NT, Gilkey MB, Katz ML, Paskett ED, Smith JS. Early adoption of the human papillomavirus vaccine among Hispanic adolescent males in the United States. Cancer 2014 Jun;120(20):3200-3207. [doi: 10.1002/cncr.28871]
54. Thompson VLS, Butler-Barnes ST, Jones BD, Wells AA, Cunningham-Williams RM, Williams S. Factors associated with human papillomavirus vaccination status at U.S. colleges and universities. Health Soc Work 2017 Feb;42(1):e1-e7. [doi: 10.1093/hsw/hlw050] [Medline: 28395066]
55. Thompson EL, Vamos CA, Vázquez-Otero C, Logan R, Griner S, Daley EM. Trends and predictors of HPV vaccination among U.S. College women and men. Prev Med 2016 May;86:92-98. [doi: 10.1016/j.ypmed.2016.02.003] [Medline: 26868093]
56. Thompson EL, Rosen BL, Maness SB. Social determinants of health and human papillomavirus vaccination among young adults, National Health Interview Survey 2016. J Community Health 2019 Feb;44(1):149-158. [doi: 10.1007/s10900-018-0565-2] [Medline: 30120681]
57. Vu M, Bednarczyk R, Escoffery C, Getachew B, Berg C. Human papillomavirus vaccination among diverse college students in the state of Georgia: who receives recommendation, who initiates and what are the reasons? Health education research. Health Educ Res 2019 Aug;34(4):415-434. [doi: 10.1093/her/cyz014]
58. Newman PA, Logie CH, Lacombe-Duncan A, Baiden P, Tepjan S, Rubincam C, et al. Parents' uptake of human papillomavirus vaccines for their children: a systematic review and meta-analysis of observational studies. BMJ Open 2018 Apr;8:e019206. [doi: 10.1136/bmjopen-2017-019206]
59. Nasserie T, Bendavid E. Systematic identification and replication of factors associated with human papillomavirus vaccine initiation among adolescents in the United States using an environment-wide association study approach. Sex Transm Infect 2022 May;98(3):203-209. [doi: 10.1136/sextrans-2021-054976] [Medline: 34039744]
60. Oh NL, Biddell CB, Rhodes BE, Brewer NT. Provider communication and HPV vaccine uptake: a meta-analysis and systematic review. Prev Med 2021 Jul;148:106554. [doi: 10.1016/j.ypmed.2021.106554] [Medline: 33857561]
61. Alexander AB, Best C, Stupiansky N, Zimet GD. A model of health care provider decision making about HPV vaccination in adolescent males. Vaccine 2015 Aug;33(33):4081-4086. [doi: 10.1016/j.vaccine.2015.06.085]

62. Gilkey MB, Calo WA, Moss JL, Shah PD, Marciniak MW, Brewer NT. Provider communication and HPV vaccination: the impact of recommendation quality. Vaccine 2016 Feb;34(9):1187-1192 [FREE Full text] [doi: 10.1016/j.vaccine.2016.01.023] [Medline: 26812078]
63. Perkins RB, Zisblatt L, Legler A, Trucks E, Hanchate A, Gorin SS. Effectiveness of a provider-focused intervention to improve HPV vaccination rates in boys and girls. Vaccine 2015 Feb;33(9):1223-1229. [doi: 10.1016/j.vaccine.2014.11.021]
64. Osazuwa-Peters N, Adjei Boakye E, Mohammed KA, Tobo BB, Geneus CJ, Schootman M. Not just a woman's business! Understanding men and women's knowledge of HPV, the HPV vaccine, and HPV-associated cancers. Prev Med 2017 Jun;99:299-304. [doi: 10.1016/j.ypmed.2017.03.014] [Medline: 28341458]
65. Adjei Boakye E, Tobo BB, Rojek RP, Mohammed KA, Geneus CJ, Osazuwa-Peters N. Approaching a decade since HPV vaccine licensure: Racial and gender disparities in knowledge and awareness of HPV and HPV vaccine. Hum Vaccin Immunother 2017 Nov;13(11):2713-2722 [FREE Full text] [doi: 10.1080/21645515.2017.1363133] [Medline: 28853980]
66. Preston SM, Darrow WW. Are men being left behind (or catching up)? Differences in HPV awareness, knowledge, and attitudes between diverse college men and women. Am J Men's Health 2019 Dec;13(6):1557988319883776 [FREE Full text] [doi: 10.1177/1557988319883776] [Medline: 31787066]
67. Radisic G, Chapman J, Flight I, Wilson C. Factors associated with parents' attitudes to the HPV vaccination of their adolescent sons : a systematic review. Prev Med 2017 Feb;95:26-37. [doi: 10.1016/j.ypmed.2016.11.019] [Medline: 27932052]
68. Gerend MA, Madkins K, Phillips G, Mustanski B. Predictors of human papillomavirus vaccination among young men who have sex with men. Sex Transm Dis 2016 Mar;43(3):185-191. [doi: 10.1097/olq.0000000000000408]
69. Meites E, Kempe A, Markowitz LE. Use of a 2-Dose schedule for human papillomavirus vaccination-updated recommendations of the Advisory Committee on Immunization Practices. MMWR Morb Mortal Wkly Rep 2016 Dec;65(49):1405-1408 [FREE Full text] [doi: 10.15585/mmwr.mm6549a5] [Medline: 27977643]
70. Meites E, Szilagyi PG, Chesson HW, Unger ER, Romero JR, Markowitz LE. Human papillomavirus vaccination for adults: updated recommendations of the Advisory Committee on Immunization Practices. MMWR Morb Mortal Wkly Rep 2019 Aug;68(32):698-702 [FREE Full text] [doi: 10.15585/mmwr.mm6832a3] [Medline: 31415491]

Abbreviations

BMHSU: Behavioral Model of Health Services Use
HPV: human papillomavirus
MMAT: Mixed Methods Appraisal Tool
MSM: men who have sex with men
Pap smear: Papanicolaou smear
STI: sexually transmitted infection

Edited by T Sanchez; submitted 05.10.21; peer-reviewed by YL Leung, J Fleming; comments to author 18.01.22; revised version received 07.03.22; accepted 01.04.22; published 26.04.22.

Please cite as:
Shin H, Jeon S, Cho I, Park H
Factors Affecting Human Papillomavirus Vaccination in Men: Systematic Review
JMIR Public Health Surveill 2022;8(4):e34070
URL: https://publichealth.jmir.org/2022/4/e34070
doi:10.2196/34070
PMID:35471242

Original Paper

The Impact of COVID-19 on Mortality in Italy: Retrospective Analysis of Epidemiological Trends

Alessandro Rovetta[1]; Akshaya Srikanth Bhagavathula[2], PhD

[1]R & C Research, Bovezzo (Brescia), Italy

[2]Institute of Public Health, College of Medicine and Health Sciences, United Arab Emirates University, Abu Dhabi, United Arab Emirates

Corresponding Author:
Alessandro Rovetta
R & C Research
via Brede T2
Bovezzo (Brescia), 25073
Italy
Phone: 39 3927112808
Email: rovetta.mresearch@gmail.com

Abstract

Background: Despite the available evidence on its severity, COVID-19 has often been compared with seasonal flu by some conspirators and even scientists. Various public discussions arose about the noncausal correlation between COVID-19 and the observed deaths during the pandemic period in Italy.

Objective: This paper aimed to search for endogenous reasons for the mortality increase recorded in Italy during 2020 to test this controversial hypothesis. Furthermore, we provide a framework for epidemiological analyses of time series.

Methods: We analyzed deaths by age, sex, region, and cause of death in Italy from 2011 to 2019. Ordinary least squares (OLS) linear regression analyses and autoregressive integrated moving average (ARIMA) were used to predict the best value for 2020. A Grubbs 1-sided test was used to assess the significance of the difference between predicted and observed 2020 deaths/mortality. Finally, a 1-sample *t* test was used to compare the population of regional excess deaths to a null mean. The relationship between mortality and predictive variables was assessed using OLS multiple regression models. Since there is no uniform opinion on multicomparison adjustment and false negatives imply great epidemiological risk, the less-conservative Siegel approach and more-conservative Holm-Bonferroni approach were employed. By doing so, we provided the reader with the means to carry out an independent analysis.

Results: Both ARIMA and OLS linear regression models predicted the number of deaths in Italy during 2020 to be between 640,000 and 660,000 (range of 95% CIs: 620,000-695,000) against the observed value of above 750,000. We found strong evidence supporting that the death increase in all regions (average excess=12.2%) was not due to chance (t_{21}=7.2; adjusted *P*<.001). Male and female national mortality excesses were 18.4% (*P*<.001; adjusted *P*=.006) and 14.1% (*P*=.005; adjusted *P*=.12), respectively. However, we found limited significance when comparing male and female mortality residuals' using the Mann-Whitney *U* test (*P*=.27; adjusted *P*=.99). Finally, mortality was strongly and positively correlated with latitude (*R*=0.82; adjusted *P*<.001). In this regard, the significance of the mortality increases during 2020 varied greatly from region to region. Lombardy recorded the highest mortality increase (38% for men, adjusted *P*<.001; 31% for women, *P*<.001; adjusted *P*=.006).

Conclusions: Our findings support the absence of historical endogenous reasons capable of justifying the mortality increase observed in Italy during 2020. Together with the current knowledge on SARS-CoV-2, these results provide decisive evidence on the devastating impact of COVID-19. We suggest that this research be leveraged by government, health, and information authorities to furnish proof against conspiracy hypotheses that minimize COVID-19–related risks. Finally, given the marked concordance between ARIMA and OLS regression, we suggest that these models be exploited for public health surveillance. Specifically, meaningful information can be deduced by comparing predicted and observed epidemiological trends.

(JMIR Public Health Surveill 2022;8(4):e36022) doi:10.2196/36022

KEYWORDS

COVID-19; deniers; excess deaths; epidemiology; infodemic; infodemiology; Italy; longitudinal analysis; mortality; time series; pandemic; public health

Introduction

Background

SARS-CoV-2 is a new beta coronavirus first identified in December 2019 in Wuhan, China. The related pathology, called COVID-19, has raged worldwide, claiming millions of victims and throwing economic and health systems into severe crises. In such a dramatic scenario, Europe is one of the most affected areas: As of December 2021, it accounted for over 30% of global official deaths (ie, approximately 1,600,000) [1]. Because the risk factors are multiple, including environmental conditions, pollution, age, gender, ethnicity, crowding, poverty, and medical comorbidities, mortality varies substantially from country to country as well as intranationally [2-4]. Indeed, the peaks in daily deaths per million inhabitants ranged from 1 (Ukraine) to over 40 (Belgium), with a median of 3.5 (IQR 2-13) [4]. The first European nation to suffer the devastating effects of COVID-19 was Italy, with mortality peaks much higher than the European median (over 15). In particular, the regions of northern Italy—especially the provinces of Bergamo and Brescia—faced a harsh first wave, reaching the highest number of deaths globally [1,5]. To date, despite a substantial reduction in mortality thanks to a massive vaccination campaign, Italy is still the second-ranking European country for official COVID-19 deaths [1,6]. Nonetheless, the debate over COVID-19 mortality has been intense during the pandemic. In the early stages, given the low testing capabilities, the calculation of mortality was subject to numerous uncertainties, which led to both overestimates and underestimates. For this reason, researchers focused their efforts on comparing 2020 data with historical death series [7].

Ordinary least squares (OLS) regression models are among the most adopted model by scientists due to their simplicity and efficacy. Specifically, OLS multiple and simple regressions have often been used to predict the course of COVID-19 cases and deaths, both individually and in conjunction with other epidemiological models such as Susceptible-Infected-Recovered (SIR) [8-10]. The literature shows that linear regression analyses are valuable short-term forecasting tools when the necessary assumptions are satisfied. However, it is not unusual for requirements such as normality of the residuals or homoskedasticity to be violated when dealing with actual epidemiological data. In these cases, the use of corrective procedures or alternative models should be considered. Among the latter, autoregressive integrated moving average (ARIMA) and SARIMA (ARIMA + seasonal component) models have shown excellent predictive capabilities. In particular, a recent study by Abolmaali and Shirzaei [11] demonstrated that the ARIMA approach could outperform other classical models such as logistic function, linear regression, and SIR. Similar findings were obtained by Alabdulrazzaq et al [12], who proved that the accuracy of the prediction of COVID-19 spread provided by their ARIMA model was both appropriate and satisfactory.

Despite more than 135,000 official deaths nationwide, some Italian conspiracy movements argue that COVID-19 is a nondangerous disease and that these numbers have been deliberately exaggerated [13]. Unfortunately, it was not uncommon even for eminent Italian scientists or other prominent personalities to have recklessly downplayed the risks of COVID-19 or favored the spread of fake news [14,15]. Thus, the infodemic question “Dead from COVID or with COVID?” soon filled social networks [13]. Indeed, such a question arises from the hypothesis that COVID-19 was noncausally correlated with the deaths recorded in Italy in 2020.

Objective

Based on this premise, this study aimed to estimate the difference between the observed and predicted numbers of deaths in Italy during 2020. In particular, we modelled all mortality trends by cause of death, sex, and age group from 2011 to 2019, predicting the best values for 2020. By doing so, causal evidence will be provided on the impact of a nonendogenous mortality factor, such as COVID-19. The results of this paper have epidemiological and infodemiological relevance since (1) 2 models widely adopted by the scientific community such as OLS linear regression and ARIMA are compared; (2) to the best of our knowledge, this is the most detailed historical and forecasting survey regarding mortality in Italy; and (3) an estimate of the statistical significance of the increase in mortality in Italy during 2020 is provided. Finally, we investigate 2 essential, but often overlooked, aspects of epidemiological and public health surveillance, namely the possible emergence of nonlinear subtrends (capable of invalidating predictions of models trained on historical global data) and the problem of multicomparison adjustment (capable of dangerously inflating false negatives).

Methods

Data Collection

For this study, we used data from the national agencies and portals of demographic and statistical research, Italy (details and references are provided in the following paragraphs). Specifically, the annual number of deaths (including deaths by sex and age groups), deaths per causes of deaths (including deaths by sex groups), and mortality (including mortality by sex and age groups) were extracted from the platforms and annual reports of the National Institute of Statistics (ISTAT) and National Health Observatory for the years 2011 to 2020 [16-18]. Demographic data (ie, population number per age group, population number and density, and per region) were gathered from Tuttitalia.it [19,20]. This portal contains all ISTAT demographic information relating to municipalities, provinces, and regions. Although the investigated period ranged from 2011 to 2020, causes of death statistics were available until 2017 as the official evaluation process takes 3 years [17]. More details on the data collection process are described in Multimedia Appendix 1.

Procedure and Statistical Analysis Key Points

Here, we provide a summary of the procedure adopted. A more detailed description is reported in Multimedia Appendix 1. We modeled regional trends in annual deaths and mortality from 2011 to 2019 through OLS linear regression. We called Δ^* the residuals’ data set from 2011 to 2019 and Δ the residuals’ data set from 2011 to 2020. Through the Grubbs 1-sided test, we

searched for high outliers in Δ* and Δ. The Grubbs test was performed using RStudio v.4.1.2 software (library: outliers). We also performed a 1-sample *t* test to assess if the regional death increases were due to chance. This was done by comparing the 2020 excess death population to a fixed null mean (ie, the expected residual). Furthermore, we calculated the difference between the model prediction and the observed value. To validate or deny any statistical anomalies in the number of deaths during 2020, we checked all the trends of the following annual statistics within the 2011-2019 time frame: male deaths by age group, female deaths by age group, male mortality by age group, female mortality by age group, deaths by causes of death, male deaths by causes of death, female deaths by causes of death. Specifically, we searched for anomalous nonlinear subtrends capable of distorting the interpretations on the cumulative data (indeed, sum of linear trends is linear). An example of this phenomenon is shown in Figure S1 in Multimedia Appendix 1. Concerning male and female deaths for age groups, we also calculated the 2020 forecast for each age group through an ARIMA (p, d, q) model using RStudio v.4.1.2 software (libraries: forecast and tseries). To facilitate the reproducibility of the analysis, we have provided all the ARIMA models in Multimedia Appendix 2. Finally, we used OLS multiple linear regression to verify any correlations with demographic and geographic statistics such as population, population density, and latitude [21].

Concerning Multicomparison Adjustment Problem

The *P* value adjustment for a multicomparison test originates from the possibility of unintentionally increasing the number of false positives [22]. However, as shown by Greenland [23], the indiscriminate and unthinking implementation of this method can lead to conclusions that are erroneous, misleading, and, when sensitive topics are touched (eg, public health), even dangerous. Indeed, a scientist is called upon to consider both the consequences and the likelihood of incurring false results [23,24]. For instance, some authors suggest it is advisable to adjust the *P* values in exploratory investigations since the chances of spurious correlations due to the look-elsewhere effect are high [24]. Conversely, adjusting *P* values can be counterproductive when hypotheses are well-targeted and false negatives carry a serious risk (eg, airport metal detector). Nonetheless, Bender and Lange [25] highlighted that it is challenging to perform a multiple test adjustment in exploratory analyses due to the possible lack of a clear structure in the multiple tests; ergo, they recommend this procedure only for well-targeted hypotheses. Such a scenario spotlights the absence of a clear consensus [26]. Additional critical issues lie in the fact that the *P* value is not the probability that the test hypothesis is true nor that chance alone produced the observed association [27]. Ergo, adopting an (un)adjusted dichotomous threshold is not suitable for assessing the statistical significance of an outcome, as *P* values should be used—at best—as graded measures of the strength of evidence against the test hypothesis [27,28]. Finally, other authors have raised further concerns about adjusted *P* values. For example, Brandt [29] pointed out the medical unreasonableness of evaluating a patient's test results based on how many tests the patient had that day. With this provocation, Brandt [29] also questioned the scientific community about the possibility of dividing the results into different studies to bypass the problem of multicomparison. In conclusion, Greenland [23] stressed that proposing a single null hypothesis represents a bias in the analysis, and *P* values test not only the degree of data compatibility with the null hypothesis but all the test's assumptions [27]. Hence, it must be admitted that every statistical interpretation or adjustment is strongly influenced by the authors' prejudices and uncertainties on the assumptions made [23,24,27]. This is also true of the so-called "robust analyses," whose complexity is further confusing. For these reasons, a scientist cannot do anything else beyond showing how methods and results vary under different conditions [23].

Our Approach

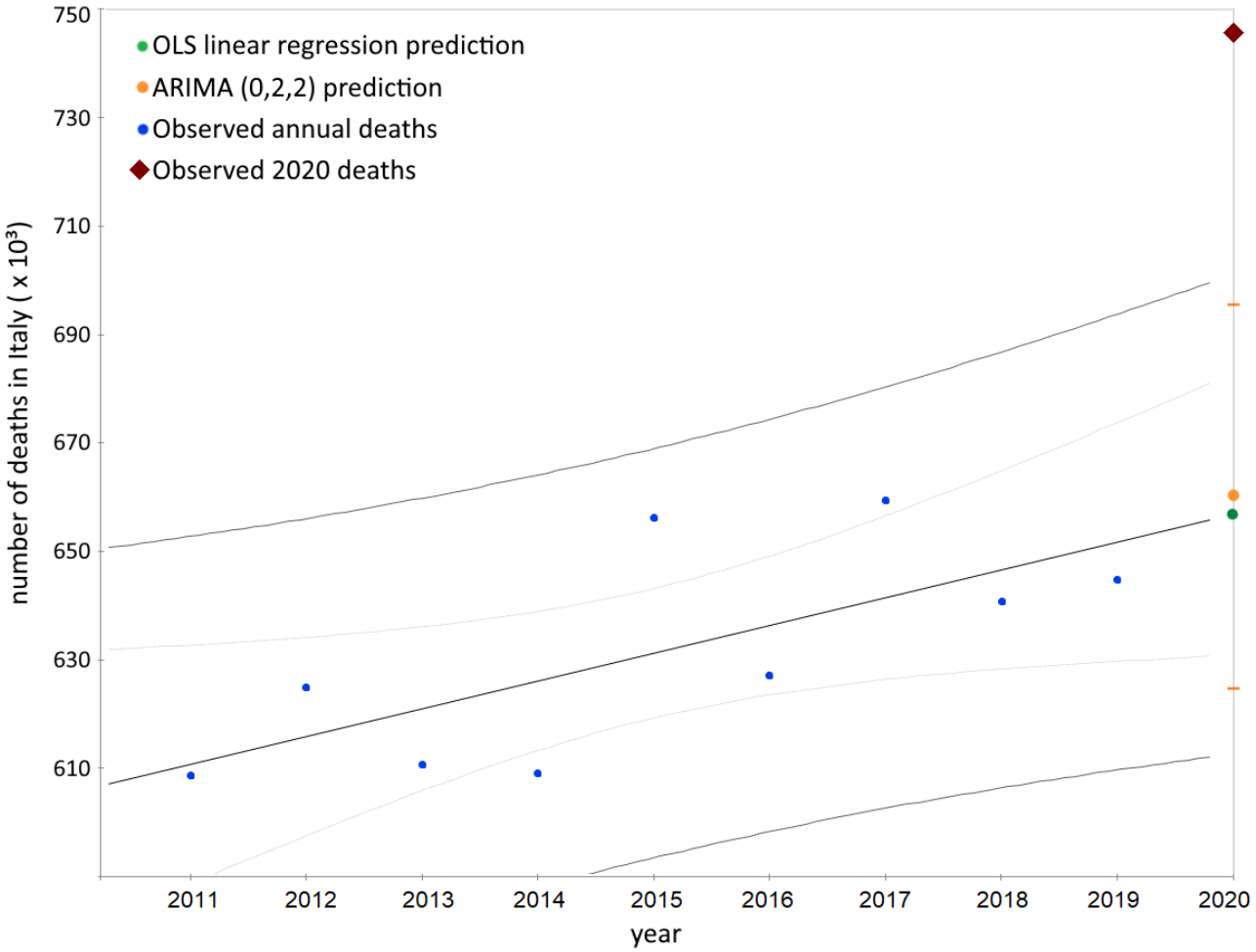
This manuscript aimed to test statistical methods to identify epidemiologically relevant anomalies in a time series and provide near-definitive evidence on COVID-19 impact on mortality in Italy. Based on the evidence summarized in the previous subsection, we concluded that the best option was to give the reader the means to conduct an independent evaluation showing how results changed under different assumptions. Specifically, we used 2 approaches: The first, proposed by Siegel [30], involves the evaluation of the significance of a global test (ie, national population by sex) and then the implementation of other subtests (ie, regional population by sex) without corrections. In particular, we believe this approach is the most suitable for the purpose of this manuscript and denote it with A1. The second approach, denoted with A2, is the more conservative Holm-Bonferroni method with number of hypotheses m=47 [31].

Results

Overall Death Excess During 2020

Compared with the OLS linear regression model prediction (Figure 1), the 2020 excess in the observed number of deaths in Italy was substantially larger (excess=89,287; % excess=13.6 [SE 5.3]). The detailed report is presented in Tables S1 and S2 in Multimedia Appendix 1. We found strong evidence supporting that the death increase in all regions was not due to chance (mean % excess=12.2 [SD 1.7]; t_{21}=7.2; adjusted *P*<.001). Figure 1 also shows the high statistical confidence between the values predicted by the OLS linear regression and ARIMA (0,2,2) model; this constitutes further proof of the goodness of the linear interpolation.

Figure 1. Annual number of deaths in Italy from 2011 to 2020: comparison between the observed value and the 2020 predictions of the ordinary least squares (OLS) linear regression and autoregressive integrated moving average (ARIMA; 0,2,2) models. The narrow bands represent the linear regression 95% CI of the mean value, while the wide bands represent the 95% CI of the observed values from 2011 to 2019. The orange dashes represent the 95% CI of the ARIMA prediction.



Male Mortality Rate During 2020

For A1, when the male mortality rate is considered, the 2020 excesses were large and highly significant in 13 of 21 regions (all $P<.005$). Moderate significant increases were observed in the other five regions ($.02 \leq P \leq .10$). A low significance was obtained only in Molise, Basilicata, and Calabria (all $P \geq .20$). Overall, the excess male mortality in Italy during 2020 was high and markedly significant ($P<.001$; excess=18.8 per 10,000; % excess=18.4 [SE 5.4]). Moreover, all regions recorded an excess male mortality between 5% (Basilicata) and 38% (Lombardy). Details of each region are provided in Table 1. Further information on the model goodness is provided in Table S3 in Multimedia Appendix 1. For A2, adjusted $P \leq .006$ was reached nationally and in 6 regions (Piedmont, Lombardy, Trento, Veneto, Liguria, Emilia Romagna), while $.02 \leq$adjusted $P \leq .08$ were reached in 5 regions (Bolzano, Friuli Venezia Giulia, Marche, Abruzzo, Apulia). Campania and Sardinia also registered a moderate significance (adjusted $P \leq .16$). Adjusted $P \geq .43$ were obtained in the remaining regions. Details of each region are provided in Table 1.

Table 1. Regional male mortality statistics: comparison of ordinary least squares (OLS) linear regression predicted mortality (predicted value) and observed mortality of 2020 (observed value) in Italy. The data are normalized to 10,000 people (deaths per 10,000).

Italian region	Predicted value	Predicted value SE	Observed value	Excess, % (SE)	*P* value[a]	Adjusted *P* value[a]
Italy	102.1	4.6	120.9	18.4 (5.4)	<.001	.006
Piemonte	105.7	5.2	132.3	25.1 (6.2)	<.001	<.001
Valle d'Aosta	114.4	11.3	136.3	19.1 (12.3)	.02	.40
Lombardia	98.5	4.1	136.2	38.3 (5.8)	<.001	<.001
Bolzano	93.6	4.9	110	17.5 (6.2)	.001	.02
Trento	90.7	4.7	121	33.4 (7)	<.001	<.001
Veneto	97.2	3.7	114.7	18 (4.5)	<.001	.002
Friuli Venezia Giulia	99.6	5.9	116.3	16.8 (7)	.002	.06
Liguria	103.4	5.6	126.5	22.3 (6.7)	<.001	.006
Emilia-Romagna	97.5	4.5	116.1	19.1 (5.6)	<.001	.006
Toscana	97.2	6	108.5	11.6 (7)	.03	.43
Umbria	94.5	6.3	105.4	11.5 (7.6)	.04	.62
Marche	96.3	5.4	111.1	15.3 (6.6)	.003	.07
Lazio	100.1	5.2	110.1	10 (5.7)	.02	.42
Abruzzo	101.9	4.5	114.6	12.5 (5)	.002	.06
Molise	107.3	7.2	113.8	6 (7.2)	.50	.99
Campania	116.1	5.5	129.9	11.9 (5.4)	.005	.12
Puglia	100.1	5.8	115.6	15.5 (6.7)	.003	.08
Basilicata	107.2	6.1	112.9	5.3 (6)	.40	.99
Calabria	106.5	6.1	113.9	6.9 (6.2)	.20	.99
Sicilia	112.7	7.2	122.9	9.1 (7)	.10	.99
Sardegna	101.2	5.3	113.7	12.3 (5.9)	.007	.16

[a]Grubbs test.

Female Mortality Rate During 2020

For A1, highly significant excess female mortality was found in the northern regions and Sardinia ($P \leq .01$, except Valle d'Aosta, $P=.02$). Moderately significant excesses were recorded in Tuscany, Marche, Molise, and Apulia ($.04 \leq P \leq .07$). Scarcely significant differences were recorded in the rest of Italy (all $P>.40$). Nevertheless, all regions experienced an excess female mortality between 4% (Basilicata) and 31% (Lombardy). Details of each region are provided in Table 2. Further information on the model goodness is provided in Table S4 in Multimedia Appendix 1. For A2, Lombardy (adjusted $P=.006$) and Trento (adjusted $P=.001$) reached the greatest statistical significance. Moderate significance ($.01 \leq$adjusted $P \leq .06$) was reached in 5 regions (Piedmont, Bolzano, Friuli Venezia Giulia, Liguria, and Emilia Romagna). Sardinia (adjusted $P=.19$) and Veneto (adjusted $P=.20$) also registered a modest significance. Low significance was observed in the remaining regions (all adjusted $P \geq .38$). Details of each region are provided in Table 2.

Table 2. Regional female mortality statistics: comparison of ordinary least squares (OLS) linear regression predicted mortality (predicted value) and observed mortality of 2020 (observed value) in Italy. The data are normalized to 10,000 people (deaths per 10,000).

Italian region	Predicted value	Predicted value SE	Observed value	Excess, % (SE)	*P* value[a]	Adjusted *P* value[a]
Italy	68.3	3.9	77.9	14.1 (6.6)	.005	.12
Piemonte	70.8	4	84.1	18.8 (6.8)	.001	.02
Valle d'Aosta	69.9	9.8	88.9	27.1 (19.3)	.02	.38
Lombardia	64.3	3.5	84.2	30.9 (7.2)	<.001	.006
Bolzano	60.5	3.6	73.9	22.1 (7.4)	<.001	.01
Trento	59.4	2.8	73.4	23.6 (5.9)	<.001	.001
Veneto	64.2	3.8	72.8	13.4 (6.9)	.009	.20
Friuli Venezia Giulia	64.2	2.8	72.6	13 (5)	.001	.05
Liguria	67.4	4.3	79.3	17.7 (7.6)	.002	.06
Emilia-Romagna	66.1	3.3	75.6	14.4 (5.7)	.002	.05
Toscana	65.4	3.7	71.2	8.9 (6.3)	.07	.80
Umbria	63.2	4	67	6.1 (6.8)	.43	.99
Marche	64	4.7	71.8	12.2 (8.5)	.05	.68
Lazio	67.8	4.7	71.9	6 (7.5)	.56	.99
Abruzzo	67.6	4.6	72	6.5 (7.4)	.43	.99
Molise	66.4	5.1	74.7	12.6 (8.9)	.06	.73
Campania	80.3	5.7	85.1	6 (7.7)	N/A[b]	N/A
Puglia	68.5	4.7	76.5	11.6 (7.8)	.04	.65
Basilicata	71.5	4.7	74.4	4.1 (7)	.40	.99
Calabria	71.7	4.4	75.1	4.8 (6.5)	.79	.99
Sicilia	78	5.9	83.4	7 (8.3)	.47	.99
Sardegna	64	3.4	71.5	11.6 (5.9)	.008	.19

[a]Grubbs test.

[b]N/A: not available.

Relationship Between Deaths and Geographical-Demographic Statistics

The linear multiregression model among the log-transformed statistics, the regional number of inhabitants (X_1), regional population density (X_2), regional latitude (X_3), and 2020 regional excess deaths (Y), returned the following equation:

$$Y=f(X_3)=k\times pow(X_3, a),$$

with $k=2.6\times10^{-7}$, a=9.9, R=0.82, adjusted P<.001.

Retrospective Analysis of Deaths

Figure 2 shows the number of deaths per cause of death from 2012 to 2017 in Italy (2018 and 2019 data were not available, as shown in Multimedia Appendix 1). Tumors and diseases of the circulatory system always accounted for over 60% of total deaths (also considering the projections for 2020). The percentages of male (female) deaths for tumors ranged from 55.6% (44.4%) to 56.3% (43.7%), while deaths related to the circulatory system were 43.1% (56.9%) to 43.7% (56.3%). All trends were markedly linear.

Finally, Figure 3 and Figure 4 show male and female deaths, respectively, by age group from 2011 to 2019. Explicitly calculating each trend for each age and sex group and summing the predictions for 2020, we obtained the best value of 648,733 deaths. All trends were markedly linear (Figures S3 and S4 in Multimedia Appendix 1). Summing up all the forecasts of the ARIMA models for each age and sex group, we obtained a total of 637,534 deaths. A similar result was obtained by summing the global trends for men and women (640,508 deaths).

Figure 2. Number of deaths per cause of death from 2011 to 2017 in Italy; the most updated National Institute of Statistics (ISTAT) data were available until 2017 (see Multimedia Appendix 1). 1: infectious and parasitic diseases; 2: tumors; 3: psychic disorders, diseases of the nervous system and organs of the senses; 4: diseases of the circulatory system; 5: diseases of the respiratory system; 6: diseases of the digestive system; 7: other morbid states; 8: poorly defined symptoms, signs, and morbid states; 9: external causes of trauma and poisoning.

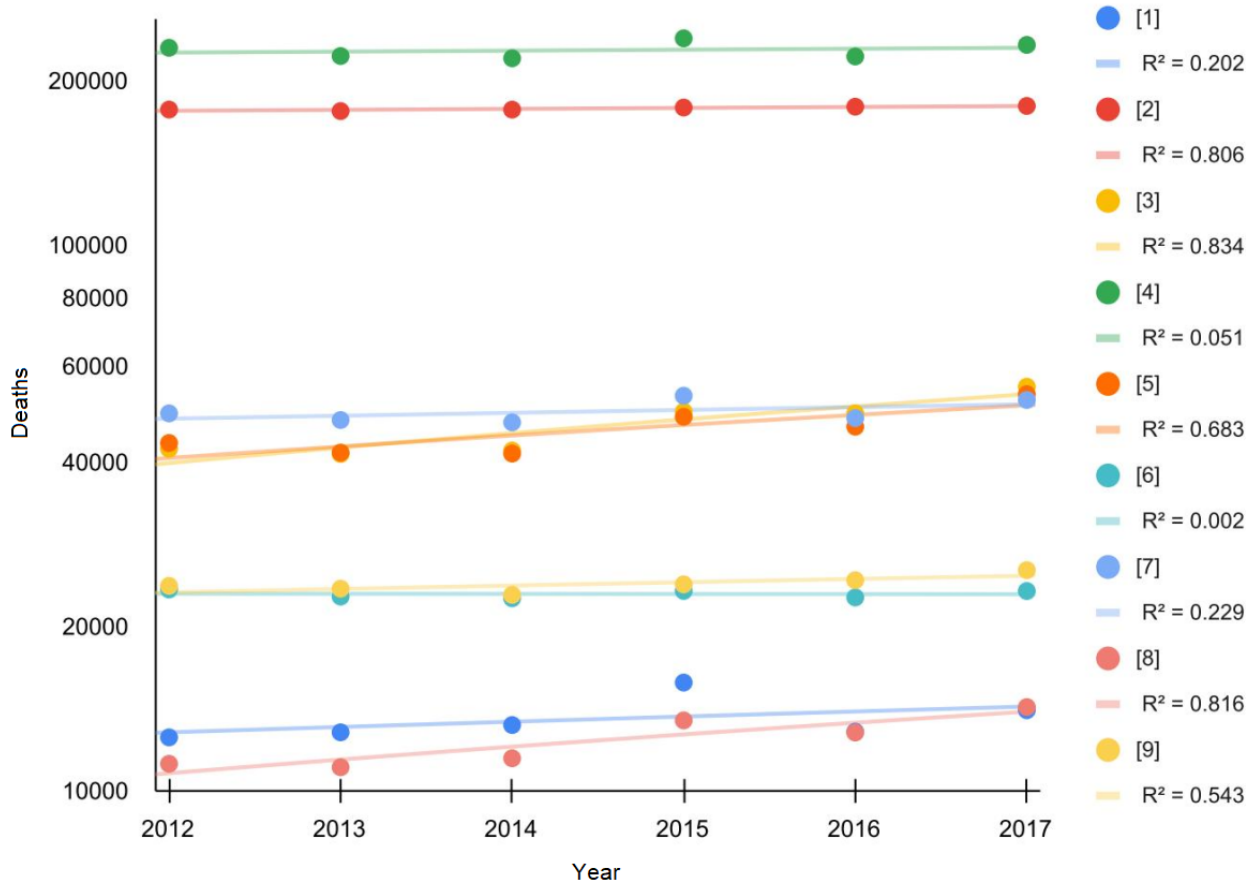


Figure 3. Male deaths per age group in Italy from 2011 to 2019 and autoregressive integrated moving average (ARIMA) predictions for 2020.

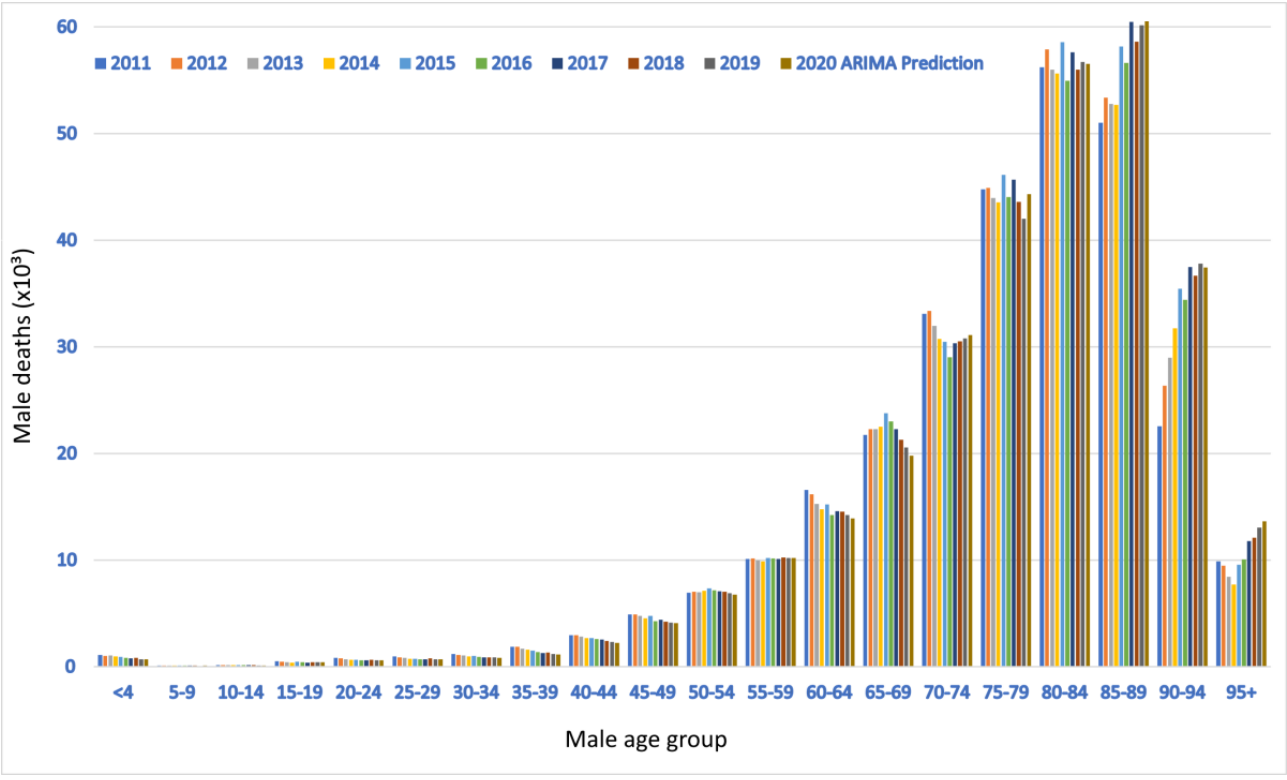
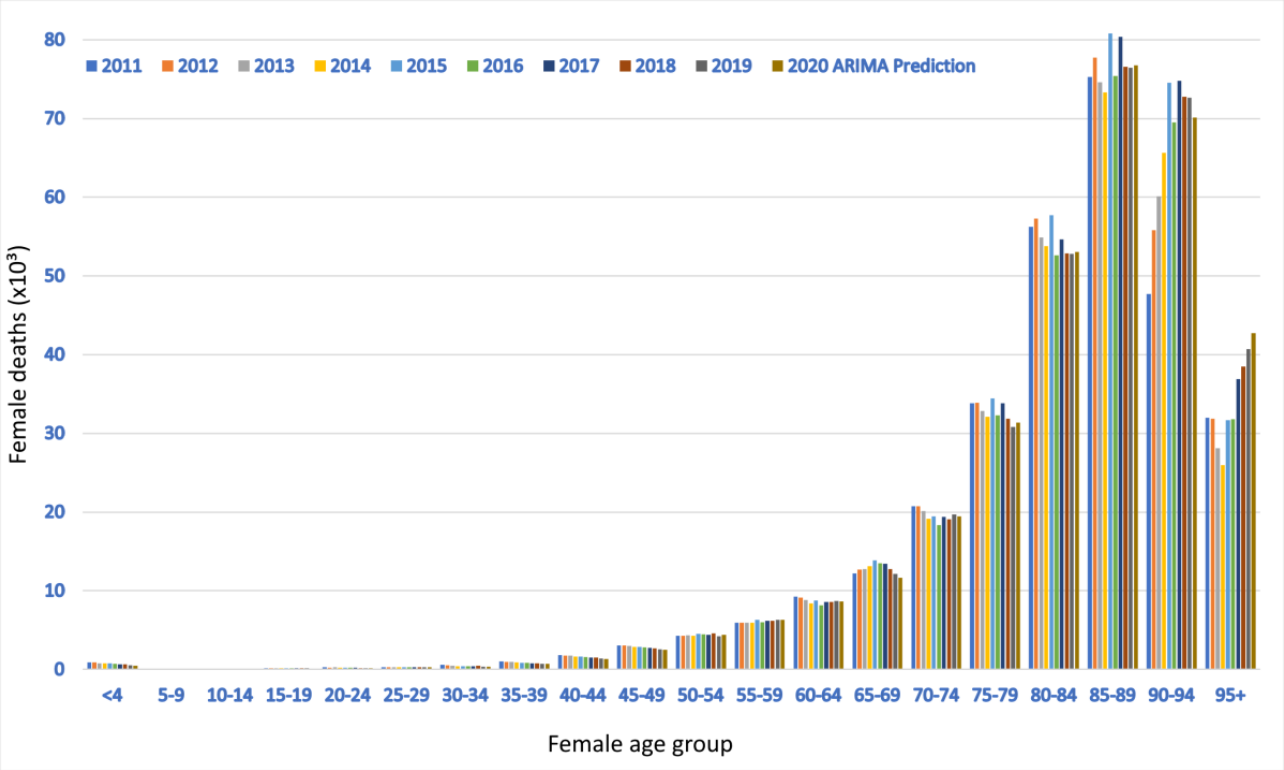


Figure 4. Female deaths per age group in Italy from 2011 to 2019 and autoregressive integrated moving average (ARIMA) predictions for 2020.



Comparison Between Male and Female Mortality

The increases in mortality were 18% (P<.001; adjusted P=.006) for men and 14% (P=.005; adjusted P=.12) for women at the national level and 16% for men compared with 13% for women, on average, at the regional level. However, we found limited significance when comparing the residuals' populations using the Mann-Whitney U test (P=.27; adjusted P=.99).

Discussion

Principal Findings

This paper provides strong evidence in favor of an anomalous mortality event during 2020 in Italy, which was not predictable based on endogenous causes such as deaths and mortality trends between 2011 and 2019. Notably, the number of total deaths observed in 2020 exceeded the linear regression model prediction by more than 89,000 (a value nearly 3 times greater than the prediction standard error) and the ARIMA prediction by more than 86,000. Grubbs and t tests confirmed that this figure was unexpected. At the national level, the increase in mortality was 18% for the male population and 14% for the female population. Nonetheless, the statistical significance of this difference was low. The total excess mortality was positively correlated with latitude, which explained the data set variability much better than demographic statistics like population number and density. All the “deceases due to causes of death” trends from 2012 to 2017 were appreciably linear or stationary; this precludes the existence of anomalous subtrends linked to the causes of death. Moreover, summing up all the 2020 death predictions by age group, we obtained a value ranging from 640,000 to 660,000 deceased, significantly far from the observed one (750,000). In conclusion, these findings confirm the absence of any confounding inner subtrends capable of explaining the excess deaths during 2020 in Italy.

Comparison With Prior Work

To the best of our knowledge, the most comprehensive and detailed study examining excess mortality during 2020 in Italy was the report redacted by the ISTAT and National Institute of Health (ISS) [32]. Their research focuses on comparing the March-December 2015-2019 and 2020 periods, starting from the assumption that COVID-19 is the cause of the discrepancies observed. On the contrary, our analysis has been more impartial since we have not introduced any hypothesis about the reasons that caused this phenomenon. Therefore, our findings provide evidence of statistical and epidemiological significance that had not been considered before. Specifically, excluding internal causes gives further strength to the theories that identify COVID-19 as the principal cause of such a tragic scenario. COVID-19 dangerousness is confirmed at the molecular-genetic level [33-36]. The strong positive correlation we found between excess mortality and latitude is compatible with greater virulence and mortality of COVID-19 in northern Italy depicted by other literature [37,38]. In this regard, an increasing number of mathematical-statistical investigations classify COVID-19 as a seasonal low-temperature infection [39-41], although the effect size of the environmental factors is still debated [42]. However, it is a fact that low temperatures can have indirect effects on the spread of infections, like the creation of indoor gatherings—with insufficient air circulation—and the weakening of the immune defenses [43,44]. Since average temperatures in northern regions are lower than the rest of the peninsula [45], this phenomenon could partially explain the Italian epidemiological scenario. A large amount of literature has also identified pollution as a relevant COVID-19 risk factor. For

instance, NO_2, PM10, and PM2.5 were causally connected with more serious situations, as they can drastically reduce the immune response and compromise respiratory functions [45-48]. This type of pollutant is widespread in the Po Valley [45,46]. Contrary to other literature, our paper did not detect a high significance in the difference between male and female national mortalities [49-58]. Nonetheless, this result is not conclusive and deserves further investigation as such a discrepancy could be more evident by considering the most affected and exposed age groups. Moreover, the COVID-19 course is influenced by numerous comorbidities, such as cancer, chronic kidney diseases, diabetes mellitus, hypertension, chronic obstructive pulmonary diseases, asthma, chronic respiratory diseases, immunocompromised state, HIV infection, heart conditions, overweight and obesity, dementia or other neurological conditions, and mental health conditions [59-61]. The majority of these pathologies are more common in older age groups, which helps explain the greater aggressiveness of the infection in some regions [17,48]. Hence, it is necessary to consider that the prepandemic epidemiological scenario has contributed to enhancing the disease damage in Italy. Nevertheless, it would be incorrect to consider only the older population as vulnerable: Phenomena such as long COVID (ie, the onset of medical complications that last weeks to months after initial recovery) are increasing in younger age groups, including children and adolescents [62,63]. The most common symptoms of long COVID are fatigue, weakness, cough, chest tightness, breathlessness, palpitations, myalgia, and difficulty focusing; their appearance is not related to the severity of the COVID-19 course [63,64]. Moreover, new variants of concern—favored by the uncontrolled spread of the virus—continuously pose new threats to all age groups [33,65,66]. In this regard, strategies such as vaccinations and nonpharmaceutical containment measures have been and continue to be fundamental to control COVID-19 diffusion, avoid hospital overcrowding, and slow down the epidemiological peaks [67-73]. Indeed, although this paper has provided evidence in favor of a high number of deaths due to COVID-19 in Italy during 2020 (before the administration of COVID-19 vaccines), lockdowns, social distancing, and masks have prevented the death toll from being numerous times higher [74-77].

Limitations

Our approach has limitations to be considered. Since statistical significance is a measure of data compatibility with the null hypothesis (including the model's assumptions), the evidence provided in this paper could vary under different initial hypotheses. However, the degree of uncertainty was reduced by targeting the tested hypotheses well. Furthermore, causal relationships have not been directly investigated. Therefore, these findings must be contextualized in light of the results of other literature. Finally, the discrepancies between the model predictions and the observed data were not weighted on the clinical characteristics of the patients.

Conclusions

This paper provides strong evidence on the absence of historical endogenous reasons capable of explaining the anomalous mortality increase recorded in Italy during 2020. Weighing these statistical results on the numerous molecular-genetic, medical, biological, virological, and epidemiological-based publications that confirmed high COVID-19 virulence, we conclude that the pandemic impact on excess deaths in Italy constitutes a scientific fact. This answers the question "Died from COVID or died with COVID?" Specifically, this manuscript can be adopted by health authorities and disclosure agencies to discredit fake news that minimizes the COVID-19 risk. Moreover, given the marked concordance between ARIMA and OLS regression models, we suggest that these methods be exploited for public health surveillance aims. In particular, considering their efficiency and effectiveness, it is possible to derive meaningful information regarding current and future epidemiological situations from the comparison between the predicted and observed trends.

Conflicts of Interest

None declared.

Multimedia Appendix 1
COVID-19 Mortality Italy, Rovetta.
[DOCX File , 814 KB - publichealth_v8i4e36022_app1.docx]

Multimedia Appendix 2
Autoregressive integrated moving average (ARIMA) models.
[ZIP File (Zip Archive), 1717 KB - publichealth_v8i4e36022_app2.zip]

References

1. Coronavirus (COVID-19) Dashboard. World Health Organization. URL: https://covid19.who.int/ [accessed 2021-12-11]
2. Assessing Risk Factors for Severe COVID-19 Illness. Centers for Disease Control and Prevention. 2020 Nov 30. URL: https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/assessing-risk-factors.html [accessed 2021-12-11]
3. Bourdrel T, Annesi-Maesano I, Alahmad B, Maesano CN, Bind M. The impact of outdoor air pollution on COVID-19: a review of evidence from , animal, and human studies. Eur Respir Rev 2021 Mar 31;30(159):1 [FREE Full text] [doi: 10.1183/16000617.0242-2020] [Medline: 33568525]

4. Jabłońska K, Aballéa S, Toumi M. Factors influencing the COVID-19 daily deaths' peak across European countries. Public Health 2021 May;194:135-142 [FREE Full text] [doi: 10.1016/j.puhe.2021.02.037] [Medline: 33892351]
5. Bontempi E. First data analysis about possible COVID-19 virus airborne diffusion due to air particulate matter (PM): The case of Lombardy (Italy). Environ Res 2020 Jul;186:109639 [FREE Full text] [doi: 10.1016/j.envres.2020.109639] [Medline: 32668559]
6. Epidemia COVID-19: Aggiornamento nazionale 1 dicembre 2021 - ore 12. Istituto Superiore di Sanità. 2021 Dec 3. URL: https://www.epicentro.iss.it/coronavirus/bollettino/Bollettino-sorveglianza-integrata-COVID-19_1-dicembre-2021.pdf [accessed 2021-12-11]
7. Meyerowitz-Katz G, Merone L. A systematic review and meta-analysis of published research data on COVID-19 infection fatality rates. Int J Infect Dis 2020 Dec;101:138-148 [FREE Full text] [doi: 10.1016/j.ijid.2020.09.1464] [Medline: 33007452]
8. Rath S, Tripathy A, Tripathy AR. Prediction of new active cases of coronavirus disease (COVID-19) pandemic using multiple linear regression model. Diabetes Metab Syndr 2020;14(5):1467-1474 [FREE Full text] [doi: 10.1016/j.dsx.2020.07.045] [Medline: 32771920]
9. Peng Z, Ao S, Liu L, Bao S, Hu T, Wu H, et al. Estimating unreported COVID-19 cases with a time-varying SIR regression model. Int J Environ Res Public Health 2021 Jan 26;18(3):1 [FREE Full text] [doi: 10.3390/ijerph18031090] [Medline: 33530563]
10. Melik-Huseynov DV, Karyakin NN, Blagonravova AS, Klimko VI, Bavrina AP, Drugova OV, et al. Regression models predicting the number of deaths from the new coronavirus infection. Sovrem Tekhnologii Med 2020;12(2):6-11 [FREE Full text] [doi: 10.17691/stm2020.12.2.01] [Medline: 34513048]
11. Abolmaali S, Shirzaei S. A comparative study of SIR model, linear regression, logistic function and ARIMA model for forecasting COVID-19 cases. AIMS Public Health 2021;8(4):598-613 [FREE Full text] [doi: 10.3934/publichealth.2021048] [Medline: 34786422]
12. Alabdulrazzaq H, Alenezi MN, Rawajfih Y, Alghannam BA, Al-Hassan AA, Al-Anzi FS. On the accuracy of ARIMA based prediction of COVID-19 spread. Results Phys 2021 Aug;27:104509 [FREE Full text] [doi: 10.1016/j.rinp.2021.104509] [Medline: 34307005]
13. De Vogli R. Morti con o per il coronavirus? L'apice degli errori sul tema è stato raggiunto in questi giorni. Il Fatto Quotidiano. 2021 Nov 11. URL: https://www.ilfattoquotidiano.it/2021/10/25/morti-con-o-per-il-coronavirus-lapice-degli-errori-sul-tema-e-stato-raggiunto-in-questi-giorni/6366640/ [accessed 2021-11-11]
14. Rovetta A, Castaldo L. Influence of mass media on Italian web users during the COVID-19 pandemic: infodemiological analysis. JMIRx Med 2021 Oct 18;2(4):e32233 [FREE Full text] [doi: 10.2196/32233] [Medline: 34842858]
15. Rovetta A. The impact of COVID-19 on conspiracy hypotheses and risk perception in Italy: infodemiological survey study using Google Trends. JMIR Infodemiology 2021;1(1):e29929 [FREE Full text] [doi: 10.2196/29929] [Medline: 34447925]
16. Tavole di mortalità. Istituto nazionale di statistica. URL: http://dati.istat.it/Index.aspx?DataSetCode=DCIS_MORTALITA1 [accessed 2021-11-01]
17. C04: Sanità e salute. Istituto nazionale di statistica. URL: https://www.istat.it/it/files//2020/12/C04.pdf [accessed 2021-11-01]
18. COVID-19. Osservatorio sulla salute. URL: https://www.osservatoriosullasalute.it/wp-content/uploads/2021/05/ro-2020-isc-covid.xlsx [accessed 2021-11-01]
19. Popolazione per età, sesso e stato civile 2019. Tutti Italia. URL: https://www.tuttitalia.it/statistiche/popolazione-eta-sesso-stato-civile-2019/ [accessed 2021-11-01]
20. Regioni italiane per densità. Tutti Italia. URL: https://www.tuttitalia.it/regioni/densita/ [accessed 2021-11-01]
21. Multiple Linear Regression Calculator. Statistics Kingdom. URL: https://www.statskingdom.com/410multi_linear_regression.html [accessed 2021-11-03]
22. Jafari M, Ansari-Pour N. Why, when and how to adjust your P values? Cell J 2019 Jan;20(4):604-607 [FREE Full text] [doi: 10.22074/cellj.2019.5992] [Medline: 30124010]
23. Greenland S. Invited commentary: the need for cognitive science in methodology. Am J Epidemiol 2017 Sep 15;186(6):639-645. [doi: 10.1093/aje/kwx259] [Medline: 28938712]
24. Greenland S, Hofman A. Multiple comparisons controversies are about context and costs, not frequentism versus Bayesianism. Eur J Epidemiol 2019 Sep;34(9):801-808 [FREE Full text] [doi: 10.1007/s10654-019-00552-z] [Medline: 31522327]
25. Bender R, Lange S. Multiple test procedures other than Bonferroni's deserve wider use. BMJ 1999 Feb 27;318(7183):600-601 [FREE Full text] [doi: 10.1136/bmj.318.7183.600a] [Medline: 10037651]
26. Reito A. Problem of multiplicity in clinical studies and inferences made when it is present: letter to the editor. Am J Sports Med 2020 Jan;48(1):NP13. [doi: 10.1177/0363546519887113] [Medline: 31877094]
27. Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole C, Goodman SN, et al. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. Eur J Epidemiol 2016 Apr 21;31(4):337-350 [FREE Full text] [doi: 10.1007/s10654-016-0149-3] [Medline: 27209009]
28. Amrhein V, Korner-Nievergelt F, Roth T. The earth is flat (> 0.05): significance thresholds and the crisis of unreplicable research. PeerJ 2017;5:e3544 [FREE Full text] [doi: 10.7717/peerj.3544] [Medline: 28698825]
29. Brandt J. 2005 INS Presidential Address: neuropsychological crimes and misdemeanors. Clin Neuropsychol 2007 Jul;21(4):553-568. [doi: 10.1080/13854040600770800] [Medline: 17613978]

30. Siegel AF. Multiple t tests: some practical considerations. TESOL Quarterly 1990;24(4):773. [doi: 10.2307/3587136]
31. Holm S. A simple sequential rejective multiple test procedure. Scand J Statist 1979;6(2):65-70 [FREE Full text] [doi: 10.2307/4615733]
32. Impact Of Covid-19 Epidemic On Total Mortality Of Resident Population Year 2020. Istituto nazionale di statistica. 2021 Mar 05. URL: https://www.istat.it/it/files//2021/03/Report-_ISS_Istat__5-marzo-2021_en.pdf [accessed 2021-11-10]
33. Scudellari M. How the coronavirus infects cells - and why Delta is so dangerous. Nature 2021 Jul;595(7869):640-644. [doi: 10.1038/d41586-021-02039-y] [Medline: 34321669]
34. Nguyen HL, Lan PD, Thai NQ, Nissley DA, O'Brien EP, Li MS. Does SARS-CoV-2 bind to human ACE2 more strongly than does SARS-CoV? J Phys Chem B 2020 Aug 27;124(34):7336-7347 [FREE Full text] [doi: 10.1021/acs.jpcb.0c04511] [Medline: 32790406]
35. Shang J, Ye G, Shi K, Wan Y, Luo C, Aihara H, et al. Structural basis of receptor recognition by SARS-CoV-2. Nature 2020 May;581(7807):221-224 [FREE Full text] [doi: 10.1038/s41586-020-2179-y] [Medline: 32225175]
36. Chen F, Zhang Y, Li X, Li W, Liu X, Xue X. The impact of ACE2 polymorphisms on COVID-19 disease: susceptibility, severity, and therapy. Front Cell Infect Microbiol 2021;11:753721 [FREE Full text] [doi: 10.3389/fcimb.2021.753721] [Medline: 34746028]
37. Cegolon L, Pichierri J, Mastrangelo G, Cinquetti S, Sotgiu G, Bellizzi S, et al. Hypothesis to explain the severe form of COVID-19 in Northern Italy. BMJ Glob Health 2020 Jun;5(6):1 [FREE Full text] [doi: 10.1136/bmjgh-2020-002564] [Medline: 32532758]
38. Fortunato F, Martinelli D, Lo Caputo S, Santantonio T, Dattoli V, Lopalco PL, et al. Sex and gender differences in COVID-19: an Italian local register-based study. BMJ Open 2021 Oct 07;11(10):e051506 [FREE Full text] [doi: 10.1136/bmjopen-2021-051506] [Medline: 34620662]
39. Fontal A, Bouma MJ, San-José A, López L, Pascual M, Rodó X. Climatic signatures in the different COVID-19 pandemic waves across both hemispheres. Nat Comput Sci 2021 Oct 21;1(10):655-665. [doi: 10.1038/s43588-021-00136-6]
40. Christophi CA, Sotos-Prieto M, Lan F, Delgado-Velandia M, Efthymiou V, Gaviola GC, et al. Ambient temperature and subsequent COVID-19 mortality in the OECD countries and individual United States. Sci Rep 2021 Apr 22;11(1):8710 [FREE Full text] [doi: 10.1038/s41598-021-87803-w] [Medline: 33888744]
41. Chen S, Prettner K, Kuhn M, Geldsetzer P, Wang C, Bärnighausen T, et al. Climate and the spread of COVID-19. Sci Rep 2021 Apr 27;11(1):9042 [FREE Full text] [doi: 10.1038/s41598-021-87692-z] [Medline: 33907202]
42. Sera F, Armstrong B, Abbott S, Meakin S, O'Reilly K, von Borries R, MCC Collaborative Research Network, CMMID COVID-19 Working Group, et al. A cross-sectional analysis of meteorological factors and SARS-CoV-2 transmission in 409 cities across 26 countries. Nat Commun 2021 Oct 13;12(1):5968 [FREE Full text] [doi: 10.1038/s41467-021-25914-8] [Medline: 34645794]
43. Fares A. Factors influencing the seasonal patterns of infectious diseases. Int J Prev Med 2013 Feb;4(2):128-132 [FREE Full text] [Medline: 23543865]
44. Mourtzoukou EG, Falagas ME. Exposure to cold and respiratory tract infections. Int J Tuberc Lung Dis 2007 Sep;11(9):938-943. [Medline: 17705968]
45. Rovetta A, Castaldo L. Relationships between demographic, geographic, and environmental statistics and the spread of novel coronavirus disease (COVID-19) in Italy. Cureus 2020 Nov 09;12(11):e11397 [FREE Full text] [doi: 10.7759/cureus.11397] [Medline: 33312795]
46. Pegoraro V, Heiman F, Levante A, Urbinati D, Peduto I. An Italian individual-level data study investigating on the association between air pollution exposure and Covid-19 severity in primary-care setting. BMC Public Health 2021 May 12;21(1):902 [FREE Full text] [doi: 10.1186/s12889-021-10949-9] [Medline: 33980180]
47. Kasioumi M, Stengos T. The effect of pollution on the spread of COVID-19 in Europe. Econ Disaster Clim Chang 2021 Oct 22:1-12 [FREE Full text] [doi: 10.1007/s41885-021-00099-y] [Medline: 34703977]
48. Pluchino A, Biondo AE, Giuffrida N, Inturri G, Latora V, Le Moli R, et al. A novel methodology for epidemic risk assessment of COVID-19 outbreak. Sci Rep 2021 Mar 05;11(1):5304 [FREE Full text] [doi: 10.1038/s41598-021-82310-4] [Medline: 33674627]
49. Raimondi F, Novelli L, Ghirardi A, Russo FM, Pellegrini D, Biza R, HPG23 Covid-19 Study Group. Covid-19 and gender: lower rate but same mortality of severe disease in women-an observational study. BMC Pulm Med 2021 Mar 20;21(1):96 [FREE Full text] [doi: 10.1186/s12890-021-01455-0] [Medline: 33743654]
50. Galasso V, Pons V, Profeta P, Becher M, Brouard S, Foucault M. Gender differences in COVID-19 attitudes and behavior: Panel evidence from eight countries. Proc Natl Acad Sci U S A 2020 Nov 03;117(44):27285-27291 [FREE Full text] [doi: 10.1073/pnas.2012520117] [Medline: 33060298]
51. L'uso E L'abuso Di Alcol In Italia. Istituto nazionale di statistica. 2015 Apr 16. URL: https://www.istat.it/it/files//2015/04/statistica_report_alcol_2014.pdf [accessed 2022-04-02]
52. Aspetti della vita quotidiana: Abitudine al fumo - età, titolo di studio. Istituto nazionale di statistica. URL: http://dati.istat.it/Index.aspx?QueryId=15513 [accessed 2022-04-02]
53. Bwire GM. Coronavirus: why men are more vulnerable to Covid-19 than women? SN Compr Clin Med 2020 Jun 4;2(7):874-876 [FREE Full text] [doi: 10.1007/s42399-020-00341-w] [Medline: 32838138]

54. Dai M, Tao L, Chen Z, Tian Z, Guo X, Allen-Gipson DS, et al. Influence of cigarettes and alcohol on the severity and death of COVID-19: a multicenter retrospective study in Wuhan, China. Front Physiol 2020;11:588553 [FREE Full text] [doi: 10.3389/fphys.2020.588553] [Medline: 33362576]
55. Zhang H, Ma S, Han T, Qu G, Cheng C, Uy JP, et al. Association of smoking history with severe and critical outcomes in COVID-19 patients: A systemic review and meta-analysis. Eur J Integr Med 2021 Apr;43:101313 [FREE Full text] [doi: 10.1016/j.eujim.2021.101313] [Medline: 33619437]
56. Patanavanich R, Glantz SA. Smoking is associated with worse outcomes of COVID-19 particularly among younger adults: a systematic review and meta-analysis. BMC Public Health 2021 Aug 16;21(1):1554 [FREE Full text] [doi: 10.1186/s12889-021-11579-x] [Medline: 34399729]
57. Wehbe Z, Hammoud SH, Yassine HM, Fardoun M, El-Yazbi AF, Eid AH. Molecular and biological mechanisms underlying gender differences in COVID-19 severity and mortality. Front Immunol 2021;12:659339 [FREE Full text] [doi: 10.3389/fimmu.2021.659339] [Medline: 34025658]
58. Hachim IY, Hachim MY, Talaat IM, López-Ozuna VM, Saheb Sharif-Askari N, Al Heialy S, et al. The molecular basis of gender variations in mortality rates associated with the novel coronavirus (COVID-19) outbreak. Front Mol Biosci 2021;8:728409 [FREE Full text] [doi: 10.3389/fmolb.2021.728409] [Medline: 34604307]
59. Ng WH, Tipih T, Makoah NA, Vermeulen J, Goedhals D, Sempa JB, et al. Comorbidities in SARS-CoV-2 patients: a systematic review and meta-analysis. mBio 2021 Feb 09;12(1):1 [FREE Full text] [doi: 10.1128/mBio.03647-20] [Medline: 33563817]
60. Gülsen A, König IR, Jappe U, Drömann D. Effect of comorbid pulmonary disease on the severity of COVID-19: A systematic review and meta-analysis. Respirology 2021 Jun;26(6):552-565 [FREE Full text] [doi: 10.1111/resp.14049] [Medline: 33955623]
61. COVID-19: People with Certain Medical Conditions. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html [accessed 2021-11-11]
62. Lopez-Leon S, Wegman-Ostrosky T, Perelman C, Sepulveda R, Rebolledo PA, Cuapio A, et al. More than 50 long-term effects of COVID-19: a systematic review and meta-analysis. Sci Rep 2021 Aug 09;11(1):16144 [FREE Full text] [doi: 10.1038/s41598-021-95565-8] [Medline: 34373540]
63. Asadi-Pooya AA, Nemati H, Shahisavandi M, Akbari A, Emami A, Lotfi M, et al. Long COVID in children and adolescents. World J Pediatr 2021 Oct;17(5):495-499 [FREE Full text] [doi: 10.1007/s12519-021-00457-6] [Medline: 34478045]
64. Raveendran AV, Jayadevan R, Sashidharan S. Long COVID: An overview. Diabetes Metab Syndr 2021;15(3):869-875 [FREE Full text] [doi: 10.1016/j.dsx.2021.04.007] [Medline: 33892403]
65. Miyakawa K, Jeremiah SS, Kato H, Yamaoka Y, Go H, Yajima S, et al. Rapid detection of neutralizing antibodies to SARS-CoV-2 variants in post-vaccination sera. J Mol Cell Biol 2022 Jan 29;13(12):918-920 [FREE Full text] [doi: 10.1093/jmcb/mjab050] [Medline: 34450642]
66. Choi JY, Smith DM. SARS-CoV-2 variants of concern. Yonsei Med J 2021 Nov;62(11):961-968 [FREE Full text] [doi: 10.3349/ymj.2021.62.11.961] [Medline: 34672129]
67. Lopez Bernal J, Andrews N, Gower C, Gallagher E, Simmons R, Thelwall S, et al. Effectiveness of Covid-19 vaccines against the B.1.617.2 (Delta) variant. N Engl J Med 2021 Aug 12;385(7):585-594. [doi: 10.1056/nejmoa2108891]
68. Bian L, Gao Q, Gao F, Wang Q, He Q, Wu X, et al. Impact of the Delta variant on vaccine efficacy and response strategies. Expert Rev Vaccines 2021 Oct;20(10):1201-1209 [FREE Full text] [doi: 10.1080/14760584.2021.1976153] [Medline: 34488546]
69. Fan Y, Chan K, Hung IF. Safety and efficacy of COVID-19 vaccines: a systematic review and meta-analysis of different vaccines at phase 3. Vaccines (Basel) 2021 Sep 04;9(9):1 [FREE Full text] [doi: 10.3390/vaccines9090989] [Medline: 34579226]
70. Pozzetto B, Legros V, Djebali S, Barateau V, Guibert N, Villard M, Covid-Ser study group, et al. Immunogenicity and efficacy of heterologous ChAdOx1-BNT162b2 vaccination. Nature 2021 Dec;600(7890):701-706. [doi: 10.1038/s41586-021-04120-y] [Medline: 34673755]
71. Epidemia COVID-19: Aggiornamento nazionale 3 novembre 2021 - ore 12. Istituto Superiore di Sanità. 2021 Nov 5. URL: https://www.epicentro.iss.it/coronavirus/bollettino/Bollettino-sorveglianza-integrata-COVID-19_3-novembre-2021.pdf [accessed 2021-11-11]
72. Grosso FM, Presanis AM, Kunzmann K, Jackson C, Corbella A, Grasselli G, Covid-19 Lombardy Working Group. Decreasing hospital burden of COVID-19 during the first wave in Regione Lombardia: an emergency measures context. BMC Public Health 2021 Sep 03;21(1):1612 [FREE Full text] [doi: 10.1186/s12889-021-11669-w] [Medline: 34479535]
73. Safety of COVID-19 vaccines. European Medicines Agency. URL: https://www.ema.europa.eu/en/human-regulatory/overview/public-health-threats/coronavirus-disease-covid-19/treatments-vaccines/vaccines-covid-19/safety-covid-19-vaccines [accessed 2021-11-11]
74. Xylogiannopoulos KF, Karampelas P, Alhajj R. COVID-19 pandemic spread against countries' non-pharmaceutical interventions responses: a data-mining driven comparative study. BMC Public Health 2021 Sep 01;21(1):1607 [FREE Full text] [doi: 10.1186/s12889-021-11251-4] [Medline: 34470630]

75. Askitas N, Tatsiramos K, Verheyden B. Estimating worldwide effects of non-pharmaceutical interventions on COVID-19 incidence and population mobility patterns using a multiple-event study. Sci Rep 2021 Jan 21;11(1):1972 [FREE Full text] [doi: 10.1038/s41598-021-81442-x] [Medline: 33479325]
76. Riccardo F, Ajelli M, Andrianou XD, Bella A, Del Manso M, Fabiani M, COVID-19 working group. Epidemiological characteristics of COVID-19 cases and estimates of the reproductive numbers 1 month into the epidemic, Italy, 28 January to 31 March 2020. Euro Surveill 2020 Dec;25(49):1 [FREE Full text] [doi: 10.2807/1560-7917.ES.2020.25.49.2000790] [Medline: 33303064]
77. Alfano V, Ercolano S. The Efficacy of lockdown against COVID-19: a cross-country panel analysis. Appl Health Econ Health Policy 2020 Aug 03;18(4):509-517 [FREE Full text] [doi: 10.1007/s40258-020-00596-3] [Medline: 32495067]

Abbreviations

ARIMA: autoregressive integrated moving average
ISS: National Institute of Health
ISTAT: National Institute of Statistics
OLS: ordinary least squares
SARIMA: ARIMA + seasonal component
SIR: Susceptible-Infected-Recovered

Edited by T Sanchez, A Mavragani; submitted 28.12.21; peer-reviewed by K Nagar, H Nguyen; comments to author 25.01.22; revised version received 31.01.22; accepted 03.03.22; published 07.04.22.

Please cite as:
Rovetta A, Bhagavathula AS
The Impact of COVID-19 on Mortality in Italy: Retrospective Analysis of Epidemiological Trends
JMIR Public Health Surveill 2022;8(4):e36022
URL: https://publichealth.jmir.org/2022/4/e36022
doi:10.2196/36022
PMID:35238784

Original Paper

Population Size Estimation From Capture-Recapture Studies Using shinyrecap: Design and Implementation of a Web-Based Graphical User Interface

Anne F McIntyre[1*], MPH, PhD; Ian E Fellows[1,2*], PhD; Steve Gutreuter[1*], PhD; Wolfgang Hladik[1*], MD, PhD

[1]Division of Global HIV & TB, Center for Global Health, Centers for Disease Control and Prevention, Atlanta, GA, United States

[2]Fellows Statistics, San Diego, CA, United States

[*]all authors contributed equally

Corresponding Author:
Anne F McIntyre, MPH, PhD
Division of Global HIV & TB
Center for Global Health
Centers for Disease Control and Prevention
1600 Clifton Rd NE
MS US1-2
Atlanta, GA, 30333
United States
Phone: 1 4044333601
Email: zat4@cdc.gov

Abstract

Background: Population size estimates (PSE) provide critical information in determining resource allocation for HIV services geared toward those at high risk of HIV, including female sex workers, men who have sex with men, and people who inject drugs. Capture-recapture (CRC) is often used to estimate the size of these often-hidden populations. Compared with the commonly used 2-source CRC, CRC relying on 3 (or more) samples (3S-CRC) can provide more robust PSE but involve far more complex statistical analysis.

Objective: This study aims to design and describe the Shiny application (shinyrecap), a user-friendly interface that can be used by field epidemiologists to produce PSE.

Methods: shinyrecap is built on the Shiny web application framework for R. This allows it to seamlessly integrate with the sophisticated CRC statistical packages (eg, Rcapture, dga, LCMCR). Additionally, the application may be accessed online or run locally on the user's machine.

Results: The application enables users to engage in sample size calculation based on a simulation framework. It assists in the proper formatting of collected data by providing a tool to convert commonly used formats to that used by the analysis software. A wide variety of methodologies are supported by the analysis tool, including log-linear, Bayesian model averaging, and Bayesian latent class models. For each methodology, diagnostics and model checking interfaces are provided.

Conclusions: Through a use case, we demonstrated the broad utility of this powerful tool with 3S-CRC data to produce PSE for female sex workers in a subnational unit of a country in sub-Saharan Africa.

(JMIR Public Health Surveill 2022;8(4):e32645) doi:10.2196/32645

KEYWORDS

population size estimation; multiple-source capture-recapture; Bayesian models; latent-class models; Shiny; HIV; key populations; epidemiology; digital health; online health application; populations; risk factors; online communities

Introduction

Background

Accurate knowledge of population size is critical in many areas of science but a challenge whenever complete counts are too difficult or expensive to be obtained. One such area is the HIV pandemic, which increasingly is driven by high-risk behaviors that define "key populations" (KP), among them, female sex workers (FSW), men who have sex with men (MSM), and people who inject drugs (PWID) [1]. Global, national, and local HIV control efforts all require knowing the size of these high-risk populations to monitor the epidemic in terms of density and distribution of populations over time and to inform effective and appropriately scaled program development, target setting, and resource allocation. Yet, there is no gold standard to derive reliable population size estimates (PSE). Instead, public health teams and stakeholders use a wide range of methods, many of which are not based on empirical data nor sound statistical concepts [2,3], potentially producing poor-quality estimates. Estimates of population sizes derived from programmatic mapping [4,5] enumerate members of the population attending venues during the exercise but often fail to account for the less socially visible, resulting in underestimates. Other nonempirical subjective methods such as Wisdom of the Crowd [6,7] and the Delphi methods [3,8] are susceptible to bias and the influence of individuals.

Capture-recapture (CRC) globally has seen wide use for PSE, including for the HIV pandemic [9-18]. The basic idea behind CRC is to engage in 2 or more encounter events or sources (these might also be referred to as samples, captures, or lists), recording which individuals appear in which events and relating the number of individuals sampled once to those sampled repeatedly. Most CRC exercises include 2 encounter events with the key assumption being that the 2 samples (2S) are independent [19]. Unfortunately, many such 2S-CRC exercises may suffer from violating this assumption resulting in overestimates (negative dependence between 2 samples) or underestimates (positive dependence between 2 samples) [3,19,20]. CRC with 3 (or more) samples (3S-CRC) relaxes this assumption, as interaction terms may be added to the statistical models to address source dependencies. Given sufficient overlap of samples and independence of samples, 3S-CRC allows for more sophisticated analysis compared with 2S-CRC [18,21], resulting in more accurate PSE. Statistical support for these analyses might not be available, creating a critical challenge for field epidemiologists to produce robust PSE.

Several statistical models satisfy the requirements to perform the aforementioned sophisticated analysis of 3S-CRC data: log-linear models, Bayesian model averaging, and Bayesian nonparametric latent-class models. Log-linear models are a classic methodology for the analysis of multiple source CRC data. Variants are implemented that allow for varying capture probabilities across events and heterogeneous capture probabilities among members of the population. Bayesian model averaging allows the analyst to flexibly account for list dependency by creating models for all possible dependencies and averaging over them in a way that is proportional to the probability that the dependence is correct. The Bayesian latent class model deals with heterogeneity in a novel way. It posits that there are unobserved subgroups in the data with different capture probabilities for each capture event. The number of these groups and their probabilities are unknown. The algorithm uses a Bayesian framework to estimate these, along with the population size. Application of these 3 types of statistical models requires computational expertise. This is a barrier to the use of CRC involving 3 or more sources, as it typically requires knowledge of specialized software [22] or programming in languages such as R [23]. To fill this need, we present a graphical user interface, shinyrecap, that guides the user through sample size estimation, data preparation and exploration, and PSE using CRC studies.

Objectives

The objectives of this paper were to describe *shinyrecap*, a free, web-based application facilitating the format and analysis of CRC data for PSE.

Methods

Overview of the Capture-Recapture Method

The application of ratio estimation for PSE from multiple encounters dates to at least 1787 [24] and gained popularity primarily among animal ecologists more than a century later [25-27], although applications abound in other areas, including epidemiology [28,29]. Early applications were restricted to sampling on 2 occasions or from 2 lists, wherein individuals encountered during the first survey are offered an identifying mark. For KP CRC, these identifiers are inexpensive but memorable unique objects or "gifts" such as brightly colored rubber bracelets or distinctive key chains. The number of individuals who accept the unique gifts are documented. The same process is repeated during a second survey, during which individuals are also asked about having received a gift during the previous capture. Estimation of the unknown number of population members from 2 samples requires the strong assumptions that (1) the population is static over the sampling interval, (2) the identifying unique objects or gifts are not lost nor misidentified, (3) individuals are sampled independently during the surveys (list independence), and (4) every population member shares a common and constant probability of encounter during the surveys (homogeneity). The first assumption is well-approximated by sampling over short time intervals. However, the remaining assumptions are unlikely to hold.

The next major innovation was the extension of estimation to data collected from 3 or more samples [30,31]. This enables relaxation of the third and fourth assumptions using statistical models that account for sampling dependence and various forms of inhomogeneity (ie, nonuniform) in encounter probabilities [27,30,32-35]. To understand why more samples allows for the assumption relaxation, consider a 3S-CRC where each capture is the same size. If the population is homogeneous and all individuals have the same probability of being captured in each sample (p_1), then the probability of being captured in all 3 samples would simply be p_1^3. On the other hand, if half the population has a capture probability of p_1 and the other half has

probability p_2, then the probability of a random person being captured in all 3 would be $0.5(p_1^3 + p_2^3)$. By comparing the counts of individuals captured in all 3 samples to what would be expected if there was homogeneity, we can measure and model it. Log-linear models, Bayesian model averaging, and Bayesian Dirichlet process mixture models (nonparametric latent-class models) and each model heterogeneity in different ways, allowing for the production of more accurate estimates in the presence of inhomogeneity.

Overview of Relevant Statistical Models

Log-Linear Models

Models for capture probabilities originated in the discipline of animal ecology [27,34]. The natural logs are modeled as linear combinations of factors representing various forms on inhomogeneity. Four general classes of models are produced, representing a wide range of model complexity: Captures have the same probability, and individuals are uniform (M_0); captures might have different probabilities, and individuals are uniform (M_t); captures have the same probability, and individuals may be heterogeneous (M_h); and captures may have different probabilities, and individuals may be heterogeneous (M_{th}). Selection of a single "best" model is typically done using either the Akaike or Bayesian information criterion (AIC and BIC, respectively) [36]. For these, lower values indicate a "better" model fit.

For heterogeneous models, log-linear models require the specification of a parametric distribution for the population's log odds of being captured. These are typically set to be either Normal, Poisson, Gamma, or Darroch. Additionally, the Chao (lower bound) correction can be used to obtain a lower bound on the population size rather than an estimate of it.

The "Normal" model incorporates heterogeneity as a Gaussian mixing distribution [37]. The Poisson, Darroch, and Gamma options incorporate different heterogeneity correction columns into the design matrix. The Darroch, and especially the Gamma, correction may produce distinctly large heterogeneity corrections and estimates of population size. Unfortunately, the correct model specifications are frequently not identifiable (roughly, parameters are not informed by the data), and so choosing based on any information criteria can lead to misspecified models [38].

Bayesian Model Averaging of Log-Linear Models

Bayesian model averaging is geared to be robust to list dependence. Ideally, one would like to have all capture events be independent draws from the population. In many cases, however, some capture events may be related. For example, in a citywide survey of PWID, it might happen that the first 2 capture events were more heavily concentrated in one area of the city than the third event, introducing potential dependence. When list dependence is present, the interactions between events should be considered.

The natural logs of expected frequencies of observable encounter combinations can be modeled as linear combinations of main and interaction effects of the sampling events [32,35]. This allows the model to flexibly account for list dependence among the various samples. Bayesian model averaging enumerates all possible models of list dependency and then puts a prior on the likelihood that each model is the true one, with more complex models typically having lower prior probability than less complex models. Combining this prior with a prior for population size allows one to calculate a posterior estimate of population size averaging over all possible models. In this posterior, estimates from each model are weighted by the posterior probability of the model, yielding a single estimate that includes model uncertainty. Some form of model averaging is important given that there may be limited information in the data available to identify the true model out of the large number of potential models [22].

The first step in the analysis is to specify a prior distribution for population size. This represents the analyst's prior knowledge about population size along with uncertainty. By default, a "noninformative" improper prior is used, which is proportional to 1 divided by the population size. Typically, analysts will have access to at least a rough idea of the range of possible population sizes from previous PSE reports or literature reviews. This information can be incorporated into the prior parameterized as a log-normal distribution with a truncation at a specified maximum population size. The "delta" parameter controls the prior, favoring simple models in the model averaging. This parameter is more difficult to interpret, and it is set to 2^{-k} by default, where k is the number of encounter events. Lower values indicate less prior weight on more complex list interactions. Once the prior is specified, the posterior probability distribution of the population size can be calculated.

Bayesian Nonparametric Latent-Class Models

Instead of assuming a parametric probability function for capture probability, as is done by traditional log-linear models, this approach posits that the population is divided into a number of groups, with members in each group having the same homogeneous capture probability. The number of homogeneous strata in a population is uncertain, and covariates that identify those classes may not be available. Thus, the strata are said to be latent, and strata identities are treated as missing data. Estimation is naturally accomplished using mixtures of distributions. A clever implementation of Bayesian nonparametric latent-class modeling can then be used to estimate population size [21]. Both the number of strata and the strata capture probabilities are inferred via Bayesian inference, with a stick-breaking Dirichlet process prior enforcing model parsimony such that models with fewer latent strata are preferred.

The degree to which fewer strata are preferred is controlled by a prior on the stick-breaking process parameterized as a Gamma distribution with shape and scale values. The relationship between the Gamma distribution and the number of latent groups is complex and mediated by a stick-breaking process. In general, the default values of 0.25 for both the shape and scale parameters result in a reasonably diffuse prior.

Estimation is based on the posterior distribution of population size, of which a sample is constructed using Markov chain

Monte Carlo (MCMC) simulation. MCMC algorithms start from initial values and produce serially correlated "chains" of samples from some distribution. That distribution converges to the joint posterior distribution only after some potentially large number of "burn-in" iterations. Therefore, valid inferences can be made only after discarding the burn-in iterations.

shinyrecap Application User Interface

shinyrecap was developed using the Shiny [23] web framework for R [39]. Shiny is a flexible, open-source toolkit used to build web applications with rich interactivity that can easily produce tables, visualize data, and create dashboards. The advantage of this framework is that it makes it easy to expose advanced algorithms and packages written in R to a noncoding audience. In *shinyrecap*, we leveraged the algorithms from the *Rcapture* [40] package for log-linear modeling, the *dga* package for Bayesian model averaging [41], and the *LCMCR* [21] package for Bayesian latent-class modeling. Whereas it would normally take substantial experience with R to use those packages, *shinyrecap* provides easy access to a wider audience with "the click of a button."

shinyrecap has been made available for public use [42] and does not require installation of or experience with R. Client-server communication occurs over a secure-sockets layer (SSL) protocol connection. Required data inputs are minimal and can be aggregate or individual-level. Any data provided to *shinyrecap* persist only for the session; neither input nor output data are saved on the web server. This provides users with security protection against third-party traffic analysis and any security intrusions into the server not concurrent with the user's session. *shinyrecap* offers a tutorial video and manual, and help buttons are presented where input information is required in each *shinyrecap* module.

Alternatively, R users can launch the interface locally from any computer by entering the following into the R console:

shinyrecap is structured in 3 parts. First, it supports the design of CRC studies by providing a tool for sample size estimation. Second, it provides a data formatting tool to assist with the data processing of CRC surveys. Finally, it provides the analysis tool to generate the estimates and outputs required for PSE.

Results

Application User Interface

Sample Size Estimation

When designing a CRC study, it is important to collect enough data to achieve sufficient precision for PSE. *shinyrecap*'s sample size estimation tool does this by allowing the user to specify population parameters such as guesstimates of the target population size and the amount of capture heterogeneity in the population, as well as sample characteristics such as the number of capture events and their expected sample sizes. It then simulates CRC studies in this population and estimates the population size using log-linear modeling for each of the simulations. Precision is estimated from the simulation results. The application supports simulation and estimation using the M_t model if homogeneity is assumed. If heterogeneity is allowed, simulation and estimation are performed using the M_{th} model with normally distributed capture probabilities.

Given the input parameters, the interface provides the user with the distribution of a log-linear population size estimator across the simulations. A table is also provided that summarizes the percent of times simulated estimates were within different ranges of accuracy. A user might find it acceptable to have their estimate within 10% of the true value 90% of the time, whereas they might choose to collect more samples if the calculator says that their estimate will only be within 10% of the true value 50% of the time.

Data Formatting

The first barrier encountered by a practitioner is putting their CRC data into the right format for analysis. *shinyrecap* is able to read 2 data types: individual and aggregate. We focus on the capture history format (aggregate data) here to demonstrate the data formatter tool. Individual-level data files have 1 row per encounter, with each column representing a sampling event (eg, 3 columns for 3S-CRC) and, within the columns, the successful encounter event result (ie, the individual accepts the unique object; individuals who refuse the object during the encounter are not counted). The usual data format used by CRC analysis programs is the capture-history format. In this format, each column should represent a successful encounter event, and each row should be an encounter history. A "1" indicates a successful encounter (capture), and "0" indicates absence, so the following history represents 80 individuals who were encountered and accepted the unique object during the 2nd events, but not during the 1st or 3rd:

When the aggregate data type is specified, the last column represents the total number of individuals with that capture history. A properly structured 3S-CRC data set would look something like Figure 1.

From the first row, we see that there were 30 individuals who were observed at event 1 but not at the 2nd and 3rd events:

There were 10 individuals captured in all 3 events, as seen in the following row:

Note that there is no row for the following history because that pertains to the unknown number of population members who were not observed at any event:

For k encounter events, there are $2^k - 1$ observable event histories and 1 unobservable history. Analysis of CRC data requires enumeration of all $2^k - 1$ observable counts (which may contain observed values of 0 but not missing values).

The capture-history data format is easily recorded from individually identifiable population members. However, in many epidemiological studies, unique individuals are not identified; rather, data are aggregated. These accumulated data files consist of counts of individuals who were encountered at each sampling event and the subsets of those who were encountered at any *preceding* sampling event(s) (Figure 2). No identifying information is collected on any subject at any event. During the 1st sampling event, only the count of individuals present and who were offered and accepted a unique (to the event) identifier is recorded. During the 2nd event, observed population members are tabulated by whether they received the identifier distributed during the first event, and those individuals are given a second (and different) aggregate identifier. At the 3rd event, the observed population members are cross-tabulated by whether they received the event-specific identifier distributed during each of the 2 previous events. We call this event-count formatted data. Although 7 counts have been recorded, the counts are aggregated differently from the required format shown in Figure 1. Note that the sum of samples should always be larger than the sum of count data.

It takes some thought to figure out how to convert the data to the required format, and the process becomes much more difficult if there are more than 3 events. The *shinyrecap* data formatting tool makes that conversion easy and reliable for any number of encounter events.

Figure 1. Example capture-history data format for 3 encounter events (3S-CRC). Absence or presence is denoted by 0 or 1, respectively.

Event 1	Event 2	Event 3	Count
1	0	0	30
0	1	0	80
1	1	0	40
0	0	1	60
1	0	1	20
0	1	1	20
1	1	1	10

Figure 2. Aggregated capture histories in event-count format for 3-source capture-recapture (3S-CRC).

Event 1		Event 2		Event 3		
Event 1	Count	Event 1	Count	Event 1	Event 2	Count
Present	100	Present	50	Present	Present	10
		Absent	100	Present	Absent	20
				Absent	Present	20
				Absent	Absent	60

Analysis

shinyrecap guides the user through the analysis process for log-linear modeling, Bayesian model averaging, and Bayesian latent-class modeling. All analyses may be exported as downloadable reports in HTML, Word, or PDF documents. To facilitate analysis transparency and reproducibility, R code to replicate the analysis is included in all reports by default.

Log-Linear Models

The log-linear section of the application has 3 sections. The first section, "Model Comparison," displays population size, standard error, AIC, and BIC for each potential model formulation. The "Model Selection" section allows the user to select one of these models and calculate a confidence interval. The "Descriptives" section provides output to help the user understand the model and diagnose potential problems. Two diagnostic plots help explore the heterogeneity structure. The first diagnostic plot displays a function of the number of units captured i times for different values of i. It should be roughly linear except in the case where the data were generated by an M_{th} model. The second diagnostic plot shows the number of individuals captured for the first time at the i_{th} sampling event. It should be linear in the case of the M_0 model and concave down in the case of an M_h model. Any other form may indicate an M_t or M_{th} model.

Bayesian Model Averaging

The first step in the Bayesian model averaging interface is to set a prior distribution for the population size. This is set to a noninformative 1/N distribution, but it is recommended to change this to something relevant to the population under study. To do this, the user can specify their beliefs for the median population size (ie, they believe that there is a 50% chance the population size is above it and 50% chance it is below) and the 90th percentile (ie, there is only a 10% chance the true population size is above this value). The application then parameterizes these beliefs as a log-normal distribution. The user may also specify a maximum population size to put an upper bound on the prior.

Once the prior is set, the user can go to the "Posterior Population Size" tab to obtain posterior estimates and credible intervals. The "Posterior Model Probabilities" tab allows the user to explore the different individual models that are averaged together and see their influence on the posterior.

Bayesian Nonparametric Latent-Class Models

The Bayesian nonparametric latent-class model is the most computationally intense analysis method. The user may control

the Gamma distribution stick-breaking prior as well as set the maximum number of latent groups. Increasing the number of latent groups increases the computation time but, since the number of groups is determined by the algorithm, does not affect the results so long as it is set sufficiently large. Although 10 is a good default value, the user can increase that to ensure that this limit does not affect the results.

There is a number of MCMC sampling options available to the user. There are 2 primary considerations that the user should be aware of. First, the MCMC process must be at equilibrium. To ensure this, the first samples generated by the algorithm should be thrown out. The number of samples thrown out is controlled by the "burn-in" option. If there are any trends in the trace plot (available in the Diagnostics tab), the burn-in period may need to be increased. Second, the sample size must be large enough that the posterior is not dominated by sampling noise. With MCMC sampling, each sample is correlated with the last sample, so the effective sample size (also in the Diagnostics tab) is often much lower than the raw number of samples generated by the process. Typically, the user should aim for an effective sample size greater than 1000. The effective sample size can be increased by increasing the number of samples generated or the number of MCMC steps taken between samples, which reduces correlation.

After specifying the prior on the number of strata and the MCMC sampling parameters, a sample from the posterior distribution is produced by pressing the "Run" button. A progress bar displays the progress of each computational operation. A posterior summary will be displayed.

Pairwise Analysis

The pairwise analysis table displays PSE, standard error, and 95% confidence limits for each possible pair of sampling events and is used as a diagnostic step to examine sampling events for homogeneity. Similar PSE across pairwise results indicate the independence assumption may have been met, whereas differences across results suggest violations of the assumption. Any of the models available in the *shinyrecap* Analysis tool may be used to incorporate such dependence into models.

Example With FSW Data

Estimates of key population size are critical for HIV program planning. For this reason, a large 3S-CRC activity was implemented in a subnational unit (SNU) of a sub-Saharan African country with high HIV burden and unmet need for HIV/AIDS treatment services. Using 3S-CRC data collected from FSW, we demonstrated the utility of the *shinyrecap* tool to estimate sample size sufficient for precision, format our data in preparation for analysis, and produce PSE using several different statistical models.

Between October 2018 and December 2018, 544 FSW hotspots in the SNU were sampled, representing 13,344 encounters with FSW over 3 sampling rounds. During encounters with FSW in hotspots, FSW distribution teams offered inexpensive and memorable objects that were unique to each of the 3 capture rounds. Eligible FSW who consented were considered enrolled in this PSE activity. In subsequent rounds, 1 week to 2 weeks apart, FSW were asked to show or describe objects they had received during previous rounds, and affirmative responses were tallied upon correct identification of the objects. Distributors recorded information on tablets and uploaded to a secure central server after each encounter. Data were aggregated into a table similar to Figure 1 for analysis.

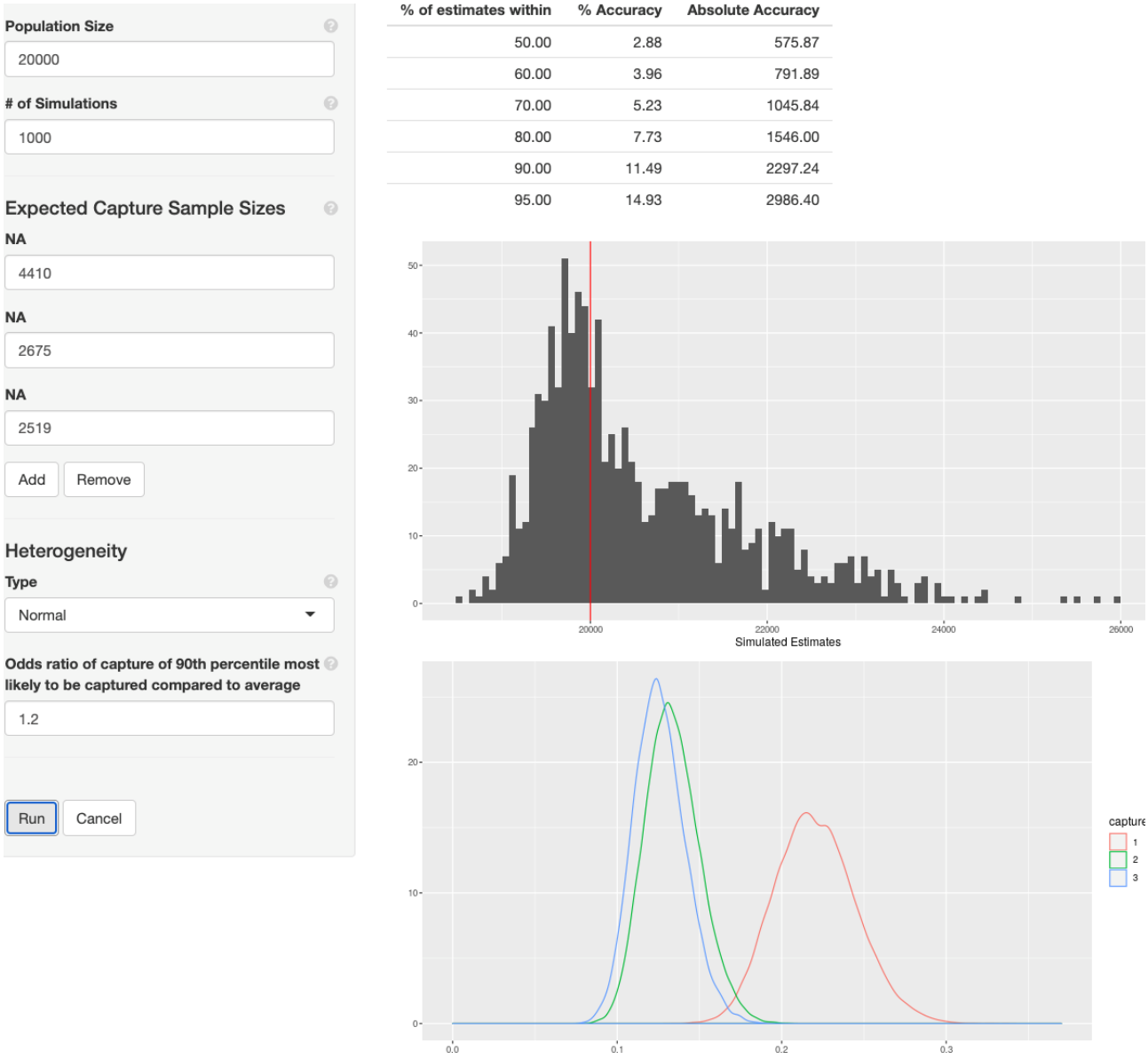
In the following sections, we work through how the *shinyrecap* application was used to assist in the planning, data management, and analysis of this study.

Sample Size Estimation

Before any study is conducted, it is wise to determine what level of precision one is likely to get out of a potential sampling plan. Figure 3 shows the result of using the sample size estimation tool in the context of the example FSW PSE study. Capture sample sizes were set at 4410, 2675, and 2519, with a theorized population size of 20,000. A moderate amount of heterogeneity was also added, such that 90% of individuals in the population had capture odds less than 1.2 times the average individual in the population.

The table in the upper right of Figure 3 summarizes the results and finds that, 80% of the time, our PSE will be within 7.73% of the true value.

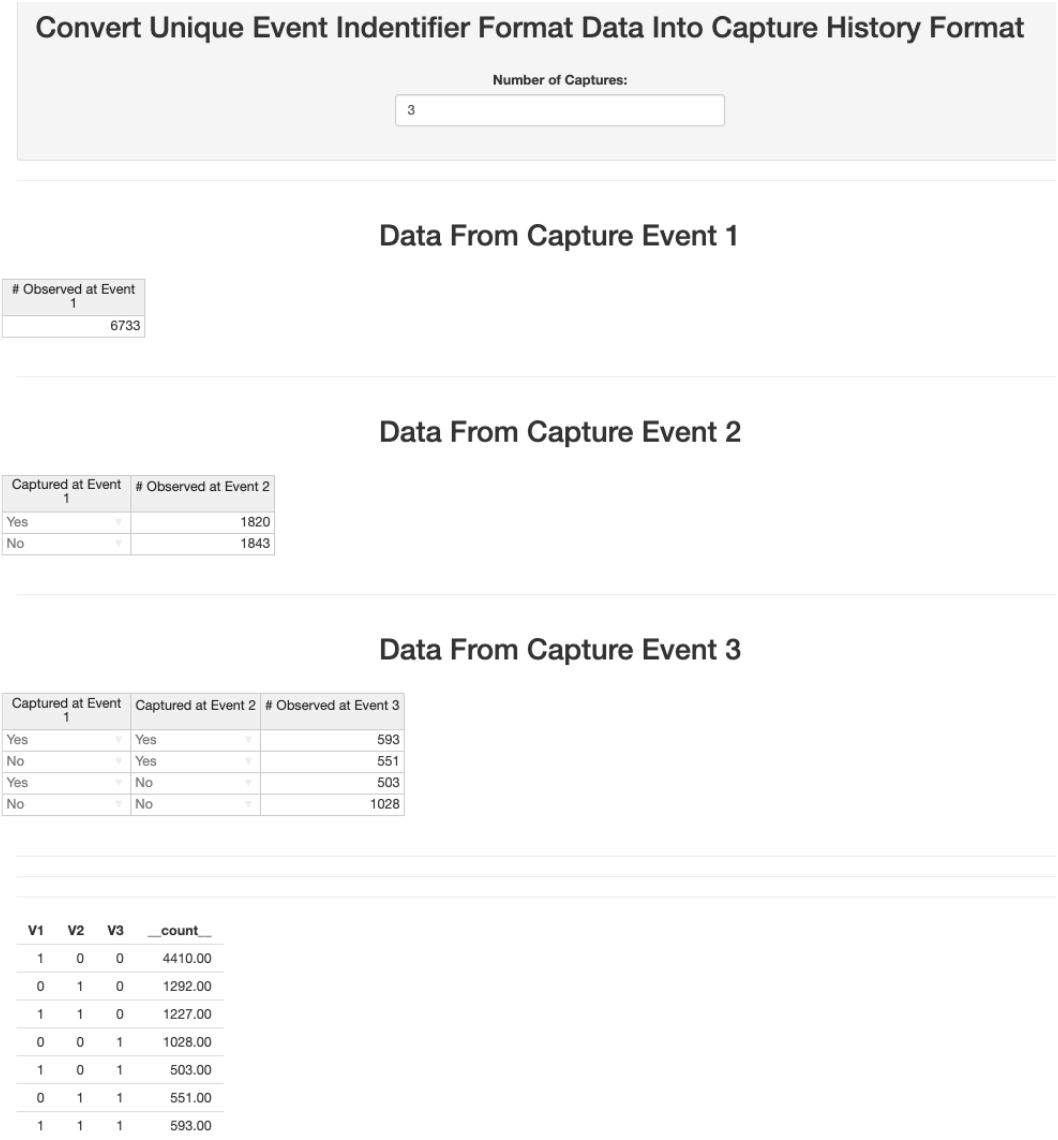
Figure 3. The sample size estimation tool applied to the example female sex worker study.



Data Format

After the data were collected, we translated it from event format to capture history format. Figure 4 shows the result of applying the data formatter to the example FSW CRC data. Once translated, the capture history data may be imported into the analysis tool for inference.

Figure 4. The data formatter tool.



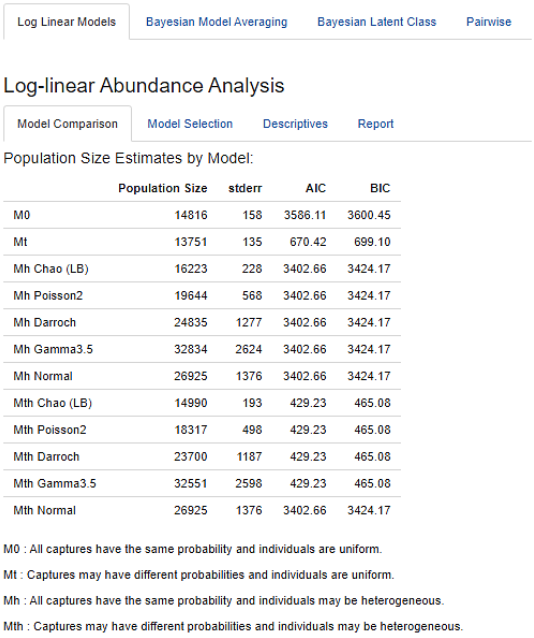
Analysis

Log-linear Models

The first class of models we apply is log-linear. Figure 5 displays the analysis tool's result table for all of the various applicable log-linear models. These may have no effects (0), effects for time (t), effects for heterogeneity (h), or both (th). Note that there are multiple listings in the figure for heterogeneous models (h and th) corresponding to different functional forms for the differing capture probabilities of individuals in the population. For most epidemiological studies, we expect capture probability to vary among individuals or over time, which means that models M_t and M_{th} are likely more appropriate than the simpler alternatives. This is consistent with the result that the AIC and BIC values are considerably lower for these compared with the M_0 and M_h models. The set of M_{th} models has the lowest AIC and BIC, indicating that there may be heterogeneity in the population.

Poisson2 induces a reasonable amount of heterogeneity and is generally a good default choice. In this case, it yields a PSE of 18,317, which, as we will see in the following sections, is consistent with other analyses.

Figure 5. Log-linear models applied to the example female sex worker study.

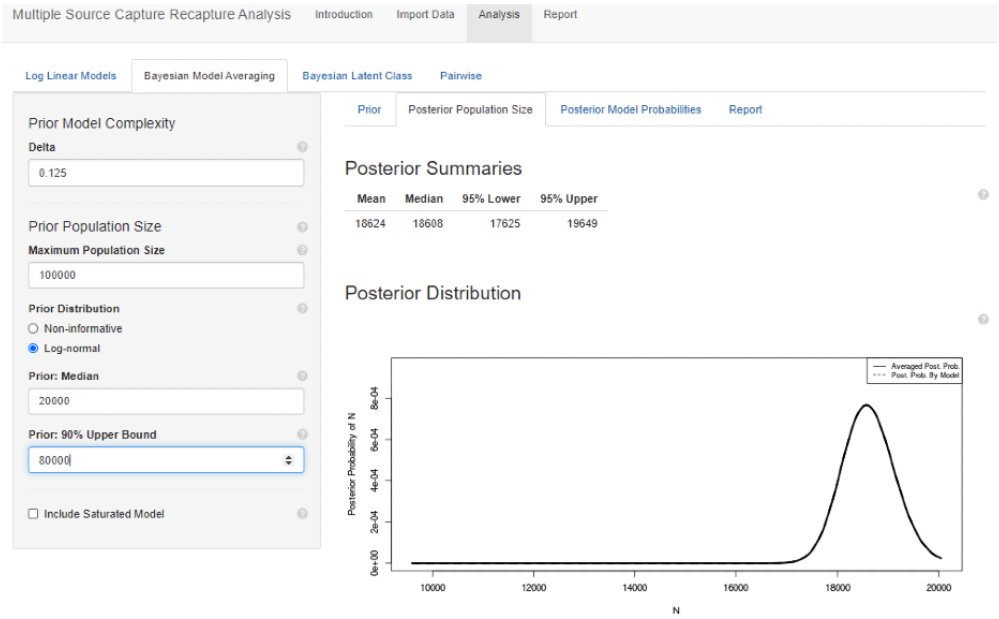


Bayesian Model Averaging

Applying a Bayesian model averaging model results in a very similar estimate compared with the Poisson2 log-linear model with a posterior estimate of 18,624 (see Figure 6). Here, we choose a diffuse prior for our analysis with a median population size of 20,000 and a 90th percentile of 80,000.

Use of the default "Noninformative" prior, which is an improper prior with mass equal to the inverse of population size, is a useful robustness check to assess the influence of our choice of prior. The posterior estimate using the noninformative prior was 18,608, which is very similar to our original result. Note that the log-normal prior median input was increased from the default of 7000 to 20,000 and the 90% upper bound was adjusted from the default of 10,000 to 80,000.

Figure 6. Bayesian model averaging applied to the example female sex worker study.

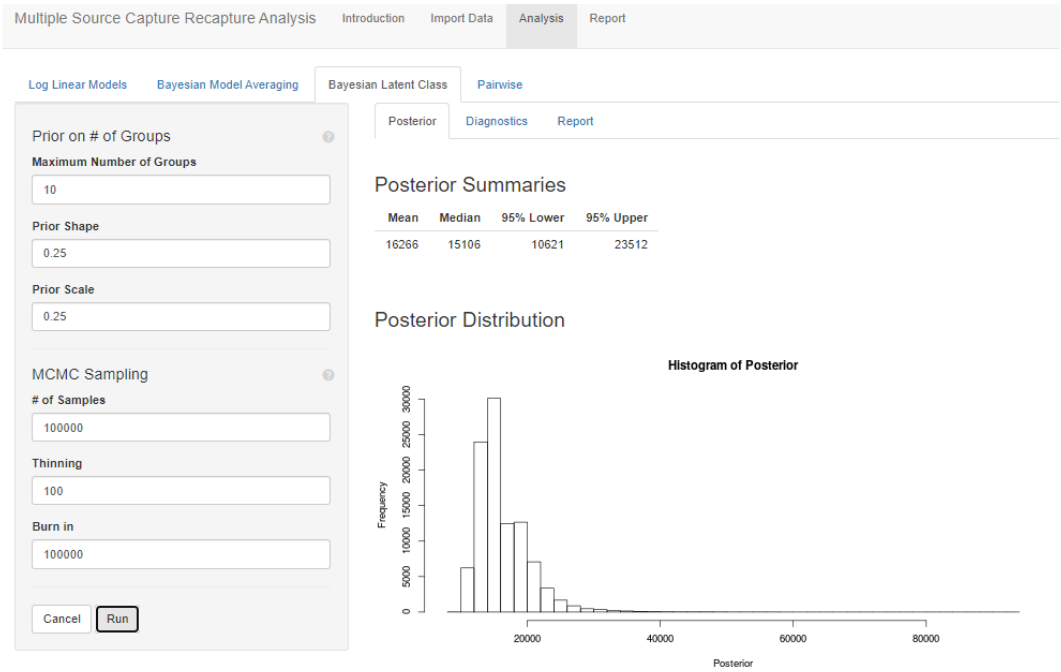


Bayesian Nonparametric Latent-Class Models

Applying the latent-class model, as in Figure 7, results in an estimate of 16,266. This is modestly lower than the other methods; however, the 95% probability interval using this method is quite wide, ranging from 10,621 to 23,512, indicating that the model's results are compatible with the other 2 methods. The latent-class model will often have intervals wider than the other 2 methods as a result of its high level of flexibility in describing the latent heterogeneity.

Note that the MCMC number of samples was increased to 100,000 from the default of 10,000, thinning was increased from the default of 10 to 100, and the burn-in was increased from the default of 10,000 to 100,000. These inputs were adjusted to increase effective sample size.

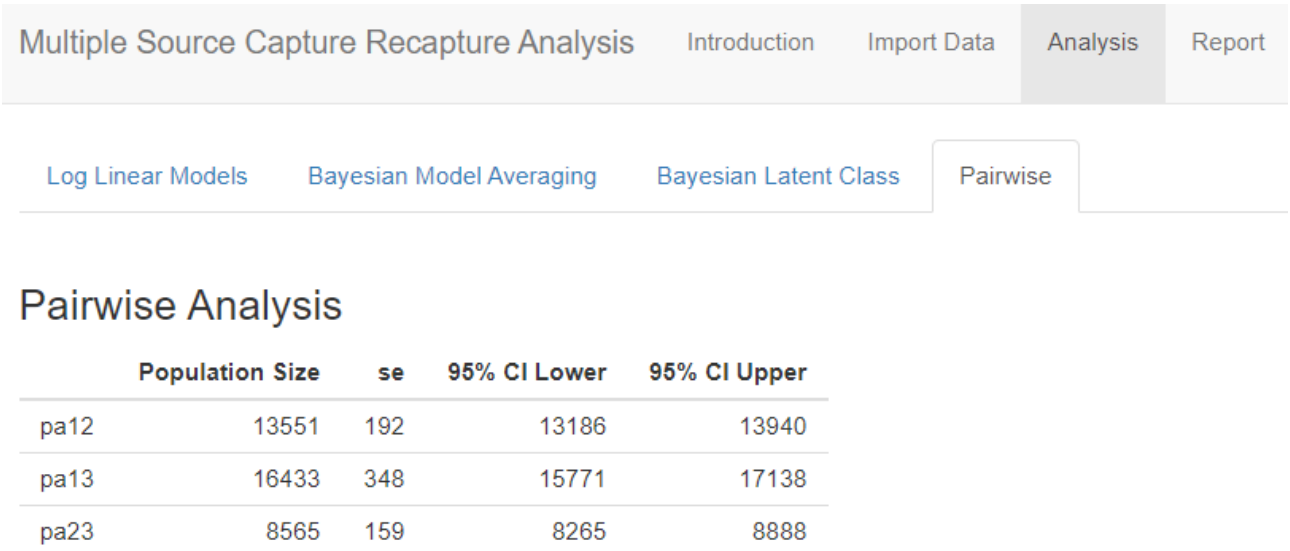
Figure 7. Bayesian latent-class modeling applied to the example female sex worker study.



Pairwise Comparison

The table in Figure 8 displays population estimates using each pair of capture events. This pairwise analysis may be helpful to review as a diagnostic step to understand the 3S-CRC data. Each row is a separate 2S-CRC analysis using only 2 of the sampling events. For example, pa12 estimates the population size using only the 1st and 2nd sampling events, pa13 estimates only the 1st and 3rd sampling events, and pa23 estimates only the 2nd and 3rd sampling events. The ideal situation is to have similar PSE for each pair, which would be consistent with independence of sampling events. Neither the 1st to 2nd nor the 1st to 3rd comparisons have intervals that overlap with the interval for the 2nd to 3rd comparison, suggesting that the independence assumption may be unreasonable for these.

Figure 8. Pairwise analysis of example female sex worker study results.



	Population Size	se	95% CI Lower	95% CI Upper
pa12	13551	192	13186	13940
pa13	16433	348	15771	17138
pa23	8565	159	8265	8888

Discussion

shinyrecap is a new Shiny application for population size estimation that is easy to use and freely accessible to anyone with an internet connection and a web browser.

The example using 3S-CRC data from FSW in an SNU of a sub-Saharan African country demonstrates how computationally intensive statistical methods are made more accessible to epidemiologists and others with *shinyrecap*. The simplicity of the sample size estimation, data formatting, and analytic tools, with supporting online manuals and tutorial videos, allows users to progress through CRC activities when statistical support might not be readily available. *shinyrecap* promotes local ownership of PSE activities, including sample size determination, formatting data for use in the *shinyrecap*, as well as using the various analytical models for estimating population size. With several key statistical models available to those

without coding expertise, local public health staff were able to test various models, compare the results, and interpret the results given their local knowledge.

Our model results using *shinyrecap* with 3S-CRC data were larger than the PSE produced from programmatic mapping and enumeration among FSW in the same SNU: 9858 in 2013 [43] and 9745 in 2015 [44]. Both these estimates were smaller than those produced by *shinyrecap*: log-linear models; for example, the M_{th} Chao lower bound was the smallest of all log-linear models, at 14,990 (14,620-15,378); the Bayesian model averaged 18,624 (17,625-19,649); and Bayesian latent-class models averaged 16,266 (10,612-23,512). The ability to produce confidence bounds is another benefit of *shinyrecap* compared with programmatic mapping and enumeration.

Shiny apps offer a solution to the problem of poor-quality estimates for key population program and policy developers and elevate the level of sophistication of analysis while building in-country capacity to implement critical surveillance activities. Recently, several Shiny apps were introduced that enhance HIV surveillance efforts to estimate awareness of HIV status over time [45], synthesize multiple PSE using the Anchored Multiplier [46], and estimate sample size for biobehavioral survey-based multiplier methods for PSE [47]. *shinyrecap* is unique among this group in that it offers multiple features in one tool to support population size estimation with 3S-CRC from sample size estimation to data formatting to multiple model options for analysis.

Our work was motivated by the needs of epidemiologists and others who require reliable tools for PSE but may not have the necessary coding experience or advanced statistical skills needed to analyze CRC data involving 3 or more samples. The application facilitates the estimation of sample sizes for captures, proper formatting of individual-level and aggregate-level data in preparation for analysis, and various options for analysis of CRC data from 3 or more sources. In addition, all output can be saved in HTML, Word, or PDF formats, with an option to include the R code used by the Shiny to produce the output. Public health teams now have a powerful tool in *shinyrecap* to produce reliable PSE for a broad range of applications without specialized computing expertise.

Conflicts of Interest

None declared.

References

1. Worldwide, more than half of new HIV infections now among key populations and their sexual partners. UNAIDS. 2019 Nov 05. URL: https://www.unaids.org/en/resources/presscentre/featurestories/2019/november/20191105_key-populations [accessed 2021-11-10]
2. Abdul-Quader AS, Baughman AL, Hladik W. Estimating the size of key populations: current status and future possibilities. Curr Opin HIV AIDS 2014 Mar;9(2):107-114 [FREE Full text] [doi: 10.1097/COH.0000000000000041] [Medline: 24393694]
3. Quaye S, Fisher Raymond H, Atuahene K, Amenyah R, Aberle-Grasse J, McFarland W, Ghana Men Study Group. Critique and lessons learned from using multiple methods to estimate population size of men who have sex with men in Ghana. AIDS Behav 2015 Feb;19 Suppl 1:S16-S23. [doi: 10.1007/s10461-014-0943-4] [Medline: 25704987]
4. Gokengin D, Aybek G, Aral SO, Blanchard J, Serter D, Emmanuel F. Programmatic mapping and size estimation of female sex workers, transgender sex workers and men who have sex with men in İstanbul and Ankara, Turkey. Sex Transm Infect 2021 Dec 29;97(8):590-595. [doi: 10.1136/sextrans-2020-054894] [Medline: 33782150]
5. Gexha Bunjaku D, Deva E, Gashi L, Kaçaniku-Gunga P, Comins CA, Emmanuel F. Programmatic mapping to estimate size, distribution, and dynamics of key populations in Kosovo. JMIR Public Health Surveill 2019 Mar 05;5(1):e11194 [FREE Full text] [doi: 10.2196/11194] [Medline: 30835241]
6. Lorenz J, Rauhut H, Schweitzer F, Helbing D. How social influence can undermine the wisdom of crowd effect. Proc Natl Acad Sci U S A 2011 May 31;108(22):9020-9025 [FREE Full text] [doi: 10.1073/pnas.1008636108] [Medline: 21576485]
7. Grasso MA, Manyuchi AE, Sibanyoni M, Marr A, Osmand T, Isdahl Z, et al. Estimating the population size of female sex workers in three South African cities: results and recommendations from the 2013-2014 South Africa health monitoring survey and stakeholder consensus. JMIR Public Health Surveill 2018 Aug 07;4(3):e10188 [FREE Full text] [doi: 10.2196/10188] [Medline: 30087089]
8. Khalid FJ, Hamad FM, Othman AA, Khatib AM, Mohamed S, Ali AK, et al. Estimating the number of people who inject drugs, female sex workers, and men who have sex with men, Unguja Island, Zanzibar: results and synthesis of multiple methods. AIDS Behav 2014 Jan;18 Suppl 1:S25-S31. [doi: 10.1007/s10461-013-0517-x] [Medline: 23709254]
9. Abeni DD, Brancato G, Perucci CA. Capture-recapture to estimate the size of the population with human immunodeficiency virus type 1 infection. Epidemiology 1994 Jul;5(4):410-414. [doi: 10.1097/00001648-199407000-00006] [Medline: 7918810]
10. UNAIDS/WHO Working Group on HIV/AIDS/STI Surveillance. Estimating the Size of Populations at Risk for HIV. In: UNAIDS https://data.unaids.org/publications/external-documents/estimatingpopsizes_en.pdf. Arlington, VA USA: Family Health International; 2003.
11. Vuylsteke B, Vandenhoudt H, Langat L, Semde G, Menten J, Odongo F, et al. Capture-recapture for estimating the size of the female sex worker population in three cities in Côte d'Ivoire and in Kisumu, western Kenya. Trop Med Int Health 2010 Dec;15(12):1537-1543 [FREE Full text] [doi: 10.1111/j.1365-3156.2010.02654.x] [Medline: 21054693]

12. Hickman M, Hope V, Platt L, Higgins V, Bellis M, Rhodes T, et al. Estimating prevalence of injecting drug use: a comparison of multiplier and capture-recapture methods in cities in England and Russia. Drug Alcohol Rev 2006 Mar;25(2):131-140. [doi: 10.1080/09595230500537274] [Medline: 16627302]
13. Paz-Bailey G, Jacobson JO, Guardado ME, Hernandez FM, Nieto AI, Estrada M, et al. How many men who have sex with men and female sex workers live in El Salvador? Using respondent-driven sampling and capture-recapture to estimate population sizes. Sex Transm Infect 2011 Jun;87(4):279-282. [doi: 10.1136/sti.2010.045633] [Medline: 21385892]
14. Héraud-Bousquet V, Lot F, Esvan M, Cazein F, Laurent C, Warszawski J, et al. A three-source capture-recapture estimate of the number of new HIV diagnoses in children in France from 2003-2006 with multiple imputation of a variable of heterogeneous catchability. BMC Infect Dis 2012 Oct 10;12(1):251 [FREE Full text] [doi: 10.1186/1471-2334-12-251] [Medline: 23050554]
15. Yu D, Calleja JMG, Zhao J, Reddy A, Seguy N. Estimating the size of key populations at higher risk of HIV infection: a summary of experiences and lessons presented during a technical meeting on size estimation among key populations in Asian countries. WPSAR 2014 Aug 11;5(3):43-49. [doi: 10.5365/wpsar.2014.5.2.008]
16. Karami M, Khazaei S, Poorolajal J, Soltanian A, Sajadipoor M. Estimating the population size of female sex worker population in Tehran, Iran: application of direct capture-recapture method. AIDS Behav 2017 Aug 16;21(8):2394-2400. [doi: 10.1007/s10461-017-1803-9] [Medline: 28623569]
17. Poorolajal J, Mohammadi Y, Farzinara F. Using the capture-recapture method to estimate the human immunodeficiency virus-positive population. Epidemiol Health 2017 Oct 10;39:e2017042 [FREE Full text] [doi: 10.4178/epih.e2017042] [Medline: 29025259]
18. Doshi RH, Apodaca K, Ogwal M, Bain R, Amene E, Kiyingi H, et al. Estimating the size of key populations in Kampala, Uganda: 3-source capture-recapture study. JMIR Public Health Surveill 2019 Aug 12;5(3):e12118. [doi: 10.2196/12118]
19. Brenner H. Use and limitations of the capture-recapture method in disease monitoring with two dependent sources. Epidemiology 1995 Jan;6(1):42-48. [doi: 10.1097/00001648-199501000-00009] [Medline: 7888444]
20. Mastro TD, Kitayaporn D, Weniger BG, Vanichseni S, Laosunthorn V, Uneklabh T, et al. Estimating the number of HIV-infected injection drug users in Bangkok: a capture--recapture method. Am J Public Health 1994 Jul;84(7):1094-1099. [doi: 10.2105/ajph.84.7.1094] [Medline: 8017531]
21. Manrique-Vallier D. Bayesian population size estimation using Dirichlet process mixtures. Biometrics 2016 Dec 08;72(4):1246-1254. [doi: 10.1111/biom.12502] [Medline: 26954906]
22. Link WA. Nonidentifiability of population size from capture-recapture data with heterogeneous detection probabilities. Biometrics 2003 Dec 11;59(4):1123-1130. [doi: 10.1111/j.0006-341x.2003.00129.x] [Medline: 14969493]
23. RStudio. URL: https://shiny.rstudio.com/ [accessed 2021-11-05]
24. Laplace PS. Sur les Naissances, les Mariages et le Morts. In: Histoire de L'Acadèmie Royale des Sciences. Paris: L'Acadèmie Royale des Sciences; 1787:693-702.
25. Dahl K. Studies of trout and trout waters in Norway. Salmon and Trout Magazine 1919;18:16-33.
26. Lincoln F. Calculating waterfowl abundance on the basis of banding returns. Washington, DC: United States Department of Agriculture; 1930:118.
27. Chapman DG, Otis DL, Burnham KP, White GC, Anderson DR. Statistical inference from capture data on closed animal populations. Biometrics 1980 Jun;36(2):362. [doi: 10.2307/2530001]
28. Hook E, Regal R. Capture-recapture methods in epidemiology: methods and limitations. Epidemiol Rev 1995;17(2):243-264. [doi: 10.1093/oxfordjournals.epirev.a036192] [Medline: 8654510]
29. Chao A, Tsay PK, Lin S, Shau W, Chao D. The applications of capture-recapture models to epidemiological data. Stat Med 2001 Oct 30;20(20):3123-3157. [doi: 10.1002/sim.996] [Medline: 11590637]
30. Darroch JN. The multiple-recapture census: I. estimation of a closed population. Biometrika 1958 Dec;45(3/4):343. [doi: 10.2307/2333183]
31. Schnabel ZE. The estimation of total fish population of a lake. The American Mathematical Monthly 1938 Jun;45(6):348. [doi: 10.2307/2304025]
32. Fienberg SE. The multiple recapture census for closed populations and incomplete 2 contingency tables. Biometrika 1972;59(3):591-603. [doi: 10.1093/biomet/59.3.591]
33. Goodman LA. Exploratory latent structure analysis using both identifiable and unidentifiable models. Biometrika 1974;61(2):215-231. [doi: 10.1093/biomet/61.2.215]
34. Pollock KH. A K-Sample Tag-Recapture Model Allowing for Unequal Survival and Catchability. Biometrika 1975 Dec;62(3):577. [doi: 10.2307/2335513]
35. Coull BA, Agresti A. The use of mixed logit models to reflect heterogeneity in capture-recapture studies. Biometrics 1999 Mar;55(1):294-301. [doi: 10.1111/j.0006-341x.1999.00294.x] [Medline: 11318172]
36. Burnham KP, Anderson DR. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. New York: Springer; 2002.
37. Rivest LP. A lower bound model for multiple record systems estimation with heterogeneous catchability. International Journal of Biostatistics 2011;7(1):23. [doi: 10.2202/1557-4679.1283]

38. Holzmann H, Munk A, Zucchini W. On identifiability in capture-recapture models. Biometrics 2006 Sep;62(3):934-6; discussion 936. [doi: 10.1111/j.1541-0420.2006.00637_1.x] [Medline: 16984340]
39. R Core Team. R: A language and environment for statistical computing. In: R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2020.
40. Baillargeon S, Rivest L. The Package: Loglinear Models for Capture-Recapture in R. J. Stat. Soft 2007;19(5):1. [doi: 10.18637/jss.v019.i05]
41. Madigan D, York J. Bayesian methods for estimation of the size of a closed population. Biometrika 1997;84(1):19-31.
42. Fellows IE. EpiApps. 2018. URL: https://epiapps.com/ [accessed 2021-05-13]
43. National Agency for the Control of AIDS (NACA). HIV Epidemic Appraisals in Nigeria: Evidence for Prevention Programme Planning and Implementation. Abuja, Nigeria: National Agency for the Control of AIDS (NACA); 2013.
44. Society for Family Health (SFH). Mapping and Characterisations of Key Populations: Evidence for Prevention Programme Planning and Implementation in Nigeria. Abuja, Nigeria: Society for Family Health (SFH); 2015.
45. Maheu-Giroux M, Marsh K, Doyle CM, Godin A, Lanièce Delaunay C, Johnson LF, et al. National HIV testing and diagnosis coverage in sub-Saharan Africa: a new modeling tool for estimating the 'first 90' from program and survey data. AIDS 2019 Dec 15;33 Suppl 3(1):S255-S269 [FREE Full text] [doi: 10.1097/QAD.0000000000002386] [Medline: 31764066]
46. Wesson PD, McFarland W, Qin CC, Mirzazadeh A. Software application profile: the Anchored Multiplier calculator-a Bayesian tool to synthesize population size estimates. Int J Epidemiol 2019 Dec 01;48(6):1744-1749 [FREE Full text] [doi: 10.1093/ije/dyz101] [Medline: 31106350]
47. Fearon E, Chabata ST, Thompson JA, Cowan FM, Hargreaves JR. Sample size calculations for population size estimation studies using multiplier methods with respondent-driven sampling surveys. JMIR Public Health Surveill 2017 Sep 14;3(3):e59 [FREE Full text] [doi: 10.2196/publichealth.7909] [Medline: 28912117]

Abbreviations

2S-CRC: 2-sample capture-recapture
3S-CRC: 3 (or more)-sample capture-recapture
AIC: Akaike information criterion
BIC: Bayesian information criterion
CRC: capture-recapture
FSW: female sex worker
MCMC: Markov chain Monte Carlo
MSM: men who have sex with men
PSE: population size estimate
PWID: people who inject drugs
SNU: subnational unit
SSL: secure-sockets layer

Edited by H Bradley; submitted 05.08.21; peer-reviewed by P Wesson, S Chabata, A Safarnejad; comments to author 08.11.21; revised version received 10.01.22; accepted 24.02.22; published 26.04.22.

Please cite as:
McIntyre AF, Fellows IE, Gutreuter S, Hladik W
Population Size Estimation From Capture-Recapture Studies Using shinyrecap: Design and Implementation of a Web-Based Graphical User Interface
JMIR Public Health Surveill 2022;8(4):e32645
URL: https://publichealth.jmir.org/2022/4/e32645
doi:10.2196/32645
PMID:35471234

Original Paper

Where Opioid Overdose Patients Live Far From Treatment: Geospatial Analysis of Underserved Populations in New York State

Kayley Abell-Hart[1], BS; Sina Rashidian[2], PhD; Dejun Teng[2], PhD; Richard N Rosenthal[3], MD; Fusheng Wang[1,2], PhD

[1]Department of Biomedical Informatics, Stony Brook University, Stony Brook, NY, United States

[2]Department of Computer Science, School of Engineering and Applied Sciences, Stony Brook University, Stony Brook, NY, United States

[3]Department of Psychiatry, Stony Brook Medicine, Stony Brook, NY, United States

Corresponding Author:
Fusheng Wang, PhD
Department of Biomedical Informatics
Stony Brook University
2313D Computer Science
Stony Brook, NY, 11794
United States
Phone: 1 631 632 2594
Email: fusheng.wang@stonybrook.edu

Abstract

Background: Opioid addiction and overdose have a large burden of disease and mortality in New York State (NYS). The medication naloxone can reverse an overdose, and buprenorphine can treat opioid use disorder. Efforts to increase the accessibility of both medications include a naloxone standing order and a waiver program for prescribing buprenorphine outside a licensed drug treatment program. However, only a slim majority of NYS pharmacies are listed as participating in the naloxone standing order, and less than 7% of prescribers in NYS have a buprenorphine waiver. Therefore, there is a significant opportunity to increase access.

Objective: Identifying the geographic regions of NYS that are farthest from resources can help target interventions to improve access to naloxone and buprenorphine. To maximize the efficiency of such efforts, we also sought to determine where these underserved regions overlap with the largest numbers of actual patients who have experienced opioid overdose.

Methods: We used address data to assess the spatial distribution of naloxone pharmacies and buprenorphine prescribers. Using the home addresses of patients who had an opioid overdose, we identified geographic locations of resource deficits. We report findings at the high spatial granularity of census tracts, with some neighboring census tracts merged to preserve privacy.

Results: We identified several hot spots, where many patients live far from the nearest resource of each type. The highest density of patients in areas far from naloxone pharmacies was found in eastern Broome county. For areas far from buprenorphine prescribers, we identified subregions of Oswego county and Wayne county as having a high number of potentially underserved patients.

Conclusions: Although NYS is home to thousands of naloxone pharmacies and potential buprenorphine prescribers, access is not uniform. Spatial analysis revealed census tract areas that are far from resources, yet contain the residences of many patients who have experienced opioid overdose. Our findings have implications for public health decision support in NYS. Our methods for privacy can also be applied to other spatial supply-demand problems involving sensitive data.

(JMIR Public Health Surveill 2022;8(4):e32133) doi:10.2196/32133

KEYWORDS

opioid use disorder; opioid overdose; buprenorphine; naloxone; geospatial analysis; epidemiology; opioid pandemic; public health

Introduction

Nonmedical opioid use, opioid use disorder, and opioid overdose are increasing problems in the United States. As of 2009, poisoning (mainly from drug overdose) surpassed motor vehicle accidents as the leading cause of injury-related death in US adults [1]. In 2016, the rate of opioid-related mortality alone surpassed firearms and motor vehicle accidents in the United States, killing over 42,000 people [2,3]. An estimated 350,000 people have died in the United States from causes related to opioids in the period from 1999 to 2016 [2].

The opioid epidemic has proved particularly severe in the Eastern United States [2]. Due to its large population, New York State (NYS) represents a large proportion of opioid overdose deaths nationally. In 2018, NYS had 3697 overdose deaths, the fifth highest of any state [4]. While other states in the Eastern United States successfully reduced the opioid prescribing rate between 2013 and 2017, prescribing in most NYS counties has remained steady [5]. Recent data show fatality rates continue to rise. On Long Island, the rate of opioid overdose increased by over 250% between 2010 and 2016. In Suffolk county alone, there were 365 opioid-related deaths in 2016 [3]. Because the number of people in NYS already experiencing opioid use disorder is large, and recent efforts to curtail prescribing may be insufficient, downstream approaches such as opioid use disorder treatment and emergency overdose treatment will continue to be essential.

Several papers have examined the spatial distribution of patients with opioid overdose or opioid use disorder in NYS. Epidemiological analyses are presented in Schoenfeld et al [3] with a focus on demographic factors and a high-resolution spatial analysis of the Long Island area. Similar analyses at the zip code level with statewide coverage are presented in Chen et al [6] and Xiang et al [7]. However, these papers rely on data from the NYS Department of Health Statewide Planning and Research Cooperative System (SPARCS) for admissions from 2004 through 2016. Since SPARCS has now released data through 2019, we report an updated zip code level map for opioid overdose similar to those in Chen et al [6].

With the exception of the most recent data, the spatial and demographic trends of opioid use disorder and opioid overdose are well studied, with many hot spots and risk factors identified. However, limited research is available in the literature regarding the spatial distribution of treatment resources relative to need. A few works have investigated spatial availability in specific regions, such as an analysis of naloxone deserts in New Jersey cities and an investigation of travel distances from sites of opioid overdose to medication-assisted treatment sites in Columbus, Ohio [8,9]. In this paper, we explore the availability of naloxone (for opioid overdose) and buprenorphine (for opioid use disorder) relative to the locations of the patients who may need them across the entirety of NYS.

The medication naloxone, often known by the brand name Narcan, is an opioid receptor antagonist that is highly effective at reversing an opioid overdose. Naloxone has been used in hospitals and emergency departments for four decades, and its safety and efficacy are well established [1,10-12]. If naloxone is administered before death, even by a layperson, opioid overdose survival approaches 100% [1,10]. However, availability of naloxone in medical settings alone may be insufficient. Opioid overdoses typically cause death in just 1 to 3 hours, and bystanders often do not call medical services for fear of police involvement [1,10,13,14]. Some studies found that medical services were called in less than 50% of incidents, even when bystanders had been trained on how to respond to an opioid overdose [1,13,14]. For these reasons, the World Health Organization recommends take-home naloxone, meaning that persons at risk of an opioid overdose and their household members or other contacts should carry naloxone in preparation for an emergency [15]. In 2018, the US Surgeon General echoed this advisory, emphasizing the role of family, friends, community members, and health care workers in preparing for opioid overdoses, which may involve misuse of prescription opioids, illicit opioids, and even high-dose prescriptions taken as directed [16].

To increase the availability of naloxone to laypeople, most US states have instituted open prescriptions known as standing orders [17,18]. Participating pharmacies (naloxone pharmacies) can offer this prescription to anyone who requests naloxone, circumventing the need for an individual visit with a physician and thereby reducing barriers associated with cost, time, and physician availability [17]. Standing orders have been significantly associated with increases in naloxone prescriptions and decreases in opioid-related deaths without affecting rates of nonmedical opioid use [17,19,20]. Standing orders may also accommodate lay prescribing (eg, through police officers or other community distribution programs), but the analysis by Gertner et al [17] showed that naloxone access laws specifically increased prescriptions from pharmacies, excluding lay prescriptions, suggesting an important role for naloxone pharmacies in fulfilling standing orders. However, only 2678 of more than 5000 NYS pharmacies are listed as participating in the naloxone standing order [21,22], suggesting a large opportunity to increase the impact of standing orders.

Another important class of medications are those that treat opioid use disorder itself by preventing opioid withdrawal and reducing cravings. Buprenorphine (brand names Suboxone, Subutex, etc) and methadone are first-line treatments for opioid withdrawal and maintenance therapy [23]. Unlike methadone, buprenorphine is a partial receptor agonist with a ceiling effect, meaning that additional doses beyond a threshold do not produce an increased effect. This feature renders buprenorphine safer than other full agonist opioids, an important factor since methadone overdose is a serious risk with methadone treatment [24-26]. There is mixed evidence for this difference in practice; one cross-sectional study found less mortality with buprenorphine than methadone [27]. Further, a major benefit of buprenorphine in the treatment of opioid use disorder is its high affinity for the mu receptor. This effect blocks the activity of other opioids, rendering concurrent use of other opioids generally ineffective and likely deterring further use [28]. Additionally, buprenorphine is a kappa receptor antagonist. Since the kappa receptor can cause dysphoria when stimulated, patients may tolerate buprenorphine better than methadone, which has some kappa agonist properties [26]. In addition to

its safety and tolerability, buprenorphine is effective in treating opioid use disorder [29]. Meta-analyses found buprenorphine similarly or slightly less effective than methadone in retaining patients in treatment. Buprenorphine was equally effective to methadone for relieving symptoms of withdrawal [26].

Despite its effectiveness and relative safety, buprenorphine remains a schedule III controlled substance. While methadone can only be prescribed for opioid use disorder through a federally licensed treatment program, individual practitioners can prescribe buprenorphine through a waiver program in accordance with the Drug Addiction Treatment Act of 2000 [30]. Physicians must complete an 8-hour course and apply for the waiver [31]. Nurse practitioners and physician assistants are also eligible but must complete a 24-hour course [32]. Although the number of waivered prescribers in the United States has risen steadily, fewer than 5000 prescribers in NYS have obtained buprenorphine waivers despite NYS having over 77,000 licensed physicians [33]. Furthermore, the waiver program limits how many prescriptions a prescriber may have at any one time. Physicians can be licensed to prescribe to a maximum of 30, 100, or 275 patients, but the vast majority (72.7% nationally) are only certified for 30 patients [34].

Despite the development of buprenorphine waivers and standing orders for naloxone, these life-saving medications are not always accessible. For instance, a report by the US Office of the Inspector General found significant geographic disparities in buprenorphine access at the county level across the United States [35]. Our work assesses where buprenorphine and naloxone are available in NYS, finding disparities at fine spatial resolution by measuring the distance to these resources for patients who have visited a hospital for an opioid overdose. In particular, this work seeks to identify regions of insufficient resources, where opioid overdose patients live in the absence of a nearby naloxone standing order pharmacy or buprenorphine-licensed prescriber. Identification of these regions is crucial for public health efforts to bridge gaps and ensure patients have access to these essential medications. Further, our methods may serve to support other states in identifying their low-access regions for improved health policy activities.

Methods

Data Sets and Preprocessing Methods

For this study, we used several public data sets and one private data set. Addresses of pharmacies participating in the naloxone standing order were downloaded from the NYS Department of Health website. The data included 2678 sites, excluding one repeated record [21]. The Python library tabula-py was used to extract the data from PDF format [36]. We used the most recent issue available at the time of writing, which was last updated in September 2019.

Addresses of buprenorphine providers in NYS were downloaded from the Buprenorphine Practitioner Locator [37]. This tracking service is provided by the US Department of Health and Human Services through the Substance Abuse and Mental Health Services Administration (SAMHSA).

To identify the locations of opioid overdose patients, private data were used from SPARCS, a statewide claims database of inpatient and outpatient medical encounters. Patient home addresses were collected from encounters from 2004 to 2019 that contained one or more *International Classification of Diseases, Ninth Revision* (ICD-9) or ICD-10 codes corresponding to the SPARCS data dictionary for opioid overdoses [38]. These codes include accidental and intentional opioid poisoning by a variety of agents such as heroin, methadone, and synthetic narcotics. Although nearly all the pharmacy and prescriber addresses were usable, many of the patient addresses in SPARCS were incomplete or otherwise not able to be geocoded, such as “Homeless,” “XX,” street intersections, and post office boxes. All patient data were accessed securely and aggregated behind a firewall in accordance with our SPARCS data use agreement and our institutional review board protocol at Stony Brook University.

All 3 data sets provide location records in human-readable format (123 my-street, my-town/county my-state, 12345). For the zip code–level map of opioid overdose rates, we used the zip code from the text address, and for the county-level summaries, we used the county name in the text address. For the other mapping, aggregation, and spatial computations, we needed locations to be represented in latitude and longitude coordinates. A geocoder is a tool for converting text addresses to latitude and longitude coordinates; we used the EaserGeocoder developed by Rashidian et al [39]. This geocoder has been shown to be more accurate than other open-source geocoders and comparable to popular commercial services such as Google and MapQuest. This geocoder was chosen because it allowed us to perform all geocoding behind a firewall and without sharing data on the web, pursuant with our data use agreement and institutional review board protocol. To format the addresses for the EaserGeocoder, further data cleaning methods were applied, such as stripping text after a comma, stripping letters that occur before the first numeral in an address, removing apartment and suite numbers, and converting one to 1.

In the naloxone pharmacy and buprenorphine prescriber data sets, we manually reviewed every address that was initially rejected by the geocoder. Many were post office boxes, street intersections, or plazas. Where possible, we searched for the business or provider name and replaced the invalid address with its corresponding street address. After review, the vast majority of resource addresses were geocoded (Table 1), resulting in 2678 pharmacy locations and 4478 prescriber locations. Due to size and data sensitivity, we did not individually review the patient addresses. Of the patient addresses, 140,219 were geocoded (Table 1). We then excluded multiple encounters with the same patient identifier, keeping only the most recent encounter that was successfully geocoded. A final total of 107,493 patient locations were available for geospatial analysis.

Table 1. Addresses geocoded in each address data set.

	Entries for New York State, n	Rejected by EaserGeocoder, n	Number geocoded, n (%)
SAMHSA[a] provider list	4484	6	4478 (99.87)
NYSDOH[b] pharmacy list	2678	0	2678 (100)
SPARCS[c] OOD[d] patient addresses	174,484	34,295	140,219 (80.34)

[a]SAMHSA: Substance Abuse and Mental Health Services Administration.

[b]NYSDOH: New York State Department of Health.

[c]SPARCS: Statewide Planning And Research Cooperative System.

[d]OOD: opioid overdose.

Methods for Resource Distance Analysis and Privacy Preservation

To assess resource availability, we sought to calculate the distance from patient residences to the nearest resources. In order to preserve privacy, we represented each point location by its enclosing census tract. SPARCS cell size policies prohibit reporting cells with fewer than 11 members. Therefore, tracts with 1 to 10 patients were merged with neighboring tracts until every polygon with a nonzero number of patients contained at least 11 patients. For each census tract or group of merged census tracts (MCTs), we calculated the distance from the polygon's centroid to the nearest resource of each type. Spatial operations were performed using a PostGIS extension of a PostgreSQL database. Specifically, PostGIS was used to find the census tract that intersected with each patient point, merge census tracts, find the centroid of the newly formed MCTs, and calculate distance from each centroid to the nearest naloxone pharmacy and the nearest buprenorphine prescriber [40].

The software ArcGIS Desktop (Esri) was used to visualize data as maps. To compute density of patients far from a given resource, we first filtered the data set to include the MCTs whose centroids were >10 km (6.2 miles) from the nearest resource. This distance was chosen because it represents a relatively substantial travel distance, especially in areas with limited public transportation. Lower distance thresholds, such as 1, 2, and 4 miles, have been used in works focusing on specific urban areas such as Baltimore City, Maryland, and Columbus, Ohio [9,41]. However, research not restricted to urban areas shows that many patients travel much farther; a study of methadone patients across the United States found that 40% of patients traveled 10 or more miles to reach an opioid treatment program, and 6% traveled more than 50 miles [42]. For our statewide analysis, we chose 10 km to identify relatively underserved suburban and rural areas. After filtering for resource-far MCTs, we then used ArcGIS's generate random points tool to generate a number of random points inside each MCT equal to the number of patients living in the MCT. Kernel density estimation was then performed on the generated points, using the kernel density tool in ArcMap with distance metric set to geodesic [43].

The programming language Python 2.7 (Python Software Foundation) was used with the library pandas for data cleaning, such as preparing the addresses for geocoding, and the Python library SQLAlchemy was used for interacting with PostGIS in Python [44].

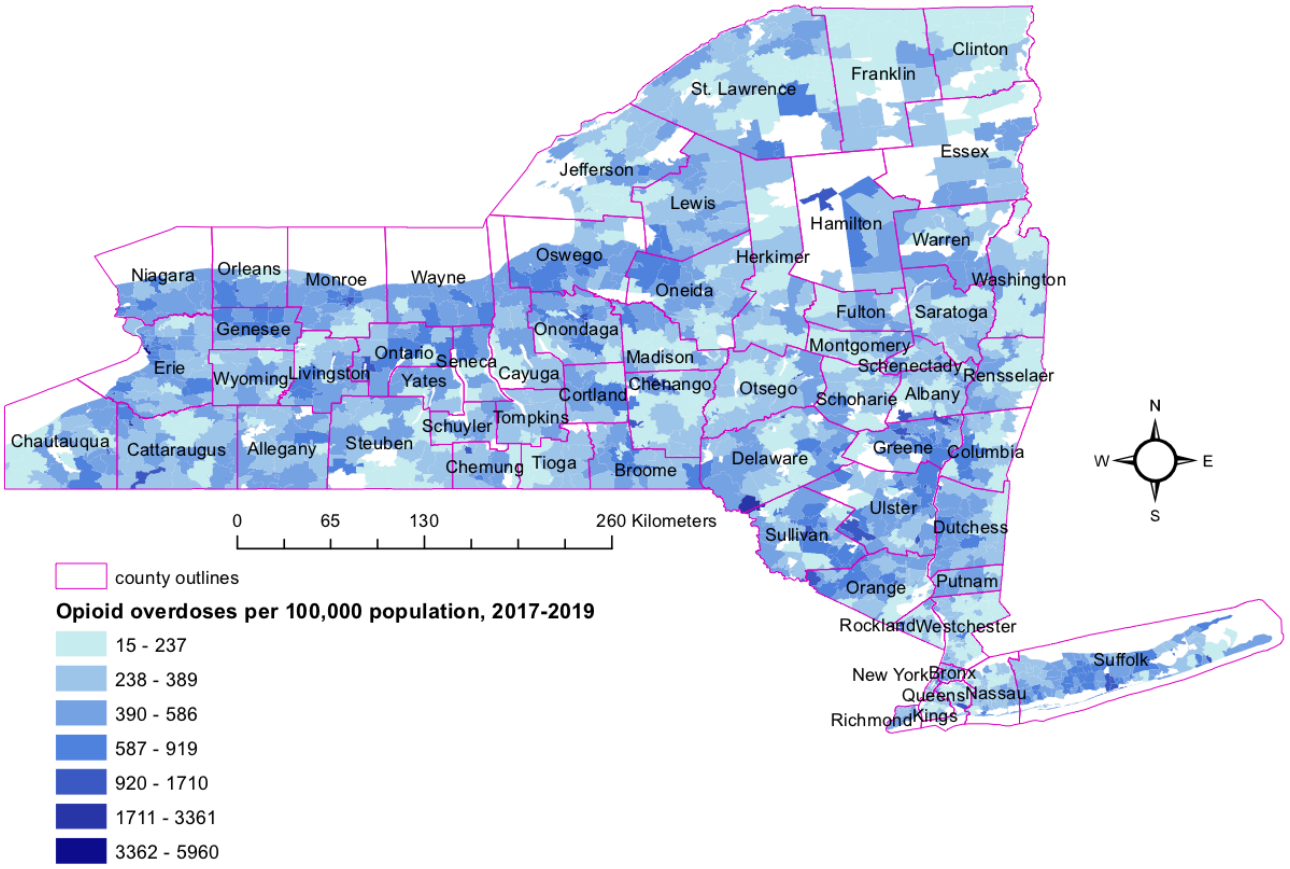
Results

Zip Code–Level Rates of Opioid Overdose for 2017-2019

First, we report an updated map of opioid overdose rates to supplement the older data published in Chen et al [6]. In this analysis, we included every opioid overdose admission with a valid zip code, without excluding multiple encounters from the same patient. Figure 1 shows the number of opioid overdose events per 100,000 residents at the zip code tabulation area level for this 3-year period. Several areas with high rates reported in Chen et al [6] continued to have high rates for recent years, such as southern central Long Island, northern Seneca county, southwestern Cattaraugus county, southern Saint Lawrence county, western Orange county, much of Greene county, and the area around the city of Buffalo. Some new hot spots are also visible, such as the area around the city of Rochester, the southwestern tip of Delaware county, and southern parts of Albany county.

If opioid overdoses were distributed randomly across zip codes, with the likelihood of each event occurring in a given zip code dependent on the zip code's population, one would still expect to observe higher rates in some areas and lower rates in others. However, many of the zip codes were observed to have much higher rates than expected by chance, as shown in Multimedia Appendix 1, Table S1. Still, given that our primary goal was to locate high numbers of patients rather than identify underlying causes, even hot spots that occurred by chance should still be worthy of attention.

Figure 1. Rate of opioid overdose per 100,000 persons at the zip code tabulation area level, 2017-2019, with county outlines overlaid for reference.



Geospatial Analysis of Distance to Buprenorphine and Naloxone Resources

Because naloxone pharmacies and buprenorphine prescribers are not distributed evenly, lack of participation disproportionately affects certain regions of NYS. We explored this issue in respect to the locations of actual patients treated for opioid overdose in NYS.

First, we offer a summary of the number of patients and resources in each county (Multimedia Appendix 1, Table S2.) Even at this low granularity, disparities are evident. Notably, the entirety of Hamilton county was found to have zero buprenorphine-waivered prescribers and zero naloxone pharmacies. Fortunately, this remote county was home to only 21 opioid overdose patients. In the other counties, the number of patients per prescriber and patients per pharmacy varied widely. Three counties had at least 100 patients per buprenorphine-waivered prescriber (Washington, Cayuga, and Cattaraugus), while 8 counties had fewer than 25 patients per prescriber. The number of patients per naloxone pharmacy was likewise highly variable, with 4 counties having over 90 patients per pharmacy (Sullivan, Putnam, Orleans, and Allegany) and 7 counties having fewer than 40 patients per pharmacy. However, a high number of resources in a county does not necessarily indicate sufficient access for the whole county, since resources are often concentrated in certain subregions of a county.

In order to find subregions with unmet needs, we present a supply and demand analysis at the high spatial resolution of the MCT. We visualized distance to resources as choropleths showing distance of MCTs to nearest resources. Figure 2 shows the distance from each MCT centroid to the nearest naloxone pharmacy, and Figure 3 shows the distance from each MCT centroid to the nearest buprenorphine provider. In both maps, a dot-density tool shows the number of patients living in the MCTs >10 km from the nearest resource.

The naloxone pharmacy map shows that most pharmacies are clustered in the urban and suburban areas, including Long Island. Naloxone pharmacies are absent in much of the northern part of the state, in the southwest, and in several of the south central areas. Although Long Island has a large number of pharmacies, its easternmost tip is lacking this resource. As for the number of patients whose residences overlap with these areas, the southwestern and south central areas appear most salient. Like the naloxone map, the buprenorphine prescriber map shows most prescribers are in urban and suburban areas, with a deficit on the eastern tip of Long Island. A large area without coverage appears in the northern central part of the state, and smaller areas appear along the border of Lake Ontario (in Oswego, Cayuga, and Wayne counties), along the southern border, and in the southern central part of the state (around Chenango and Cortland counties). Overlapping patients appear most numerous in Wayne and Oswego counties. However, in both maps, visual identification of the densest areas is limited by the fact that the dots obscure each other in dense areas. For this reason, we directly calculate and visualize density in the next section.

Figure 2. Distance in kilometers of merged census tract centroids to nearest naloxone pharmacy and density of opioid overdose patients (admissions 2004-2019) in each merged census tract. OOD: opioid overdose; MCT: merged census tract.

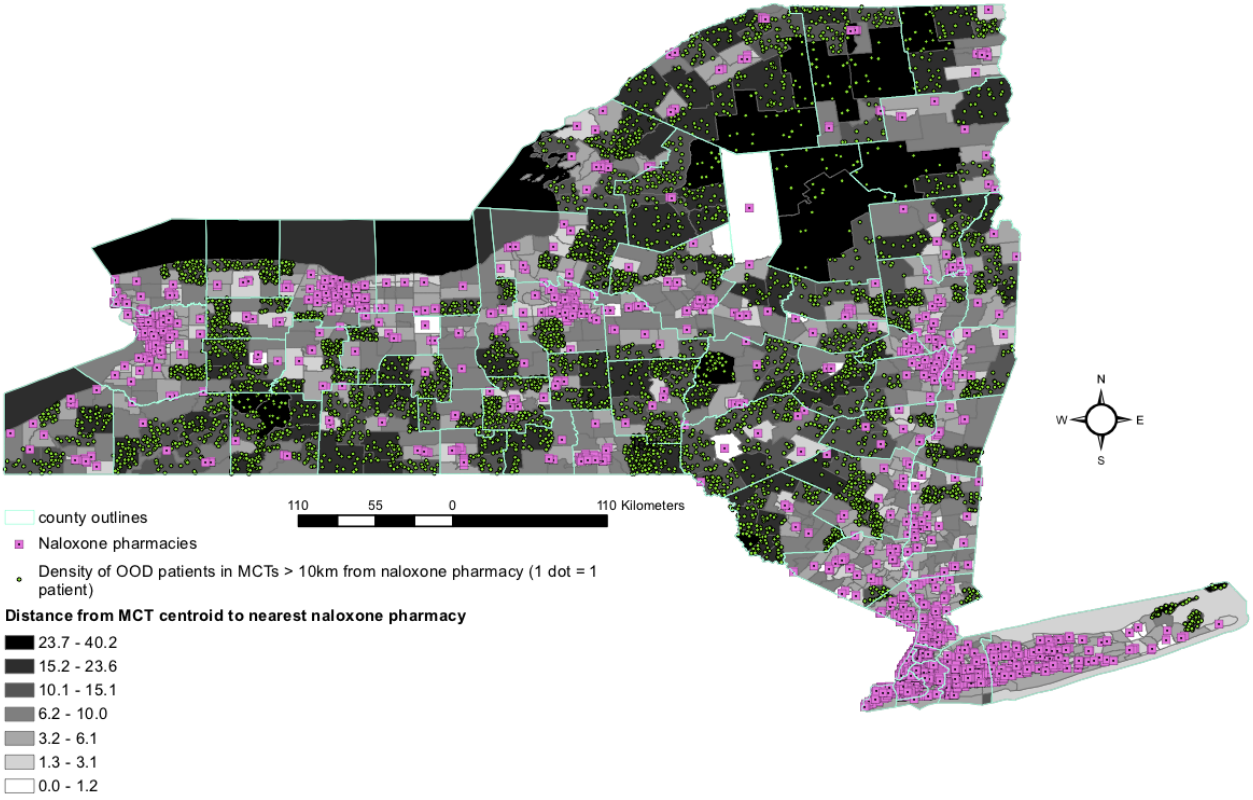
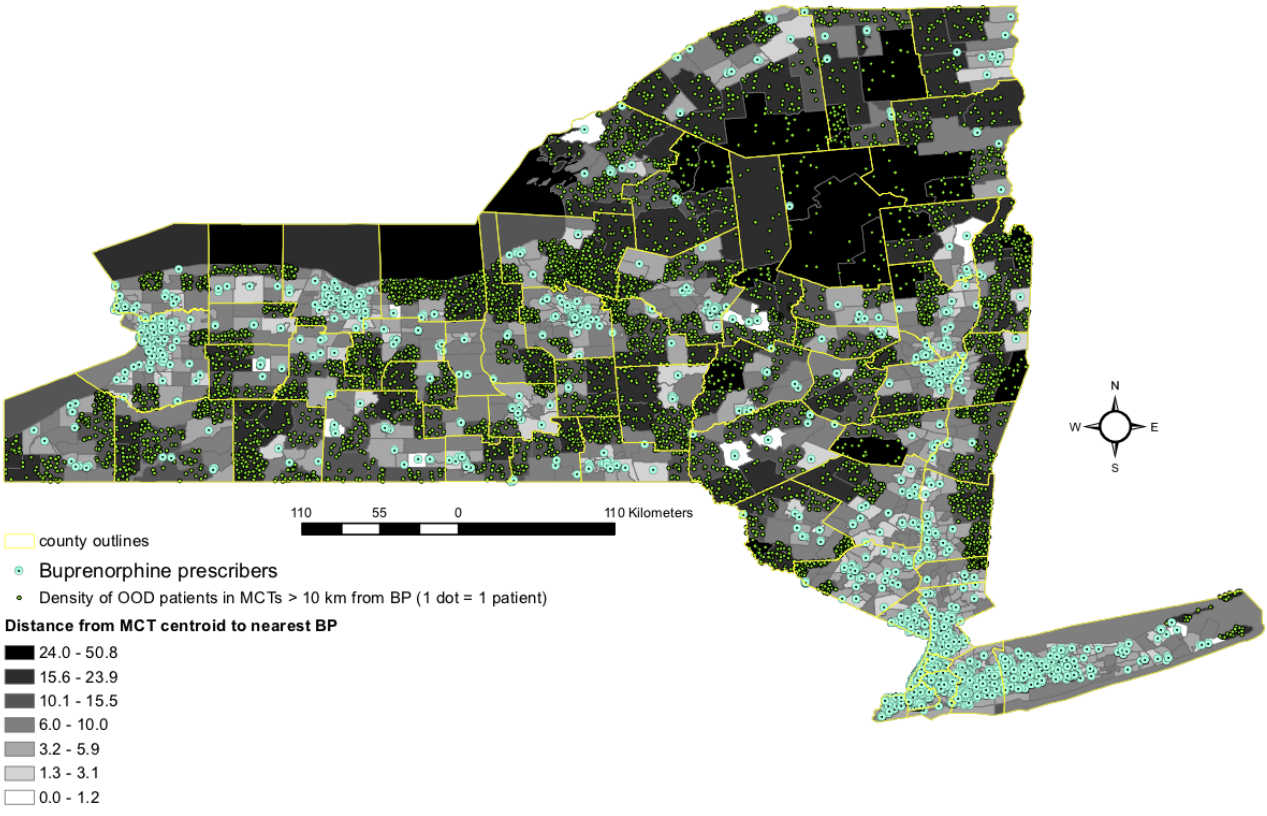


Figure 3. Distance in kilometers of merged census tract centroids to nearest buprenorphine-waivered prescriber and density of opioid overdose patients (admissions 2004-2019) in each merged census tract. OOD: opioid overdose; MCT: merged census tract; BP: buprenorphine prescriber.



Kernel Density Estimation Reveals Hot Spots of High Need Relative to Resources

Figures 2 and 3 highlight regions of NYS that are far from opioid use disorder and opioid overdose treatment resources but do not fully reflect patient density in these regions. To examine density of resource-far patients, we performed kernel density estimation. We visualized the density of patients living in MCTs >10 km from the nearest naloxone pharmacy or >10 km from the nearest buprenorphine-waivered prescriber (Figure 4 and Figure 5, respectively). For naloxone distance, the highest density region was found in eastern Broome county (Figure 4). Although a number of naloxone pharmacies exist in Broome county, they are all in the western part of the county, mostly near the city of Binghamton. For buprenorphine distances, the densest regions were in the northwestern part of the state, particularly Oswego and Wayne counties (Figure 5). Together, these data outline significant geographic disparities existing in targetable subregions of NYS.

Figure 4. Density of opioid overdose patients living in merged census tracts >10 km from the nearest naloxone pharmacy in New York State. KDE: kernel density estimation.

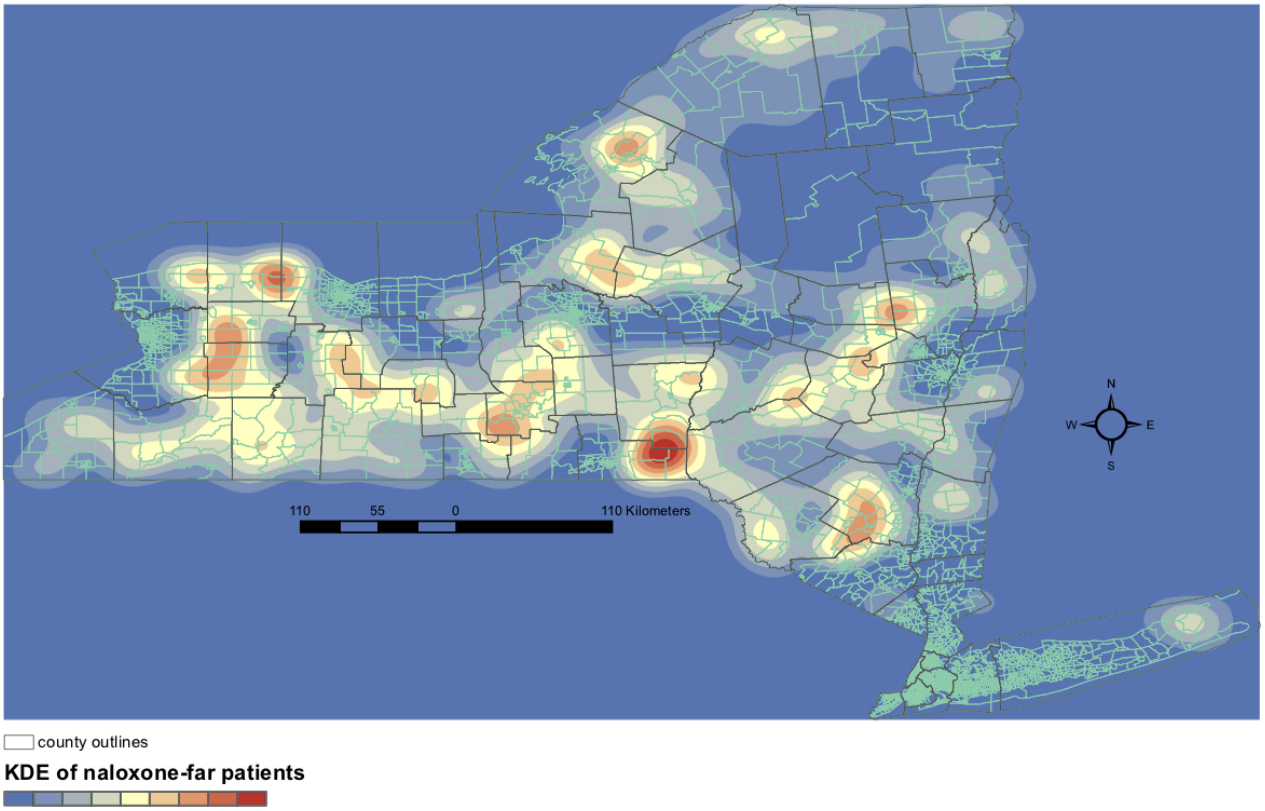
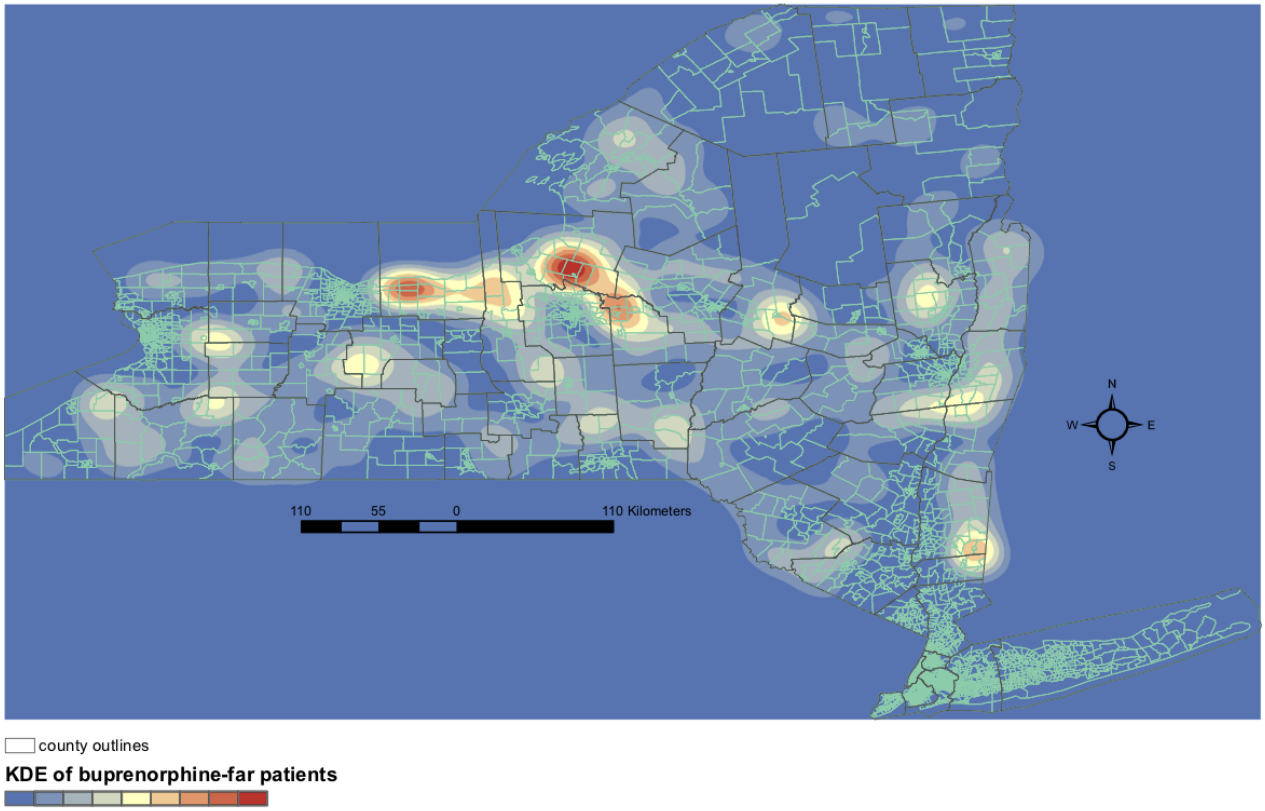


Figure 5. Density of opioid overdose patients living in merged census tracts >10 km from the nearest New York prescriber with a buprenorphine waiver. KDE: kernel density estimation.



Discussion

Summary and Comparison With Related Studies

Our updated map shows that opioid overdose continues to be frequent across NYS, with particular burden in certain areas. Our analysis of treatment resource locations shows that access to naloxone and buprenorphine is far from universal. In Broome county and several other regions, areas without nearby naloxone pharmacies overlap with the residences of a large number of opioid overdose patients. In regions such as Oswego county, areas without buprenorphine providers likewise overlap with the residences of many patients. These findings have a variety of implications for targeted interventions.

Our analysis of new SPARCS data (2017-2019) provides an update to the existing research detailing opioid overdose epidemiology in NYS [3,6]. Past studies have examined naloxone and buprenorphine resources using a variety of methods, often focusing on a small geographic area. Lozo et al [8] investigated pharmacies' participation in naloxone programs in 10 cities in New Jersey in conjunction with the rate of opioid-related hospital visits. Hyder et al [9] calculated distances from the sites of opioid overdose events in Columbus, Ohio, to the nearest medication for opioid use disorder treatment sites. Guerrero and Kao [45] examined the relationship between integrated treatment providers for substance abuse and neighborhood demographics such as income and race in Los Angeles county, California. Our work provides an important new contribution, assessing both naloxone and buprenorphine resources across a large geographic area.

Study Interpretation and Implications

It is not surprising for rural areas to have reduced proximity to resources that are often concentrated in municipal areas. However, our findings show that the overlap of patients with low-resource areas is not a uniform phenomenon across rural regions; instead, hot spots exist in particular subregions. We hope that the identification of these regions will help public health agencies to prioritize them through targeted interventions.

In the case of naloxone, it may be possible to recruit pharmacies that do not yet participate in the naloxone standing order. However, some areas might lack any pharmacies, with or without naloxone. A Google Maps search for pharmacy in the area of the Broome county hot spot shows one nearby pharmacy, on the border of Broome and Delaware counties, in the town of Deposit. This pharmacy is not listed in the directory of naloxone pharmacies, so its absence is not a geocoding error. It may be an ideal candidate for recruitment, yet having naloxone on standing order at a single pharmacy in the area might still be insufficient, and some other hot spots might not have any nearby pharmacies. One far-reaching solution is to make naloxone available over the counter (OTC), meaning that it could be stocked in any store and not limited to pharmacies [12]. The US Food and Drug Administration (FDA) has already promoted OTC naloxone as an essential step for improving naloxone accessibility. When potential manufacturers cited OTC labeling requirements as a barrier, the FDA developed a prototype label and completed their own comprehensibility testing, essentially greenlighting the process in 2018 [46,47]. Given the continued lack of OTC products, there may be other barriers or simply a

lack of financial interest for pharmaceutical companies; in this case, a government contract could bridge the gap.

As for buprenorphine, it may likewise be possible to recruit prescribers in the regions we identified. However, training clinicians and obtaining waivers does not necessarily mean that they will be able or willing to accept new patients. It may be preferable to expand psychiatric care resources in general, perhaps incentivizing the establishment of new practices in these areas, and even incentivizing young professionals to enter addiction psychiatry. Telehealth modalities are also an important route for improved care access, especially in remote areas. However, the federal government requires completion of an in-person physical to initiate buprenorphine treatment. This requirement was temporarily lifted during the COVID-19 pandemic in order to reduce in-office visits, leading to multiple successful telehealth programs [48-50]. Simply making this change permanent could improve buprenorphine accessibility for patients in the regions we identified.

Extremely remote areas represent a further challenge. Remote areas with low numbers of patients potentially spread over a large geographic area (eg, Herkimer and Hamilton counties) may be difficult to reach with spatially targeted interventions. These patients further underscore the need for telehealth prescribing, mail delivery of prescriptions, and nonpharmacy naloxone, as well as improved case finding and community outreach strategies.

Finally, spatial availability of resources such as naloxone and buprenorphine is only one of many obstacles to overdose prevention and recovery. The opioid epidemic is a complex crisis with many drivers; the spatial decision support suggested by this work is in no way intended as a sole solution. Other important resource types include methadone, naltrexone, fentanyl testing strips, therapy, and peer support. Given that social determinants of health play a significant role in the risk for opioid use disorder and opioid overdose, novel public education, identification, and engagement strategies might be implemented differentially targeting the described “landscapes of despair” that likely overlap with the census tracts underresourced with naloxone and medication for opioid use disorder prescribers [51]. Further, programs and policies must work to address the widespread financial and emotional distress that has worsened during the COVID-19 pandemic.

Limitations

For this study, we processed a large amount of real-world data. These data provide a powerful picture of opioid resource need and availability. However, we faced several limitations in working with these data. Even after intensive data cleaning, about 20% of patient addresses in SPARCS were not able to be geocoded. This problem might disproportionately affect certain populations, particularly homeless patients. In theory, a homeless patient should still have their address recorded as the place they reside, such as the address of a homeless shelter, park, or street corner. However, many addresses were recorded for opioid overdose patients in SPARCS such as undomiciled or homeless shelter, making geocoding impossible. This limitation is especially important given the increased risk of overdose in homeless individuals [52]. Our analyses are also unable to count opioid overdose patients who did not go to a hospital, which could be even more common for rural patients.

A further limitation of SPARCS data is the inevitable temporal lag; SPARCS data takes time to be compiled and released, so analyses cannot reflect the latest trends. In the Overdose Detection Mapping Application Program (ODMAP), overdoses are reported by first responders in a mobile app so that they can be compiled in almost real time [53]. The ODMAP website shows that there are participating agencies in every New York county. However, this fact does not necessarily indicate complete coverage of overdose events, since a single county has many agencies that would all need to participate. For example, the ODMAP website only lists one participating police department in Broome county, the City of Binghamton Police Department. Government agencies that have access to these data should use them to complement the less current but more complete analyses like the ones presented in this work.

The data for prescriber and pharmacy locations present some unique challenges as well. One limitation is that because NYS data were used, low-resource areas might appear exaggerated near the borders of the state. In particular, a high-density region of patients far from buprenorphine providers appears on the eastern border north of New York City (Figure 5), but these patients may be able to access resources across the border in Danbury, Connecticut, depending on their insurance coverage. Although the Broome county hot spot occurs near the southern border of NYS, it borders on a low-density region of Pennsylvania, so this still appears to be a high-need area. An important complication of prescriber data is that some prescribers work at multiple locations but do not always register every location in SAMHSA’s buprenorphine locator. It is also unknown how many NYS buprenorphine prescribers chose not be listed publicly in SAMHSA’s directory. One analysis of administrative records found that just over half of waivered prescribers had chosen to be listed publicly, although the number could be higher in NYS [54]. However, unlisted providers may be more difficult for prospective patients to find, since they are not in the directory, and unlisted providers may be less likely to be accepting new opioid use disorder patients. Therefore, gaps in the availability of listed providers may still represent important resource deficits.

Additionally, geocoding is not 100% accurate; even commercial geocoders such as Google and MapQuest only agree on about 95% of test addresses [39]. Geocoding error explains why a buprenorphine prescriber appeared in Hamilton county, when the county-level summary of prescribers showed zero buprenorphine prescribers in Hamilton county. Because the MCTs surrounding this erroneous point are so large, it so happens that their centroids are still >10 km away, so this point did not materially affect our results.

Last, our privatization methods necessarily introduce some error. For example, it is possible to have an MCT whose centroid is far from resources, even though parts of the MCT are not, especially for large or unusually shaped MCTs. If most patients actually live in the part of the MCT that is nearer to resources, a misleading hot spot could appear. To privately address the possibility of misleading hot spots, we provide close-up maps

of the regions we highlighted, showing the boundaries of each MCT shape, the associated patient counts, and the nearby resources (Multimedia Appendix 2-4). Ideally, one would calculate the distance from each patient address coordinate to the nearest resource instead of generalizing their location to the MCT. Such analysis may be feasible for public health planning operations that have access to these data. However, due to the possibility of reverse geocoding, such results cannot be made public.

Conclusions

Geospatial analysis of naloxone and buprenorphine resources revealed areas of need across NYS. The locations of these subregions should be informative to other researchers and to NYS health agencies. Rather than trying to provide for all remote areas, public health efforts can prioritize these subregions and reach high-need patients. In addition, our approach may be helpful to other states in identifying targets for resource application to address their opioid epidemic with more local efficiency. It may also be applicable to other spatial resource problems involving sensitive data.

Acknowledgments

This work was supported in part by grants Advanced CyberInfrastructure 1443054 and Information and Intelligent Systems 1350885 from the National Science Foundation, and seed grant 1158484-63845-6 from the Stony Brook University Office of the Vice President for Research (FW). This work was also supported by grant T32GM127253 from the Scholars in BioMedical Sciences Training Program and the Science Training and Research to Inform Decisions fellowship via the National Science Foundation Research Traineeship program (KAH).

Authors' Contributions

KAH performed the analyses and visualization of results and wrote the manuscript. FW was responsible for supervising the team and editing the manuscript. DT assisted in development of algorithms using PostGIS. SR was responsible for extracting the SPARCS data and assisted in querying, geocoding, and editing the manuscript. RNR was involved in conceptualization and interpretation of results and editing the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables.
[DOCX File , 23 KB - publichealth_v8i4e32133_app1.docx]

Multimedia Appendix 2
Close-up of Broome county and surrounding areas, with merged census tract polygons labeled with the number of patients in that polygon. Pink lines show original census tract boundaries that were removed during merging.
[PNG File , 402 KB - publichealth_v8i4e32133_app2.png]

Multimedia Appendix 3
Close-up of western New York State, with merged census tract polygons labeled with the number of patients in that polygon. Pink lines show original census tract boundaries that were removed during merging.
[PNG File , 438 KB - publichealth_v8i4e32133_app3.png]

Multimedia Appendix 4
Close-up of northwestern New York State, with merged census tract polygons labeled with the number of patients in that polygon. Pink lines show original census tract boundaries that were removed during merging.
[PNG File , 479 KB - publichealth_v8i4e32133_app4.png]

References

1. Clark A, Wilder C, Winstanley E. A systematic review of community opioid overdose prevention and naloxone distribution programs. J Addict Med 2014;8(3):63. [doi: 10.1097/ADM.0000000000000034]
2. Kiang MV, Basu S, Chen J, Alexander MJ. Assessment of changes in the geographical distribution of opioid-related mortality across the United States by opioid type, 1999-2016. JAMA Netw Open 2019 Feb 01;2(2):e190040 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.0040] [Medline: 30794299]

3. Schoenfeld ER, Leibowitz GS, Wang Y, Chen X, Hou W, Rashidian S, et al. Geographic, temporal, and sociodemographic differences in opioid poisoning. Am J Prev Med 2019 Aug;57(2):153-164 [FREE Full text] [doi: 10.1016/j.amepre.2019.03.020] [Medline: 31227281]
4. 2018 Drug overdose death rates. Atlanta: Centers for Disease Control and Prevention; 2020. URL: https://www.cdc.gov/drugoverdose/data/statedeaths/drug-overdose-death-2018.html [accessed 2020-09-07]
5. Medicare Part D opioid prescribing mapping tool. Centers for Medicare & Medicaid Services. 2020. URL: https://cms-oeda.maps.arcgis.com/apps/MapSeries/index.html?appid=5390718d875d4c049b1ac5976a9ff083 [accessed 2020-09-07]
6. Chen X, Hou W, Rashidian S, Wang Y, Zhao X, Leibowitz GS, et al. A large-scale retrospective study of opioid poisoning in New York State with implications for targeted interventions. Sci Rep 2021 Mar 04;11(1):5152 [FREE Full text] [doi: 10.1038/s41598-021-84148-2] [Medline: 33664282]
7. Xiang A, Hou W, Rashidian S, Rosenthal R, Abell-Hart K, Zhao X, et al. Association of opioid use disorder with 2016 presidential voting patterns: cross-sectional study in New York State at census tract level. JMIR Public Health Surveill 2021 Apr 21;7(4):e23426 [FREE Full text] [doi: 10.2196/23426] [Medline: 33881409]
8. Lozo KW, Nelson LS, Ramdin C, Calello DP. Naloxone deserts in NJ cities: sociodemographic factors which may impact retail pharmacy naloxone availability. J Med Toxicol 2019 Apr 25;15(2):108-111 [FREE Full text] [doi: 10.1007/s13181-019-00700-7] [Medline: 30805884]
9. Hyder A, Lee J, Dundon A, Southerland LT, All D, Hammond G, et al. Opioid treatment deserts: concept development and application in a US Midwestern urban county. PLoS One 2021 May 12;16(5):e0250324 [FREE Full text] [doi: 10.1371/journal.pone.0250324] [Medline: 33979342]
10. Giglio RE, Li G, DiMaggio CJ. Effectiveness of bystander naloxone administration and overdose education programs: a meta-analysis. Inj Epidemiol 2015 Dec;2(1):10 [FREE Full text] [doi: 10.1186/s40621-015-0041-8] [Medline: 27747742]
11. Boyer EW. Management of opioid analgesic overdose. N Engl J Med 2012 Jul 12;367(2):146-155. [doi: 10.1056/nejmra1202561]
12. Evoy KE, Hill LG, Davis CS. Considering the potential benefits of over-the-counter naloxone. Integrated Pharm Res Pract 2021 Feb;10:13-21. [doi: 10.2147/iprp.s244709]
13. Sherman SG, Gann DS, Scott G, Carlberg S, Bigg D, Heimer R. A qualitative study of overdose responses among Chicago IDUs. Harm Reduct J 2008 Jan 24;5(1):2 [FREE Full text] [doi: 10.1186/1477-7517-5-2] [Medline: 18218071]
14. Lankenau SE, Wagner KD, Silva K, Kecojevic A, Iverson E, McNeely M, et al. Injection drug users trained by overdose prevention programs: responses to witnessed overdoses. J Community Health 2013 Feb 31;38(1):133-141 [FREE Full text] [doi: 10.1007/s10900-012-9591-7] [Medline: 22847602]
15. McDonald R, Strang J. Are take-home naloxone programmes effective? Systematic review utilizing application of the Bradford Hill criteria. Addiction 2016 Jul 30;111(7):1177-1187 [FREE Full text] [doi: 10.1111/add.13326] [Medline: 27028542]
16. Adams J. US Surgeon General's advisory on naloxone and opioid overdose. 2018. URL: https://www.hhs.gov/surgeongeneral/priorities/opioids-and-addiction/naloxone-advisory/index.html [accessed 2021-06-24]
17. Gertner AK, Domino ME, Davis CS. Do naloxone access laws increase outpatient naloxone prescriptions? Evidence from Medicaid. Drug Alcohol Depend 2018 Sep 01;190:37-41. [doi: 10.1016/j.drugalcdep.2018.05.014] [Medline: 29966851]
18. Xu J, Davis CS, Cruz M, Lurie P. State naloxone access laws are associated with an increase in the number of naloxone prescriptions dispensed in retail pharmacies. Drug Alcohol Depend 2018 Aug 01;189:37-41. [doi: 10.1016/j.drugalcdep.2018.04.020] [Medline: 29860058]
19. McClellan C, Lambdin BH, Ali MM, Mutter R, Davis CS, Wheeler E, et al. Opioid-overdose laws association with opioid use and overdose mortality. Addict Behav 2018 Nov;86:90-95. [doi: 10.1016/j.addbeh.2018.03.014] [Medline: 29610001]
20. Rees D, Sabia J, Argys L, Latshaw J, Dave D. With a little help from my friends: the effects of naloxone access and good samaritan laws on opioid-related deaths. Cambridge: National Bureau of Economic Research; 2017. URL: https://www.nber.org/system/files/working_papers/w23171/w23171.pdf [accessed 2021-06-25]
21. Directory of pharmacies dispensing naloxone with standing orders. New York State Department of Health. 2019 Sep. URL: https://www.health.ny.gov/diseases/aids/general/opioid_overdose_prevention/docs/pharmacy_directory.pdf [accessed 2020-09-09]
22. NYS pharmacy:license statistics. 2020. URL: http://www.op.nysed.gov/prof/pharm/pharmcounts.htm# [accessed 2020-12-14]
23. Dunlap B, Cifu AS. Clinical management of opioid use disorder. JAMA 2016 Jul 19;316(3):338-339. [doi: 10.1001/jama.2016.9795] [Medline: 27434445]
24. Walsh SL, Preston KL, Stitzer ML, Cone EJ, Bigelow GE. Clinical pharmacology of buprenorphine: ceiling effects at high doses. Clin Pharmacol Ther 1994 May;55(5):569-580. [doi: 10.1038/clpt.1994.71] [Medline: 8181201]
25. Walsh SL, Eissenberg T. The clinical pharmacology of buprenorphine: extrapolating from the laboratory to the clinic. Drug Alcohol Depend 2003 May 21;70(2 Suppl):S13-S27. [doi: 10.1016/s0376-8716(03)00056-5] [Medline: 12738347]
26. Whelan PJ, Remski K. Buprenorphine vs methadone treatment: a review of evidence in both developed and developing worlds. J Neurosci Rural Pract 2012 Jan 26;3(1):45-50 [FREE Full text] [doi: 10.4103/0976-3147.91934] [Medline: 22346191]

27. Connock M, Juarez-Garcia A, Jowett S, Frew E, Liu Z, Taylor R, et al. Methadone and buprenorphine for the management of opioid dependence: a systematic review and economic evaluation. Health Technol Assess 2007 Mar;11(9):1-171 [FREE Full text] [doi: 10.3310/hta11090] [Medline: 17313907]
28. Mishlen K, Barbieri VA, Sullivan MA, Bisaga A. Opioid use following an outpatient detoxification and induction onto XR-NTX: testing the blockade as a predictor of retention in treatment. Drug Alcohol Depend 2017 Feb;171:e146. [doi: 10.1016/j.drugalcdep.2016.08.403]
29. Rosenthal RN, Goradia VV. Advances in the delivery of buprenorphine for opioid dependence. Drug Des Devel Ther 2017;11:2493-2505 [FREE Full text] [doi: 10.2147/DDDT.S72543] [Medline: 28894357]
30. Bliley T. H.R. 2634: Drug Addiction Treatment Act of 2000. 2000 Jul 27. URL: https://www.congress.gov/bill/106th-congress/house-bill/2634 [accessed 2020-09-15]
31. Buprenorphine waiver management. Substance Abuse and Mental Health Services Administration. 2019. URL: https://www.samhsa.gov/medication-assisted-treatment/training-materials-resources/buprenorphine-waiver [accessed 2020-09-07]
32. Become a buprenorphine waivered practitioner. Substance Abuse and Mental Health Services Administration. 2020. URL: https://www.samhsa.gov/medication-assisted-treatment/become-buprenorphine-waivered-practitioner [accessed 2020-09-15]
33. NYS medicine: license statistics. 2020. URL: http://www.op.nysed.gov/prof/med/medcounts.htm# [accessed 2020-12-14]
34. Practitioner and program data. Substance Abuse and Mental Health Services Administration. URL: https://www.samhsa.gov/medication-assisted-treatment/practitioner-resources/DATA-program-data [accessed 2020-09-08]
35. Office of the Inspector General. Geographic disparities affect access to buprenorphine services for opioid use disorder. US Department of Health and Human Services. 2020 Jan. URL: https://oig.hhs.gov/oei/reports/oei-12-17-00240.pdf [accessed 2020-12-31]
36. Ariga A. GitHub: chezou/tabula-py. 2021. URL: https://github.com/chezou/tabula-py [accessed 2021-07-16]
37. Buprenorphine practitioner locator. Substance Abuse and Mental Health Services Administration. 2020. URL: https://www.samhsa.gov/medication-assisted-treatment/practitioner-program-data/treatment-practitioner-locator [accessed 2020-09-07]
38. SPARCS: Opioid-Related Facility Visits in New York State: Beginning 2010 DATA DICTIONARY. Office of Quality and Patient Safety Bureau of Health Informatics. 2016. URL: https://www.health.data.ny.gov/w/rxm6-fp54/fbc6-cypp?cur=QXIiWf3r6CB&from=-y6YY3DkBV2 [accessed 2022-03-22]
39. Rashidian S, Dong X, Avadhani A, Poddar P, Wang F. Effective scalable and integrative geocoding for massive address datasets. SIGSPATIAL '17 Proc of the 25th ACM SIGSPATIAL Int Conf Adv Geographic Inf Syst 2017 Nov:1-10. [doi: 10.1145/3139958.3139986]
40. PostGIS: spatial and geographic objects for PostgreSQL. 2021. URL: https://postgis.net/ [accessed 2021-07-16]
41. Beardsley K, Wish ED, Fitzelle DB, O'Grady K, Arria AM. Distance traveled to outpatient drug treatment and client retention. J Substance Abuse Treatm 2003 Dec;25(4):279-285. [doi: 10.1016/s0740-5472(03)00188-0]
42. Rosenblum A, Cleland CM, Fong C, Kayman DJ, Tempalski B, Parrino M. Distance traveled and cross-state commuting to opioid treatment programs in the United States. J Environ Public Health 2011:1-10 [FREE Full text] [doi: 10.1155/2011/948789] [Medline: 21776440]
43. ArcMap: kernel density. Esri. 2020. URL: https://desktop.arcgis.com/en/arcmap/latest/tools/spatial-analyst-toolbox/kernel-density.htm [accessed 2020-12-31]
44. The Python SQL toolkit and object relational mapper. SQLAlchemy. 2021. URL: https://www.sqlalchemy.org/ [accessed 2021-07-16]
45. Guerrero EG, Kao D. Racial/ethnic minority and low-income hotspots and their geographic proximity to integrated care providers. Subst Abuse Treat Prev Policy 2013 Sep 23;8(1):34. [doi: 10.1186/1747-597x-8-34]
46. Statement from FDA Commissioner Scott Gottlieb, M.D., on unprecedented new efforts to support development of over-the-counter naloxone to help reduce opioid overdose deaths. US Food and Drug Administration. 2019 Jan 17. URL: https://tinyurl.com/4s4m2haz [accessed 2021-06-25]
47. Comprehension for OTC Naloxone (CONFER) Pivotal Label Comprehension Study. US Food and Drug Administration. 2018. URL: https://www.fda.gov/media/119745/download [accessed 2021-06-25]
48. FAQs: Provision of methadone and buprenorphine for the treatment of Opioid Use Disorder in the COVID-19 emergency. SAMHSA. 2020 Apr 21. URL: https://www.samhsa.gov/sites/default/files/faqs-for-oud-prescribing-and-dispensing.pdf [accessed 2021-06-25]
49. Samuels E, Clark S, Wunsch C, Jordison KL, Reddy N, Vanjani R, et al. Innovation during COVID-19: improving addiction treatment access. J Addict Med 2020 Jul;14(4):e9. [doi: 10.1097/adm.0000000000000685]
50. Thornton J, Varisco T, Bapat S, Downs C, Shen C. Impact of COVID-19 related policy changes on buprenorphine dispensing in Texas. J Addict Med 2020 Nov;14(6):e374. [doi: 10.1097/adm.0000000000000756]
51. Monnat SM, Brown DL. More than a rural revolt: landscapes of despair and the 2016 presidential election. J Rural Stud 2017 Oct;55:227-236 [FREE Full text] [doi: 10.1016/j.jrurstud.2017.08.010] [Medline: 29269990]
52. Yamamoto A, Needleman J, Gelberg L, Kominski G, Shoptaw S, Tsugawa Y. Association between homelessness and opioid overdose and opioid-related hospital admissions/emergency department visits. Soc Sci Med 2019 Dec;242:112585 [FREE Full text] [doi: 10.1016/j.socscimed.2019.112585] [Medline: 31634808]
53. Overdose Detection Mapping Application Program. URL: http://www.odmap.org/#agency [accessed 2021-11-05]

54. Ghertner R, Ali MM. Increases in providers with buprenorphine waivers in the United States from 2016 to 2019. Psychiatr Serv 2020 Sep 01;71(9):971-971. [doi: 10.1176/appi.ps.201900635] [Medline: 32460681]

Abbreviations

FDA: Food and Drug Administration
ICD-9: International Classification of Diseases, Ninth Revision
MCT: merged census tract
NYS: New York State
ODMAP: Overdose Detection Mapping Application Program
OTC: over-the-counter
SAMHSA: Substance Abuse and Mental Health Services Administration
SPARCS: Statewide Planning And Research Cooperative System

Edited by H Bradley; submitted 16.07.21; peer-reviewed by S Linder, A Pengput; comments to author 25.10.21; revised version received 08.11.21; accepted 08.02.22; published 12.04.22.

Please cite as:
Abell-Hart K, Rashidian S, Teng D, Rosenthal RN, Wang F
Where Opioid Overdose Patients Live Far From Treatment: Geospatial Analysis of Underserved Populations in New York State
JMIR Public Health Surveill 2022;8(4):e32133
URL: https://publichealth.jmir.org/2022/4/e32133
doi:10.2196/32133
PMID:35412467

Original Paper

Influence of Maternal Exposure to Mass Media on Growth Stunting Among Children Under Five: Mediation Analysis Through the Water, Sanitation, and Hygiene Program

Shutong Huo[1,2*], MPH; Kai Wang[1*], MSc; Zongchao Liu[3], MPH; Yuao Yang[3], MPH; Jia Yi Hee[1], MClinEpi; Qiwei He[1], PhD; Rie Takesue[4], MPH; Kun Tang[1,5], MSc, MBBS, DPhil

[1]Vanke School of Public Health, Tsinghua University, Beijing, China

[2]Program in Public Health, University of California, Irvine, CA, United States

[3]Department of Biostatistics, Mailman School of Public Health, Columbia University, New York, NY, United States

[4]Health Section Programme Division, UNICEF Headquarters, New York, NY, United States

[5]Institute of Healthy China, Tsinghua University, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Kun Tang, MSc, MBBS, DPhil
Vanke School of Public Health
Tsinghua University
4th Floor Mingli Building
Haidian District
Beijing, 100083
China
Phone: 86 010 62781788
Email: tangk@tsinghua.edu.cn

Abstract

Background: The issue of malnutrition in the Democratic Republic of Congo is severe. Meanwhile, the Water, Sanitation, and Hygiene program has been demonstrated to be effective in reducing the rates of growth stunting among children.

Objective: We aimed to explore the association between maternal exposure to mass media and stunting in children through water, sanitation, and hygiene behaviors.

Methods: Mediation analysis was conducted using data from the 2018 Multiple Indicators Cluster Surveys.

Results: Mothers' exposures to television and the internet in the Democratic Republic of Congo significantly decreases the risk of stunting in children by 5% and 10%, respectively, mediated by household water, sanitation, and hygiene facilities and practices.

Conclusions: These findings could inform interventions and policies to reduce the rate of stunting rate children by promoting water, sanitation, and hygiene through mass media, especially through the internet and television.

(JMIR Public Health Surveill 2022;8(4):e33394) doi:10.2196/33394

KEYWORDS

water, sanitation and hygiene; mass media; malnutrition; Democratic Republic of Congo; DRC; mediation analysis; children; pediatric; stunting; television; internet; sanitation; hygiene

Introduction

Undernutrition resulted in approximately 45% of deaths in children under 5 years, while stunting affected approximately 149 million children under 5 years in 2020 globally [1], posing significant global health issues. The United Nations Children's Fund (UNICEF) defines *stunting rate* as the percentage of children aged between 0 to 59 months whose height is between 2 standard deviations (moderate and severe stunting) and 3 standard deviations (severe stunting) below the median for their age [2]. World Health Organization (WHO) data indicate that sub-Saharan Africa has the highest prevalence of stunting globally—approximately 32.3% children under 5 years in 2020 [3]. In sub-Saharan Africa, the largest percentage of children with stunted growth, approximately 43% in 2010, were from the Democratic Republic of Congo [2]. This issue gained the

attention of the World Health Assembly, and 6 global nutrition targets were set for 2025 [3]; in 2012, the first target—a 40% reduction in the number of children under 5 years with stunted growth—was met. However, in 2017, 42% of children in the Democratic Republic of Congo had stunted growth [4]. The prevalence of stunting continues to increase in the Democratic Republic of Congo, despite actions being implemented to reduce stunting in children in the Democratic Republic of Congo [4,5]. Stunting has been demonstrated to negatively affect cognitive performance, educational performance, and maternal reproductive outcomes.

WHO identified poor maternal health and nutrition, inadequate infant- and child-feeding practices, and infections as the main causes of stunting in children under 5 years. In addition, unsafe water sources and poor handwashing practices are also major factors that cause stunting [3,6]. Safe water sources and improved hygiene through good handwashing practices are the primary components of the WHO Water, Sanitation, and Hygiene (WASH) program, which is a primary health service initiative to provide access to healthy and safe water, and sanitation facilities including soap and water for proper handwashing [7]. By building community-based interventions, such as the WHO WASH program, children will be better guarded against diarrheal diseases, malaria, intestinal worms, and environmental causes of subclinical infection, and prevention of these illnesses will lower the risk of stunting in children [3]. Several experimental studies [8-13] have demonstrated that increased access to water, sanitation, and hygiene is significantly associated with lower risks of stunting. A randomized experiment in India reported that the implementation of a sanitation program resulted in reduced stunting [8], and a cluster-randomized controlled trial in rural Mali found that sanitation intervention using behavioral modification gradually increased the availability of and accessibility to latrines, and reduced stunting in children [9]. Furthermore, a number of programs have been executed globally, such as the US Centers for Disease Control and Prevention Global Water, Sanitation, and Hygiene program, which resulted in a 25% decrease in childhood diarrhea, and the UNICEF WASH Program, which has provided access to healthy water to 14 million people for daily use in cooking, drinking, and personal hygiene [14-16].

Because only approximately 50% of the population in the Democratic Republic of Congo have access to clean water sources, UNICEF supplies safe drinking water, maintains sufficient sanitation facilities, and promotes hygienic practices during critical periods in communities and schools [17]. In the Democratic Republic of Congo, several water, sanitation, and hygiene awareness-raising sessions have been organized with campaigners who disseminate information by directly addressing the public or through mass media (eg, radio, television, and the internet) [2,18,19]. Regionally, the Democratic Republic of Congo WASH Consortium Program improved household health and reduced water-associated illnesses in 2020 [20]. The program used social media, such as Twitter and Facebook, radio broadcasts, and newsletters to disseminate knowledge and evidence about water, sanitation, and hygiene [21]. Literature has shown that mass media directly influences people's behavior, and the effect increases quickly by improving mass media [22,23]. In addition, studies [24-26] conducted in other countries have demonstrated that media access is associated with better water, sanitation, and hygiene behaviors, especially in resource-poor settings. The use of mass media has also been associated with reduced rates of stunting. In sub-Saharan Africa, stunting is significantly associated with mass media exposure [27-29]. Studies conducted in other low- or middle-income countries, such as Bangladesh [30], Indonesia [31], and India [32] also found that the exposure of mother to the mass media is associated with stunting. In India, a community-led and community-managed programs using technology and mass media to change community behavior reduced stunting rates in children under 2 years [3].

Although published research has explored whether water, sanitation, and hygiene facilities and practices are associated with stunting; whether mass media are associated with water, sanitation, and hygiene facilities and practices; and whether mass media are associated with stunting; there is little research on associations between mass media, stunting, and water, sanitation, and hygiene–related behaviors as an integral. Hence, we aimed to investigate whether maternal exposure to mass media influences stunting rates and water, sanitation, and hygiene facilities and practices and to identify which type of mass media exposure works best to improve water, sanitation, and hygiene and nutrition behaviors.

Methods

Data Source and Study Design

Data from the UNICEF Multiple Indicator Cluster Surveys (MICS) were utilized. MICS, which consists of 6 rounds of surveys that focus mainly on maternal health and child development, has become the largest source of statistically sound and internationally comparable data on women and children worldwide [33]. Ethics approval for the survey was provided by individual review boards within each participating country at the time of survey implementation. Surveys were conducted in 26 provinces throughout urban and rural areas in the Democratic Republic of Congo during 2017 to 2018, in which a multistage clustering sampling strategy was used to select households. Type of residence was retained for stratification. With sampling units at different levels, a total of 21,630 households were drawn. From the original data, a sub–data set was formalized, in which 20,019 households remained after matching for household indices across 3 correlated data sets. The sub–data set documented basic and health-related information for women of reproductive age as well as their children under the age of 5 years.

Measures and Outcomes

Basic Characteristics

Data on maternal education level, household wealth, child's sex and age, residency, and Democratic Republic of Congo province of residence were utilized for this study. *Maternal education level* was coded as an ordinal discrete variable with 3 categories (no formal education or did not complete primary school, completed primary school, completed secondary school).

Household wealth was categorized by quintiles of wealth index (low, low-middle, middle, high-middle, and high). *Child's age* in years was coded as an ordinal categorical variable ranging from 0 to 4. *Sex* (male or female) and *residency* (urban or rural) were both binary variables, and *province* of residence was coded as a nominal categorical variable (26 provinces).

Primary Outcome

The primary outcome of this study was whether a child in a given household was considered stunted, characterized with the binary variable *stunting*.

Primary Exposure

The primary exposure of this study was maternal exposure to mass media. Exposure included variables that indicated whether a mother was exposed to magazines or newspapers, radio, television, and the internet. As families in Democratic Republic of Congo often have limited access to these items, these variables were coded as binary variables rather than with multilevel values indicating detailed frequencies.

Mediator

The potential mediator of this study was coded as a composite variable that summarized information on household water, sanitation, and hygiene facilities and practices. Data from questions in the survey were grouped into handwashing practice, appropriate point-of-use water treatment, improved toilet, and improved drinking water source. These questions were addressed in our study either by household - level spot check observations or by self - reporting. After filtering valid responses, data from the 4 items were summarized into one composite variable, with 3 ordinal levels. Level 1 indicated that the household met none of the criteria, level 2 indicated that the household met at least one but not all of the criteria, and level 3 indicated that all of the criteria were met. A higher level indicated better water, sanitation, and hygiene facilities and practices.

Among 20,019 households in the sub–data set, 19,397 were included in the analysis based on responses to the questions and child's age in the household. Given the limited information about water, sanitation, and hygiene facilities and practices covered by the survey and conceiving reasonable associations among the primary outcome, exposure, and the potential mediator, only a subset of the original survey questions was taken into account.

Statistical Analysis

Descriptive and regression analyses were performed to investigate the causal relationships between water, sanitation, and hygiene; stunting; and mother's exposure to mass media. Descriptive analysis was conducted on baseline characteristics, primary outcome and exposure, and potential mediator. Categorical variables were described as proportions, and continuous variables were described with mean with standard deviation. Mediation analysis (Figure 1) was performed using binary logistic regression:

where Y denotes an event that whether the child is stunted as a binary outcome, β_1 denotes the coefficient for mass media use, and β_0 denotes the intercept of the fitted model, which also includes a vector of control variables X with their coefficients phi (or Z with their coefficients theta), and ordinal logistic regression,

where Y denotes the water, sanitation, and hygiene practice composite variable with an ordinal outcome of m levels ($1 \leq i \leq (m-1)$), β_{1j} denotes the coefficients for mass media use, β_{0j} denotes the intercept of the fitted model, and the model also includes a set of control variables Z.

Model 1 was a logistic regression model that included the mass media use, maternal education level, child's age and sex, family wealth, residency, province, and mother's age as covariates and stunting as the outcome. Model 2 was an ordinal regression that included the mass media use, maternal education level, family wealth, residency, province, and mother's age as covariates and the composite variable as the outcome. The hypothesized mediating role of the composite variable between mass media exposure and stunting was tested for each media type. Combining model 1 and model 2 together,

where Y denotes the stunting outcome, Z denotes the mass media use, M denotes the water, sanitation, and hygiene facilities and practices, and δ denotes the average causal mediation effect, which can be achieved by subtracting Equation 3 from Equation 4;

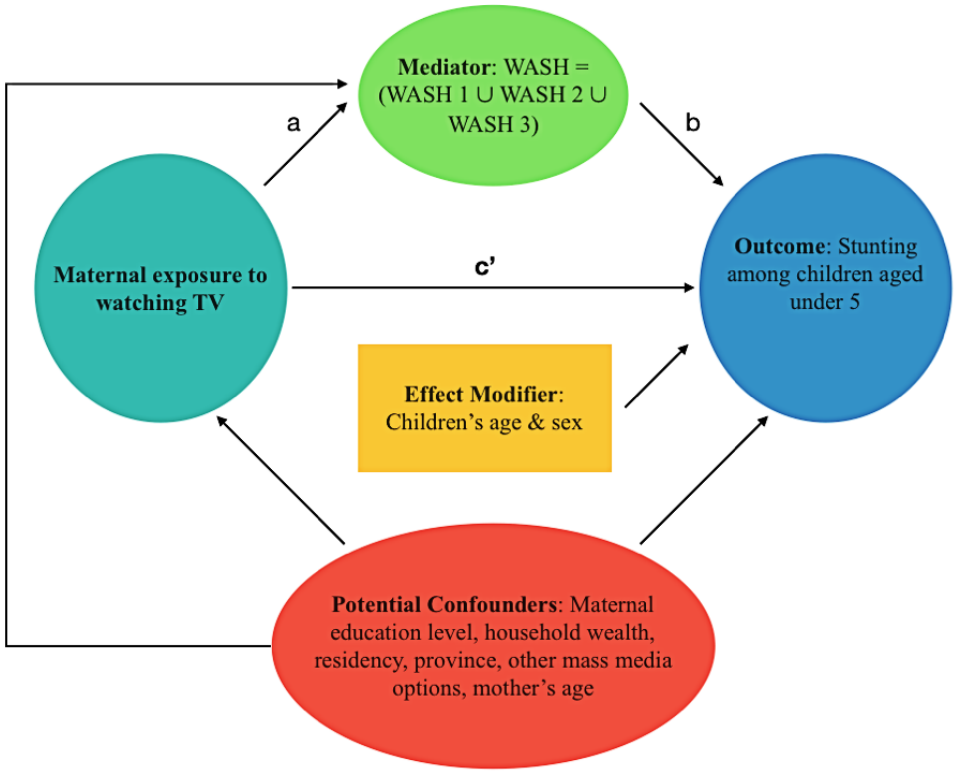
where Y denotes the stunting outcome, Z denotes the mass media use, M denotes the water, sanitation, and hygiene facilities and practices practice, and ϕ denotes the average direct effect, which can be achieved by subtracting Equation 5 from Equation 6; and total effect in the causal pathway was estimated using a general model-based approach [34].

Total effect = Average causal mediation effect + Average direct effect

All statistical analyses were performed using R (version 3.6.1). The level of statistical significance was set at 5% (P<.05) for all statistical tests. Results were reported as odds ratio (OR) and 95% confidence intervals. Interpretation for the coefficients of the ordinal regression model was slightly different from that of the logistic regression models, despite the same link function—a general model-based approach to mediation analysis [35,36] was used to examine the mediating effects, implemented using an R package [34]. The point estimates and 95% confidence intervals for average causal mediation effect, average direct effect, total effect, and the proportion of mediation were reported for each type of media. The number of Monte Carlo draws for quasi-Bayesian approximation was set to 1000. The White

heteroskedasticity-consistent estimator for the covariance matrix was used to attain robust standard errors.

Figure 1. Potential mediation effects in causal pathways. a and b: average causal mediation effect (ACME); c': average direct effect (ADE); WASH: Water, Sanitation, and Hygiene.



Results

The prevalence of stunting was found to be 44.43% (8619/19,397) in children under 5 years (Table 1). The majority of children were from households without improved drinking water sources (12,288/19,397, 63.35%), appropriate water treatment (18,961/19,397, 97.75%), improved toilets (17,213/19,397, 88.74%), and appropriate handwashing practices (17,130/19,397, 88.31%). Most children resided in rural areas (14,191/19,397, 73.16%), and more than half of the mothers had primary education levels or less.

Radio had the highest access rate (3943/19,397, 20.33%) among all the mass media methods, followed by television (2300/19,397, 11.86%), magazine or newspaper (1084/19,397, 5.96%), and internet (342/19,397, 1.76%). In the descriptive analysis, compared to children of mothers who have never been exposed to mass media, children of mothers who have been exposed to mass media had a higher-value composite variable (ie, better water, sanitation, and hygiene practices), lower prevalence of stunting, higher household wealth level, higher mother's education level, and were more likely to reside in urban areas. Provincial information can be found in Multimedia Appendix 1.

Mothers' exposure to television or internet significantly improved household water, sanitation, and hygiene practices, after adjusting for mother's education level, household wealth level, residency, province of residence, and mother's age (Table 2). For children whose mother had ever watched television, the odds of the combined high and middle water, sanitation, and hygiene practices versus that of the low level were estimated to be 1.64 times those of children whose mother had never watched television. For children whose mother had ever used the internet, the odds ratio of the same was estimated to be 2.89. The logistic model in which stunting was regressed on the composite variable showed a significantly decreased effect of water, sanitation, and hygiene practices variable on the odds of stunting, when given other covariates were controlled for. For children whose household had intermediate water, sanitation, and hygiene practices, the odds of being stunted were 1.01 (95% CI 0.94-1.09) times those of children in household with the lowest level water, sanitation, and hygiene practices. For children whose household had the highest versus the lowest water, sanitation, and hygiene practices level, the odds ratio of being stunted was estimated to be 0.41 (95% CI 0.18-0.85).

The estimated coefficients of total effects and average causal mediation effects (Table 3) can be interpreted as the increased or decreased risk of stunting among children whose mothers have ever been exposed to a type of mass media compared with those whose mothers have never been exposed to this type of mass media. Mothers' exposure to magazines or newspapers had no significant total effect on stunting (P=.75). In comparison, mothers' exposure to any of the other 3 types of media had significantly positive effects on decreasing the risk of child stunting, with 2% (P=.02), 5% (P<.001), and 10% (P<.001) decreased risk of stunting for radio, television, and internet, respectively. These total effects from the mediation analysis package are consistent with the total effects from logistic regression (model 1). For mothers' exposure to television or internet, there was a significant effect on decreasing

stunting risk mediated by the composite water, sanitation, and hygiene variable, indicating a partial mediation effect with proportions of mediation of 18% (*P*=.03) and 19% (*P*=.02), respectively.

Table 1. Characteristics of the study participants by exposure to mass media.

Characteristics	Total (n=19,397)	Television (n=2300)	Magazine or newspaper (n=1084)	Radio (n=3943)	Internet (n=342)
Stunting, n (%)	8619 (44.43)	627 (27.26)	374 (34.50)	1468 (37.23)	74 (21.64)
Mother's age (years), mean (SD)	29.99 (7.17)	29.74 (6.72)	29.41 (7.13)	29.89 (6.99)	28.90 (6.12)
Household water, sanitation, and hygiene, n (%)					
Improved drinking water source	7109 (36.65)	1838 (79.91)	599 (55.26)	2085 (52.88)	284 (83.04)
Appropriate point-of-use water treatment	426 (2.20)	201 (8.74)	70 (6.46)	174 (4.41)	68 (19.88)
Improved toilet	2184 (11.26)	440 (19.13)	165 (15.22)	621 (15.75)	90 (26.32)
Appropriate handwashing practice	2267 (11.69)	700 (30.43)	282 (26.01)	826 (20.95)	151 (44.15)
Child's age (years), n (%)					
0	4175 (21.52)	450 (19.57)	224 (20.66)	839 (21.28)	76 (22.22)
1	3945 (20.34)	503 (21.87)	241 (22.23)	830 (21.05)	76 (22.22)
2	3764 (19.41)	462 (20.09)	212 (19.56)	753 (19.10)	65 (19.01)
3	3914 (20.18)	464 (20.17)	227 (20.94)	775 (19.66)	69 (20.18)
4	3599 (18.55)	421 (18.30)	180 (16.61)	746 (18.92)	56 (16.37)
Mother's education, n (%)					
Below primary school or no formal schooling	4480 (23.10)	115 (5.00)	48 (4.43)	429 (10.88)	13 (3.80)
Primary school	7672 (39.55)	346 (15.04)	128 (11.81)	1207 (30.61)	41 (11.99)
Secondary and above	7245 (37.35)	1839 (79.96)	908 (83.76)	2307 (58.51)	288 (84.21)
Family wealth, n (%)					
Low	5859 (30.21)	220 (9.57)	224 (20.66)	816 (20.69)	45 (13.16)
Low-middle	4394 (22.65)	280 (12.17)	226 (20.85)	782 (19.83)	45 (13.16)
Middle	3759 (9.38)	548 (23.83)	224 (20.66)	767 (19.45)	45 (13.16)
High-middle	3023 (15.58)	636 (27.65)	203 (18.73)	781 (19.81)	74 (21.64)
High	2362 (12.18)	616 (26.78)	207 (19.10)	797 (20.21)	133 (38.89)
Child sex, n (%)					
Male	9564 (49.31)	1132 (49.22)	523 (48.25)	1939 (49.18)	187 (54.68)
Female	9833 (50.69)	1168 (50.78)	561 (51.75)	2004 (50.82)	155 (45.32)
Residency, n (%)					
Urban	5206 (26.84)	1879 (81.70)	605 (55.81)	1903 (48.26)	274 (80.12)
Rural	14191 (73.16)	421 (18.30)	479 (44.19)	2040 (51.74)	68 (19.88)

Table 2. Effects of mothers' exposure to mass media on the association between stunting and water, sanitation, and hygiene.

Media type	Model 1: Total effects on stunting		Model 2: Effects on water, sanitation, and hygiene		Model 3: Direct effects on stunting	
	OR[a] (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Magazine or newspaper	1.01 (0.87, 1.17)	.90	0.91 (0.76, 1.08)	.28	1.01 (0.88, 1.17)	.86
Radio	0.90 (0.82, 0.98)	.01	0.99 (0.90, 1.10)	.88	0.89 (0.82, 0.98)	.01
Television	0.78 (0.69, 0.89)	<.001	1.64 (1.40, 1.93)	<.001	0.79 (0.69, 0.90)	<.001
Internet	0.60 (0.45, 0.79)	<.001	2.89 (2.02, 4.12)	<.001	0.63 (0.47, 0.82)	<.001

[a]OR: odds ratio.

Table 3. The mediating effect of the composite water, sanitation, and hygiene variable on mothers' exposure to mass media and stunting.

Media type	Total effect		Average causal mediation effect		Proportion of mediation[a]	
	RD[b] (95% CI)	*P* value	RD (95% CI)	*P* value	RD (95% CI)	*P* value
Magazine or newspaper	0.00 (–0.02, 0.03)	.75	0.00 (–0.00, 0.01)	.28	0.05 (–1.95, 1.68)	.76
Radio	–0.02 (–0.04, –0.00)	.02	0.00 (–0.00, 0.00)	.92	–0.00 (–0.21, 0.12)	.92
Television	–0.05 (–0.08, –0.03)	<.001	–0.01 (–0.02, –0.00)	.03	0.18 (0.02, 0.39)	.03
Internet	–0.10 (–0.14, –0.06)	<.001	–0.02 (–0.03, –0.00)	.02	0.19 (0.03, 0.42)	.02

[a]Proportion of mediation = Average causal mediation effect / Total effect.

[b]RD: risk difference.

Discussion

In the Democratic Republic of Congo, we found that mothers' exposure to television and the internet could significantly decrease their children's risk of having stunted growth through the mediation effect of the composite water, sanitation, and hygiene practices variable. In addition, we also found that mothers' exposure to television and the internet could increase their household water, sanitation, and hygiene practices, and that children with better household water, sanitation, and hygiene have a lower risk of stunting. To the best of our best knowledge, this is the first pathway analysis study about the pathway from media exposure to stunting. Our analysis was conducted with a well-established mediation method in a counterfactual framework that provides a sound theoretical basis for causal inference. Moreover, we used a comprehensive range of indicators to represent water, sanitation, and hygiene facilities and practices of households in the Democratic Republic of Congo, and because the study utilizes data from a national survey, it is representative of the whole population.

Our results were consistent with those from some previous studies. Low- and middle-income countries in Asia, such as India and Bangladesh, found that children were more likely to have severely stunted growth if their mothers had never been exposed to mass media [37-40]. A plausible explanation could be that mothers are able to gain more information about nutrition and childcare from mass media. Moreover, other studies [24,31,41] have indicated media access is associated with water, sanitation, and hygiene–related knowledge and behaviors. Poor water, sanitation, and hygiene is related to the substantial global burden of disease and disability due to subsequent malnutrition [42,43]. For instance, improving access to water, sanitation, and hygiene can reduce the burden of infectious diseases, such as diarrheal diseases, which is associated with the risk of stunting [43,44]. Evidence shows that poor water, sanitation, and hygiene is responsible for approximately half of mothers and children who are underweight, because of the synergy, wherein one increases vulnerability to the other, between diarrheal diseases and undernutrition [45].

Given that our findings suggest that there is a positive association between media access and water, sanitation, and hygiene practice, it is possible that mass media could be used as the primary means of health intervention and health education. Published literature has discussed how mass media can be used to promote health knowledge and behaviors [46]—radios were reportedly the most frequently used channel, followed by television [24,47]. This is consistent with our results, which indicated that radio usage has the highest usage. These traditional forms of media—radio, and television—have been used to increase water, sanitation, and hygiene knowledge and behaviors among target groups [24]. In addition, we found that exposure to the internet had the largest effect on decreasing the risk of stunting. Findings from the Indonesia National Nutrition Communication Campaign revealed that the use of social media, such as Facebook, Twitter, Instagram, and YouTube, extended the campaign's reach and strengthened messaging from other sources [48]. Many campaigns also have demonstrated that the use of mass media is effective in increasing child survival [46]. In addition to water, sanitation, and hygiene–related behaviors, there were many other child survival health behaviors that mass media could impact, such as oral rehydration therapy [49], bed net use [50,51], and vaccination [52].

In addition, we found that although the internet is the most effective type of mass media to influence stunting rates through water, sanitation, and hygiene practices, it had a minor rate of use. On the other hand, radio was the most popular media type but only had a small mediation effect of water, sanitation, and hygiene practices on the association between mothers' exposure to mass media and the children's stunting status. This may be partly attributed to the convenience of receiving information from the internet for mothers as individuals [53]. Instead of sharing the radio or television channels with other family numbers, women can access health information on their own through the internet (eg, child health care and feeding). Moreover, unlike watching television and reading newspapers, new media technologies are more interactive, which allows more engagement [54]. For example, many websites and apps employ a computer algorithm to target users by their preference. By searching or reading nutrition-related pages on the internet, users are prompted with similar articles by websites. Conventional mass media do not share these characteristics. The Democratic Republic of Congo WASH Consortium also reported that the visits to its website had a higher rate of use than those of other communication tools [21]. In addition, exposure to magazines or newspapers was not significant in our model (*P*=.28). The reason for this might be the higher costs of and consequently limited access to magazines and newspapers compared with the convenience of website and social media [55]. However, to identify why the internet is more effective than traditional methods of mass media in decreasing stunting

risk, additional studies are needed in order to obtain more data. Meanwhile, it might increase information disparity if any campaign focuses only on the internet, because the internet access rate is low on average but relatively higher among the households with high socioeconomic status.

Our study also had some limitations. First, data were from a cross-sectional study [33] whose observational nature without temporal sequence makes it challenging to make strong causal claims. Although we controlled for potential confounders measured in our data sets, there still could exist unmeasured exposure–outcome, exposure–mediator, and mediator–outcome confounders. For instance, paternal factors were not considered in the models due to the limitation of the MICSs data set. As a result, we were unable to differentiate the effects of maternal exposure to mass media from those of paternal source. Therefore, our exposure of interest in this study should be interpreted as an approximation of the household's exposure to mass media rather than strictly maternal exposure. Second, we did not control for the increase in familywise error rate for multiple statistical tests. Overall, we consider this study to be preliminary and encourage further investigation in this field. Lastly, mediation model warrants further refinement through interdisciplinary collaboration on questions such as whether and how information not directly related to water, sanitation, and hygiene in mass media can affect household water, sanitation, and hygiene practices.

We used causal mediation analysis to reveal the pathway from media exposure to stunting status in the Democratic Republic of Congo. As stunting in children continues to be a severe issue in the Democratic Republic of Congo, the findings of this study can inform and guide interventions or policies to reduce the rate of stunting among children by promoting water, sanitation, and hygiene through mass media. Mass media campaigns, using internet and television, may improve water, sanitation, and hygiene practices and help prevent stunting.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Characteristics of the study participants by exposure to mass media.
[DOCX File , 17 KB - publichealth_v8i4e33394_app1.docx]

References

1. Nutrition. World Health Organization. URL: https://www.afro.who.int/health-topics/nutrition [accessed 2021-02-24]
2. Improving child nutrition: the achievable imperative for global progress. United Nations Children's Fund. 2013 Apr. URL: https://data.unicef.org/resources/improving-child-nutrition-the-achievable-imperative-for-global-progress/ [accessed 2022-03-15]
3. Global nutrition targets 2025: stunting policy brief. World Health Organization. 2014 Dec 20. URL: https://apps.who.int/iris/handle/10665/149019 [accessed 2022-02-13]
4. Democratic Republic of the Congo: nutrition profile. United States Agency International Development. 2021 Jun. URL: https://www.usaid.gov/sites/default/files/documents/tagged_DRC-Nutrition-Profile.pdf [accessed 2022-02-13]
5. Kismul H, Acharya P, Mapatano MA, Hatløy A. Determinants of childhood stunting in the Democratic Republic of Congo: further analysis of Demographic and Health Survey 2013-14. BMC Public Health 2017 Aug 01;18(1):74 [FREE Full text] [doi: 10.1186/s12889-017-4621-0] [Medline: 28764669]
6. Kwami CS, Godfrey S, Gavilan H, Lakhanpaul M, Parikh P. Water, sanitation, and hygiene: linkages with stunting in rural Ethiopia. Int J Environ Res Public Health 2019 Oct 09;16(20):3793 [FREE Full text] [doi: 10.3390/ijerph16203793] [Medline: 31600942]
7. Water, sanitation and hygiene in health care facilities: status in low-and middle-income countries and way forward. World Health Organization. 2015. URL: https://www.who.int/water_sanitation_health/publications/qa-wash-hcf.pdf [accessed 2021-02-26]
8. Hammer J, Spears D. Village sanitation and children's human capital: evidence from a randomized experiment by the Maharashtra Government. The World Bank 2013 Nov 18:1 [FREE Full text] [doi: 10.1596/1813-9450-6580]
9. Pickering AJ, Djebbari H, Lopez C, Coulibaly M, Alzua ML. Effect of a community-led sanitation intervention on child diarrhoea and child growth in rural Mali: a cluster-randomised controlled trial. Lancet Glob Health 2015 Nov;3(11):e701-e711 [FREE Full text] [doi: 10.1016/S2214-109X(15)00144-8] [Medline: 26475017]
10. Spears D, Ghosh A, Cumming O. Open defecation and childhood stunting in India: an ecological analysis of new data from 112 districts. PLoS One 2013 Sep 16;8(9):e73784 [FREE Full text] [doi: 10.1371/journal.pone.0073784] [Medline: 24066070]
11. du Preez M, Conroy RM, Ligondo S, Hennessy J, Elmore-Meegan M, Soita A, et al. Randomized intervention study of solar disinfection of drinking water in the prevention of dysentery in Kenyan children aged under 5 years. Environ Sci Technol 2011 Nov 01;45(21):9315-9323. [doi: 10.1021/es2018835] [Medline: 21936492]

12. Luby SP, Agboatwalla M, Painter J, Altaf A, Billhimer WL, Hoekstra RM. Effect of intensive handwashing promotion on childhood diarrhea in high-risk communities in Pakistan: a randomized controlled trial. JAMA 2004 Jun 02;291(21):2547-2554. [doi: 10.1001/jama.291.21.2547] [Medline: 15173145]
13. Luby SP, Agboatwalla M, Painter J, Altaf A, Billhimer W, Keswick B, et al. Combining drinking water treatment and hand washing for diarrhoea prevention, a cluster randomised controlled trial. Trop Med Int Health 2006 Apr;11(4):479-489 [FREE Full text] [doi: 10.1111/j.1365-3156.2006.01592.x] [Medline: 16553931]
14. CDC at work: global water, sanitation and hygiene (WASH). Centers for Disease Control and Prevention. URL: https://www.cdc.gov/healthywater/global/programs/index.html [accessed 2021-02-26]
15. Water, sanitation and hygiene (WASH). UNICEF. URL: https://www.unicef.org/wash [accessed 2021-02-26]
16. Clean and safe water projects: UNICEF's mission. UNICEF. URL: https://www.unicefusa.org/mission/survival/water [accessed 2021-03-30]
17. Water, sanitation and hygiene. UNICEF. URL: https://www.unicef.org/drcongo/en/what-we-do/water-sanitation-and-hygiene [accessed 2021-02-26]
18. Tackling COVID-19 (coronavirus) with water, sanitation and hygiene in DRC. World Bank. URL: https://www.worldbank.org/en/news/feature/2020/04/20/tackling-covid-19-coronavirus-with-water-sanitation-and-hygiene-in-drc [accessed 2021-02-26]
19. DRC: Three WASH projects for 200,000 beneficiaries. Reliefweb. URL: https://reliefweb.int/report/democratic-republic-congo/drc-three-wash-projects-200000-beneficiaries [accessed 2021-02-26]
20. DRC WASH Consortium - key documents. Concern Worldwide. URL: https://www.concern.net/insights/drc-wash-consortium-key-documents [accessed 2021-02-26]
21. Rinck D, Rhodes B, Mushagalusa Ciza A. DRC WASH Consortium final evaluation report. Concern Worldwide. URL: https://tinyurl.com/2p8t9vz6 [accessed 2022-12-14]
22. Enikolopov R. Mass media and its influence on behaviour. Centre de Recerca en Economia Internacional. 2017. URL: https://www.crei.cat/wp-content/uploads/2018/03/Opuscle-44_ENG.pdf [accessed 2022-03-17]
23. 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] [Medline: 32113495]
24. Alexander CC, Shrestha S, Tounkara MD, Cooper S, Hunt L, Hoj TH, et al. Media access is associated with knowledge of optimal water, sanitation and hygiene practices in Tanzania. Int J Environ Res Public Health 2019 Jun 03;16(11):1963 [FREE Full text] [doi: 10.3390/ijerph16111963] [Medline: 31163573]
25. Galiani S, Gertler P, Orsola-Vidal A. Promoting handwashing behavior: the effect of mass media and community level interventions in Peru. Water and Sanitation Program. 2012 Nov. URL: https://www.wsp.org/sites/wsp/files/publications/WSP-Peru-HWWS-Impact-Evaluation-Research-Brief.pdf [accessed 2021-02-26]
26. Tidwell JB, Gopalakrishnan A, Lovelady S, Sheth E, Unni A, Wright R, et al. Effect of two complementary mass-scale media interventions on handwashing with soap among mothers. J Health Commun 2019 Mar 26;24(2):203-215. [doi: 10.1080/10810730.2019.1593554] [Medline: 30912707]
27. Adedokun ST, Yaya S. Factors associated with adverse nutritional status of children in sub-Saharan Africa: Evidence from the Demographic and Health Surveys from 31 countries. Matern Child Nutr 2021 Jul 07;17(3):e13198 [FREE Full text] [doi: 10.1111/mcn.13198] [Medline: 33960678]
28. Ntenda PAM. Association of low birth weight with undernutrition in preschool-aged children in Malawi. Nutr J 2019 Sep 02;18(1):51 [FREE Full text] [doi: 10.1186/s12937-019-0477-8] [Medline: 31477113]
29. Letamo G, Navaneetham K. Prevalence and determinants of adult under-nutrition in Botswana. PLoS One 2014 Jul 23;9(7):e102675 [FREE Full text] [doi: 10.1371/journal.pone.0102675] [Medline: 25054546]
30. Sarma H, Khan JR, Asaduzzaman M, Uddin F, Tarannum S, Hasan MM, et al. Factors influencing the prevalence of stunting among children aged below five years in Bangladesh. Food Nutr Bull 2017 Sep 30;38(3):291-301. [doi: 10.1177/0379572117710103] [Medline: 28758423]
31. Hanson C, Allen E, Fullmer M, O'Brien R, Dearden K, Garn J, et al. A national communication campaign in Indonesia is associated with improved WASH-related knowledge and behaviors in Indonesian mothers. Int J Environ Res Public Health 2020 May 25;17(10):3727 [FREE Full text] [doi: 10.3390/ijerph17103727] [Medline: 32466193]
32. Sharma AK, Baig VN, Yadav AK, Bharadwaj AK, Singh R. Prevalence and risk factors for stunting among tribal under-five children at south-west, Rajasthan, India. National J Comm Med 2016:461-467.
33. Khan S, Hancioglu A. Multiple indicator cluster surveys: delivering robust data on children and women across the globe. Stud Fam Plann 2019 Sep;50(3):279-286 [FREE Full text] [doi: 10.1111/sifp.12103] [Medline: 31486080]
34. Tingley D, Yamamoto T, Hirose K, Keele L, Imai K. mediation:R package for causal mediation analysis. J Stat Soft 2014;59(5):1-38. [doi: 10.18637/jss.v059.i05]
35. Imai K, Keele L, Tingley D, Yamamoto T. Unpacking the black box of causality: learning about causal mechanisms from experimental and observational studies. Am Polit Sci Rev 2011 Nov 10;105(4):765-789. [doi: 10.1017/s0003055411000414]
36. Imai K, Keele L, Yamamoto T. Identification, inference and sensitivity analysis for causal mediation effects. Statist Sci 2010 Feb 1;25(1):51-71. [doi: 10.1214/10-sts321]
37. Rahman A. Significant risk factors for childhood malnutrition: evidence from an Asian developing country. Sci J Pub Health 2016 Jan;4(1):16-17 [FREE Full text]

38. Rajaram S, Zottarelli LK, Sunil TS. Individual, household, programme and community effects on childhood malnutrition in rural India. Matern Child Nutr 2007 Apr;3(2):129-140 [FREE Full text] [doi: 10.1111/j.1740-8709.2007.00086.x] [Medline: 17355445]
39. Rahman A, Chowdhury S, Hossain D. Acute malnutrition in Bangladeshi children: levels and determinants. Asia Pac J Public Health 2009 Jul 29;21(3):294-302. [doi: 10.1177/1010539509335399] [Medline: 19403548]
40. Rahman A, Chowdhury S. Determinants of chronic malnutrition among preschool children in Bangladesh. J Biosoc Sci 2006 Mar 28;39(2):161-173. [doi: 10.1017/s0021932006001295]
41. Parvez SM, Azad R, Rahman M, Unicomb L, Ram PK, Naser AM, et al. Achieving optimal technology and behavioral uptake of single and combined interventions of water, sanitation hygiene and nutrition, in an efficacy trial (WASH benefits) in rural Bangladesh. Trials 2018 Jul 06;19(1):358 [FREE Full text] [doi: 10.1186/s13063-018-2710-8] [Medline: 29976251]
42. Prüss-Üstün A, Bos R, Gore F, Bartram J. Safer water, better health : costs, benefits and sustainability of interventions to protect and promote health. World Health Organization. 2008. URL: https://apps.who.int/iris/handle/10665/43840 [accessed 2022-02-14]
43. Cumming O, Cairncross S. Can water, sanitation and hygiene help eliminate stunting? current evidence and policy implications. Matern Child Nutr 2016 May;12 Suppl 1:91-105 [FREE Full text] [doi: 10.1111/mcn.12258] [Medline: 27187910]
44. Black R, Allen L, Bhutta Z, Caulfield L, de Onis M, Ezzati M, Maternal and Child Undernutrition Study Group. Maternal and child undernutrition: global and regional exposures and health consequences. Lancet 2008 Jan 19;371(9608):243-260. [doi: 10.1016/S0140-6736(07)61690-0] [Medline: 18207566]
45. Bhutta Z, Ahmed T, Black R, Cousens S, Dewey K, Giugliani E, Maternal and Child Undernutrition Study Group. What works? interventions for maternal and child undernutrition and survival. Lancet 2008 Feb 02;371(9610):417-440. [doi: 10.1016/S0140-6736(07)61693-6] [Medline: 18206226]
46. Naugle DA, Hornik RC. Systematic review of the effectiveness of mass media interventions for child survival in low- and middle-income countries. J Health Commun 2014 Sep 10;19 Suppl 1(sup1):190-215 [FREE Full text] [doi: 10.1080/10810730.2014.918217] [Medline: 25207453]
47. Anjorin SS, Uthman OA, Ameyaw EK, Ahinkorah BO, Chattu VK, Bishwajit G, et al. Undernutrition, polygynous context and family structure: a multilevel analysis of cross-sectional surveys of 350 000 mother-child pairs from 32 countries. BMJ Glob Health 2020 Oct 15;5(10):e002637 [FREE Full text] [doi: 10.1136/bmjgh-2020-002637] [Medline: 33060094]
48. Hall C, Syafiq A, Crookston B, Bennett C, Hasan MR, Linehan M, et al. Addressing communications campaign development challenges to reduce stunting in Indonesia. Health 2018;10(12):1764-1778. [doi: 10.4236/health.2018.1012133]
49. Rao KV, Mishra VK, Retherford RD. Knowledge and use of oral rehydration therapy for childhood diarrhoea in India: effects of exposure to mass media. Scholarspace. 1998 Nov 10. URL: http://scholarspace.manoa.hawaii.edu/handle/10125/3478 [accessed 2022-12-14]
50. Bowen HL. Impact of a mass media campaign on bed net use in Cameroon. Malar J 2013 Jan 25;12(1):36 [FREE Full text] [doi: 10.1186/1475-2875-12-36] [Medline: 23351674]
51. Ankomah A, Adebayo SB, Arogundade ED, Anyanti J, Nwokolo E, Inyang U, et al. The effect of mass media campaign on the use of insecticide-treated bed nets among pregnant women in Nigeria. Malar Res Treat 2014 Mar 20;2014:694863-694867 [FREE Full text] [doi: 10.1155/2014/694863] [Medline: 24778895]
52. Jung M, Lin L, Viswanath K. Effect of media use on mothers' vaccination of their children in sub-Saharan Africa. Vaccine 2015 May 21;33(22):2551-2557. [doi: 10.1016/j.vaccine.2015.04.021] [Medline: 25896379]
53. The future of the public's health in the 21st century. NCBI. 2002. URL: https://www.ncbi.nlm.nih.gov/books/NBK221224/ [accessed 2022-12-14]
54. Williams G, Hamm MP, Shulhan J, Vandermeer B, Hartling L. Social media interventions for diet and exercise behaviours: a systematic review and meta-analysis of randomised controlled trials. BMJ Open 2014 Feb;4(2):e003926 [FREE Full text] [doi: 10.1136/bmjopen-2013-003926] [Medline: 24525388]
55. Vance K, Howe W, Dellavalle RP. Social internet sites as a source of public health information. Dermatol Clin 2009 Apr;27(2):133-6, vi. [doi: 10.1016/j.det.2008.11.010] [Medline: 19254656]

Abbreviations

MICS: Multiple Indicators Cluster Surveys
WASH: Water, Sanitation, and Hygiene program
WHO: World Health Organization
UNICEF: United Nation Children’s Fund

Edited by Y Khader; submitted 05.09.21; peer-reviewed by C Ma, B Sindelar; comments to author 19.11.21; revised version received 24.12.21; accepted 28.12.21; published 06.04.22.

Please cite as:
Huo S, Wang K, Liu Z, Yang Y, Hee JY, He Q, Takesue R, Tang K
Influence of Maternal Exposure to Mass Media on Growth Stunting Among Children Under Five: Mediation Analysis Through the Water, Sanitation, and Hygiene Program
JMIR Public Health Surveill 2022;8(4):e33394
URL: https://publichealth.jmir.org/2022/4/e33394
doi:10.2196/33394
PMID:35384848

Original Paper

Design of a Vaccine Passport Validation System Using Blockchain-based Architecture: Development Study

Hsiu An Lee[1,2], PhD; Wei-Chen Wu[3], PhD; Hsin-Hua Kung[1,2], MS; Jai Ganesh Udayasankaran[2,4], MSc, MBA; Yu-Chih Wei[5], PhD; Boonchai Kijsanayotin[2,6,7], MSc, MD, PhD; Alvin B Marcelo[2,8], MD; Chien-Yeh Hsu[2,9,10], PhD

[1]National Institute of Cancer Research, National Health Research Institutes, Tainan, Taiwan

[2]Standard and Interoperability Lab, Smart Healthcare Center of Excellent Taiwan, Taipei, Taiwan

[3]Department and Graduate Institute of Finance, National Taipei University of Business, Taipei, Taiwan

[4]Sri Sathya Sai Central Trust, Prasanthi Nilayam, Puttaparthi, India

[5]Department of Information and Finance Management, National Taipei University of Technology, Taipei, Taiwan

[6]Department of Clinical Epidemiology and Biostatistics, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, Thailand

[7]Thai Health Information Standards Development Center, Health System Research Institute, Ministry of Public Health, Bangkok, Thailand

[8]University of the Philippines, Manila, Philippines

[9]Department of Information Management, National Taipei University of Nursing and Health Sciences, Taipei, Taiwan

[10]Master Program in Global Health and Development, Taipei Medical University, Taipei, Taiwan

Corresponding Author:
Chien-Yeh Hsu, PhD
Department of Information Management
National Taipei University of Nursing and Health Sciences
No.365, Ming-te Road
Peitou District
Taipei, 112
Taiwan
Phone: 886 0939193212
Email: cyhsu@ntunhs.edu.tw

Abstract

Background: COVID-19 is an ongoing global pandemic caused by SARS-CoV-2. As of June 2021, 5 emergency vaccines were available for COVID-19 prevention, and with the improvement of vaccination rates and the resumption of activities in each country, verification of vaccination has become an important issue. Currently, in most areas, vaccination and reverse transcription polymerase chain reaction (RT-PCR) test results are certified and validated on paper. This leads to the problem of counterfeit documents. Therefore, a global vaccination record is needed.

Objective: The main objective of this study is to design a vaccine passport (VP) validation system based on a general blockchain architecture for international use in a simulated environment. With decentralized characteristics, the system is expected to have the advantages of low cost, high interoperability, effectiveness, security, and verifiability through blockchain architecture.

Methods: The blockchain decentralized mechanism was used to build an open and anticounterfeiting information platform for VPs. The contents of a vaccination card are recorded according to international Fast Healthcare Interoperability Resource (FHIR) standards, and blockchain smart contracts (SCs) are used for authorization and authentication to achieve hierarchical management of various international hospitals and people receiving injections. The blockchain stores an encrypted vaccination path managed by the user who manages the private key. The blockchain uses the proof-of-authority (PoA) public chain and can access all information through the specified chain. This will achieve the goal of keeping development costs low and streamlining vaccine transit management so that countries in different economies can use them.

Results: The openness of the blockchain helps to create transparency and data accuracy. This blockchain architecture contains a total of 3 entities. All approvals are published on Open Ledger. Smart certificates enable authorization and authentication, and encryption and decryption mechanisms guarantee data protection. This proof of concept demonstrates the design of blockchain architecture, which can achieve accurate global VP verification at an affordable price. In this study, an actual VP case was established and demonstrated. An open blockchain, an individually approved certification mechanism, and an international standard vaccination record were introduced.

Conclusions: Blockchain architecture can be used to build a viable international VP authentication process with the advantages of low cost, high interoperability, effectiveness, security, and verifiability.

(JMIR Public Health Surveill 2022;8(4):e32411) doi:10.2196/32411

KEYWORDS

COVID-19; vaccine passport; global border control; health policy; international infectious disease strategy; vaccine; policy; strategy; blockchain; privacy; security; testing; verification; certification; Fast Healthcare Interoperability Resource

Introduction

Background

COVID-19 is an ongoing pandemic caused by SARS-CoV-2. The disease was discovered at the end of 2019 [1] in Wuhan City, Hubei Province, People's Republic of China, which rapidly spread to many countries worldwide in early 2020 and eventually became a global pandemic. More than 170 million confirmed cases had been reported in countries and regions worldwide as of June 2021, with over 3.7 million deaths [2-4], making COVID-19 1 of the largest epidemics in human history. There are at least 287 COVID-19-preventive vaccines available in the world [5], with 5 vaccines (Pfizer/BioNTech, Astrazeneca-SK Bio, Serum Institute of India, Janssen, and Moderna) available for emergency use in June 2021, but none have completed clinical trials. Nonetheless, as vaccines are developed, disease conditions in various countries are gradually being controlled.

Currently, all vaccination or reverse transcription polymerase chain reaction (RT-PCR) test results are certified and verified on paper by hospitals or testing institutions. Meanwhile, falsified documents pose a significant risk in confirming vaccines or conducting testing. At the end of 2020, the European Union (EU) began discussions on the design of vaccine passports (VPs). The official policies and measures were announced in June 2021 and went into effect in July 2021. The EU Digital COVID Certificate is available in both digital and paper formats, and it includes an official quick response (QR) code to ensure its authenticity. As a result, Japan began to design VPs as well. Although, based on the current strategy, the VP may become an important key to future global activities, the problems of document verification and data use among countries have yet to be resolved.

The Vaccination Certificate Program (VCI) is a consortium of 12 information technology and health care organizations, including Microsoft, Oracle, and the Mayo Clinic [6-8]. The project's goal is to provide a secure and decentralized solution for open licenses that is interoperable and widely adopted across multiple platforms. The software they proposed enables the exchange of verifiable clinical data via smart health cards or digital health wallets. As a result, it will require an ecosystem of interoperable applications, data, and processes. The VCI provides digital access to vaccination records by utilizing smart health cards open specifications based on W3C Verifiable Credential and Health Level Seven International (HL7) Fast Healthcare Interoperability Resource (FHIR) standards. The VCI solution will necessitate the use of a token (smart card) and membership in an alliance, both of which are not required in our blockchain architecture.

Our team proposed “a blockchain-based global infectious disease surveillance and case-tracking model for COVID-19” [9] by the end of 2020. Although the primary goal of the COVID-19 defense was to prevent the virus from spreading across borders, we demonstrated how the blockchain can help with defense and ensure that information is correct. Another study conducted by our team demonstrated the viability of blockchain-based global information exchange [10,11]. Gstrein [12] proposed using blockchain to establish immunity certificates. The authors also suggested using immutable blockchain technology to prevent the spread of fraudulent information and reports. Furthermore, this research tries to address the issues of candidate privacy and anonymity.

Objectives

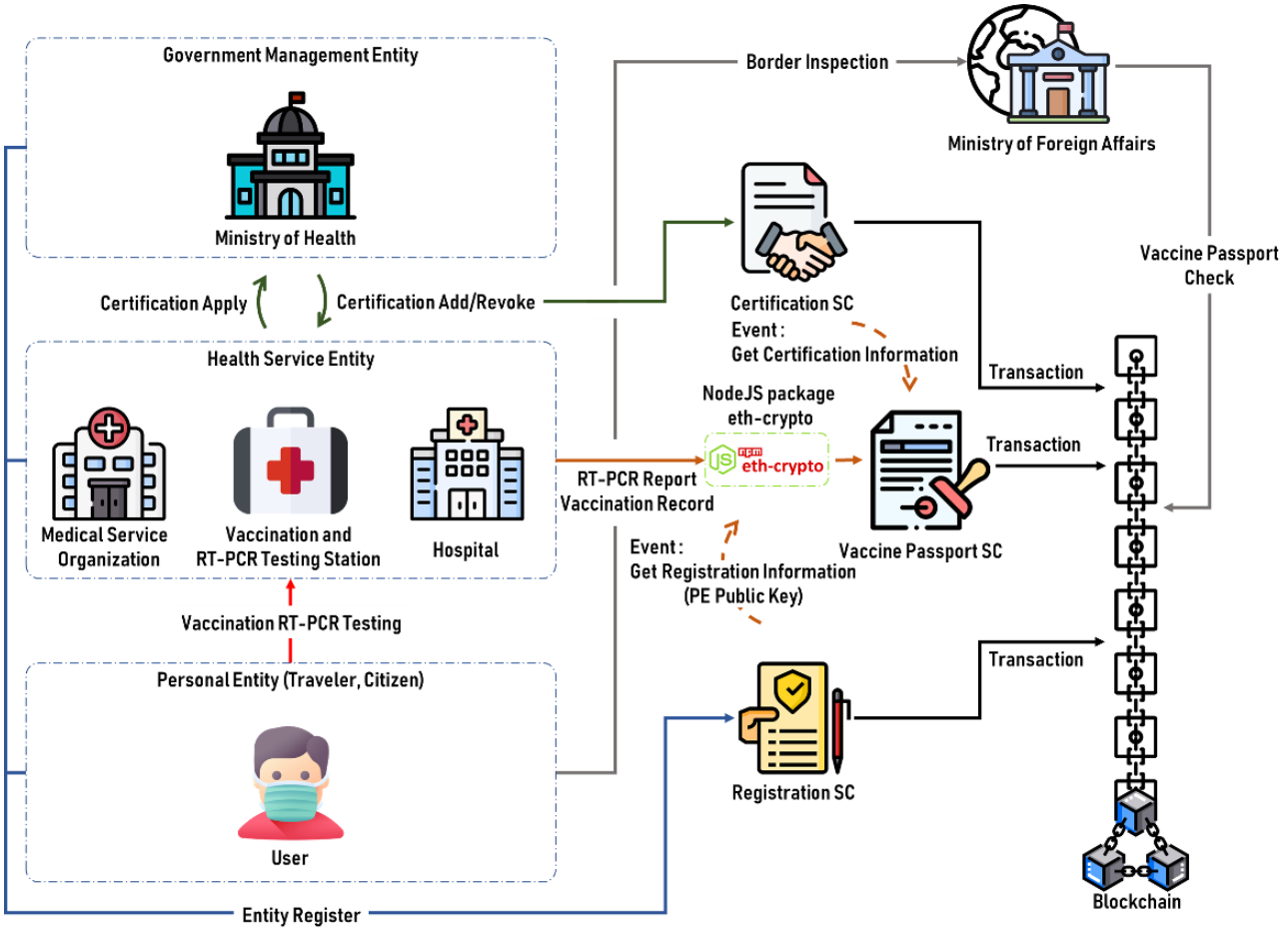
The worldwide VP verification mechanism is currently incomplete. There is no universal standard for digital VPs, and paper VPs are suspect of fabrication. Many developing nations cannot afford the worldwide data integration structure, since it requires a large number of resources and money. To address these issues, this study provides an innovative study of a global architecture that uses blockchain to complete the verification and distribution of VPs, such as RT-PCR testing results or vaccination records. The blockchain's public verification is utilized for passport verification, while smart contracts (SCs) with various purposes verify that a user's VP is valid. Finally, the public chain technique can be used to save building costs and create a worldwide VP chain.

Methods

Study Design

The architectural design application is divided into 2 parts: entity architecture and blockchain architecture. Basically, the system is an architecture of multientities interacting with each other using SCs built on the blockchain. Government-managed entities (Ministry of Health [MoH]), health service entities (eg, medical service organizations, medical stations, and hospitals), personal entities (eg, passengers, citizens, and patients), and border control entities (BCEs) are all classed according to their duties (Ministry of Foreign Affairs). The blockchain includes SCs and the public chain. Individual functions are discussed later. An overview of the blockchain architecture for VPs is shown in Figure 1.

Figure 1. Overview of the blockchain architecture for VPs. PE: personal entity; RT-PCR: reverse transcription polymerase chain reaction; SC: smart contract; VP: vaccine passport.



Entities' Identification and Application Process

Government-Managed Entity

In this study, a government-managed entity (GME) is a government agent that issues certificates to health service organizations to identify that the vaccination certificates or testing results submitted by the organizations can be trusted at border inspections. In addition, the identification of a GME, including the MoH, the Centers for Disease Control and Prevention (CDC), and the COVID-19 Central Command Center, is recognized by every country's government.

Health Service Entity

A health service entity (HSE) is described as a facility that can provide a VP. Medical service organizations, medical stations, and hospitals are all included. Every HSE should request a certificate from the local GME to demonstrate that the former's reports can be trusted. The certificate will then be recorded in the blockchain through the certificate SC. When a personal entity (PE) accepts the HSE vaccine or RT-PCR testing services, the latter should get the PE's blockchain address. After the VP is completed, it is encrypted using *eth-crypto*, a Node.js-supported encryption software program. The encrypted VP is then recorded in the blockchain via the VP SC. The complete procedure is outlined later.

Personal Entity

A PE is described as a normal person seeking a VP via vaccination or RT-PCR testing in an HSE. When a PE registers on the blockchain via the SC, their blockchain address is generated. The identity of a PE includes travelers, citizens, patients, and others who want to get a VP. When a PE obtains a vaccine or RT-PCR testing from an HSE, the PE must supply the HSE with a blockchain address. Upon completion, the VP is transferred into the blockchain through the VP SC. The private key of the PE can then be used to decode the VP contained in the chain.

Border Control Entity

A BCE is defined as the unit in each country responsible for border control. The Ministry of Foreign Affairs is part of a BCE's identity. During border inspections, a BCE gets encrypted VPs of PEs from the blockchain, where the PEs are asked to provide a private key to decrypt their individual encrypted VPs. The problem of falsified documents is handled because all VPs come from the blockchain and are provided by verified HSEs. Furthermore, countries can reach an agreement to allow BCEs to accept validated VPs from various HSEs as long as the HSHs are certified. The public chain is used by this framework to facilitate interoperability. The BCE does not require blockchain registration, since the BCE only needs to retrieve data. Additionally, in this multientity architecture, any entity can become a BCE in order to authenticate a PE's VP.

Vaccine Passport Content and Format

In this study, the World Health Organization's (WHO) interim guidelines for establishing a Smart Vaccination Certificate release was followed [13-16]. Vaccination information, the vaccination hospital, date, time, and the RT-PCR testing report should all be included in the VP. Only a certified or recognized HSE can complete a VP. In addition, the FHIR standard format was used in this research. In recent years, the FHIR has become the most widely used medical data standard and framework. FHIR adoption can increase data interoperability and usability, allowing users to quickly integrate the VP blockchain into their systems.

In the FHIR architecture, various resources are used to integrate the entire VP, which include patient (PE), organization (HSE), practitioner (administrative or medical staff), immunization (first or second dose injection records), immunization recommendation (second dose of vaccine reservation information), and observation (RT-PCR/quick screening and testing report). The FHIR resource structure of the VP is provided in Multimedia Appendix 1.

Blockchain Architecture and the Functions of a Smart Contract

Blockchain Architecture

To reduce resource consumption and VP usage restrictions, this study used the proof-of-authority (PoA) consensus technique, which allows authorized nodes to join as block-establishing nodes. The number of nodes in the PoA consensus process is unlimited, but the number of validators is limited. The nodes are in charge of synchronizing the blockchain ledger, while the verifier is in charge of confirming transactions and packaging blocks. The blockchain of the PoA consensus method has the potential to outperform conventional and decentralized public chains, such as Bitcoin and Ethereum, in terms of efficiency and scalability, due to the limited number of validators.

Authorized nodes are set as government agencies of various countries, while countries worldwide jointly maintain the verification of blocks. The blockchain architecture was Ethereum's public chain, and the Ethereum protocol was used to move VPs to the blockchain architecture, create a new block, and connect it to the blockchain using Geth (Go Ethereum). Due to the regulations of different countries or organizations, it is acceptable if certain BCE-permitted VPs for PEs are created by an HSE that has not obtained a certificate from a GME.

SCs of Different Functions

A total of 3 SCs are used in the research architecture: registration, VP, and certification SCs.

The registration SC is used to store the public keys of all users, which are used to encrypt VPs.

The VP SC is used by HSE to generate VPs, including VP encryption and storage of HSE certification information.

The certification SC is used to issue the certification authority of the organization, and the main purpose is that each country can develop and certify its own recognized HSE. When the BCE checks the blockchain VP information, it can quickly confirm the HSE certification.

Results

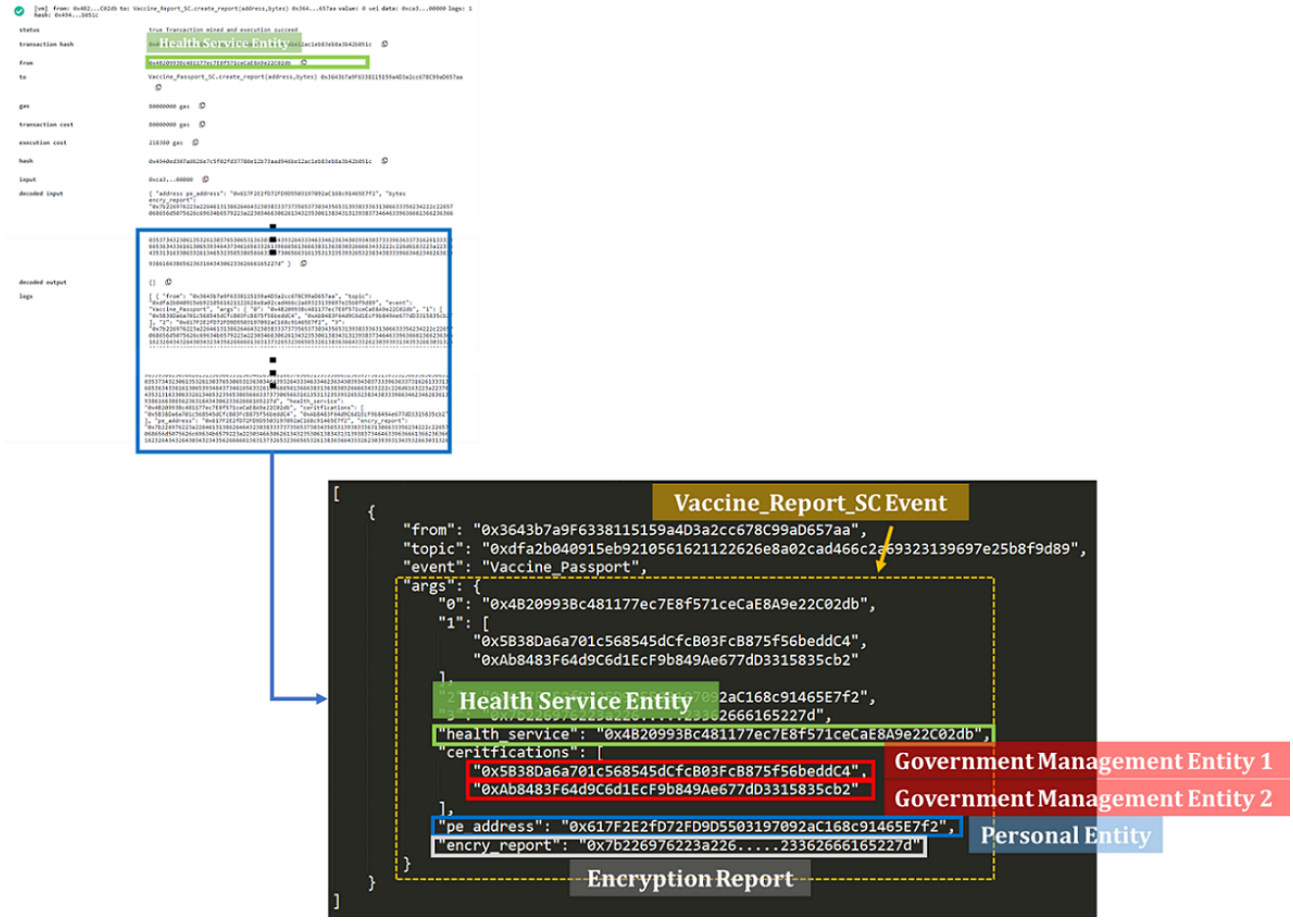
In the scenario, there are 3 entities: a PE seeking a VP, an HSE who can administer vaccinations, and a GME representing the MoH.

VP Verification

A PE wants to travel to other countries. The BCE system uses the blockchain to retrieve the encrypted VP. The PE's private key is required to decrypt the VP after passing HSE authentication. After that, the PE goes through VP verification to pass the border check. Figure 2 shows the PE's VP data recorded in the blockchain, which includes the HSE, HSE certificate, PE, and encrypted VP.

After obtaining the block data, the PE should provide a private key to unlock the encrypted VP. Then, the decrypted VP is presented in FHIR format, which can be read through the front-end graphical user interface (GUI).

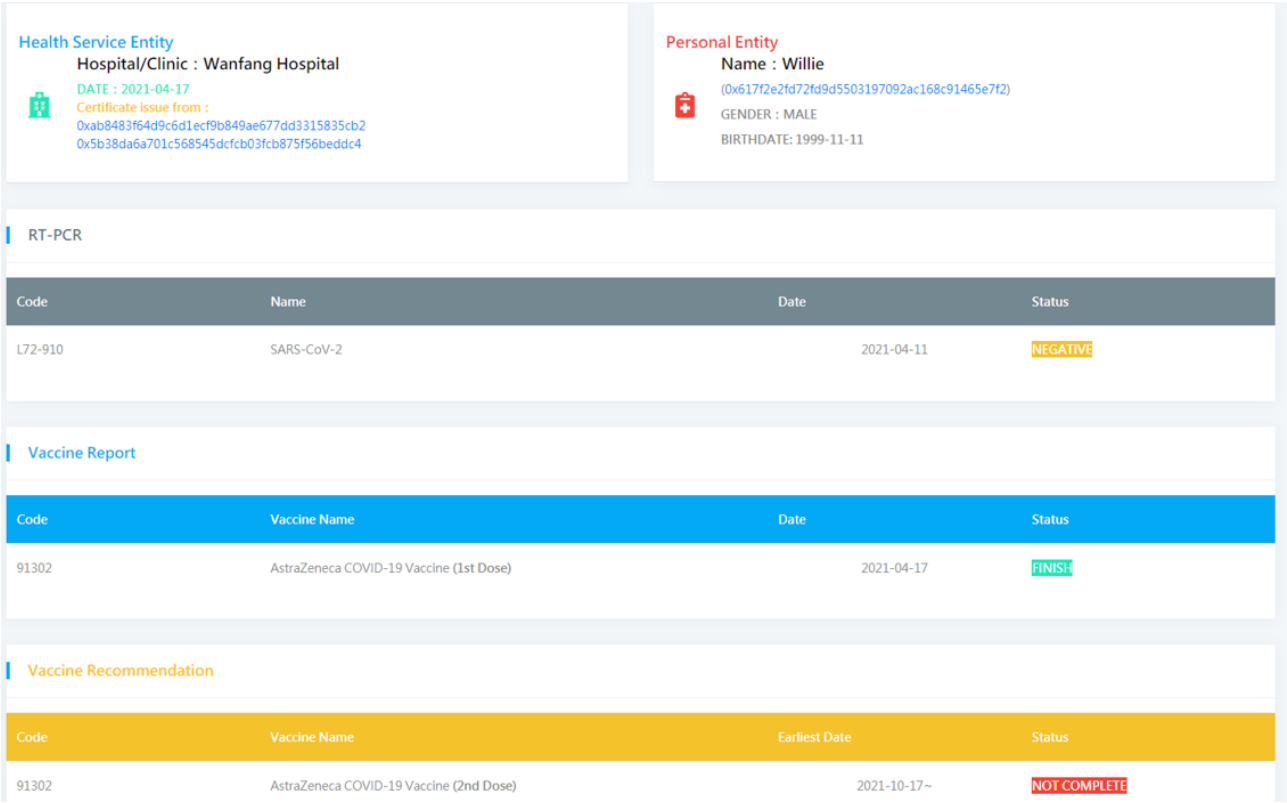
Figure 2. The blockchain content of a VP. VP: vaccine passport.



VP Content and International Standard

In this study, the VP contained (1) patient information (eg, name, national ID number/residence permit number/passport number, biological sex, nationality, birthday), (2) vaccine information (eg, vaccine type, vaccine brand, manufacturer, number of doses, vaccination date, name/code of the medical institution responsible for vaccination, name/code of the medical staff who administered the vaccine, country where the vaccination was given, date of next vaccination record), and (3) RT-PCR test reports (eg, disease name [COVID-19 or SARS-CoV-2], inspection date, report output date, testing method [needs to be tested by molecular biology nucleic acid], and check report results, including negative, positive, and undetectable). To verify the VP, the inspection unit should double-check its content and the information obtained in the block. The VP presented is shown in Figure 3.

Figure 3. The GUI of an electronic VP for border inspection. GUI: graphical user interface; RT-PCR: reverse transcription polymerase chain reaction; VP: vaccine passport.



Testing Feedback From a Different Country's User

This study proposed a blockchain architecture for the authentication process of an executable international VP, with the advantages of low cost, high interoperability, effectiveness, security, and verifiability. In Southeast Asian countries, a regional travel protocol has been designed and cross-nationally tested. The findings reveal that the proposed architecture is capable of quickly verifying the VP for international travel.

The international standard format of the HL7 FHIR was designed in this architecture to ensure interoperability. The BCE can easily obtain a traveler's validated VP, which has been verified by the source country's supervisory unit.

Testing was performed on a simulated public chain environment. A total of 3 testing scenarios were tested in this study. In the testing process, 3 identities were registered. Both the certification SC and the VP SC required the signature of the private key. The private key was also used to decrypt the VP stored on the chain. All the users can use the web-based GUI shown in Figure 3.

The testing process was designed to simulate the scenario of a traveler using a VP. The simulation process was as follows:

1. The GME (country X MoH), the HSE (hospital A), and the PE (traveler) register a VP blockchain account.
2. Hospital A applies for authorization from the country X MoH through an SC.
3. The country X MoH announces the blockchain account to foreign BCEs for foreign verification of the permission (from the country X MoH) of hospital A.
4. The traveler goes to hospital A receiving vaccination and provides the blockchain account to hospital A.
5. Hospital A produces a VP for the traveler after vaccination and uploads the traveler's VP to the VP blockchain via the SC. The traveler receives an index of the VP on the blockchain.
6. The traveler travels abroad and provides the index of the VP to the foreign BCE.
7. The BCE obtains encrypted VP data.
8. The traveler provides a private key (obtained at step 1) for VP decryption.
9. The BCE confirms the VP content and the authorization of hospital A.
10. The verification is completed, allowing the traveler to enter the country.

The GUI of step 5 (VP production) is shown in Figure 4. The GUI of steps 7-9 is shown in Figure 5.

In the first scenario, this was performed according to our test process. A web-based user portal was used to present the contents of the VP, as shown in Figure 3. The data of the VP were encrypted and stored in a block, which needed to be decrypted using the private key of the PE.

In the second scenario, we attempted to create a fake VP. With the design of the SC and decryption mechanism, if the PE tries to authenticate with another person's VP, the encrypted VP will not be decrypted due to a private key error.

In the third scenario, we tested VPs generated by the HSE that have not obtained any GME certificates. When decrypting such VPs, the hospital certificate field is emptied. Therefore, the VP is considered an invalid VP, when checked.

The platform is currently in its testing stage, and there is a small number of users on the architecture chain. The users' comments are summarized as follows:

- Blockchain architecture helps to verify the correctness of data.
- The construction cost of the blockchain architecture is low, and it can be quickly deployed and expanded to local areas.
- The protection provided by blockchain technology can make users believe that the system is secure.
- VP-issuing units and verification units can be clearly displayed in the VP to improve the safety of border control.

Figure 4. The GUI of VP production. FHIR: Fast Healthcare Interoperability Resource; GUI: graphical user interface; HSE: health service entity; VP: vaccine passport.

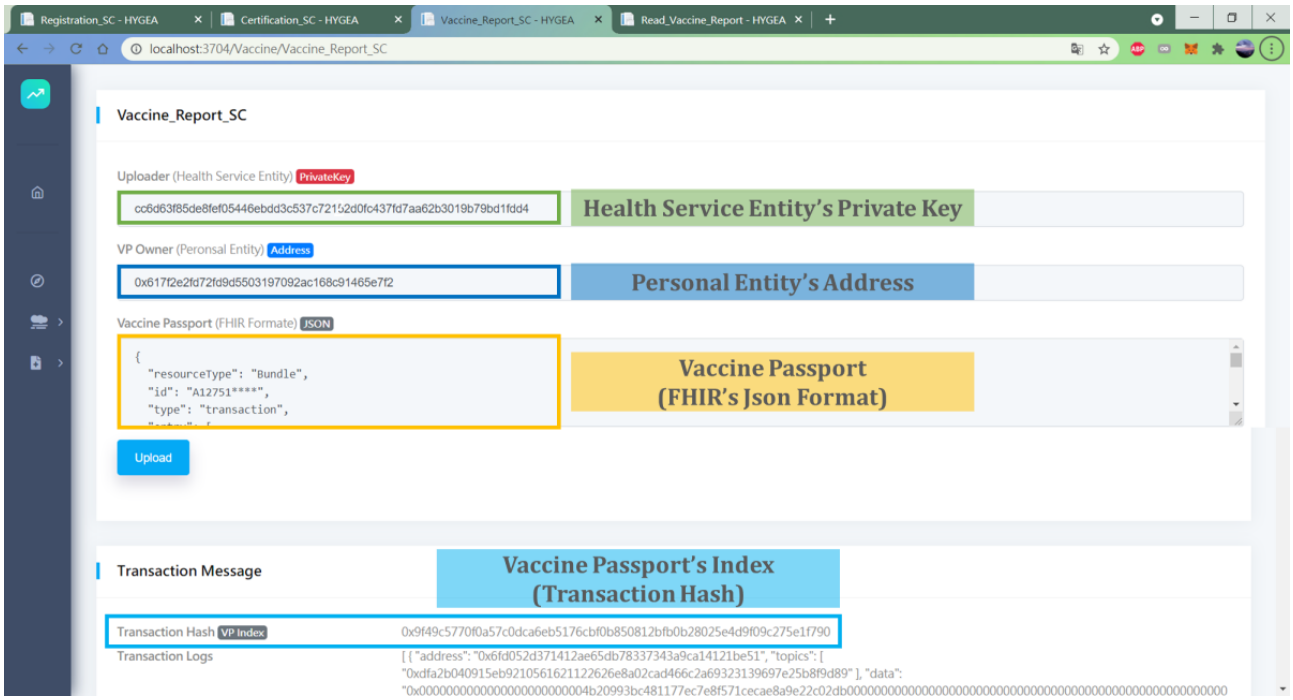
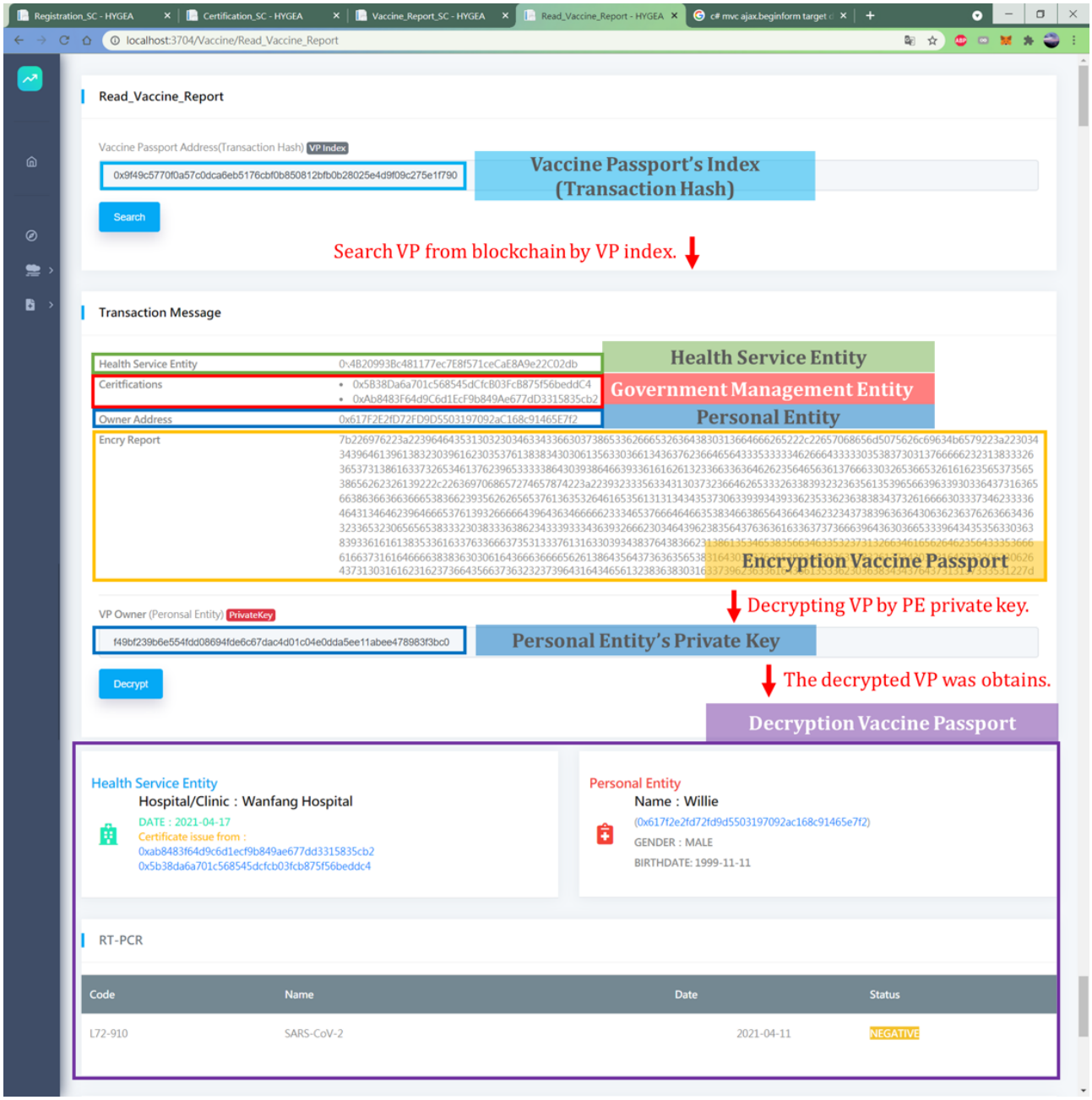


Figure 5. The GUI of VP decryption and VP content. GUI: graphical user interface; PE: personal entity; RT-PCR: reverse transcription polymerase chain reaction; VP: vaccine passport.



Efficiency and Cost of Public Blockchain Architecture

In the traditional authentication mechanism, facilities such as servers, users, and encrypted secure networks are required. The cost includes a hardware environment, software platform development, and personnel maintenance costs. However, the blockchain architecture under the PoA architecture can save these costs (Table 1).

In this study, a set of blockchain test environments was constructed for comparison.

The PoA blockchain is built on 3 basic computers (because it is PoA, it does not require supercomputing power). The BCE needs a general internet connection to obtain the encrypted VP on the blockchain. The user only needs to provide the private key to decrypt the personal VP.

Table 1. The cost comparison table between blockchain and traditional architectures.

Facility	Blockchain-based architecture	Traditional architecture
Server	3 general servers	All VP[a] providers must have advanced servers (each country at least 1 server).
Internet	General internet	High-security protection.
System maintainer	At least 1 for each service server	At least 1 for each service server.

[a]VP: vaccine passport.

Discussion

Principal Findings

The major goal of this study was to design a general VP verification architecture for international use. The decentralized mechanism of the blockchain was used to build an open and unforgeable VP information platform. The blockchain architecture proposed in this research is a single standardized vaccine verification system framework that can be jointly created by the public and private sectors. The content of the VP is recorded using the FHIR international standard, and the SC of the blockchain is used for authorization and authentication to achieve hierarchical management of different international hospitals and individuals who receive the injection. The blockchain stores the encrypted VP, which is managed by the user who manages private keys. The blockchain uses the public chain of the PoA, and all information can be accessed on the designated chain, achieving the goal of low development cost and high efficiency in VP management, so countries in different economic states can use it. Additionally, the vaccine holder only needs to bring the private key—text, barcode, or cold wallet document—to complete the VP verification. This free authentication mechanism convinces government agencies, border control, airlines, hotels, department stores, restaurants, educators, and others to believe that these data are authentic and reliable and can design different authentication procedures to issue authorization verification certificates freely.

The FHIR is a standard issued by HL7. It evolved from the HL7 v2, HL7 v3, and HL7 Clinical Document Architecture (CDA) standards and aims to be easier to implement [1]. However, the interoperability between different systems in the medical field is poor. There are many problems that need to be solved for system interoperability. HL7 developed the FHIR as the basis for achieving interoperability. The FHIR is a widely used health information standard that describes the data format and data elements of electronic health records (EHRs). Schleyer et al [2] developed an FHIR-based medical dashboard that integrates clinical data and EHRs from the hospital information system (HIS). The test results after use by medical institutions show that the integrated dashboard is useful. Based on the FHIR, clinical information can be effectively integrated [2]. The FHIR uses a widely adopted network technology, is friendly to the same standard system and independent of users, and provides useful applications and format frameworks for EHR providers, health care providers, and public health [3]. Recently, many studies and industry products have demonstrated how the FHIR can be used for health care data integration [4-11] and have discussed how the FHIR can achieve interoperability between different health care systems.

Public and private organizations worldwide are looking for ways to fight COVID-19 in the hopes of finding a strategy and solution that can help society return to “normal.” Vaccination is the most critical issue at present. When herd immunity is achieved, what follows is how to restore “normal” activities worldwide.

The political issues and policies of different countries create problems when exchanging VPs between them. For example, the EU Digital COVID Certificate Regulation [12] was available since July 1, 2021. EU residents and citizens can have their EU Digital COVID Certificates in QR code or paper format issued and verified across the EU countries. That is the first solution proposed to prove an individual’s vaccination status. Presently, the United States uses handwritten paper certificates as proof of vaccination. Some states also have digital passes, but they have not constructed a complete verification and international use strategy. However, although paper documents are usually illegible or incorrect, they are also easy to lose and forge. Moreover, the mobile app–based authentication system itself is not complete, proving ineffective on a global scale amidst the majority’s lack of access to smartphones.

These verification certificates can allow organizations to plan on their own, such as identifying which organization accepts the VP, who has been certified by the organization, and whether the VP is authentic and has not been revoked. Therefore, the blockchain and public ledgers can be used as the frontline application of the VP verification system. If properly operated, the public ledgers cannot be forged and can reliably [13] prove an individual’s vaccination status.

The blockchain was originally used in the medical field and faced challenges related to transparency and confidentiality because “everyone can see everything” on the blockchain network [14-16]. The increasing transparency and decline of confidentiality, such as information on transmissions, are generally considered limitations of the blockchain. However, this limitation is an advantage in the process of using VPs. As VP holders need to prove to everyone that they have been vaccinated, such change gives the application of the blockchain a good advantage and the blockchain can be quickly deployed at low cost. Furthermore, different application SCs are developed in this research framework. Through SCs, entities of different levels can make individual authentication policies, which greatly improves usability and interoperability. Furthermore, cryptocurrency is inherently contained in the blockchain VP system of our design. If there is a need for payment for VP exchange and verification, the function of payment is available in our system.

Future Directions

By comparison with prior work, there are several important points that need to be studies in the future. First, blockchain technology ensures data security and privacy and has been successfully used in different domains. However, more study is needed to demonstrate the benefits of using the blockchain for VPs for international users. Second, for the blockchain architecture of VPs to work smoothly, it is preferred that a cross-country VP alliance service system be developed for international C0VID-19 defense.

Conclusions

Although many countries endeavor to reopen their economies and allow inward and outward tourism, there is still uncertainty on the state of health of citizens from different countries with varying degrees of policies and types of vaccination. The use of VPs is 1 of the ways to promote the recovery of global activities. Unfortunately, at present, VPs are still difficult for many countries to implement. The verification, trust, and effectiveness of paper documents must be considered. The simple blockchain architecture proposed in this research can be implemented collaboratively by public and private entities and be rapidly expanded. The open nature of the blockchain contributes to establishing transparency and data accuracy. The smart certificate enables authorization and authentication, while the encryption and decryption mechanism ensures data protection. To make it globally available and accessible, the FHIR international data standard was adopted in this research.

The principal findings in this study are as follows: First, blockchain architecture was used to build the authentication process of an executable international VP, with advantages of low cost, high interoperability, effectiveness, security, and verifiability. Second, the international data standard FHIR was adopted in this research. Third, this PoA demonstrated the design of blockchain architecture that, when adopted, can accurately achieve global VP verification and at a cost any country can afford. Fourth, the platform has been tested by several users in different countries in the Asia eHealth Information Network (AeHIN) and has shown that it is a suitable platform for VP verification.

Acknowledgments

This project received funding from the Ministry of Education, Taiwan (Projects 107EH12-404 and 107EH12-402), and the Ministry of Science and Technology, Taiwan (Projects 109-2221-E-227-003-MY2 and MOST110-2321-B-195-001).

Authors' Contributions

The work presented in this paper was carried out in collaboration among all authors. HAL, HHK, and CYH conceptualized the study and research design. HAL, WCW, HHK, and YCW designed the architecture of the system. HAL and JGU carried out the literature review and system analysis. HHK put a lot of effort in the development of the system. HAL drafted the manuscript, and CYH made significant revisions. WCW, YCW, BK, and ABM remotely tested the system. CYH, JGU, BK, and ABM supervised the methods of the testing on a vaccine passport (VP) and suggested valuable improvements. All authors approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The FHIR resource structure of a VP, encrypted VP contents, and functions of an SC in a blockchain. FHIR: Fast Healthcare Interoperability Resource; SC: smart contract; VP: vaccine passport.
[DOCX File , 2143 KB - publichealth_v8i4e32411_app1.docx]

References

1. HL7. Home - FHIR DSTU 1 (v0.0.82). URL: http://wiki.hl7.org.cn:81/index.php?title=Home_-_FHIR_DSTU_1_(v0.0.82) [accessed 2022-04-20]
2. Schleyer TKL, Rahurkar S, Baublet AM, Kochmann M, Ning X, Martin DK, FHIR Development Team, et al. Preliminary evaluation of the Chest Pain Dashboard, a FHIR-based approach for integrating health information exchange information directly into the clinical workflow. AMIA Jt Summits Transl Sci Proc 2019;6:656-664. [doi: 10.1007/978-3-319-93414-3_5]
3. Goodman A, Braun P, Braunstein M. Pediatrics 2018;141:12. [doi: 10.1542/peds.141.1MA1.12]
4. Storck M, Hollenberg L, Dugas M, Soto-Rey I. Interoperability improvement of Mobile Patient Survey (MoPat) implementing Fast Health Interoperability Resources (FHIR). Stud Health Technol Inform 2019;258:141-145. [doi: 10.4018/ijehmc.2019010105]
5. Giordanengo A, Bradway M, Grøttland A, Hartvigsen G, Årsand E. A FHIR-based Data Flow Enabling Patients with Diabetes to Share Self-Collected Data with the Norwegian National Healthcare Systems and Electronic Health Records Systems. URL: https://ehealthresearch.no/files/documents/Postere/Poster_2018-02_16-9-ATTD-FHIR-based-Data-Flow.pdf [accessed 2022-04-20]

6. Kiourtis A, Mavrogiorgou A, Menychtas A, Maglogiannis I, Kyriazis D. Structurally mapping healthcare data to HL7 FHIR through ontology alignment. J Med Syst 2019 Feb 05;43(3):62. [doi: 10.1007/s10916-019-1183-y] [Medline: 30721349]
7. Gøeg KR, Rasmussen RK, Jensen L, Wollesen CM, Larsen S, Pape-Haugaard LB. A future-proof architecture for telemedicine using loose-coupled modules and HL7 FHIR. Comput Methods Programs Biomed 2018 Jul;160:95-101. [doi: 10.1016/j.cmpb.2018.03.010] [Medline: 29728251]
8. Alper B, Mayer M, Shahin K, Richardson J, Schilling L, Tristan M, et al. 20 Achieving evidence interoperability in the computer age: setting evidence on FHIR. BMJ Evidence-Based Med 2019;24(Suppl 1):A15. [doi: 10.1136/bmjebm-2019-ebmlive.28]
9. Lackerbauer AM, Lin AC, Krauss O, Hearn J, Helm E. A model for implementing an interoperable electronic consent form for medical treatment using HL7 FHIR. Eur J Biomed Inform 2018;14(3):37-47. [doi: 10.24105/ejbi.2018.14.3.6]
10. Dixon BE, Taylor DE, Choi M, Riley M, Schneider T, Duke J. Integration of FHIR to facilitate electronic case reporting: results from a pilot study. Stud Health Technol Inform 2019 Aug 21;264:940-944. [doi: 10.3233/SHTI190362] [Medline: 31438062]
11. Hussain MA, Langer SG, Kohli M. Learning HL7 FHIR using the HAPI FHIR server and its use in medical imaging with the SIIM dataset. J Digit Imaging 2018 Jun 3;31(3):334-340 [FREE Full text] [doi: 10.1007/s10278-018-0090-y] [Medline: 29725959]
12. Gstrein OJ. The EU Digital COVID Certificate: a preliminary data protection impact assessment. Eur J Risk Regul 2021 Jul 06;12(2):370-381. [doi: 10.1017/err.2021.29]
13. Kuo TT, Kim HE, Ohno-Machado L. Blockchain distributed ledger technologies for biomedical and health care applications. J Am Med Inform Assoc 2017 Nov 01;24(6):1211-1220 [FREE Full text] [doi: 10.1093/jamia/ocx068] [Medline: 29016974]
14. Greenspan G. MultiChain Private Blockchain — White Paper. URL: https://www.multichain.com/download/MultiChain-White-Paper.pdf [accessed 2022-04-20]
15. Moser M. Anonymity of bitcoin transactions: an analysis of mixing services. 2013 Presented at: Münster Bitcoin Conference; July 17-18, 2013; Münster, Germany.
16. De Filippi P. The interplay between decentralization and privacy: the case of blockchain technologies. J Peer Prod 2016(7):1-19.

Abbreviations

BCE: border control entity
EU: European Union
FHIR: Fast Healthcare Interoperability Resource
GME: government-managed entity
GUI: graphical user interface
HL7: Health Level Seven International
HSE: health service entity
MoH: Ministry of Health
PE: personal entity
PoA: proof of authority
QR: quick response
RT-PCR: reverse transcription polymerase chain reaction
SC: smart contract
VCI: Vaccination Certificate Program
VP: vaccine passport

Edited by T Sanchez; submitted 27.07.21; peer-reviewed by A Khurshid, M Hölbl, J Liu; comments to author 25.08.21; revised version received 10.09.21; accepted 01.04.22; published 26.04.22.

Please cite as:
Lee HA, Wu WC, Kung HH, Udayasankaran JG, Wei YC, Kijsanayotin B, Marcelo AB, Hsu CY
Design of a Vaccine Passport Validation System Using Blockchain-based Architecture: Development Study
JMIR Public Health Surveill 2022;8(4):e32411
URL: https://publichealth.jmir.org/2022/4/e32411
doi:10.2196/32411
PMID:35377316

Original Paper

Health-Related Quality of Life of HIV-Positive and HIV-Negative Pregnant Women in an Impoverished Area: Cross-sectional Study

Shuiling Qu[1,2], PhD; Ailing Wang[2*], MM; Xiaoyan Wang[2], MM; Yehuan Yang[2], BM; Xiaoping Pan[2], PhD; Tong Zhang[3*], PhD

[1]Chinese Center for Disease Control and Prevention, Beijing, China

[2]National Center for Women and Children's Health Chinese Center for Disease Control and Prevention, Beijing, China

[3]Capital Institute of Pediatrics, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Tong Zhang, PhD
Capital Institute of Pediatrics
No 2 Yabao Road
Chaoyang District
Beijing, 100020
China
Phone: 86 10 85695188
Email: zt@chinawch.org.cn

Abstract

Background: Liangshan prefecture of Sichuan province was an impoverished mountainous area in China, where the annual number of HIV-positive pregnant women accounted for approximately 10% of China's total population in the decades before 2020. In general, pregnant women living here are likely to be physically and mentally different from those in other places.

Objective: This study aims to explore the health-related quality of life (HRQoL) of pregnant women living with HIV in an impoverished area.

Methods: From December 2018 to January 2019, HIV-positive and HIV-negative parturients within 18 months after delivery were recruited in Liangshan Prefecture, Sichuan Province. Questionnaires were designed to collect their demographic data, while the EuroQol 5-Dimension, 3-Level questionnaire was used to measure their HRQoL when they were in the second trimester from 4 to 6 months of pregnancy, and their quantitative health scores were converted to corresponding healthy utility values by using the Chinese Utility Value Integral System (time trade-off coefficient).

Results: A total of 250 pregnant women (133 HIV-positive and 117 HIV-negative) were enrolled in the study. Among them, 55 (41.35%) and 75 (64.10%) of HIV-positive and HIV-negative pregnant women self-reported full health (healthy state 11111), respectively. The median health utility value of the 250 pregnant women was 0.961 (IQR –0.046 to 0.961), and those of the HIV-positive and HIV-negative pregnant women were 0.875 (0.424-0.961) and 0.961 (IQR –0.046 to 0.961), respectively. We observed a significant difference only in the dimension of anxiety or depression between the two groups (P=.002) and no significant difference in the distribution of health utility indices between the two groups in terms of maternal age, education level, occupation, annual household income, prenatal care visits, family size, and medical insurance category. Multivariate ordinal logistic regression analysis showed that age (odds ratio [OR] 0.62, P<.05) and prenatal care visit (OR 0.29, P<.01) were independent risk factors for health status.

Conclusions: Most pregnant women self-reported satisfactory HRQoL in this impoverished mountainous area. HIV-negative pregnant women had an edge over HIV-positive pregnant women, and there were significant differences in anxiety or depression dimensions between the two groups.

(JMIR Public Health Surveill 2022;8(4):e29906) doi:10.2196/29906

KEYWORDS

health-related quality of life; EQ-5D-3L; HIV; impoverished area; public health; pregnant women; depression; anxiety

Introduction

Generally, pregnant women have a poorer health-related quality of life (HRQoL) owing to the impact of pregnancy on their physiological and mental health [1]. Apart from limited physical activities in early pregnancy, most pregnant women are likely to experience pregnancy syndromes such as nausea, vomiting, dizziness, depression, nervousness, and anxiety throughout their pregnancy [2]. Moreover, compared with HIV-negative pregnant women, HIV-positive pregnant women have also been reported to have poorer overall physique [3], higher susceptibility to depression [3], and poorer mental health [4]. As evidenced by previous research, educational background, number of children, pregnancy symptoms, and occupation are all key factors influencing the HRQoL of pregnant women [2,5]. Labor losses and indirect costs for HIV and HRQoL are strongly associated with severity [6]. HIV/AIDS is also associated with a significant economic burden for caregivers living with HIV [7]. Nowadays, with 1.3 million (range 970,000 to 1.6 million) HIV-positive pregnant women worldwide in 2020 [8], HIV continues to be a major global public health issue. Accordingly, improvement of patient care—by evaluating HRQoL in HIV-positive pregnant women—is of great importance to inform decision-making, resource allocation, and health policy formulation.

HRQoL could be measured by using not only generic scales such as the World Health Organization Quality of Life-100 (WHOHRQOL-100), the EuroQoL 5-Dimension (EQ-5D) questionnaire, and the 36-Item Short Form survey (SF-36), but also specific scales such as the Medical Outcomes Study HIV Health Survey (MOS-HIV). The EQ-5D is used as a multi-attribute utility instrument for measuring HRQoL, which comprises a short, cognitively undemanding questionnaire that takes only a few minutes to complete [9,10].

In China, the general rate of mother-to-child transmission (MTCT) of HIV decreased from 34.8% before 2005 to 3.6% in 2020 over the years [11]. In Liangshan Prefecture, the highest HIV-affected epidemic region in Sichuan Province, China, where the annual number of local HIV-positive pregnant women is estimated to be 10% of China's total in recent decades, the rate of MTCT of HIV was 4.3%. The per capita GDP of Liangshan Prefecture, an impoverished mountainous area, is far lower than that of Sichuan Province [12]. Previous studies on HIV-positive pregnant women in Liangshan focused on sexual behaviors [13], changing modes of HIV transmission [14], perceptions of social norms [15], HIV prevalence [16], etc. Hence, this cross-sectional study aimed to explore the HRQoL of HIV-positive pregnant women in the Liangshan Prefecture, Sichuan Province.

Methods

Study Design and Participants

From December 2018 to January 2019, in Liangshan Prefecture, Sichuan Province, HIV-positive and -negative parturients within 18 months after delivery were recruited to the clinic for a questionnaire survey one by one. The questionnaire contained questions about demographic data including age, marital status, education level, and annual household income. The HRQoL of pregnant women was evaluated using the EQ-5D-3L. The inclusion criteria for participant recruitment were being registered residents of Liangshan Prefecture, having given birth in Liangshan prefecture, and having no severe mental or neurological disease.

Previous studies have shown that most HIV-positive women in Liangshan Prefecture are from the Yi minority, with a lower education level, and speak in Yi language [17,18]. In this study, we sought local doctors to help participants interpret the questionnaire and translate it to us.

Since this was a cross-sectional study, the following formula was used to estimate the sample size:

As a part of study on cost-effectiveness analysis, the personal indirect costs were set as the SD at CNY 500 (US $78.56), the allowable error (δ) was set at CNY 100 (US $15.71), and (α=.05) was 1.96. The minimum required sample size of HIV-positive pregnant women was 96.

Assessment

The EQ-5D-3L questionnaire was composed of five dimensions: mobility (MO), self-care (SC), usual activities (UA), pain or discomfort (PD), and anxiety or depression (AD); each dimension was divided into three levels: no problems, some problems, and extreme problems. The participants were asked to indicate their healthy state by ticking the box next to the most appropriate statement for each of the 5 dimensions. The healthy state could be either converted into corresponding healthy utility value by using the Chinese Utility Value Integral System (time trade-off coefficient; Table 1) [19] or graded into 4 healthy state levels in accordance with the EQ-5D-3L [20]. The formula for healthy utility value is as follows: F=1 – C - MO- SC – UA – PD – AD - N3, where N3 equals 0.000 or 0.022 if any dimension is at level 3.

Steps for grading a healthy state are as follows: first, the distance between each healthy state and full health is calculated (11111; for example, the distance between 12321 and 11111 is equal to 1 + 2 + 3 + 2 + 1 – 1 – 1 – 1 – 1 – 1 = 4). The mild healthy state "2" indicates that the distance is between 1 and 4, no dimension is at level 3, and there are at most three dimensions at level 2. The severe healthy state "4" implies that the distance is between 7 and 9, no dimension is at level 1, and at least two dimensions are at level 3. The others are in a moderate healthy state "3."

Table 1. Chinese Utility Value Integral System (time trade-off coefficient) for the EuroQol 5-Dimension questionnaire.

Dimension	Mobility			Self-care			Usual activities			Pain or discomfort			Anxiety or depression			C	N3
Level	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3		
Coefficient	0	0.099	0.246	0	0.105	0.208	0	0.074	0.193	0	0.092	0.236	0	0.086	0.205	0.039	0.022/0.000

Statistical Analysis

The chi-square test or the Fisher exact test was used to test the difference in the proportion of education level, occupation, and health insurance category between HIV-positive and HIV-negative pregnant women. Moreover, the Fisher exact test was applied to test the differences between the two groups under various health dimensions and factors influencing HRQoL, respectively. Multivariate ordinal logistic regression analysis for categorical data analysis was used for analysis of risk factors for health status. R packages vcd and MASS were used for the chi-square test, the Fisher exact test, and multivariate ordinal logistic regression analysis. The acquired data were then compared using the 2-tailed *t* test, with the level of statistical significance determined at α=.05.

Ethical Procedures

This study has been approved by the Ethics Review Committee of the National Center for Women and Children's Health, Chinese Center for Disease Control and Prevention (FY2018-06) and that of the Chinese Center for Disease Control and Prevention (201922).

Results

Baseline Characteristics of the Recruited Subjects

In total, 250 pregnant women (133 HIV-positive and 117 HIV-negative) were recruited in this study (Table 2). The average age was 30.52 (range 18-53) years, with a significant difference between HIV-positive (mean 32.37, SD 5.62 years) and HIV-negative (mean 28.41, SD 7.07 years) pregnant women (t_{248}=4.85, P<.001).

Table 2. Baseline characteristics of pregnant women.

Characteristics	HIV-positive pregnant women, n (%)	HIV-negative pregnant women, n (%)	Chi square (*df*)	*P* value
Age (years)			7.709 (1)	.005
18-35	86 (64.66)	95 (81.20)		
35-53	47 (35.34)	22 (18.80)		
Education level			1.692 (1)	.19
Illiterate	114 (85.72)	92 (78.63)		
Primary or junior high school	19 (14.28)	25 (21.37)		
Occupation			0.032 (1)	.86
Farmer	118 (88.72)	102 (87.18)		
Migrant worker	15 (11.28)	15 (12.82)		
Annual household income (CNY)[a]			1.692 (1)	.19
<8000	70 (52.63)	51 (43.59)		
≥8000	63 (47.37)	66 (56.41)		
Prenatal care visits			3.797 (3)	.28
≤2	48 (36.09)	56 (47.86)		
3	34 (25.56)	25 (21.37)		
4	20 (15.04)	16 (13.68)		
≥5	31 (23.31)	20 (17.09)		
Family size (persons)			0.265 (2)	.88
3	18 (13.53)	16 (13.68)		
4-5	70 (52.63)	58 (49.57)		
≥6	45 (33.84)	43 (36.75)		
Medical insurance			0.005 (1)	.95
Insured	126 (94.74)	112 (95.73)		
Uninsured	7 (5.26)	5 (4.27)		

[a]1CNY=US $0.16.

Health Utility Measurements

The median health utility value was 0.961 (IQR –0.046 to 0.961), and those of the HIV-positive and HIV-negative pregnant women were 0.875 (IQR 0.424-0.961) and 0.961 (IQR –0.046 to 0.961), respectively. A significant difference was only found on the AD dimension between the two groups (*P*=.002) and not on other dimensions. Regarding SC, 98% of pregnant women had “no problems,” 5 (1.88%) had “some problems,” and 0 had “serious problems.” Approximately 90% and 85% of the pregnant women had “no problems” in the MO and UA dimensions, respectively, whereas for the PD dimension, 75% of the pregnant women had “no problems” and 24% of them had “some problems” (Table 3).

In total, 130 (52.00%) of pregnant women were in full health (healthy state, 11111), including 55 (41.35%) HIV-positive pregnant woman and 75 (64.10%) HIV-negative pregnant women. Only 1 HIV-negative pregnant woman transitioned to grade 4 owing to the presence of other serious diseases after further inquiry. Consequently, this subject was excluded from the subsequent survey of health grade comparison between groups, per the chi-square test. A significant difference in health status between HIV-positive and HIV-negative pregnant women was observed (χ^2_2=12.5, *P*=.002). Furthermore, HIV-positive pregnant women had a relatively poorer health status than HIV-negative pregnant women (Table 4).

Table 3. Health status of pregnant women under different EuroQol 5-Dimension dimensions.

Dimensions	HIV-positive pregnant women (n=133), n (%)	HIV-negative pregnant women (n=117), n (%)	Fisher exact test (*P* value)
Mobility			.45
1	119 (89.47)	107 (91.45)	
2	14 (10.53)	9 (7.69)	
3	0 (0)	1 (0.86)	
Self-care			<.001
1	130 (97.74)	115 (98.29)	
2	3 (2.26)	2 (1.71)	
Usual activities			.14
1	109 (81.95)	104 (88.89)	
2	23 (17.29)	11 (9.40)	
3	1 (0.75)	2 (1.71)	
Pain or discomfort			.53
1	98 (73.68)	91 (77.78)	
2	34 (25.56)	24 (20.51)	
3	1 (0.75)	2 (1.71)	
Anxiety or depression			.002
1	77 (57.89)	90 (76.92)	
2	52 (39.1)	22 (18.80)	
3	4 (3.01)	5 (4.28)	

Table 4. Grading of healthy states of pregnant women (N=250).

Grading of healthy states	HIV-positive pregnant women (n=133), n (%)	HIV-negative pregnant women (n=117), n (%)	Total, n
Full health	55 (41.35)	75 (64.10)	130
Mild health	45 (33.83)	22 (18.80)	67
Moderate health	33 (24.81)	19 (16.24)	52
Severe health	0 (0)	1 (0.86)	1

Analysis of Factors Influencing Health Utility

Based on the results of the chi-square or the Fisher exact test, no significant difference was observed in pregnant women's age, education level, occupation, annual household income, prenatal care visits, family size, and type of health insurance category on the distribution of health grade. However, a lower proportion (25.00%) of the uninsured than insured pregnant women was observed in grade 1 (Table 5).

Multivariate ordinal logistic regression analysis showed that age (odds ratio [OR] 0.62, $P<.05$) and prenatal care visits (OR 0.29, $P<.01$) were independent risk factors for health status (Table 6).

Table 5. Independent analysis of health status and relevant factors.

Characteristics	Health grade n (%)				Chi-square (*df*)	*P* value
	1	2	3	4[a]		
Age (years)					5.170 (2)	.75
20-35	102 (56.35)	65 (35.91)	14 (7.73)	0 (0)		
35-53	28 (40.58)	35 (50.72)	5 (7.25)	1 (1.45)		
Education level					N/A[b]	.90[c]
Illiterate	107 (51.94)	83 (40.29)	15 (7.28)	1 (0.49)		
Primary or junior high school	23 (52.27)	17 (38.64)	4 (9.09)	0 (0)		
Occupation					N/A	.57[c]
Farmer	113 (51.36)	90 (40.91)	16 (7.27)	1 (0.45)		
Migrant worker	17 (56.67)	10 (33.33)	3 (10.00)	0 (0)		
Annual household income (CNY)[d]					1.557 (2)	.46
<8000	62 (51.24)	52 (42.98)	7 (5.79)	0 (0)		
≥8000	68 (52.71)	48 (37.21)	12 (9.30)	1 (0.78)		
Prenatal care visits					N/A	.16[c]
≤2	61 (58.65)	36 (34.62)	6 (5.77)	1 (0.96)		
3	29 (49.15)	26 (44.07)	4 (6.78)	0 (0)		
4	20 (55.56)	15 (41.67)	1 (2.78)	0 (0)		
≥5	20 (39.22)	23 (45.10)	8 (15.69)	0 (0)		
Family size (person)					N/A	.88[c]
3	19 (55.88)	12 (35.29)	3 (8.82)	0 (0)		
4-5	69 (53.91)	50 (39.06)	9 (7.03)	0 (0)		
≥6	42 (47.73)	38 (43.18)	7 (7.95)	1 (1.14)		
Health insurance					N/A	.05[c]
Insured	127 (53.36)	91 (38.24)	19 (7.98)	1 (0.42)		
Uninsured	3 (25.00)	9 (75.00)	0 (0)	0 (0)		

[a]N/A: not applicable.

[b]*P* values obtained using the Fisher exact test.

[c]In view of the distribution and professional judgment, the sample of health grade 4 was deleted during statistical analysis.

[d]1CNY=US $0.16.

Table 6. Multivariate ordinal logistic regression analysis of risk factors for health status.

Factor	B (SE)	*t* test (df)	*P* value	OR (95% CI)
Intercept 1|2	1.16 (0.39)	2.98 (1)	<.01	N/A[a]
Intercept 2|3	3.63 (0.46)	7.86 (1)	<.001	N/A
Age	0.62 (0.28)	2.23 (248)	<.05	0.62 (0.16-1.08)
Prenatal care visit	0.29 (0.11)	2.61 (248)	<.01	0.29 (0.11-0.47)

Discussion

Principal Findings

Our results indicate that 41.35% of HIV-positive pregnant women and 64.10% HIV-negative pregnant women self-reported full health. The median health utility value of the 250 pregnant women was 0.961. A significant difference was only observed in the dimension of anxiety or depression between the two groups.

This serious ceiling effect [21] may be the negative impact of the sensitivity of health measurement, which differed from the discovery of a good health utility measurement for HIV-positive

pregnant women in Yunnan Province [22]. The difference may be due to the following reasons. First, 89.01% of households in Yunnan Province had an annual income of more than 10,000 CNY (US $1571.29), which was higher than that of the participants in this study [22]. The average annual food expenditure of a family and the annual household income in this study were CNY 4000 (US $628.52) and CNY 8000 (US $1257.03), which translated into an Engel coefficient of 57.14%. The Engel coefficient indicated that the people were in the stage of barely meeting daily needs in 2017 according to the Engel law. Second, participants in Yunnan Province had better education than participants in this study. Third, most of the local residents are very sturdy as they live in a mountainous area of Liangshan. If the HIV-positive pregnant women in this study were compared with HIV/AIDS cases, the HRQoL was superior to that of those without antiviral treatment (0.801) [23] and those with antiviral treatment (0.82) [24].

As reported herein, health utility value of HIV-negative pregnant women was higher than that of HIV-positive pregnant women. Based on the analysis on the five dimensions in the EQ-5D-3L, a significant difference was found only in the AD dimension between the HIV-positive and HIV-negative pregnant women, but not in the MO, SC, UA, and PD dimensions. Similarly, it was reported [25,26] that HIV-positive patients had a higher risk of developing depression, which suggested that HIV-positive pregnant women may have psychological burden, and psychological intervention could exert a significant positive effect. Theoretically, the status of health may be affected by the degree of education, occupation, annual household income, and other factors, but this study was not able to validate this point just yet. Nevertheless, HRQoL was still found to be significantly positively correlated with annual household income, degree of education, and occupation [27]. The difference may be explained by sampling. As described above, Liangshan Prefecture was an impoverished area, with a homogenous income source structure, education, and occupation. In this study, age, and prenatal care visits were independent risk factors for health status. Mobile health or SMS text messaging could be used for improving the quality of antenatal care [28,29].

Limitations

This study still has the following limitations. First, it was designed as a retrospective analysis, serving as a part of the health economics evaluation of the prevention of MTCT of HIV in Liangshan Prefecture. There may be a recall bias among the pregnant women when recalling their HRQoL during their second trimester of pregnancy. Studies on general HRQoL assessment with the EQ-5D-3L appear largely free of recall bias within follow-up visits of 2-12 months [30], and agreement of HRQoL determined using the EQ-5D-3L between conventional (1 week) and retrospective change (3 months later) is fair [31], indicating that the recall bias could be accepted. Second, there may be a considerable selection bias due to the low feasibility of a random sampling owing to the particularity of HIV-positive pregnant women themselves.

In addition, HIV-positive pregnant women surveyed in this study experienced more anxiety or depression during pregnancy than HIV-negative ones. Further communication revealed that the anxiety was generated from the fear of MTCT of HIV. Hence, there is a need for further emphasis, research, and intervention to improve the mental health of HIV-positive pregnant women.

Conclusions

An interactive voice response tool may be a choice for people living with HIV, and the higher usage of the tool showed greater improvements in quality of life [32]. Mobile phones were found both to be acceptable and feasible in the collection of maternal and child health data from women living with HIV in South Africa. Accordingly, publicity and education are necessary to achieve full awareness of the prevention of MTCT of HIV to increase the confidence of the population group.

Authors' Contributions

TZ and AW designed all aspects of the study, including the study protocol. SQ analyzed the data and drafted the manuscript. YY participated in field investigation for collecting costs data. XP revised the manuscript. XW participated in data management. TG and AW had the final responsibility for the decision to publish. All the authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Sut HK, Asci O, Topac N. Sleep Quality and Health-Related Quality of Life in Pregnancy. J Perinat Neonatal Nurs 2016;34(4):302-309. [doi: 10.1097/JPN.0000000000000181] [Medline: 27223099]
2. Davoud A, Abazari M. The Relationship between Quality of Life and Physical Activity, Worry, Depression, and Insomnia in Pregnant Women. Iran J Psychiatry 2020 Apr;15(2):159-168 [FREE Full text] [Medline: 32426012]
3. Aurpibul L, Tongprasert F, Wichasilp U, Tangmunkongvorakul A. Depressive Symptoms Associated with Low Quality of Life Among Pregnant and Postpartum Women Living with HIV in Chiang Mai, Thailand. Int J MCH AIDS 2020;9(3):421-429 [FREE Full text] [doi: 10.21106/ijma.375] [Medline: 33163291]
4. Bernatsky S, Souza R, de Jong K. Mental health in HIV-positive pregnant women: results from Angola. AIDS Care 2007 May;19(5):674-676. [doi: 10.1080/09540120601012705] [Medline: 17505929]

5. Oviedo-Caro MA, Bueno-Antequera J, Munguía-Izquierdo D. Explanatory factors and levels of health-related quality of life among healthy pregnant women at midpregnancy: A cross-sectional study of The PregnActive Project. J Adv Nurs 2018 Dec;74(12):2766-2776. [doi: 10.1111/jan.13787] [Medline: 29989189]
6. Lopez-Bastida J, Oliva-Moreno J, Perestelo-Perez L, Serrano-Aguilar P. The economic costs and health-related quality of life of people with HIV/AIDS in the Canary Islands, Spain. BMC Health Serv Res 2009 Mar 30;9:55 [FREE Full text] [doi: 10.1186/1472-6963-9-55] [Medline: 19331682]
7. Katana PV, Abubakar A, Nyongesa MK, Ssewanyana D, Mwangi P, Newton CR, et al. Economic burden and mental health of primary caregivers of perinatally HIV infected adolescents from Kilifi, Kenya. BMC Public Health 2020 Apr 16;20(1):504 [FREE Full text] [doi: 10.1186/s12889-020-8435-0] [Medline: 32299411]
8. Estimated percentage of pregnant women living with HIV who received antiretrovirals for preventing mother-to-child transmission. World Health Organization. URL: https://www.who.int/data/gho/data/indicators/indicator-details/GHO/estimated-percentage-of-pregnant-women-living-with-hiv-who-received-antiretrovirals-for-preventing-mother-to-child-transmission [accessed 2021-11-12]
9. Devlin NJ, Brooks R. EQ-5D and the EuroQol Group: Past, Present and Future. Appl Health Econ Health Policy 2017 Apr;15(2):127-137 [FREE Full text] [doi: 10.1007/s40258-017-0310-5] [Medline: 28194657]
10. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med 2001 Jul;33(5):337-343. [doi: 10.3109/07853890109002087] [Medline: 11491192]
11. Wang A, Song L. Protecting the beginning of Life Safeguarding Health of Maternal child -- Review prospect of the Prevention of mother-to-child transmission of HIV, syphilis Hepatitis B in China in the past 20 years. Chin J AIDS STD 2021;27(7):677-679 [FREE Full text] [doi: 10.13419/j.cnki.aids.2021.07.01]
12. Statistical Bureau of Sichuan, NBS Survey Office in Sichuan. Sichuan Statistical Yearbook 2020. Beijing: China Statistics Press; 2020.
13. Pei R, Ji-Ke C, Yu G, Yang Y, Nan L, Liao Q, et al. Sexual behaviors related to HIV infection in Yi women of childbearing age in rural areas of southwest China. AIDS Care 2020 Nov;32(11):1388-1392. [doi: 10.1080/09540121.2020.1784445] [Medline: 32583673]
14. Xiao C, Jike C, Liu D, Jia P, Xu X, Xiao L, et al. The changing modes of human immunodeficiency virus transmission and spatial variations among women in a minority prefecture in southwest China: An exploratory study. Medicine (Baltimore) 2020 Mar;99(6):e18776 [FREE Full text] [doi: 10.1097/MD.0000000000018776] [Medline: 32028390]
15. Yang S, Jike C, Pei R, Liu D, Yu G, Wang J, et al. Perceptions of social norms played an important role in the occurrence of casual sex among Yi minority residents in China: a population-based study. AIDS Care 2021 May 20:1-8. [doi: 10.1080/09540121.2021.1929814] [Medline: 34011235]
16. Yang S, Yang C, Liao Q, Zhai W, Yu G, Xiao L, et al. Analysis of HIV prevalence among pregnant women in Liangshan Prefecture, China, from 2009 to 2015. PLoS One 2017;12(9):e0183418 [FREE Full text] [doi: 10.1371/journal.pone.0183418] [Medline: 28880917]
17. Yu G, Yao YN, Yang SJ, Wang QX, Liao Q, Nan L, et al. HIV recent infection among women in Liangshan Yi autonomous prefecture during 2011-2014. Chin J AIDS STD 2018;24(1):70-73+83. [doi: 10.13419/j.cnki.aids.2018.01.19]
18. Liao Q, Gong YH, Wang QX, Bian SY, Yu G, Wang K, et al. Analysis of HIV/AIDS sentinel surveillance results of pregnant and parturient women in Sichuan Liangshan Yi Autonomous Prefecture during 2011 and 2015. Chin Gen Pract 2016;19(33):4115-4119. [doi: 10.3969/j.issn.1007-9572.2016.33.019]
19. Liu GG, Wu H, Li M, Gao C, Luo N. Chinese time trade-off values for EQ-5D health states. Value Health 2014 Jul;17(5):597-604 [FREE Full text] [doi: 10.1016/j.jval.2014.05.007] [Medline: 25128053]
20. Wu H. A study on health utility value measurement. Shenyang Pharmaceutical University. (In Chinese) 2012:1-167.
21. Whitehurst DGT, Bryan S. Another study showing that two preference-based measures of health-related quality of life (EQ-5D and SF-6D) are not interchangeable. But why should we expect them to be? Value Health 2011 Jun;14(4):531-538 [FREE Full text] [doi: 10.1016/j.jval.2010.09.002] [Medline: 21315635]
22. Wang X, Guo G, Zhou L, Zheng J, Liang X, Li Z, et al. Health-related quality of life in pregnant women living with HIV: a comparison of EQ-5D and SF-12. Health Qual Life Outcomes 2017 Aug 30;15(1):158 [FREE Full text] [doi: 10.1186/s12955-017-0731-8] [Medline: 28851384]
23. Castro R, De Boni RB, Luz PM, Velasque L, Lopes LV, Medina-Lara A, et al. Health-related quality of life assessment among people living with HIV in Rio de Janeiro, Brazil: a cross-sectional study. Qual Life Res 2019 Apr;28(4):1035-1045. [doi: 10.1007/s11136-018-2044-8] [Medline: 30415417]
24. Miners A, Phillips A, Kreif N, Rodger A, Speakman A, Fisher M, ASTRA (Antiretrovirals, Sexual TransmissionAttitudes) Study. Health-related quality-of-life of people with HIV in the era of combination antiretroviral treatment: a cross-sectional comparison with the general population. Lancet HIV 2014 Oct;1(1):e32-e40 [FREE Full text] [doi: 10.1016/S2352-3018(14)70018-9] [Medline: 26423814]
25. Do AN, Rosenberg ES, Sullivan PS, Beer L, Strine TW, Schulden JD, et al. Excess burden of depression among HIV-infected persons receiving medical care in the united states: data from the medical monitoring project and the behavioral risk factor surveillance system. PLoS One 2014;9(3):e92842 [FREE Full text] [doi: 10.1371/journal.pone.0092842] [Medline: 24663122]

26. Arseniou S, Arvaniti A, Samakouri M. HIV infection and depression. Psychiatry Clin Neurosci 2014 Mar;68(2):96-109 [FREE Full text] [doi: 10.1111/pcn.12097] [Medline: 24552630]
27. Mafirakureva N, Dzingirai B, Postma MJ, van Hulst M, Khoza S. Health-related quality of life in HIV/AIDS patients on antiretroviral therapy at a tertiary care facility in Zimbabwe. AIDS Care 2016 Jul;28(7):904-912. [doi: 10.1080/09540121.2016.1173639] [Medline: 27098289]
28. Benski AC, Schmidt NC, Viviano M, Stancanelli G, Soaroby A, Reich MR. Improving the Quality of Antenatal Care Using Mobile Health in Madagascar: Five-Year Cross-Sectional Study. JMIR Mhealth Uhealth 2020 Jul 08;8(7):e18543 [FREE Full text] [doi: 10.2196/18543] [Medline: 32673263]
29. Watterson JL, Castaneda D, Catalani C. Promoting Antenatal Care Attendance Through a Text Messaging Intervention in Samoa: Quasi-Experimental Study. JMIR Mhealth Uhealth 2020 Jun 02;8(6):e15890 [FREE Full text] [doi: 10.2196/15890] [Medline: 32484446]
30. McCann AC, Phillips KM, Trope M, Caradonna DS, Gray ST, Sedaghat AR. Characterising the potential for recall bias in anchor-based MCID calculation of patient-reported outcome measures for chronic rhinosinusitis. Clin Otolaryngol 2020 Sep;45(5):768-774. [doi: 10.1111/coa.13589] [Medline: 32449602]
31. Haagsma JA, Spronk I, de Jongh MAC, Bonsel GJ, Polinder S. Conventional and retrospective change in health-related quality of life of trauma patients: an explorative observational follow-up study. Health Qual Life Outcomes 2020 May 27;18(1):157 [FREE Full text] [doi: 10.1186/s12955-020-01404-1] [Medline: 32460896]
32. Byonanebye DM, Nabaggala MS, Naggirinya AB, Lamorde M, Oseku E, King R, et al. An Interactive Voice Response Software to Improve the Quality of Life of People Living With HIV in Uganda: Randomized Controlled Trial. JMIR Mhealth Uhealth 2021 Feb 11;9(2):e22229 [FREE Full text] [doi: 10.2196/22229] [Medline: 33570497]

Abbreviations

AD: anxiety or depression
EQ-5D-3L: EuroQol 5-Dimension, 3-Level
HRQoL: health-related quality of life
MO: mobility
MOS-HIV: Medical Outcomes Study HIV Health Survey
MTCT: mother-to-child transmission
PD: pain or discomfort
SC: self-care
SF-36: 36-Item Short Form survey
TTO: time trade-off
UA: usual activities
WHOHRQOL-100: World Health Organization Quality of Life-100

Edited by H Bradley; submitted 25.04.21; peer-reviewed by J Lei, P Akande, P Nguyen, AAS Sawitri; comments to author 30.07.21; revised version received 13.10.21; accepted 21.02.22; published 05.04.22.

Please cite as:
Qu S, Wang A, Wang X, Yang Y, Pan X, Zhang T
Health-Related Quality of Life of HIV-Positive and HIV-Negative Pregnant Women in an Impoverished Area: Cross-sectional Study
JMIR Public Health Surveill 2022;8(4):e29906
URL: https://publichealth.jmir.org/2022/4/e29906
doi:10.2196/29906
PMID:35380543

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org